UNIVERSIDADE
BEIRA INTERIOR

A Predictive Maintenance Model based on Multivariate Analysis with Artificial Intelligence

João Carlos Antunes Rodrigues

Tese para obtenção do Grau de Doutor em
Engenharia e Gestão Industrial
(3º Ciclo de Estudos)

Dezembro de 2023

A Predictive Maintenance Model based on Multivariate Analysis with Artificial Intelligence

João Carlos Antunes Rodrigues

Tese para obtenção do Grau de Doutor em

Engenharia e Gestão Industrial

(3º Ciclo de Estudos)

Orientador: Professor Doutor António João Marques Cardoso
Co-Orientador: Professor Doutor José Manuel Torres Farinha

Provas de Doutoramento: 8 de novembro de 2023

Júri:

PRESIDENTE:
Prof. Doutor Joaquim Mateus Paulo Serra

VOGAIS:
Prof. Doutor Diego Galar
Prof. Doutor Luís António de Andrade Ferreira
Prof. Doutor Inácio de Sousa Adelino da Fonseca
Prof. Doutor Jorge Manuel dos Reis Gama

Dezembro 2023

Declaração de Integridade

Eu, **João Carlos Antunes Rodrigues**, que abaixo assino, estudante com o número de inscrição D2737 do Doutoramento em Engenharia e Gestão Industrial da Faculdade de Engenharia, declaro ter desenvolvido o presente trabalho e elaborado o presente texto em total consonância com o **Código de Integridades da Universidade da Beira Interior**.

Mais concretamente afirmo não ter incorrido em qualquer das variedades de Fraude Académica, e que aqui declaro conhecer, que em particular atendi à exigida referenciação de frases, extratos, imagens e outras formas de trabalho intelectual, e assumindo assim na íntegra as responsabilidades da autoria.

Universidade da Beira Interior, Covilhã, 6 de dezembro de 2023.

Assinado por: **João Carlos Antunes Rodrigues**
Num. de Identificação: 14756494
Data: 2023.12.06 09:23:18 +0000

Dedication

I would like to thank my grandparents who, in their own way, were able to give me very enriching advice and goals that always guided me so that I could see that life has to be lived, but it has to be, above all, fought with justice, dedication and honesty.

To the following, my parents, I owe and am grateful for my birth, which, despite being late for them, they always knew with a lot of work and pride, to guide me in what they both thought was the best for their youngest son. They were always fighting so that I lacked nothing and that my goals at all levels were being fulfilled.

To my brothers, including my brother-in-law Rui, I thank them for their unconditional support, encouragement, love, friendship, and patience shown in the most difficult moments. To my sister, I must make a very special note. She was always the biggest driver for me to follow this course and achieve many of the goals, never letting me get discouraged or relax, making me continue and keep fighting and studying, focused on the goal in the face of adversity. To the rest of the family, who knew how to back up and support my choices and those of my closest family, making us feel supported in the most difficult moments.

I would like to thank Professor/Advisor, Doctor António João Marques Cardoso for all the guidance and confidence that was given to me throughout my PhD, it was undoubtedly a great pride to have him as my supervisor.

To Professor/Co-Advisor, Doctor José Torres Farinha, I thank him for the guidelines, the advice, his great selflessness in helping me and above all I praise his friendship and ever-present dedication, without him this project would not be possible.

To Professor, Doctor Mateus Mendes, I thank him for his dedication, patience and attention, allowing me to evolve and better understand what seemed impossible.

To Professor Doctor Ricardo Mateus, I thank all the support, teachings and trust given along this journey.

To Doctor Sylvie Santos Angelo, thank you for always being willing to help me with anything I needed.

I also leave a word of appreciation to my friends, to those I already had and to those I made, for the encouragement and great camaraderie provided on this journey.

Resumo

A manutenção e a gestão de ativos têm um papel preponderante no sucesso de qualquer indústria.

O desenvolvimento da sensorização e o aumento da capacidade de armazenamento e processamento de dados, aliados à Inteligência Artificial, vieram permitir uma melhoria significativa nas técnicas de manutenção e gestão do ciclo de vida dos ativos, contribuindo para uma maior disponibilidade e eficiência dos mesmos com menores custos de manutenção.

O projeto doutoral proposto descreve modelos de manutenção preditiva baseados em Inteligência Artificial.

A fiabilidade e o desempenho dos motores Diesel dependem bastante da qualidade e condição dos seus óleos lubrificantes. A presente tese descreve modelos para classificar automaticamente a condição do óleo, utilizando Redes Neuronais Artificiais e Análise de Componentes Principais. Os resultados dos modelos classificadores de lubrificantes foram comparados com as classificações de peritos humanos. A comparação mostra que os modelos de classificação desenvolvidos são credíveis.

A presente tese de doutoramento apresenta modelos de previsão de valores de sensores a curto, médio e longo prazo, ambos usando redes neuronais. O modelo de previsão de longo prazo é capaz de prever o valor de sensores até 90 dias de antecedência.

Usaram-se métodos de aprendizagem supervisionados e não supervisionados para criar modelos de classificação do estado de uma máquina industrial. O principal objetivo era determinar quando o ativo se encontrava no seu estado de funcionamento normal ou fora desta zona, correndo assim o risco de falha. Os resultados mostraram que é possível classificar e prever o estado de máquinas industriais utilizando redes neuronais artificiais.

Os modelos propostos apoiam a monitorização e manutenção de ativos, sendo que as principais implicações são a melhoria da disponibilidade operacional, incremento de qualidade, menor impacto ambiental, mais segurança e racionalização de custos.

Palavras-chave

Manutenção Preditiva, Disponibilidade, Ativos Físicos, Redes Neuronais Artificiais, Previsão, Análise de Componentes Principais, Inteligência Artificial, Análise de Dados, Previsão, Classificação

Resumo alargado

A manutenção de ativos físicos assume cada vez mais um papel preponderante no sucesso das empresas, sejam elas industriais ou de serviços. A racionalização de custos, aliada a uma rigorosa política de manutenção, conferem às empresas vantagens competitivas num mercado cada vez mais exigente.

Esta tese de doutoramento surgiu com o objetivo de resolver alguns problemas de Manutenção Preditiva. Esta investigação visa responder a lacunas identificadas no estado da arte, como a previsão baseada em sensores e classificação do estado de equipamentos a longo prazo.

O presente trabalho descreve contribuições inéditas para o estado da arte dos óleos lubrificantes. Os resultados mostram que é possível criar bons modelos utilizando Redes Neuronais Artificiais (RNA) para classificar óleos levando em consideração todas as variáveis. Os modelos podem até classificar lubrificantes com um erro bastante baixo. Por meio da Análise de Componentes Principais (ACP) foi determinada a relevância de cada variável para a análise do óleo, proporcionando assim uma melhor perceção da importância de cada parâmetro em análise. Os resultados também mostram que um modelo neuronal não precisa de usar todas as variáveis.

A Análise de Componentes Principais também permitiu a criação de um algoritmo que calcula a percentagem de degradação de um óleo lubrificante a partir das referências padrão do fabricante, de modo a que esse algoritmo funcione para qualquer lubrificante industrial.

Ressalta-se que ambos os classificadores de lubrificantes (ACP, RNA e Peritos Humanos) foram comparados entre si, tendo convergido aproximadamente em 90%, o que confirma que os modelos de classificação desenvolvidos são fiáveis.

Os algoritmos desenvolvidos podem auxiliar as indústrias, de uma forma geral, uma vez que fornecem informação de fácil interpretação e as auxiliam na tomada de decisão sobre o momento mais adequado para substituir o óleo nos ativos.

Por meio de uma pesquisa exaustiva do estado da arte na área de previsão industrial concluiu-se que, até ao momento, não existe um modelo publicado de previsão de falhas com um intervalo de tempo tão longo, o que demonstra a inovação e contribuição deste estudo para a ciência e para a competitividade da indústria. Refira-se que os algoritmos desenvolvidos já foram testados e aplicados, tendo apresentado erros de previsão em geral inferiores a 10%.

O projeto atual possui um modelo de previsão de curto prazo e outro modelo de previsão de longo prazo, ambos utilizando redes neuronais. O modelo de previsão de longo prazo é capaz de prever falhas de ativos com 90 dias de antecedência, o que permite que as indústrias façam paragens programadas dos seus ativos, evitando assim perdas decorrentes de paragens não programadas.

A criação de vetores de entrada em Redes Neuronais Artificiais utilizando janelas deslizantes ao longo de séries temporais melhorou muito o treino, chegando-se à conclusão que janelas sobrepostas permitem que a rede aprenda em menos iterações. Janelas maiores facilitam a captura de valores de pico, mas o tamanho ideal da janela precisa de ser determinado experimentalmente.

Em relação à previsão de curto prazo, foi demonstrado que a reamostragem de dados pode tornar o processo de previsão mais rápido, pois reduz consideravelmente o conjunto de dados de entrada na rede.

Também foi desenvolvido um algoritmo para classificar os valores previstos do sensor. Desta forma, o algoritmo poderá classificar os ativos no futuro em operação normal, alerta ou mau funcionamento.

Esta tese oferece bastantes contributos para a área industrial, especialmente nas áreas da manutenção, segurança, qualidade, sustentabilidade e eficiência, pois maximizará a disponibilidade de ativos, contribuindo para o sucesso dos Sistemas de Gestão da Qualidade e da Manutenção. Tais incrementos positivos em diversos setores terão como principal consequência a redução de custos, o que se tornará um diferencial competitivo para as indústrias, pois poderão aproximar-se do mercado com preços mais competitivos.

Importa referir que parte do trabalho foi publicado em artigos científicos e foi apresentado em diversos congressos, tendo recebido a distinção da melhor apresentação no congresso TEPEN 2021& IncoME-VI na China. Recebeu também o 2º Prémio Inovação Jovem Engenheiro- PIJE 2021, da Ordem dos Engenheiros.

Palavras-chave

Manutenção Preditiva, Disponibilidade, Ativos Físicos, Redes Neuronais Artificiais, Previsão, Análise de Componentes Principais, Inteligência Artificial, Análise de Dados, Previsão, Classificação

Abstract

The maintenance of physical assets is increasingly assuming a leading role in the success of companies, whether industrial or services. The pressure of budgets, combined with a strict maintenance policy, gives companies competitive advantages in an increasingly demanding market.

This PhD thesis emerged with the aim of solving some Predictive Maintenance problems. The research also aims to respond to gaps identified in the state of the art as the prediction and classification of the long-term state of equipment.

The present work describes novel contributions to the state of the art of lubricating oils. The results show that it is possible to create good models using Artificial Neural Networks (ANN) to classify oils considering all variables. Models can even possibly rank lubricants with a small error. Using Principal Component Analysis (PCA), the relevance of each variable for oil analysis was determined, thus providing a better insight into the importance of each parameter under analysis. The results also show that a neural model does not need to use all variables.

Principal Component Analysis also allowed the creation of an algorithm that calculates the percentage of degradation of a lubricating oil, from the manufacturer's standard references so, this algorithm works for any industrial lubricant.

It is noteworthy that lubricant classifiers (PCA, RNA and Human Experts) were compared with each other, having converged in more than 90%, which confirms the reliability of the classification.

The developed algorithms can support industries, in a general way, since they provide information that is easy to interpret, and helps them to make decisions about the most appropriate time to replace oil in the assets.

Through exhaustive research of the state of the art in prediction and industrial forecast, it was concluded that, to date, there is no published model for predicting failures with such a long-time span, which demonstrates the innovation and contribution of the present research for science and for the competitiveness of the industry. It should be noted that the developed algorithms have already been tested and applied, showing in general a prediction error below 10%.

The current project has a short-term prediction model and another long-term prediction model, both using neural networks. The long-term forecasting model can predict asset failures 90 days in advance, which allows industries to make scheduled stops on their assets, thus avoiding losses resulting from unscheduled stops.

Adequate feature input vectors in Artificial Neural Networks using sliding windows along time series greatly improved the training, leading to the conclusion that overlapping windows allow the network to learn in less iterations. Larger windows make it easier to capture peak values, but the optimal window size needs to be determined experimentally.

Regarding short-term forecasting, it has been shown that data resampling can make the forecasting process faster, as it considerably reduces the input data set in the network.

An algorithm was also developed to determine the expected equipment state through classification of the predicted sensor values. This way, the algorithm will be able to classify the probable state of the assets in the future in normal operation, alert or malfunction.

This PhD thesis is very important for the industrial area, especially in the areas of maintenance, safety, quality, sustainability and efficiency, as it will maximize the availability of assets, contributing to the success of the Quality Management Systems and Maintenance Management Systems. Such positive increments in several sectors will have as main consequence the reduction of costs, increase of equipment availability and improvement of quality, what will become a competitive differentiator for the industries, because they will be able to approach the market with more competitive prices and quality.

Part of the work was published in scientific articles and presented at several congresses and received the distinction of best presentation award in TEPEN 2021 & IncoME-VI congress in China. It also received the 2nd Young Engineer Innovation Award- PIJE 2021, from Ordem dos Engenheiros, Portugal.

Keywords

Predictive Maintenance, Availability, Physical Assets, Artificial Neural Networks, Forecasting, Principal Component Analysis, Artificial Intelligence, Data Analysis, Prediction, Classification

Contents

List of figures

List of tables

Acronyms and abbreviations

ANN	Artificial Neural Network
ARIMA	Auto Regressive Integrated Moving Average
AVG	Average
CA	Critical Analysis
FMEA	Failure Mode and Effect Analysis
FMECA	Failure Mode, Effects and Criticality Analysis
FF	Feedforward
FFNN	Feedforward Neural Network
IIoT	Industrial Internet of Things
IoT	Internet of Things
ITER	Iterations
MAPE	Mean Absolute Percentage Error
MLP	Multilayer Perceptron
MSE	Mean Square Error
NN	Neuronal Networks
PC	Principal Component
PCA	Principal Component Analysis
PCI	Principal Component Index
RF	Random Forest
RNN	Recurrent Neural Network
TAN	Total Acid Number
TPR	True Positive Rate
UBI	Universidade da Beira Interior
WS	Window Size

Chapter 1

Introduction

Advanced detection technology, combined with high-performance computing, helps industries operate reliably. Industries are striving to constantly improve industrial processes and equipment. Maintenance plays a key role in this field, with its main objective being to prevent physical asset failure.

1.1 Background

Detection techniques are increasingly cheaper, more accurate, and less invasive. The recent evolution of detection technology opens new ways to train prediction algorithms with a higher degree of reliability. Some algorithms require large amounts of data and long processing times to generate predictions with low error. Therefore, it is necessary to have a reliable data history. For this to happen, it is essential to have correctly calibrated sensors, installed and connected to a functional data system [1],[2].

Data processing capacity constantly evolves, enabling industries to improve many of their production, maintenance, and even logistics processes. The new industrial generation requires changes in its processes in virtually all areas, including monitoring and forecasting of production [3], quality control, or maintenance based on operating conditions [4].

The industrial maintenance sector is in a phase of reorganization, exploration, and research, where Artificial Intelligence (AI) is increasingly used. The main advantages of using AI are the reduction of maintenance costs and the increase of the asset's availability, due to the use of intelligent machine learning algorithms to solve complex problems.

Industrial processes must adapt as technology changes. That is currently the case with Industry 4.0, which may require process changes in all areas, including tracking products [5], monitoring and prediction data [3], quality control [6], or condition-based maintenance [7], among other uses of sensor networks and algorithms.

Machine learning can be useful in managing and controlling quality, reducing maintenance costs, and improving the overall manufacturing process. It can make all the difference in modern industries.

Maintenance is a combination of technical and administrative activities required to maintain equipment, facilities, and other physical assets. The goal is to maintain those assets in the desired operational condition or restore them, so that they can fulfil their function with quality [8],[9],[10]. The main objectives of a good maintenance policy are safety, quality, cost reduction, and availability [11]. The optimization of those four objectives at the same time is challenging since they often conflict with each other. In those cases, it is the maintenance management's responsibility to find the best compromise solution based on the company's strategic objectives.

Predictive maintenance is one of the fastest growing types of maintenance in the industry nowadays [12]. Its aim is to predict the occurrence of faults before they happen, using algorithms and sensor data. The algorithms are trained based on historical data, the operating condition of the assets is monitored, and the trends are predicted in near real time.

Industrial systems currently use tens, hundreds, or thousands of sensors to collect data to be used primarily to monitor processes and equipment condition [1], [13].

Due to developments in data processing, along with storage algorithms and hardware, it is currently possible to store and process large quantities of data to predict the future behaviour of equipment, thus making it possible to forecast failures in advance [14].

The asset's behaviour, after being observed and analysed, can be predicted with state-of-the-art algorithms [15]. It should also be noted that predictive maintenance promotes environmental sustainability, as it contributes to reduce industrial downtimes, unnecessary maintenance interventions, production surpluses, and non-conforming products.

1.2 Problem description

This dissertation is entitled " A Predictive Maintenance Model based on Multivariate Analysis with Artificial Intelligence ", because this is the area of focus in all the tasks carried out during this doctoral program.

The first phase of the project consisted of a study on the degradation of the lubricating oils of two bus companies, in order to identify the best maintenance policy adopted by the companies concerned.

When analysing the state of the art of this problem, a gap was identified in the classifications of lubricants, since at that time there were few works that had their results validated and compared by human experts. After the identification of this gap, progress was made to create a calculation formula for the degradation of lubricating oils from Diesel engines, and the results of this formula would have to be compared with neural networks and human experts.

At the conclusion of this first task, it became clear how important lubrication is to a physical asset. Therefore, it was verified that a more in-depth study of parameters in industrial machines was necessary, as well as a new challenge that would give the opportunity to study and, perhaps solve a chronic fault in a chip pump, which had a life cycle much shorter than expected.

In addition to the resolution of the chip pump problem, the main objective is to anticipate the behaviour of the chip pump in the long term, to alert the operators when a risk is anticipated.

Afterwards, it was intended to do something similar, but with a completely different machine. This time the problem was to predict the short, medium, and long-term behaviour of a pulp paper press, also using shallow feedforward networks.

The last problem proposed was to automatically classify the state of the pulp paper press, using the prediction data to classify the predicted future state into one of three categories: normal operation, alert or malfunction.

1.3 Research motivation

Research in the area of maintenance assumes an increasing importance in the industrial area, as proven by the growing number of articles published in recent years, in the area of predictive maintenance and condition monitoring. Some of those works are referred in Chapter 3.

The main motivation comes from the potential of these studies in the areas of maintenance, safety, quality, sustainability, and efficiency.

The market is increasingly competitive and global so, one of the major objectives of companies and industries nowadays is to reduce their costs to the minimum. Knowing that the management and maintenance of assets assumes a large part of the costs of a company, it is a priority to try to reduce this cost as much as possible, offering the maximum performance and availability of equipment. Therefore, the impact of the implementation of these projects in the industries in terms of cost reduction was one of the motivations for carrying out this thesis, because this way companies can approach the market with more competitive prices and less quality issues.

The development of technology, processing capacity and information systems offered industries in general and the maintenance department in particular the opportunity to improve their maintenance policies. Nowadays, predictive maintenance is one of the most sought types of maintenance in the industry.

One of the main objectives of predictive maintenance is to determine, in advance, the need for maintenance services. However, according to the state of the art, these services or failures were foreseen in the short to medium-term. Therefore, performing long-term prediction and classification was a great motivation to carry out this doctoral project.

It is important to highlight that only the study described in Chapter 5 used Diesel Bus Engines. All the other studies and models presented in this doctoral thesis are limited to a particular paper industry. Hence, some problems of this industry were used as case studies and research. This way, the studies presented here will have a practical industrial application.

1.4 Objectives and research questions

This research work aims to answer six research questions, described below. The first research question is answered in Chapter 5, the second and third research questions are discussed in Chapter 6. The fourth question is answered in Chapter 7 and, finally, Chapter 8 provides answers to the fifth and sixth research questions.

Lubrication plays a fundamental role in the life cycle of a Diesel engine and its reliability. Due to the importance that a lubricating oil assumes in the life cycle of an engine and its reliability, the first research question of this doctoral thesis arises.

- **Research question 1:** Using mathematical models and based on analyses of lubricating oils, manufacturer reference values and human expert ratings, is it possible to create a calculation formula to determine the percentage degradation of a lubricating oil in a Diesel engine?

The remaining research questions are limited to the paper industry.

Chip pumps are indispensable equipment in the production of pulp, as they are mainly responsible for transporting the wood chips. Because their availability must be the maximum possible, when studying the life cycle of the system, it was noticed that the shaft of the third chip pump was cracking much earlier than expected. Hence the second research question.

- **Research question 2:** Which causes are making the chip pump shaft fail so frequently? Which methods can be used to assess all possible causes? How can these systematic failures be avoided in the future?

After identifying the causes that led to the systematic cracking of the chip pump shaft, it is imperative to know and predict its behaviour, thus the third research question arises.

- **Research question 3:** What models can be created to predict the short and long-term operating conditions of a chip pump using shallow networks, more specifically the multilayer perceptron (MLP)?

It is known that pulp paper presses must work well, so that the pulp has the ideal quality. Then, there is the need to develop algorithms that help them to have the maximum availability, and thus appears the research question 4.

- **Research question 4:** What algorithms can be developed to predict the behaviour of a pulp paper press in the short, medium and long-term? Can these algorithms be developed using XGBoost or feedforward Neural Networks? Does data overlap influence learning process and the prediction results?

Through the prediction of the sensed values, the long-term behaviour of the assets can be known. However, it is necessary to classify the future state of the equipment in order to help the maintenance department in the decision making, so that the need for the fifth research question arises.

- **Research Question 5:** Can the state of a pulp press be classified using feedforward neural networks? What is the number of states indicated to classify the state of the press?

Due to the great correlation between the state of an equipment and its lubrication quality, it is imperative to know the state of the pulp paper press lubricating oil, to decide if the lubricants are in good condition or need to be replaced. Thus, the last research question of this thesis arises.

- **Research question 6:** Can lubricating oil from a cellulosic paper press be classified into two or more classes using feedforward neural networks? Do the classification results converge with the results of human experts?

In order to answer the questions mentioned above, the following six main objectives were defined:

- **Objective 1:** Create a method that determines the percentage of degradation state of a Diesel engine oil using PCA and based on analysis of lubricating oils, manufacturer references and human expert ratings.
- **Objective 2:** Identify the causes of systematic cracking of the chip pump shaft by analysing the system and using Ishikawa diagram and FMECA analysis.
- **Objective 3:** Develop algorithms to predict the short and long-term sensor values of a chip pump and create a warning system for the risk of equipment failure. The MAPE of all the predicted variables must be less than 10%.
- **Objective 4:** Develop algorithms to predict the short, medium and long-term performance of a pulp press. In general, the MAPE of the predicted variables should be less than 10%. Analyse the effect of data overlap on the feature vectors for the machine learning models. Compare the results of the evolved algorithms with the results of recurrent neural networks (RNN).
- **Objective 5:** Develop an algorithm to classify the state of a pulp paper press. Estimate the optimal number of press operating states using a clustering method.
- **Objective 6:** Develop an algorithm using feedforward neural networks to classify the state of a lubricating oil of a pulp press into two or more classes. The results of the classifier network should converge more than 95% with human expert classifications.

1.5 Thesis structure

This doctoral thesis is divided into 10 chapters plus appendices. The first chapter makes an introductory approach to the thesis, presenting a series of relevant information about the contents covered throughout the document.

Maintenance is the focus of the second chapter. This part of the thesis addresses topics such as the history of maintenance, its evolution, and some theoretical definitions of it.

The third chapter presents a robust state of the art, containing references to several research works already developed and published, involving neural networks for prediction and classification. Some works on oil condition assessment for Diesel engines and condition monitoring of engine oil are also presented.

Data Analysis and Forecast Methods is the subject of the fourth chapter. Here information about most of the methods and techniques used in the developed works is exposed.

The fifth chapter describes the first part of the practical investigation. There is exposed the study that was made to the lubricating oils of Diesel engines until reaching the final degradation formula. The chapter ends with a comparison between the results of the formula and the classifications made by artificial neural networks and human experts.

The study on chip pumps is described in Chapter 6. In this chapter the methodology to reach the detection of the chronic failure of the chip pump shaft is presented. This chapter ends with a short and long-term prediction of the sensor values of the chip pump, using shallow networks. These values will later be used for a failure alert system.

Chapter 7 is about the study of the behaviour of a pulp paper press. Different prediction algorithms were developed, analysed, and compared in the short, medium, and long-term. In this chapter, the effect of sliding windows with and without overlap in the final forecast result is also studied.

The classification of the status of the pulp paper press is the subject of the eighth chapter. It is in this part of the document that the entire process to classify the status of the physical asset in the future is presented.

Chapter 9 presents the main contributions of the studies carried out throughout this doctorate, scientific publications carried out and awards won in the same period.

The conclusions of these studies are presented in chapter 10. This chapter ends with some ideas for future developments. Afterwards, there are bibliographic references, followed by appendices.

Chapter 2

Maintenance

This chapter focus maintenance. The goal is to describe the various types of maintenance and expose some theoretical definitions in this area. A bit of the history and evolution over the years are also presented.

2.1 Types of maintenance

European Standard 13306:2017 defines the generic terms used for all types of maintenance and organization of the maintenance, regardless of asset type considered. It is in accordance with standard 13306:2017 that the following definitions and Figure 1 are presented [8].

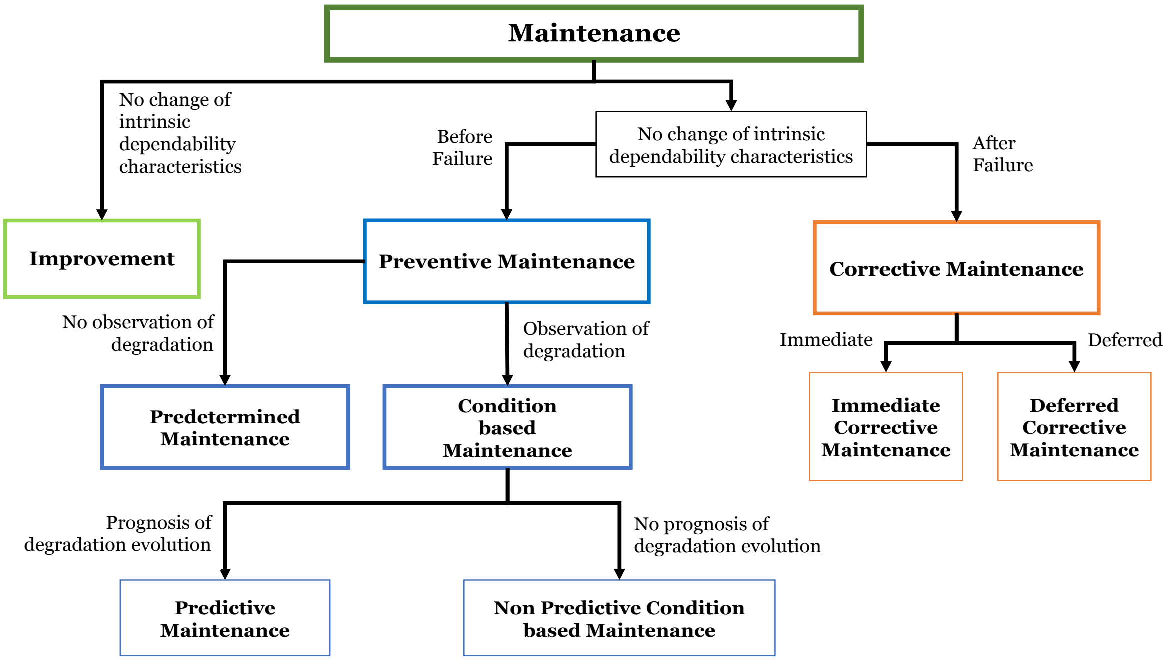

Figure 1 - Maintenance's overall views.

It is important to select the most appropriate type of maintenance for each equipment, because it is well known that good maintenance can have a significant impact on a company's performance. The three main categories of maintenance are, in order, as follows:

- Improvement;
- Preventive Maintenance;
- Corrective Maintenance.

2.1.1 Improvement

Improvement is a combination of all technical, administrative, and managerial actions, intended to ameliorate the intrinsic reliability and/or maintainability and/or safety of an item, without changing the original function.

An improvement may also be introduced to prevent misuse in operation and to avoid failures.

2.1.2 Preventive maintenance

Preventive Maintenance is carried out intended to assess and/or to mitigate degradation and reduce the probability of failure of an item.

Predetermined maintenance

Predetermined maintenance is Preventive maintenance carried out in accordance with established intervals of time or number of units of use but without previous condition investigation. Intervals of times or number of units of use may be established from knowledge of the failure mechanisms of the item.

Condition based maintenance

Condition based maintenance is preventive maintenance that includes evaluation of the physical conditions, analysis, and any potential follow-up maintenance procedures.

The condition assessment could involve operator observation, inspection, testing, condition monitoring of system parameters, etc., all of which could be done on demand or continuously, or according to a schedule.

Predictive maintenance

Predictive maintenance is condition-based maintenance carried out following a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item.

2.1.3 Corrective maintenance

Corrective maintenance is maintenance carried out after fault recognition and intended to restore an item into a state in which it can perform a required function.

Deferred

Deferred maintenance consists of asset and infrastructure repairs that are postponed or rescheduled for the future due to limited resources. These are maintenance tasks that have to be completed eventually, either to avoid safety problems, breakdowns or damage.

Immediate

Immediate maintenance is corrective maintenance that is performed right away after a problem is found to prevent unfavourable outcomes.

2.2 The history of maintenance

The term "maintenance" appeared for the first time in the 16th century in Central Europe. Then, it gained popularity with the advent of the mechanical clock and the emergence of the first technicians engaged in assembly and assistance.

Although various authors assert that maintenance has gone through various stages over time; in general, their perspectives converge to state that maintenance has gone through three generations and is now beginning its fourth, in which monitoring plays a crucial role. The four generations are:

- i First Generation - Mechanization;
- ii Second Generation - Industrialization;
- iii Third Generation - Automation;
- iv Fourth Generation - Condition monitoring.

The meaning of the word "maintenance" as it first appeared in the military lexicon was "maintaining the material and the combat units at a constant and operational level."

The first generation of maintenance continued up until the middle of the 20th century. Up until that point, the tools were straightforward and frequently oversized for the job. Since this time period, economy placed a strong emphasis on productivity and efficiency. Systematic maintenance was not required. Instead, corrective maintenance was used primarily on equipment [16].

The attitude toward failures and equipment malfunction during this time was that "all the equipment suffered so, having failures or breaks was normal." The only preventive maintenance procedures used in the first generation were cleaning and lubricating. As a result, it can be said that corrective maintenance characterises the first generation. In the 1920s, businesses started to recognise the value that maintenance could have on their machinery [17].

In order to increase production after the First World War, industries formed specialised teams to repair the damage as quickly as possible [18].

Industry needed to produce more than ever during the Second World War so, these businesses focused on avoiding the appearance of equipment damage in addition to wanting to fix flaws as quickly as possible. The development of maintenance engineering was necessary for the development of commercial aviation in the 1940s, largely due to the safety of goods and, ultimately, the safety of people, which necessitates a more frequent use of preventive maintenance.

The second generation of maintenance begun when industries started implementing production lines. That happened for the first time in the 1950s and industries realised how crucial it was to be autonomous and automated because of the potential effects that damage could have on the price of the finished product. Due to the high demand for production in the 1960s of the 20th century, businesses were forced to work two or three shifts instead of the usual one, which relieved them of the need to fix problems at night or on weekends. Consequently, it was crucial to prevent and anticipate industrial equipment malfunctions. The evolution of computing, which enabled the application of mathematical and statistical processes which were more reliable than was previously possible, was related to this need for production.

Microelectronics also made significant progress at the same time, enabling the accurate construction of high-quality measuring devices, and reducing measurement errors. With the development of more precise measuring tools and reliable mathematical techniques, it became possible to anticipate future failures, giving rise to predictive maintenance or condition maintenance and the third generation of maintenance.

According to Husband, a new, more comprehensive maintenance concept, known as terotechnology, first emerged in Europe in the 1970s [19]. This concept combines sound maintenance, management, financial, and engineering practises with the constant goal of minimising the cost of the economic cycle of physical assets. This novel idea in maintenance management aids businesses in choosing the most effective ways to repair, prevent, or replace physical assets.

2.3 The economic point of view of maintenance

Maintenance was often seen as a source of unnecessary cost by industry so, it was often overlooked by companies. Nowadays, the role of maintenance is better understood. It is considered a key factor for the success of companies, helping them to reduce production costs and, consequently, increase profits [20].

Although applying predictive maintenance policies may involve significant costs, these are often less than the benefits generated from a well-planned system [21].

Most industrial machines involve an expensive hardware network, formed by many sensors for data collection and storage. In addition to hardware, predictive maintenance requires additional costs for training staff, as well as analysing data and developing and training prediction and classification methods.

By enabling more efficient, sustainable, and higher quality production, the application of predictive maintenance also affects the company's image in the market and contributes to increase its value.

Predictive maintenance can be applied to almost all industrial equipment. However, due to its high implementation costs, technical and economic analyses must be performed before proceeding to modelling and deployment, namely determining the criticality of the equipment in case of failure or anomaly, and the potential economic losses for the company.

According to François Monchy, the more expensive the unavailability of an equipment, the more important its maintenance [22]. In other words, direct and indirect costs of equipment unavailability along with the value generated by the equipment are the most important factors to consider when choosing a maintenance policy.

The greatest advantage of predictive maintenance is that it can assess the current condition of any machine and predict when it needs maintenance before a fault happens. With a properly implemented and updated maintenance policy, it is possible to schedule equipment maintenance for times that will have the least impact in production schedule and deadlines, minimizing disruptions in production lines and improving the quality of the items produced by the factory, contributing to the profitability and sustainability of any company's business.

2.4 Diagnosis and predictive maintenance

Based on the identification of the physical asset's weakest parts, predictive maintenance seeks to maximize the system's availability [23].

A failure, which may be full or partial, is defined by the European Standard EN 13306:2017 as the loss of an item's capacity to carry out a necessary function following failure [8].

Currently, predictive maintenance makes extensive use of hardware for data collection and storage, as well as software for analysis. Farinha provides a summary of the topic [24]. Predictive maintenance's goal is to make it possible to schedule corrective work in advance, preventing unforeseen equipment failures.

Due to the significant trend in simulation-based optimization, maintenance optimization is a top priority [25]. To reduce overall maintenance costs or to increase asset productivity and availability, the best maintenance strategies are currently relentlessly sought after [26],[27],[28]. The necessity of enhancing this sector is emphasized by the fact that maintenance expenses might approach 50% of production costs [29],[30].

Since visual inspection was the initial approach of predictive maintenance, it has evolved. The development of sensors and computing power has led to the usage of a number of sophisticated signal processing approaches based on algorithms for pattern recognition, classification, grouping, and prediction [29].

It is possible to identify a variety of failure modes, causes, effects, and asset criticality using the FMECA reliability theory process [30].

The fundamental goal of predictive maintenance is to prevent the same failures by foreseeing them after identifying all potential failures using the Ishikawa Diagram and, subsequent FMECA analysis. These diagnostic tools are explained in Chapter 4.

2.5 Industrial maintenance in industry 4.0

The Internet of Things is increasingly present in industries, proof of this are the number of papers published using the IoT [31],[32].

The IoT is becoming more prevalent in the market, as hardware costs fall and computational power increases [33],[34]. Making processes more predictable, straightforward, controllable, and effective, is crucial for minimizing the costs associated with equipment manufacturing and maintenance [33],[35].

The technological revolution has led to Industry 4.0, and predictive maintenance is part of it [36],[37]. It is constantly required to make decisions on people and equipment in a market that is global and very competitive. Such predictive maintenance decisions typically rely on vast volumes of data [38],[39]. One of the key issues in this discipline is predicting with minimal error the need to carry out maintenance operations on the assets at a specific future point in the medium and long-term [38].

Hashemian categorizes condition-based maintenance methods for machinery and industrial processes into three groups [39].

2.6 Data analysis in maintenance

Industry 4.0 is a consequence of scientific and technological advances, including predictive maintenance.

The amount of data extracted from industrial processes has exponentially increased due to the rise of non-invasive sensing technologies and decreasing hardware costs. However, it is essential to calibrate the sensors correctly, so that the data acquired are reliable [40],[41],[42]. Poor or incorrect data do not add value and can lead to prediction errors [43],[44].

Analysis of reliable data with predictive computational techniques can avoid unnecessary equipment changes, save costs, and improve safety, availability, and efficiency of processes [39].

2.7 The 5 maintenance's levels

There are 5 levels of maintenance [43], which are the following:

1st Level - Simple adjustments are required instead of disassembling or replacing the equipment. The equipment operator does this.

2nd Level - The less complex interventions that fall under level 2 maintenance have easy-to-follow procedures. Furthermore, during these operations, the equipment in question does not need to be completely disassembled in order to replace parts. The task is carried out by a technician or the equipment operator.

3rd Level - Diagnosis, identification of faults, simple mechanical fixes, or replacement of damaged or non-functional parts. Task carried out by a maintenance crew or specialised technician.

4th Level - Heavy or complex mechanical repairs, intricate preventive maintenance schedules, or complex corrective maintenance tasks. Tasks carried out by a maintenance crew.

5th Level - Equipment renewal tasks or repairs with a very high degree of complexity or importance in the company. Task performed by a highly skilled and versatile maintenance crew.

2.8 Bathtub curve

The "bathtub curve" is a graph that is used to better understand the failures that develop in a specific piece of equipment over the course of its lifespan. This graph, which illustrates how the reliability concept is applied in the creation of the "bathtub curve," is better understood with the aid of Figure 2.

The probability that a device or system component will operate within the quality parameters specified during a specific time period, under pre-established operating conditions, is known as reliability [18].

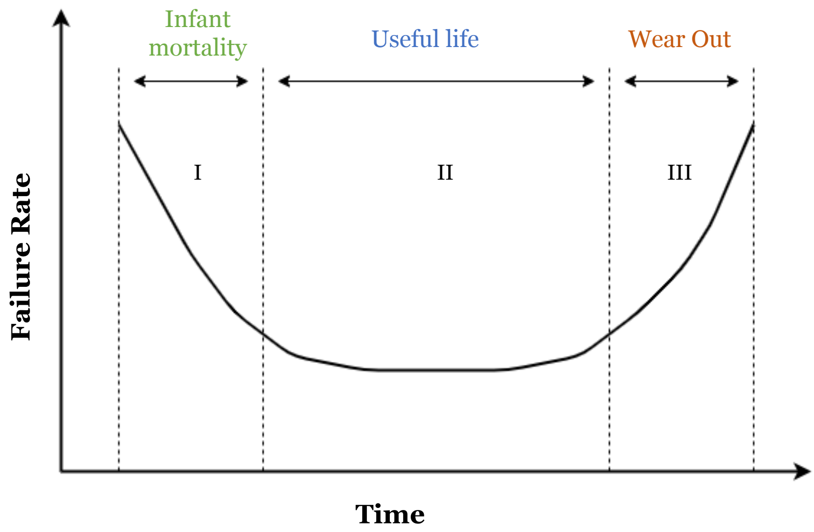

Figure 2 - A typical Bathtub Curve with a Time-dependent Failure Rate.

The curve illustrates the notion that there are three distinct periods that make up a population of devices' operation [44].

In the first phase of the curve there is an initially very high failure rate. However, it tends to decrease over time. This is justified due to the various mishaps caused by defects or errors that exist in the equipment in its manufacture and installation. This first phase is called infant mortality. The type of maintenance that must be performed during this life span of the equipment must undergo corrective maintenance.

In the period of useful life or maturity, there is a low rate of malfunctions that remains constant. The malfunctions are usually the result of accidents, failures in the use of the equipment or inadequate operations. Sellitto advises that predictive maintenance, during this period of the equipment's life, helps to detect the beginning of the wear phase. In the last stage of the bathtub curve, there is an increase in the rate of damage to the equipment, largely due to the wear and tear of the equipment. Therefore, Sellitto in the same paper proposes preventive maintenance [45].

2.9 Failures and faults

According to European Standard BS EN 13306:2017, a failure is the loss of the ability of an item to perform a required function after failure of the item, which may be complete or partial. It should be noted that the concept as it is defined does not apply to items consisting of software [8].

The way an item is unable to carry out a necessary function is known as the failure mode, but it can also be defined by the function that was lost or the state change that took place.

The events that lead to failure during specification, design, manufacture, installation, use, or maintenance are referred to as the failure cause.

Failure is an occurrence, while fault is a state, with fault being the inability to perform a necessary function, except for when performing planned maintenance or other planned actions, or when lacking outside resources.

Since wear out is a physical phenomenon that causes material loss, deformation, or change, the likelihood that an item will fail due to wear out rises, as it is used more frequently or for longer periods of time and as associated stresses are applied.

Aging is a physical phenomenon that modifies the material's physical and/or chemical properties. As a result, aging failures are those whose likelihood of occurrence increases over time, independent from operating time.

The cause of misuse failure is the application of stresses during use that go beyond the item's stated capabilities and the design specifications.

A numerical index of a failure's or fault's severity along with the likelihood or regularity of its occurrence serves as the definition of criticality of failure. The severity of a failure may depend on factors such as environment, cost, quality, availability, and safety.

2.10 Summary

The industry is undergoing major changes, triggered by the rapid advancement of digital technologies and a reduction in associated technology costs. As a consequence, many research efforts are focused on discovering and improving the industries' performance.

Maintenance is one of the areas that can most improve the performance of an industry, in economic, operational and productive terms. Therefore, there is an effort of the industrial maintenance departments to keep up with this change, which is only possible by investing in new technologies, sensing and investing in new skills.

Predictive maintenance, according to several authors, works by routinely checking the condition of components rather than replacing them, which results in better data, increased production output, and fewer catastrophic failures. In industrial plants, predictive maintenance can raise safety, quality, and availability.

Due to all these advantages, predictive maintenance, when implemented properly, can be a giant step towards the success of an industry.

Chapter 3

Related Work

3.1 Oil condition assessment for Diesel engines

Industrial lubricants are essential for maintaining the quality and extending the life of machinery and equipment. Lubrication reduces damage from high temperatures, vibrations, corrosion and friction when used correctly and with lubricants within recommended specifications.

Oil condition monitoring for Diesel engines has been studied using a variety of techniques, including machine learning techniques due to the previously mentioned importance.

A study on condition monitoring based on oil in the Diesel engines of a fleet of urban buses is presented by Raposo *et al.* [46] The study develops a predictive maintenance strategy for oil replacement and charts the progression of oil degradation. The methodology presented by the authors considers only some variables of the oils, showing very interesting results about following the P-F curve. The P-F curve is a graphical representation of the interval between the detection of a Potential failure and the actual Failure — that is, the interval where a maintenance team should intervene to prevent a potential failure from becoming an actual failure.

Heavy transport systems are the focus of a study by Gajewski & Valis [47]. Several dozen heavy crawler engines were used to collect the different types of oils. The study employs neural networks and these data to find patterns that represent system deterioration.

An ANN is used by Parlak *et al.* [48] to forecast particular fuel consumption and Diesel engine temperature.

3.2 Condition monitoring of engine oil

Another long-term objective is to monitor the oil in the engine in real time. Oil degradation is influenced by many different factors, some of which may result in faster or slower degradation, including working time, kilometres driven, driving habits, motor type, and age of the motor. As a result, in order to replace the oil at the best time, an analysis must be performed to ascertain the oil's level of deterioration. The main issue with laboratory analysis is that it is a time-consuming process that necessitates human involvement. Even though it is only required every few kilometres, the goal is to automate the process to cut costs and lower the likelihood of human error.

Accurate online monitoring has two key benefits: first, it minimises the amount of downtime required to inspect the oil and, second, it increases the likelihood that the oil will be changed at the ideal time - neither too early nor too late. The drawback is that it calls for sufficient sensors that need to be installed inside engines and come into contact with oil and offline sampling. The sensors must be reliable enough to withstand the operating circumstances without breaking down and precise enough to provide accurate readings.

Lubricating oil plays an important role in vehicle maintenance, and good lubrication can extend engine life as well as reduce maintenance costs. Le *et al.* [49], through machine learning models, classified the condition of military vehicle engine lubricating oils. Oil condition is classified into three categories: normal, degraded, and unsuitable.

A fair review of modern sensors and techniques for online oil condition monitoring is provided by Zhu *et al.* [50]. According to these authors, the sensors are divided into four groups: electromagnetic, physical, chemical, and optical. The oil's dielectric constant is measured by electromagnetic sensors. Another type of sensor measures the conductivity of the oil. The third type measures the magnetic susceptibility. A fourth type measures the viscosity of the oil. Zhu *et al.* mention the viscometer, ultrasound, thermal conductivity sensor, and ferrography when speaking of physical techniques. The chemical methods that were reviewed include the thin-film contaminant monitor and pH measurement. There are two optical methods: reflectometry and infrared absorption.

A method for online oil engine condition monitoring is proposed by Kumar *et al.* [51]. It is based on an optical sensor that converts oil darkness into electrical resistance. One of the elements that directly relates to oil quality is how the colour of the oil changes over time. The authors contend that by keeping an eye on the oil's colour, it is possible to predict when the oil is degrading and replace it. The technique is somewhat similar to that used by Yonghui *et al.* [52], who combine the use of an inductive sensor and a fibre optic transducer. While the optical sensor picks up small oil-contaminating particles, the inductive sensor detects large ferrous and some non-ferrous wear debris.

El-Hag *et al.* [53] monitor the oil condition in power transformers using features extracted from acoustic and radio frequency partial discharge signals. The level of oil degradation is determined by training a neural network on the pulse width, rise time, and frequency components.

Zhu *et al.* suggest a technique for keeping track of the condition of wind turbine oil using dielectric constant sensors and commercially available oil viscosity sensors. Based on the correlation between particle concentration and oil degradation, the sensor readings are calibrated. By measuring the inductance change of two planar coils wound around a pair of ferrite cores. Zhu *et al.* propose a technique to condition monitor oil using a sensor that detects wear debris [54].

The technique shares some similarities with Du *et al.*'s [55] approach, which also measures metallic, ferrous, and non-ferrous particles in lubricant oil using inductive sensors.

3.3 Condition monitoring using PCA and ANN

One multivariate analysis technique that is very good for data mining is principal component analysis. Data transformation, feature extraction, and the identification of a dataset's most crucial variables are statistical processes. Therefore, it is possible to predict, through PCA analysis, which variables need to be monitored and which variables are candidates for removal without reducing predictive power. The relationship between fuel parameters and the quantity of particles in Diesel motor emissions is determined by Westerholm and Li using PCA [56]. Capone *et al.* calculate the quantity of unburned fuel in lubricating oil using PCA [57].

Transmission line faults are common in long-distance power transmission systems, so their determination and classification are crucial. Mukherjee *et al.* [58] propose a method for classifying faults in transmission lines using an approach based on PCA. This study extracts failure characteristics in terms of a Principal Component Index (PCI), followed by a threshold-based analysis of PCI values. The development of two threshold values helps to segregate the three distinct levels of fault disturbance in terms of PCI values, thus developing fault signatures for classification. According to the authors, this classification method presented a 99.78% accuracy.

Additionally, various Artificial Neural Network architectures have been used to learn about and to predict the condition of the oil. Shaban *et al.* predict transformer oil parameters using a cascade of artificial neural networks [59]. The effectiveness of ANN and Support Vector Machines for predicting motor emissions is compared by Niu *et al.* [60]. Ghobadian *et al.* [61] model the efficiency of a Diesel engine using waste oil through an ANN. Li X *et al.* and Li Y *et al.* use convolution neural networks to detect gear faults based on different signals, namely sounds produced by the gears [62] and operational parameters [63].

3.4 ANN in predictive maintenance and forecast

In this section some works are presented, whose aim is to predict the values of the sensors installed in the equipment, stressing the importance of this research field for predictive maintenance using AI.

Kanawaday *et al.* [64] took advantage of the machine data generated by various sensors applying different data analysis algorithms to obtain information that helps in making decisions. The data captured by the sensors were always accompanied by the date and time, both of which are vital parameters for predictive modelling.

Short-term forecasting work in maintenance has also been carried out by other authors. However, it should be noted that those studies are only focused on short-term forecasting, which shows a clear limitation in long-term forecasting. An example of this type of study is the work presented below.

Kolocas *et al.* [65] presented a predictive maintenance methodology to predict possible equipment failures of an industrial equipment in real time, using data from process sensors of operation periods. The alert period for the failure of the asset is forecasted in short-term, since a forecast gap was defined around 5-10 minutes before the incident occurred.

The following review section demonstrates a promising way of research in the use of neural networks in predictive maintenance.

Tian [66] developed an Artificial Neural Network (ANN) based method designed to achieve more accurate remaining life prediction of equipment subject to condition monitoring. The proposed ANN method is validated using vibration monitoring data collected from pump bearings. The ANN model has as input to the network the age of the equipment, current condition measurement values and inspection performed. The network gives a percentage of the asset's life as an output.

Rafiee *et al.* [67] used a 2-layer perceptron neural network to detect gear and bearing failures and identify gearboxes using a new feature vector updated by the standard deviation of wavelet packet coefficients of vibration signals. Synchronization of vibration signals used cubic Hermite interpolation by parts.

Heidarbeigi *et al.* [68] developed a neural network built to predict gearbox failures. In this project a backpropagation learning algorithm and a multilayer network were used. The network has three classification outputs, which are: worn, broken teeth of gear and faultless condition. The ideal Multilayer Perceptron neural network (MLP) selected for classification exhibited a 489-10-3 layer structure and had 87% accuracy. The model shown works based on vibration differences, so it can be used in other applications.

Karpenko [69] developed a neural network pattern classifier to diagnose and identify failures in an actuator of a Fisher-Rosemount 667 industrial process valve. The network is trained with experimental data obtained from the asset. The test results show that the resulting multilayer feedforward network can detect and identify various types of failure.

Wang [70] presents an artificial intelligence algorithm based on neural networks to identify failures in Diesel engine lubrication pumps using vibration data. The algorithm has been tested on more than fifty lube pumps that have proven its effectiveness.

The studies mentioned above show that neural networks using monitoring data such as vibration and temperature can detect and even anticipate failures. That is useful in the diagnosis of faults with high reliability, as well as foreseeing potential failures and preventing them from happening. The research carried out also shows that there is a gap in a long-term forecast, specially predicting with 3 months advance. Nonetheless, this should be a research goal, because industries often need several weeks to prepare and carry out complex maintenance operations with minimum downtime.

3.5 Neural networks for prediction and classification

This section reviews relevant works using neural networks for prediction and classification, namely in the field of predictive maintenance.

Effective maintenance is essential to keep assets at maximum availability and accident free. For these reasons, Bukhsh *et al.* [71] developed a model to predict the need for railway maintenance. Elhag and Wang [72] presented an application of artificial neural networks to assess bridge risk by computing their risk scores and categories.

Balluff *et al.* [73] developed a model to predict wind speed and pressure through recurrent neural networks.

Deepika and Prakash [74] predicted the power consumption of a virtual machine with the help of backwards predictive analytics using a multilayer perceptron, achieving a 91% accuracy.

Hongxiang *et al.* [75] developed an algorithm using artificial neural networks (ANNs) to analyse spectroscopy data from lubricant oils. Results proved that ANNs can be used to classify distinct types of lubricants and to distinguish routine conditions of a Diesel engine from operating conditions.

Okoh *et al.* [76] presented an approach to determine when a system needs to undergo maintenance, repair, and overhaul, before a failure occurs. One of the main innovations of this project is that forecasts were made in the long-term.

One of the main challenges of maintenance is to increase the availability of equipment and, hence, it is important to predict failures before they happen. Makridis *et al.* [77] presented a machine learning approach for detecting anomalies from data collected through sensors installed on vessels, predicting the condition of specific parts of the vessels' main engine.

In 2021, Zhagparov *et al.* [78] proposed a solution to automate the prediction of grain yield based on machine learning using the XGBRegressor algorithm on the territory of the Republic of Kazakhstan. Comparisons were made with linear regression and decision tree regressor algorithms.

Dong *et al.* [79], in 2020, developed a prediction model based on the XGBoost algorithm that considers all potential influential factors simultaneously; the objective of this model was to predict the electrical resistivity based on an experimental database.

Due to the search for a more sustainable world, wind energy emerges as one of the most important sources of energy production. Zang *et al.* [80] propose a fault detection method for main bearing wind turbines based on SCADA data using an RNA artificial neural network. This algorithm makes it possible to identify the initial stage of main bearing failures, allowing for early intervention.

According to several authors, neural networks can improve support in decision making in the fields of maintenance and condition monitoring, as well as in many other fields where machine learning models can be applied to extract models from data [81],[82], [83],[84],[85].

3.6 Condition monitoring of paper presses

Nowadays, approximately 90% of the pulp and paper used in the world are made from wood. The cellulose is converted into pulp, which is used in the production of paper. In the past 40 years, the consumption of paper has increased by a factor of about four and is still increasing [86],[87].

In the process of paper production, the presses are used to press the paper pulp until the product reaches the required dry content and, consequently, has the ideal quality for use. In addition, the presses help to obtain other important properties for paper quality, such as absorption capacity and volume.

The use of an inefficient press causes the humidity not to be removed from the paper, decreasing the product's resistance and causing an increase in sheet breakage. If the material has a moisture content higher than recommended, it needs to be subjected to steam in the dryer section for longer, which increases production costs.

Therefore, the use of good presses ensures that the paper has the desired quality and also reduces costs for the manufacturer. Thus, the press is an indispensable piece of equipment in the industrial production of paper, as it increases the quality of the final product and can reduce production costs for the company when working under perfect conditions. Knowing how important monitoring and predictive models are to the company's success, they are indispensable for increasing asset availability and avoiding breakdowns. Condition monitoring plays a central role in the maintenance of paper machines. The main objective is to maximize the availability and reduce the costs of these manufacturing units and to prevent unexpected damage or mechanical breakdowns.

The results of the tests by Suomela *et al.* [88] in 2002 make it clear that thermal imaging combined with adaptive drive has great potential for monitoring paper machine components.

The work by Bissessur *et al.* [89] features the ability to detect faults and provide early warning of impending problems based on collected vibration data and pre-processing spectra. These data processed by a neural network provide an instant decision about the state of the felt that is monitored. This method can be extended to diagnose faults in a wide range of mechanical and rotating equipment in industries.

Mateus *et al.* developed predictive models based on long-term deep neural networks applied to a dataset of sensor readings. The results show that it is possible to predict future behaviour up to a month in advance with reasonable confidence (errors in general inferior to 10%) using long short-term memory and gated recurrent unit deep neural networks [81,85].

3.7 Summary

Industries are increasingly focused on the long-term, and the maintenance department has to be fully aligned with this long-term vision.

Given the above, it is important to study and investigate the long-term risk assessment of equipment to determine the probable long-term status of a physical asset.

The classification of an equipment’s status works found in the literature only performed classifications of failure in the present, short or medium-term, which is a limitation of the state of the art. To close this gap, the methods proposed in this PhD thesis aim to perform predictions and classifications in the short, medium and long-term.

With regard to the state of lubricating oils, when evaluating the state of the art, it was concluded that there is a need to develop an oil classification model that can aid human expert decision making.

Chapter 4

Data analysis and forecast methods

This section aims to briefly explain the main theoretical principles behind the techniques used in this thesis.

4.1 Time series

Time series can be used for forecasting in the industrial field [90],[91]. A time series is a series of observations recorded sequentially over time [92],[93]. They are applied in several areas of engineering and mathematics.

The analysis of time series aims to understand the dependence of current and past data and exploit them. Therefore, it is useful to predict the future behaviour [93]. For example, in the Control of Industrial Processes, the forecast methods are essential to characterize products, to understand the quality of the processes and, thus to adjust the processes in order to avoid over production capacity [94]. In the general economy, various institutions, such as governments, financial institutions and political organizations, make decisions based on forecasts of economic variables, such as Gross Domestic Product, the unemployment rate or inflation for exports and imports [95].

A stochastic process can be defined as a finite or infinite sequence of random variables and a time series can be understood as one of the infinite realizations of a stochastic process [96].

In most time series techniques, it is assumed that the data are stationary. Thus, given this characteristic, it is important to know that a stationary process is the one in which the mean, the variance and the autocorrelation structure do not change over time.

Auto Regressive Moving Average (ARMA) provides a parsimonious description of a stationary stochastic process in terms of two polynomials, one for the Auto Regression (AR) and the second for the Moving Average (MA). Autoregressive Integrated Moving Average (ARIMA) is a generalization of the autoregressive model of moving averages (ARMA). Both models are adjusted to the time series data for a better understanding of the data or to predict future points in the series. ARIMA models are applied in some cases where the data show evidence of no stationarity, in which an initial differentiation step (corresponding to the "integrated" part of the model) can be applied one or more times to eliminate the non-stationarity [90].

The main disadvantage of the ARIMA model is that it assumes a linear relationship between the dependent and independent variable, which is not always the case [97].

4.2 Artificial intelligence

The work of Charles Spearman, who suggested that intelligence should be understood as a general skill that was brought together as a set of intellectual tasks with specific skills, served as the foundation for the modern study of intelligence [98]. According to Robert J. Sternberg, intelligence is the capacity for learning from experience as well as for adapting to, modifying, and choosing environments [99].

Artificial intelligence, according to John McCarthy, is "the science and engineering of creating intelligent machines, especially intelligent computer programs." In recent years, artificial intelligence has primarily been used for three purposes: cloud computing, big data and machine learning [100].

A machine with artificial intelligence must be programmed to learn and solve a wide range of problems, whether they are straightforward or complex. It should be noted that a machine must pass the Turing test in order to be classified as intelligent. Alan Turing, a computer scientist and mathematician, deserves credit for the test's name because of his significant contributions to the advancement of computing [101]. The idea of computation is abstracted by the Turing machine. It can simulate any type of computation, making it able to resolve any problem that is algorithmic and therefore can be solved by computation [102]. The Turing Test is a research method in artificial intelligence to determine whether a computer is capable of thinking like a human being or not [103], [104], [105], [106].

There was a gestation period between 1943 and 1955 before the concept of artificial intelligence was introduced. McCulloch and Pitts [107],[108],[109] developed a model for the first artificial neuron during this time, and Minsky later produced the first artificial neural network with 40 neurons [108]. It was suggested, in 1955, that a team of 10 carefully chosen scientists work on artificial intelligence for two months at Dartmouth College in Hanover, New Hampshire, during the summer of 1956. The authors of this proposal thought it would be possible to make a significant advancement in artificial intelligence during that summer. They listed several AI-related issues, including Automatic Computers, How Can a computer be Programmed to Use a Language, Neuron Nets, Theory of the Size of a Calculation, Self-promotion, Abstractions, Randomness and Creativity, before AI itself was created.

There was a lot of initial excitement for artificial intelligence in the 1950s and 1960s because there were high expectations for it. Little has been accomplished despite the great enthusiasm, but many important advancements were achieved along the years.

4.3 Machine learning methods

Machine learning methods can be classified as supervised, unsupervised and reinforcement. Supervised learning requires a training set of inputs and outputs to learn a function that minimizes its prediction error. Unsupervised Learning tries to discover more compact representations from a dataset, without knowing their outputs. And Reinforcement learning prescribes decisions based on the feedback obtained after the actions performed.

4.4 Artificial neural networks for prediction

Optimization of neural networks is a challenging problem, and it has been the topic of many works [110], [111], [112].

As the computing power and big data increase, deep learning models are becoming more popular in several fields of science. Deep models are characterized by containing several layers, while shallow models rarely have more than three layers. For instance, deep networks are the preferred architecture in object detection and complex classification problems. Shallow neural networks are more adequate for classification problems where the features are easy to identify. Despite many clear distinctions between deep and shallow neural networks, some techniques developed for deep learning can help improving shallow models [111].

The importance of the present work is reinforced by several authors that have emphasized the necessity to change the focus from short-term (15 days) maintenance policies to long-term ones (90 days). The importance of these contributions corresponds to the increase of equipment's availability, which permits increased productivity and, at last, the success of the companies [113], [114], [115], [116].

4.5 Feedforward neural networks

A Feedforward Neural Network (FFNN) is a type of artificial neural network in which there are no loops in the connections between the nodes. A recurrent neural network, in which particular pathways with feedback are established, is different from a feedforward neural network. Since information is only processed in one direction, the feedforward model is the simplest type of neural network. Although, the data may move through several hidden nodes, it always proceeds forward and never backward.

The input layer, which consists of the input features, is the first layer. The output layer, which contains the network's output, is the last layer [115]. Since we cannot see the inputs or outputs of the intermediate layers, they are referred to as hidden layers.

A single layer perceptron is a common example of a feedforward neural network in its most basic configuration. A number of inputs are introduced into the layer in this model and multiplied by the weights. The weighted input values are then added together to produce a total. The value produced is frequently 1 and, if the sum of the values is below the threshold, the output value is -1. The threshold is typically set at zero.

An important feedforward neural network model, that is frequently used for classification tasks, is the single layer perceptron. Single layer perceptrons also incorporate some features of machine learning. Through training, the feedforward neural network can adjust its weights to produce more accurate output values by comparing the results [116].

This learning and training process uses generally a gradient descent process. Although the process of updating weights in multi-layered perceptrons is nearly analogous, it is more specifically known as back-propagation. In these circumstances, the network's hidden layers are adjusted in accordance to the output values generated by the top layer. The error between the output desired and the output obtained is backpropagated and the network's weights are adjusted using the gradient descent algorithm.

Multilayer Perceptron (MLP) is a type of FF neural network. It usually comprises three types of layers: one input layer, several hidden layers, and one output layer. The main applications of MLP networks are pattern classification, recognition, and prediction.

Figure 3 represents a single-layer feedforward neural network and Figure 4 represents a multilayer feedforward neural network.

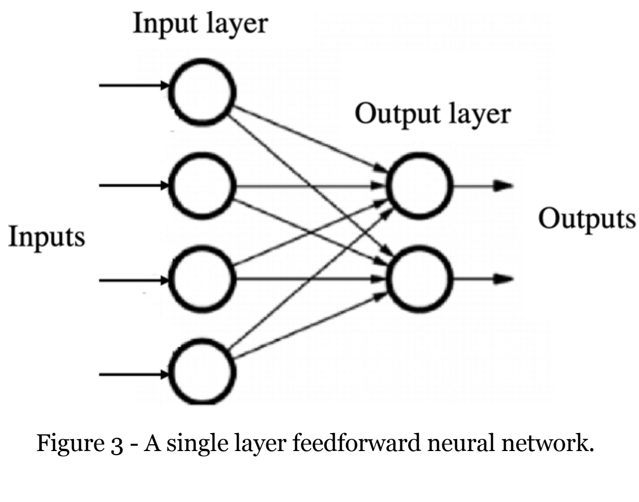

Figure 3 - A single layer feedforward neural network.

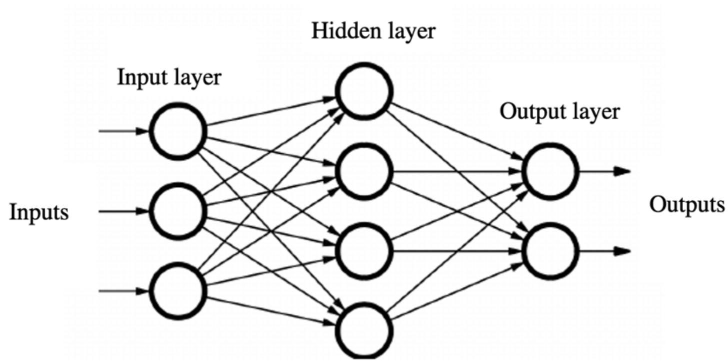

Figure 4 - A multilayer feedforward neural network.

4.6 Perceptron

The simplest unit used for supervised learning is called a perceptron. Binary classifiers can be built with one or more perceptrons. They determine whether an input, which is typically represented by a collection of vectors, falls under a given category.

A perceptron is a single-layer neural network [117], to put it simply. Input values, weights and bias, net sum, and an activation function are the four main components [118].

A perceptron works as follows: all the input values are multiplied by their weights to start the process; the weighted sum is then calculated by adding the sum of all these multiplied values; and the output of the perceptron is then generated by applying the weighted sum to the activation function. The vital function of the activation function is to make sure that the output is mapped between necessary values like (0,1) or (-1,1). It is significant to remember that a node's strength can be determined by the weight of an input. Similar to this, the bias value of an input allows for the adjustment of the activation function curve.

Perceptrons are crucial for binary classification. In other words, binary classification of data using the perceptron is very common. For this reason, perceptrons are also sometimes referred to as linear binary classifiers. Figure 5 shows an example of a perceptron with a continuous output from 0 to 1. This model was chosen as an example because in this thesis a sigmoid transfer function was used, and therefore the output is a floating point number between 0 and 1.

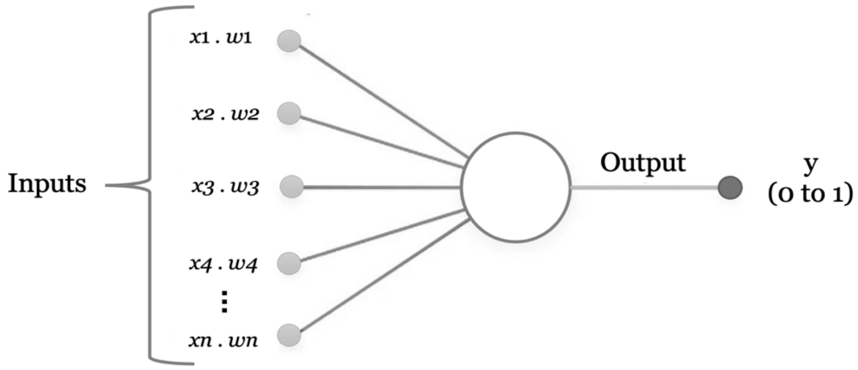

Figure 5 – Example of perceptron.

4.7 XGBoost

The name XGBoost comes from eXtreme Gradient Boosting and represents a category of algorithm based on Decision Trees with Gradient Boosting [119].

XGBoost is a scalable and highly accurate implementation of gradient boosting that pushes the limits of computing power for boosted tree algorithms, being built largely for energizing machine learning model performance and computational speed [79].

Gradient boosting is a machine learning technique for regression and classification problems, which produces a predictive model in the form of an ensemble of weak predictive models, usually decision trees.

Decision trees are techniques where a function returns an output decision from an input vector of values (of attributes). A decision tree goes through a series of steps or tests to get its output value, creating several branches along the way.

In this decision tree, each node represents a distinct choice. An attribute's relative importance in the model increases the more decisions that involve it.

4.8 Recurrent neural networks (LSTM & GRU)

A Recurrent Neural Network (RNN) is a type of artificial neural network which uses sequential data or time series data.

Long Short-Term Memory (LSTM) network is an advanced RNN, a sequential network, that allows information to persist. In general, it can handle the vanishing gradient problem faced by RNN. LSTM extracts patterns from sequential data and stores these patterns in internal state variables. Each cell can retain important information for a longer period when it is used. Such information properties allow the LSTM to perform well in predicting dynamic sequences [120], [121],[122], [123].

The Gated Recurrent Unit (GRU) was designed by Cho *et al.* [121]. The closed recurrent unit is a special type of optimized recurrent neural network based on LSTM [124],

[125]. The difference is that the GRU combines the input port and the forgetting port in the LSTM into a single update port [126],[127]. It is important to note that these neural networks are not explained in detail because predictive models have not been developed using this type of neural network. These types of neural networks have been used by other authors and their results are compared with the predictive models developed in Chapter 7.

4.9 Random forest

Random forest is a popular and effective ensemble machine learning algorithm. It is widely used for classification and regression predictive modelling problems [126].

The Breiman's work about random forest was decisively influenced by the early search of Amit and Geman, in 1997, on geometric feature selection: the random subspace method of Ho, in 1998, and the random split selection approach of Dietterich, in 2000 [127], [128].

Breiman throughout his various papers has shown that decision trees can produce substantial gains in accuracy when used together in regression and classification tasks. These two types of tasks represent a large percentage of machine learning systems [129].

In a regular decision tree, a single decision tree is built; however, in a random forest, many decision trees are built and the number of trees is usually defined by the user [130], [131].

In a joint process, each tree casts a unit vote and the decision trees are used to choose the final class. In random forests, a sample of substituted data is used to build a decision tree along with the subset of variables. The sampling and subset are recorded randomly [132],[133].

4.10 Data grouping (clustering)

Cluster analysis is an unsupervised learning technique used to group elements into clusters so, that elements within the same group (cluster) are as similar as possible, while elements from distinct groups are as different as possible.

To define the similarity – or difference – between the elements, a distance function is used, which needs to be defined considering the context of the problem being solved.

Methods of this type have applications in various fields, such as data visualization, pattern recognition, learning theory, computer graphics, identification, or classification.

4.11 K-means

K-means is one of the most popular unsupervised learning algorithms used for data clustering [132]. This algorithm assigns each data point to one of K groups (clusters) that minimize the square of the distance between that point and the centroid of each cluster.

The application of K-means suffers from some difficulties, such as the requirement that the number K of groups to be formed or their sensitivity to the initial conditions are provided *a priori*. Therefore, they must be determined experimentally during the data analysis process.

4.12 Principal component analysis

Principal Component Analysis, also known as PCA, is a multivariate statistical test that makes it possible to transform a set of variables correlated with each other into a new set of uncorrelated variables that are called principal components. The principal components are calculated in decreasing order of importance, with the last component being the one that makes the least contribution to the explanation of the total variance of the original data; on the contrary, the first main component will be the one that makes the greatest contribution. Each principal component identified in the test is a linear combination of all the original variables, which are independent of each other and estimated to retain, in order of estimation, as much information as possible, in terms of the total variation of the data [133] .

The description of practical computational methods came to be developed by Hotelling (1933-1936), who used it to determine and analyse correlation structures. It should be noted that in PCA, the mathematical representation of linear combinations does not allow the detection of cause-effect relationships in the initial variables, if they exist [134].

In summary, the main objective of PCA is, essentially, to reproduce the matrix of correlations with a smaller number of orthogonal factors, losing the minimum information.

4.13 FMEA and FMECA

Criticality analysis is based on failure mode evaluation analysis. Criticality means the failure probability of the equipment. A minor failure of critical equipment may lead to a severe impact on the performance of that equipment. So, critical equipment needs a higher degree of maintenance activity and maintenance frequency to prevent any failure. Among the variables that allow the evaluation of equipment's criticality, the safety factor represents the associated risk in case of failure. If the probability of injury (man or machine) is high, in the case of equipment failure, this value is given by the safety factor and the lower the chances of injury, the lower is the safety factor's value [135].

Facing a systematic analysis approach, Failure Mode and Effects Analysis (FMEA), and Failure Mode Effects and Criticality Analysis (FMECA), allow the identification of existing problems in the process, that may arise, given the examination of previous errors, through a hierarchy of potential failures, and making recommendations to avoid them through maintenance techniques.

4.14 Ishikawa diagram

The Ishikawa Diagram used to find, organize, classify, document and graphically display the causes of a given problem, grouped by categories, which facilitates brainstorming ideas and analysing the occurrence. As the causes are hierarchical, it is possible to identify the root origins of a problem.

Brainstorming with the team can encourage an in-depth analysis of the problem, which involves most of the possible causes of a problem, as it promotes discussion and, consequently, the improvement of the process. The Ishikawa Diagram is also known by several other names, such as:

- Fishbone Diagram;
- Cause and Effect Diagram;
- PDCA fishbone;
- Tree Causes Fishbone.

The Ishikawa diagram can be actively used in clinical settings and mental/behavioural health. It offers a methodical, structured approach for locating and compiling potential causes of an effect [136].

Pros and cons of Ishikawa diagrams

Like all methods, the Ishikawa Diagram also has advantages and disadvantages, as shown below.

Advantages

- Easy to construct;
- Flexible and generic;
- Helps find the root cause of problems.

Disadvantages

- Can be oversimplified;
- Does not prioritize problems by importance;
- Subjective.

4.15 Model evaluation metrics

Two metrics were used to evaluate the forecast models. Equation 1 presents the formula for calculating the Mean Absolute Percent Error (MAPE) and Equation 2 presents the formula for calculating the Mean Squared Error (MSE) [137].

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2 \tag{1}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{|Y_t|} \tag{2}$$

where Y_t represents the actual value, $\hat{Y}_t$ the predicted value, t is discrete instant time that varies between 1 and n, and n is the total number of data point samples.

Precision (Equation 3) quantifies the number of positive class predictions that belong to the positive class.

Recall (Equation 4) quantifies the number of positive class predictions made from all positive examples in the dataset. Recall can also be called True Positive Rate (TPR), or Sensitivity.

Accuracy (Equation 5) is a metric that summarizes the performance of a classification model as the number of correct predictions divided by the total number of predictions.

The F1 (Equation 6) score is commonly used to measure performance of binary classification. It is a harmonic mean between precision and recall.

The next equations show the formulas for these metrics, where *TP* is True Positives, *FN* is False Negatives, *FP* is False Positives and *TN* is True Negatives.

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

$$Accuracy = \frac{CorrectPredictions}{TotalPrediction} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

$$F1Score = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \tag{6}$$

4.16 Summary

The Industrial Internet of Things (IIoT) is a recent idea that is applied directly to the industrial sector as a result of the significance that IoT has acquired recently in industry and maintenance.

Industrial sensors must be timely calibrated and certified in order to produce an accurate forecast. This is crucial because, without metrology's support based on measurement quality, there may be assessment mistakes and discrepant data, which could lead to prediction errors, inadequate forecasts, risks, high costs, and, ultimately, a loss of market confidence.

Nowadays, with the increase of computational power and cheaper technology, it is possible to use artificial intelligence to improve quality, reduce failures and improve productivity of the industry.

In this study, the most important tools were feedforward neural networks, as it was through them, using multilayer perceptron (MLP), that it was possible to predict the sensory values and classify the lubricants and the state of the Pulp Paper Press. It should be mentioned that XGBoost also allowed predicting the sensory data of a Pulp Paper Press.

The role of PCA is also highlighted, as through PCA it was possible to develop a formula to calculate the degradation state of the oil. The FMECA and the Ishikawa Diagram were crucial to identify the cause of the systematic failure of the Chip Pump shaft.

Regarding the metrics, the most important in this study were the MSE and the MAPE, because it was through these metrics that the various prediction algorithms were evaluated.

Chapter 5

Classification of engine oils

This chapter describes the work to classify engine oils. Part of this work was published in a scientific journal paper [138].

5.1 Introduction

Engine oil condition monitoring is a strategic area in the maintenance management industry. Early oil replacement would result in unneeded unavailability, as well as financial and environmental costs that could be avoided. Late replacement can reduce the oil's ability to protect the engine, increasing the risk of engine damage and early ageing as well as the possibility of accidents that could endanger people, equipment, or vehicles in crowded areas. Making the right decision at the right time can help protect the environment, a company's profits, and the safety of people and property. Modern tools from data mining and artificial intelligence can help.

The methodology for developing models to simplify the process of oil analysis is discussed in this chapter and tested using a dataset for Diesel engine oil from urban passenger buses. The neural network developed by Ines Gonçalves was used in this study to validate the results of the PCA formula [138],[139]. In this study, multivariate systems analysis, specifically Principal Component Analysis, was used to compare the results with neural model improvements. Each variable's relevance was demonstrated by PCA to be unique, and some of the variables may even have a detrimental effect on the ANN's predictive ability.

Two passenger bus companies provided the data that were used in the experiments. Each company offered a dataset that included the findings of laboratorial analysis of the oils and a classification of those oils made by experts in the field.

5.2 Datasets used

Two datasets, referred to as A and B for short, were obtained from two different public transportation bus companies. They served as the basis for the current study. The datasets include the measurements of 21 parameters of the oil analysis performed in a lab using oil collected from buses at various stages. The bus mileage and oil mileage are also included in each sample. There are 23 parameters in total that are used as input variables for this analysis.

The 23 parameters include the bus's mileage, the oil's mileage, the amount of antifreeze found in the oil, the percentage of fuel, the Finacheck water content, smoothness, nitration, oxidation, sulfation, TBN, viscosity at 100°C, and the parameters Al, Cr, Cu, Fe, Mo, Na, Ni, Pb, Si, Sn, V, and PQ. The variables were then used as inputs for the PCA analysis and neural networks after being normalized.

The datasets also include the specialized company's decision, denoted as 1, 2, or 3. Decision 1 denotes the business' determination that the oil is in good shape and can be maintained for regular bus operation. Decision 2 indicates that it is time to replace the oil because it is worn out. Decision 3 indicates that the oil is beyond the point at which it should have been changed and that the bus must be stopped right away for safety reasons.

A total of 47 samples are included in Dataset A. They were taken from various Company A buses. Four samples each from twenty-two different buses make up the 88 total samples in Dataset B. The inability to obtain larger datasets for the current study was a barrier that could not be transposed. Nevertheless, there was consistency in the outcomes for both the PCA and neural models. They have been repeated numerous times and are repeatable under similar conditions.

In the train and test sets, many neural models showed good performance and minimal error. The ones that performed better on the test set were those that were chosen for analysis, demonstrating the generality of the model. This demonstrates that the findings are reliable and that a larger dataset could be handled by the method. Since PCA is a factor-analysis method, the Kaiser-Meyer-Olkin (KMO) test was used to determine whether the datasets were adequate for PCA [140]. The KMO test yields a score between 0 and 1, with higher scores generally indicating that there is sufficient sample diversity in the dataset to perform factor analysis.

On the other hand, a low score indicates high correlations between the variables and unreliable factoring process results. For dataset A, the KMO test yields a score of 0.35, which is extremely low. For dataset B, the score is 0.636, which is acceptable, indicating that dataset B's factor analysis results are more reliable.

5.3 Principal component analysis loadings

Table 1 shows results of the PCA analysis for datasets A and B. As the table shows, Si, Fe, Al and Cr contents are the four most important variables, with loadings above 0.7, for dataset A. For dataset B, Fe, Soot and Cr are the top three variables, and the only ones with score above 0.7. Five of the top ten variables are related to oil status and five are related to wear and contamination.

TABLE 1 -PCA LOADINGS

	A		B	
ORDER OF RELEVANCE	**VARIABLE**	**LOADING**	**VARIABLE**	**LOADING**
1	**Si Content**	0.872	**Fe Content**	0.889
2	**Fe Content**	0.864	**Soot**	0.835
3	**Al Content**	0.789	**Cr Content**	0.781
4	**Cr Content**	0.729	Viscosity at 100 ºC	0.689
5	**Sn Content**	0.668	**Sn Content**	0.682
6	**PQ Index**	0.551	**Cu Content**	0.611
7	Ni Content	0.425	Pb Content	0.571
8	**Soot**	0.421	**Sulfation**	0.507
9	**Oxidation**	0.412	Nitration	0.496
10	V Content	0.376	**Oxidation**	0.488
11	**Cu Content**	0.282	**Al Content**	0.482
12	**Sulfation**	0.266	**Si Content**	0.423
13	Mo Content	0.166	**PQ Index**	0.320
14	Pb Content	0.134	Na Content	0.166
15	Fuel Content	0.132	Antifreeze	0.162
16	Na Content	0.118	Water Content	0.127
17	Viscosity at 100 ºC	0.089	Mo Content	0.072
18	TBN	0.069	V Content	0.020
19	Nitration	- 0.003	Ni Content	- 0.008
20	Water Content	- 0.140	Fuel Content	- 0.134
21	Antifreeze	-0.142	TBN	- 0.395

5.4 Degradation formula

The principal components are calculated by decreasing order of importance. The first component is the most important, the last is the less important explanatory variable. Each principal component identified is a linear combination of all the original variables. PCA was applied to the datasets presented above, in order to understand the companies' policies, the state of the oils when the samples were collected for chemical analysis, as well as to determine which variables are more important to measure for correct assessment of the situation of the oil. The PCA experiments and analysis were performed using R Studio software. The formula used to calculate the percentage of oil degradation is Equation 7.

$$LD = \sum_{j=1}^{n} CP_j \times \frac{VA_j}{MF_j} \quad (7)$$

Where,

CPj = Loading first principal component parameter

MFj = Limit reference of lubricant parameter according to manufacturer

VA*j* = Value of the analysed parameter

LD = Lubricant degradation (%)

n = Number of parameters

CPj arises from the first principal component. The loadings are from a numerical point of view, equal to the coefficients of the variables, and provide information about which variables give the largest contribution to the components.

- Loadings are between -1 and 1.
- An absolute value that is close to 1 or -1 indicates that the variable has a significant impact on the component. Values that are relatively close to *0* show that the variable only has a minor impact on the component.
- Whether a variable and a principal component are positively or negatively correlated is shown by the sign of a loading (+ or -).

MF*j* comes from the limit recommendations that the manufacturer indicates for each lubricating oil, and VA*j* is the value of each parameter taken from the oil sample in question. The result of this formula will show the percentage degradation status of the oil. The results from this formula are in column "Det. %" of Table 4.

5.5 Neural network results

Shallow Feedforward Neural Networks with an output layer and a hidden layer of variable width (number of neurons) were the neural models employed in the current study. The output was kept within the range [0, 1] thanks to the sigmoid transfer function used by the hidden neurons. To enable a greater output amplitude and speed up the learning process, the output neuron used a linear transfer function (Relu).

A new dataset named AB was created. It contained a total of 176 samples, 88 from each dataset. All the neural models used for experiments with the AB dataset contained seven neurons in the hidden layer. The training was carried out with the Levenberg-Marquardt method. Training was performed with 70% of the samples, validation with 15% and test with the remainder 15%.

This is desirable because it can be used to gauge the oil's quality: the lower the value, the higher the oil's quality. To obtain a model that can be compared to the classification of the human experts, as described in Section 5.3, it is also crucial to map the output to 1, 2, or

3. As a result, it was mapped using the following rules in the discrete interval [1, 3]: Anything less than 1.50 was mapped to 1; everything between 1.50 and 2.50 was mapped to 2; and anything higher than 2.50 was mapped to 3.

The model was trained with 10+2 variables. That is, the 10 variables are those with the higher loadings of the first principal component of the AB set. These variables are in bold and underlined in Table 1. The other two variables are the mileage of the bus and the mileage of the oil.

The model trained with just 12 input variables produces a total of 14 prediction errors: 7 for each company.

Table 2 shows the results of the classifier neural network in more detail.

Table 3 shows the confusion matrices of the errors of the model trained.

TABLE 2 - NEURAL NETWORK RESULTS

# INPUTS	R (TRAIN TEST)	R (VALIDATION SET)	R (TEST SET)	R (ALL DATASET)	MSE (ALL DATASET)
10+2	0.96	0.86	0.84	0.94	0.15

TABLE 3 - CONFUSION MATRICES OF THE ERRORS OF THE MODEL TRAINED

# INPUTS	PREDICTED	COMPANIES A&B		
10 + 2	**3**	0	0	**47**
	2	4	**34**	7
	1	**81**	1	2
	ACTUAL	**1**	**2**	**3**

Observing Table 3, it can be seen that in state 3 the neural network does not present any classification error and in state 1 it presents only 3 errors. It should be noted that these are the most important states, since they indicate when the oil is actually in good condition or should be replaced.

5.6 Results comparison

Model predictions are sometimes different from business decisions, but some of the errors may be due to poor human decisions.

Through the PCA it was confirmed that, in both companies, it is possible to identify oils that were analysed very early and oils that were analysed with a much more advanced degree of deterioration.

Table 4 compares the results obtained for the AB dataset. Table 4 shows the results of the first 20 oil analyses of the database using the different types of classification. It shows the classification of the oil by the human experts, the classification given by the artificial neural network trained with 12 variables and the percentage of oil degradation calculated by PCA Formula.

TABLE 4 - CLASSIFICATIONS OF THE FIRST 20 OIL ANALYSES USING EXPERTS, ANN AND PCA FORMULA.

OIL ANALYSIS ID	EXPERTS	ANN	PCA FORMULA
1	1	1	31.7%
2	1	1	13.6%
3	2	2	49.9%
4	1	1	18,80%
5	1	1	52.1%
6	3	3	150.6%
7	1	1	35.6%
8	1	1	17.6%
9	1	1	13.4%
10	3	3	79.1%
11	1	1	51.6%
12	2	2	54.8%
13	1	1	35.7%
14	1	1	31.7%
15	1	1	13.6%
16	2	2	49.9%
17	1	1	18.8%
18	1	1	52.1%
19	3	3	150.6%
20	1	1	35.6%

Table 4 shows that there is convergence between the classifications of the expert, the ANN and the PCA formula.

Comparing the results of all the oil analyses in the AB database, it can be concluded that, the average deterioration of class 1 for human experts is 37.86%, and for ANN it is 37.03%. It was 52.07% for humans and 52.44% for ANN for class 2. The average degradation for class 3 is 73.05% for humans and 74.14% for ANN.

The average deteriorations demonstrate that the ANN was successful in classifying the oils into the three states of most and least deterioration.

The main limitation of this formula is that it presents a percentage result so, when this percentage result has to be passed to a qualitative classification there may be a divergence depending on the user's interpretations.

Another limitation is the need to know the reference limits for the parameters of the lubricant being analysed according to the manufacturer.

The main advantage of this formula is that in addition to allowing an automatic and reliable classification of the lubricating oil, it allows to assess whether the lubricants are fit to continue in the equipment avoiding unnecessary replacements, thus avoiding unneeded costs. The result of lubricant oil degradation is quantitative, which allows much more informed and adaptable decision making to the maintenance policy adopted.

5.7 Summary

Monitoring engine oil condition is very important to prolong engine life, avoid unnecessary pollution and accidents due to engine overheating or other failures.

The present chapter described a formula for calculating the level of deterioration of the lubricating oil of an engine using PCA, which allows classifying the state of the oils with high precision.

To validate the created calculation formula, the results obtained were compared with the results of neural models and classifications of human experts. The results converged by more than 90%, which shows that this formula is viable, reliable and can classify lubricating oils with high precision.

The present analysis can be useful to help companies make the best decisions at the best time, or even decide which variables are more important to monitor. Future research includes testing other lubricant classification techniques.

Chapter 6

Chip pump behaviour prediction

This chapter describes the work to predict a chip pump's behaviour. Most of this work was published in a scientific journal paper [141].

6.1 Introduction

This chapter provides a case study of a wood chip pump system working in an industrial paper firm, where data analysis is carried out and a predictive system is created. This asset experienced frequent failures on one axis. The life cycle of the pump shaft and its entire attachment system was substantially shorter than predicted by the manufacturer. The shaft gradually developed fissures. The goal of the analysis was to identify the root cause of the failure and any other possible prospective failures.

The Ishikawa Diagram and Failure Mode, Effects, and Criticality Analysis (FMECA) were both utilized to find every potential source of malfunction. In order to increase the reliability of the system, sensors were placed to monitor the main condition variables of the system's equipment.

A comprehensive analysis of the information gathered from the sensors put in place in each piece of equipment is provided, together with the lowest and maximum expected values. In addition to studying the behaviour of the variables, forecasting techniques based on time series and Artificial Neural Networks were applied.

Based on the widely used method of exponential smoothing, a short-term prediction model with a gap of 5 days was put into practice. A 3-month gap long-term prediction model based on artificial neural networks was developed. Five days is enough time for the company to prepare minor interventions. The larger gap gives the business time to properly plan and schedule maintenance interventions, preventing production loss and maximizing downtime. The length of the intervals was chosen to maximize production time and minimize maintenance downtime in order to gain a competitive advantage.

A dashboard with certain alarms presented through semaphores, as well as some numerical and graphical data, was created.

The system is made to prevent unanticipated breakdowns and cut expenses as much as possible, which are two of the key goals of an effective maintenance strategy [142].

The case study in this chapter explores the use of a variety of diagnostic and forecasting techniques in tandem with improved maintenance outcomes and increased equipment availability. Both the proposed prediction approach and the methodology for fault diagnosis are adaptable to other equipment. Any type of equipment can use the defect diagnosis method, whereas machine learning techniques can be used to any dataset with the right modifications and training.

6.2 Chip pump system diagnosis

In Figure 6 the chip pump serial/parallel diagram is shown. It consists of three chip pumps, each of which is mechanically connected to one asynchronous motor. The system inputs are liquor and wood chips. They are combined to create the finished item.

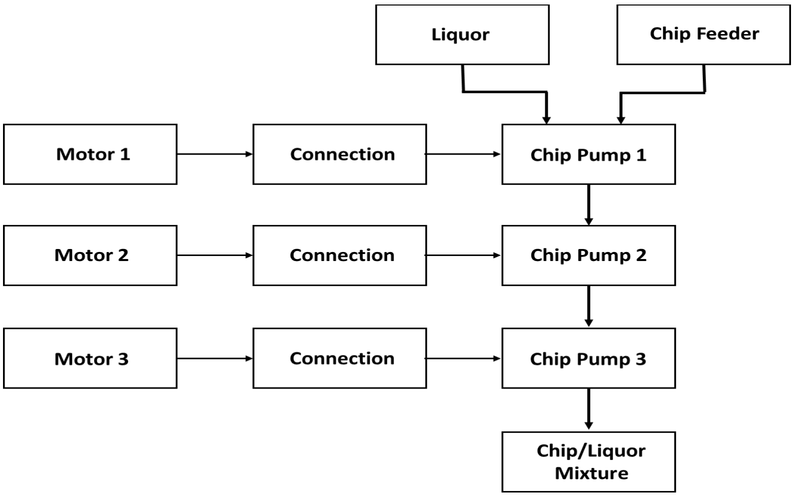

Figure 6 - Chip pump diagram.

The company realized that the shaft of chip pump 3 had a life span shorter than anticipated. The fixation cones on that chip pump had been harmed by shaft cracks brought on by the chip pump's repeated failures.

After several observations, it was found that the pressure increases throughout the system. The maximum pressure reached was at the outlet of Chip Pump 3. Figure 7 shows this increase.

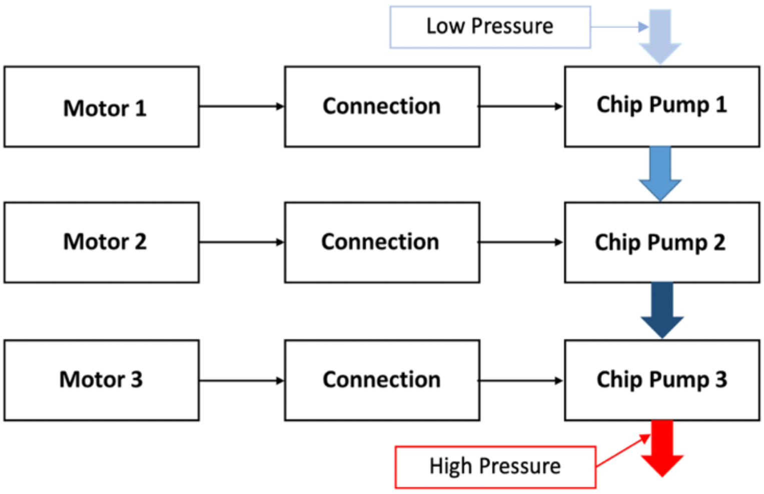

Figure 7 - Diagram of the pressure increase in the system.

Ishikawa diagram helps the analysis of the potential root causes of equipment flaws [143]. The chip pump cone and the crack or break of the shaft are represented in the Ishikawa Diagram in Figure 8.

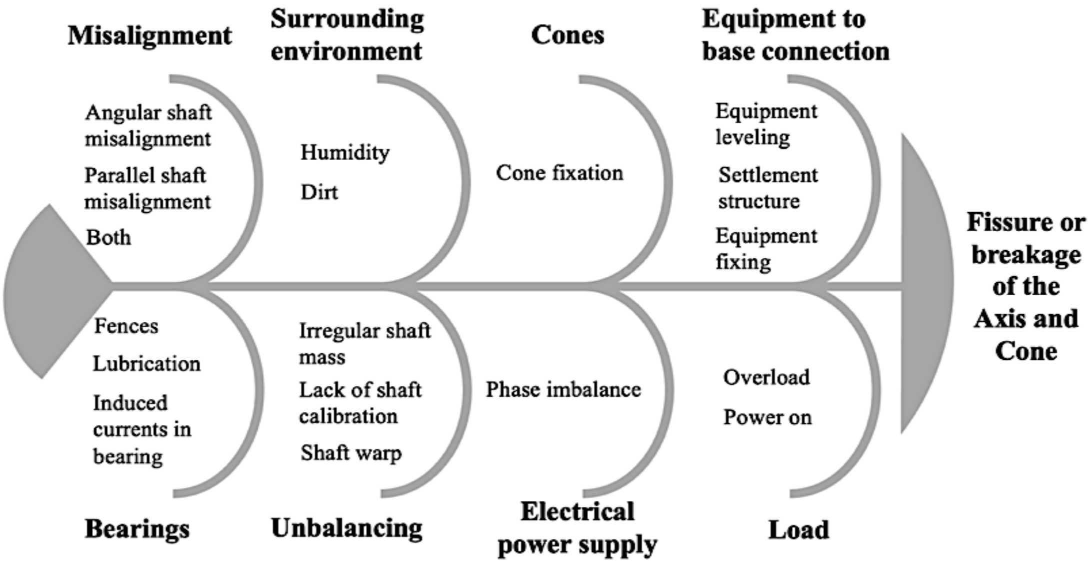

Figure 8 - Ishikawa diagram about fissure or breakage of the shaft and cone.

The preceding root-cause methodology was supplemented with an FMECA that followed the IEC 60812:2018 guidelines [144].

The main potential issues with the asset can be found using FMECA. A hierarchy of probable failures and a list of suggestions for preventing them through maintenance approaches can be used to produce this kind of study.

FMECA allows for the development of a working method, evaluation of failure modes and their effects, organization of the failure modes, identification of the points of failure and confirmation of the system's integrity, expedited failure resolution and, finally, the definition of test and verification requirements for the preventive maintenance plan. A failure analysis can be used to understand the asset's failure mechanism. FMECA includes Failure Mode and Effect Analysis (FMEA) and the Criticality Analysis (CA) [145],[146].

According to the FMECA matrix, shown in Figure 9, the "fissure or breakage of shaft and cone" is the primary issue.

Equipment	Chip Pump				Prepared by		Team Company					
Team	Company				Date		2021					
Equipment Module	**Function**	**Failure Mode**	**Failure Effect**	**Severity**	**Potential Cause of Failure**	**Occurrence**	**Preventive Action**	**Detection Action**	**Detection**	**RPN**	**Recommended Actions**	**Responsible and Deadline**
Chip Pump (412-306)	Mechanical Traction / Mechanical Transmission	Crack or breakage of shaft and cone	Stop system for Chip Pump (412-306) and production system		**Misalignment**							
				2	Angular shaft misalignment	1	Angular alignment	Vibration Analysis	3	6	Perform alignment	
				2	Angular shaft misalignment	1	Angular alignment	Vibration Analysis	3	6	Perform alignment	x
				3	Angular and parallel shaft misalignment	2	Angular and parallel alignment	Vibration Analysis	3	18	Perform alignment	
					Imbalance							
				3	Irregular shaft mass	1	Regulate the mass	Vibration Analysis	4	12	Replacement	
				3	Lack of shaft calibration	1	Calibration	Vibration Analysis	3	9	Perform calibration	
				3	Wash Shaft	1	Replacement	Vibration Analysis	3	9	Replacement	
					Cones							
				4	Cone fixing	1	Fix the cone	Vibration Analysis	3	12	Perform predictive inspection	
					Connecting the equipment to the base							
				4	Leveling of equipment	2	Leveling the equipment	Vibration analysis and leveling check	3	24	Inspect settlement	
				4	Equipment laying structure	4	Fix the equipment	Visual displacement of the equipment	1	16	Inspect settlement	
				4	Fixing the equipment	4	Fix the equipment	Visual displacement of the equipment	1	16	Using standard screws	
					Environment							
				2	Humidity	1	Correct infiltrations	Existence of fungi	1	2	Perform isolation	
				2	Dirtiness	1	Check the cleaning of the equipment	Existence of dirt	1	2	Perform a clean-up	
					Bearings							
				2	Seals	4	Leak control or replacement	Leak checking	1	8	Perform lubrication	
				2	Lubrication	1	Lubricate bearings	Excessive friction in bearings	2	4	Perform lubrication	
				2	Induced currents in bearing	1	Improve housing insulation	Measurement of the current in the rotor	4	8	Improve housing insulation	
					Electric Power							
					Engine windings temperature		Download load	Temperature Measurement of windings			Do not exceed the recommended load	
				3	Phase imbalance	2	Balance phases	Phase measurement	1	6	Systemic phase control	
					Load							
				4	Overload	4	Respect the maximum recommended load	Analysis of Vibrations, Temperature and Electric Currents of the Motor	3	48	Follow the equipment standard	
				4	Start	4	Start at the proper speed in sequence and in conjunction with the start-up of the previous pumps	Tachometers / Voltimeters and Amperimeters	4	64	Follow the manufacturer's procedures	

Figure 9 - FMECA analysis of fissure or breakage of shaft and cone.

Based on the Ishikawa diagram, the FMECA analysis, and a subsequent vibration analysis, it was possible to draw the conclusion that the chip pump machine's inadequate seating was the real cause of the defects. This seating problem resulted in excessive vibration, which cracked the shaft and harmed the cones.

6.3 Chip pump system monitoring

The company made the decision to build a monitoring system over the crucial variables noted in the Ishikawa and FMECA analyses once the issue was fixed.

The system is equipped with the following sensors to keep an eye on its health: accelerometers, temperature sensors in the motor windings, oil circuits, and roller bearings, load sensors, pressure sensors, and flow meters. Every minute, sensor measurements are recorded.

Figure 10 provides a comprehensive view of the variables that are always being tracked.

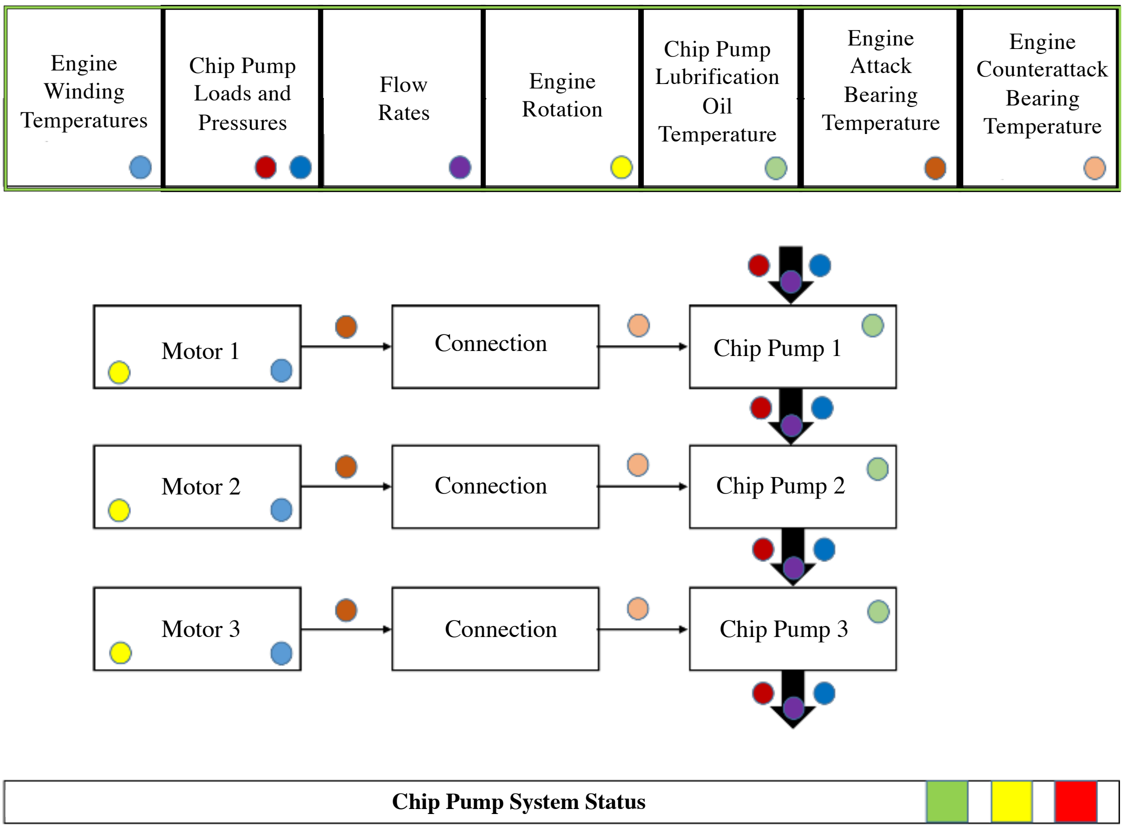

Figure 10 - Global vision of the variables that are continuously monitored.

The current objectives are limited to monitoring the state of the equipment and predicting future values. A short-term forecast is made five days in advance to predict future values. The forecast for a longer period is made for the next three months.

When peaks or ebbs in the predicted parameters are noticed, the company can anticipate malfunctions based on the forecast results. The company lowers its operating and maintenance expenses by anticipating and preventing these failures.

6.4 Condition variable global analysis

The average and amplitude of the condition monitoring variables were the subject of the initial analysis. All the following variables were examined: vibration; temperature of the attack and counterattack bearings; oil level; motor windings temperature; load; pressure; flow; and rotational speed. The average, minimum, and maximum values, as well as the times at which the two latter occurred, were all calculated.

The temperature and pressure of the three chip pumps are presented in tables in this section (Tables 5-8). The system's pressure significantly increases as the mixture's density increases.

TABLE 5 - ANALYSIS OF PRESSURES BEFORE AND AFTER EACH CHIP PUMP

	PRESSURE BEFORE CHIP PUMP 1		PRESSURE AFTER CHIP PUMP 1		PRESSURE AFTER CHIP PUMP 2		PRESSURE AFTER CHIP PUMP 3	
Year	**Average Value (kPa)**	**Max. Value (kPa)**	**Average Value (kPa)**	**Max. Value (kPa)**	**Average Value (kPa)**	**Max. Value (kPa)**	**Average Value (kPa)**	**Max. Value (°C)**
2017	-	-	357.67	1031.45	678.12	1492.63	1007.29	1201.84
2018	46.57	160.16	357.19	565.89	685.23	1547.28	995.44	1187.94
2019	48.31	162.68	361.64	558.97	676.26	909.47	1023.69	1244.60

Table 6 presents a comparison of engine winding temperatures from May 2017 to August 2019.

TABLE 6 - CHIP PUMP LUBRICATING OIL TEMPERATURE

	CHIP PUMP 1		CHIP PUMP 2		CHIP PUMP 3	
Year	**Average Value (°C)**	**Max. Value (°C)**	**Average Value (°C)**	**Max. Value (°C)**	**Average Value (°C)**	**Max. Value (°C)**
2017	36.1	43.11	37.47	51.42	36.56	44.56
2018	-	-	42.54	108.19	42.71	64.87
2019	54.18	61.72	54.07	62.5	54.28	62.23

TABLE 7 - TEMPERATURE ANALYSIS OF THE DRIVE PUMP BEARING FOR THE CHIP PUMP

	CHIP PUMP 1		CHIP PUMP 2		CHIP PUMP 3	
Year	Average Value (ºC)	Max. Value (ºC)	Average Value (ºC)	Max. Value (ºC)	Average Value (ºC)	Max. Value (ºC)
2017	53.58	76.87	58.99	83.36	68.53	95.58
2018	63.75	94.62	71.69	93.29	72.16	105.78
2019	62.02	89.37	64.33	95.71	68.46	105.32

TABLE 8 - TEMPERATURE ANALYSIS OF THE COUNTERATTACK BEARING TO THE CHIP PUMP MOTOR

	CHIP PUMP 1		CHIP PUMP 2		CHIP PUMP 3	
Year	Average Value (ºC)	Max. Value (ºC)	Average Value (ºC)	Max. Value (ºC)	Average Value (ºC)	Max. Value (ºC)
2017	27.36	46.22	27.71	53.90	24.45	57.38
2018	27.83	55.95	27.93	58.86	25.60	56.68
2019	27.98	48.32	28.78	50.77	26.70	53.67

6.5 Dataset smoothing, normalization, and filtering

A dataset provided by the company was used to train the models and predict the parameters. The dataset includes sensor data with a 1-minute sampling period from 2017 to 2020.

The dataset included the following 11 variables: vibration, pressure, velocity, temperature of the U, V, and W windings, oil temperature, flow, temperature of the attack-roller bearing, temperature of the counterattack-bearing, and load. It should be noted that, since the load is only used as an input to the system, there will not be a forecast for it.

The last known value for each variable in the dataframe was used to replace any missing or null values found throughout the entire dataset.

Then, using a sliding window with the predetermined window width (wn, in samples), a median filter was applied. Finally, using the Python StandardScaller library, the data of all the study variables were normalized; [0, 1] was the normalization range used.

6.6 Short-time forecast

The short-term forecast is based on a self-adaptive exponential smoothing model using Equation 8.

$$S_{t+1} = \alpha_t \times X_t + (1-\alpha_t)\, S_t \tag{8}$$

Where:

St+1 is the expected value for time t+1

αt is the the Auto Adaptive Smoothing Coefficient for time t (0≤ αt ≤1)

Xt is the variable value at time t

St is the expected value for time t

The Auto Adaptive Smoothing Coefficient αt is calculated through Formula:

$$\alpha_{t+1} = \mathrm{Min}(1, k_t) \tag{9}$$

Where prediction error Et is given by:

$$E_t = X_t - S_t \tag{10}$$

and,

$$k_t = |A_t/M_t|, \text{ if } M_t>0 \textit{ otherwise} \tag{11}$$

$$A_t = \beta\times E_t+(1-\beta) \times A_{t-1},\ 0\leq\beta\leq1 \tag{12}$$

$$M_t = \beta\times|E_t|+(1-\beta) \times M_{t-1},\ 0\leq\beta\leq1 \tag{13}$$

Et is the forecast error for time t. β is a parameter of the algorithm – a larger value will result in faster response of the filter.

Figure 11 shows the result of the short-time prediction algorithm for the vibration variable.

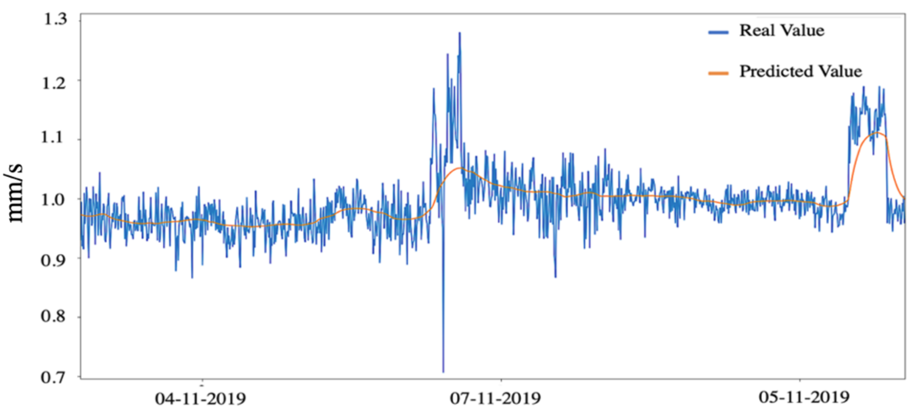

Figure 11 - Short-time prediction algorithm for the vibration variable.

6.7 Input vector creation

To create the input vector for the neural network, a sliding window of width *wn* is applied.

Table 9 represents a sliding window *W*, of width *wn*, applied to the time series *u*, so that samples *wn* of the sequence *u* are selected to create the input to the neural network.

TABLE 9 - A SLIDING WINDOW *W*, WITH SIZE *WN*, IS APPLIED TO THE TIME SERIES *U*

u[n - ...]	u[n-*wn*-2]	u[n-*wn*-1]	u[n-wn]	...	u[n-1]	u[n]
			W[*wn*]	...	*W*[2]	*W*[1]

Applying the sliding window *W* to sequence *u*, *wn* samples, from *u[n]* to *u[n-wn],* are selected to create the input vector to the neural network.

Once the *wn* samples are selected, a signature *Sn* of the window is calculated to feed as input to the neural network.

The signature *Sn* comprises the mean value of the window (m*w*), the Standard Deviation (std*w*), the median (med*w*) of the *wn* samples, and the Power Spectrum Density (psd*w*), as represented in Equation 14.

$$Sn(n) = [m_w, std_w, med_w, psd_w] \tag{14}$$

Once the sequence of signatures of each window is created, a transformed dataset is constructed, with the structure represented in Table 10.

TABLE 10 - TRANSFORMED DATASET, CONTAINING THE SIGNATURES OF EACH WINDOW *WN*

Sn[n - ...]	*Sn[n-5]*	*Sn[n-4]*	*Sn[n-3]*	*Sn[n-2]*	*Sn[n-1]*	*Sn[n]*

To train the model to predict future values, a time gap *g*, in samples, is applied to create the desired output vector. The vector is introduced, then the predicted value *p* for time *n+g* is a function of *Sn[n],* as shown in Equation 15.

$$p\,[n + g] = f(Sn[n]) \tag{15}$$

To train the model to make the predictions, it is used the correspondence between each sample of the sequence of signatures, *Sn[n],* in the dataset, and the value of the sequence a number *g* of samples ahead in the future, *Sn[n+g]*. In the experiments, *g* was the number of samples in 90 days (*g = number of samples stored in 3 months*). Table 11 represents the prediction model, where the signal signature at time *n-3* is used to predict the value at time *n*.

TABLE 11 - REPRESENTATION OF THE PREDICTION MODEL WITH A TIME GAP OF 3. SIGNAL AT TIME *N-3* IS USED TO PREDICT SIGNAL AT TIME *N*

wn[n - ...]	*wn[n-5]*	*wn[n-4]*	*wn[n-3]*	*wn[n-2]*	*wn[n-1]*	*wn[n]*
p[n - ...]	*p[n-2]*	*p[n-1]*	*p[n+0]*	*p[n+1]*	*p[n+2]*	*p[n+3]*

6.8 Long-time forecast

Eighty percent of the dataset was used for training, and twenty percent was used for testing. On a machine with an Intel Xeon E5-2680v2 CPU, each training iteration takes six to eight hours.

The machine learning model used to make the predictions was an Artificial Neural Network, namely MLPRegressor from Sklearn library. The neural network, after several training procedures, achieved good results.

Figure 12 shows the results of the counterattack bearing temperature prediction. Figure 13 shows the results of the attack bearing temperature prediction. The signal is in blue, the prediction in orange.

These results were obtained using a multilayer neural network with two hidden layers, with 200 and 10 neurons, respectively, using the ReLU activation function. The sliding window applied on the data is 7 days. In order to improve the stability of the predicted values, they were smoothed using a median filter with a 20-sample window.

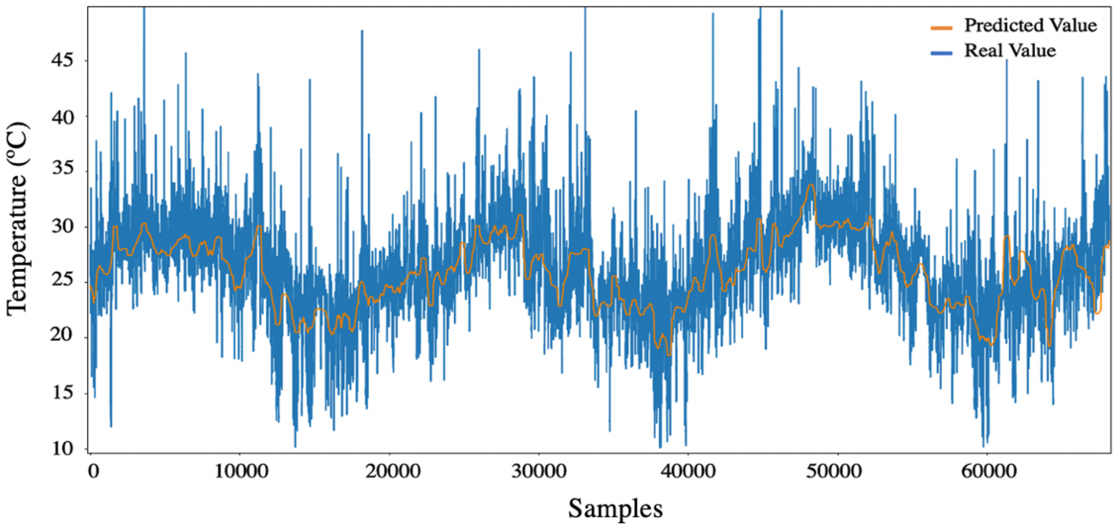

Figure 12 - Results of the counterattack bearing temperature prediction.

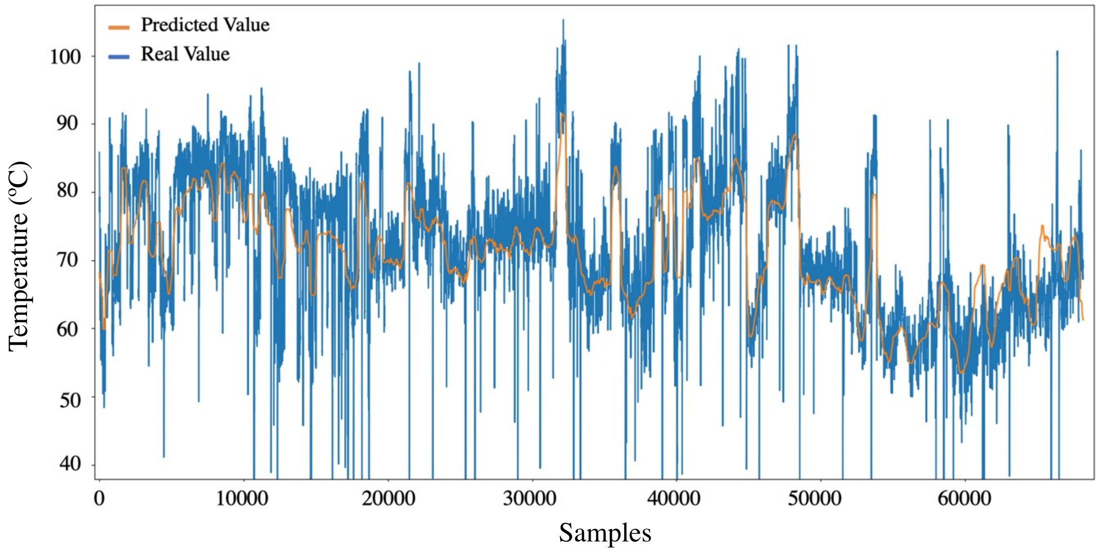

Figure 13 - Results of the attack bearing temperature prediction.

Table 12 shows the Mean Absolute Percentage Errors of all the predicted variables so that it is possible to better understand the neural networks' effectiveness.

TABLE 12 - MAPE OF ALL THE PREDICTED VARIABLES

VARIABLE	MAPE
Vibration	9.47
Pressure	1.59
Velocity	1.32
Winding temperature U	4.26
Winding temperature V	4.32
Winding temperature W	4.47
Oil	5.34
Flow	3.35
Temperature Attack	6.63
Temperature Against Attack	9.72

With an average error rate of less than 10%, the equipment status can be predicted using the neural network's prediction of the variables listed above in the table.

6.9 User interface

The end user interface was developed using semaphores, numerical values, and graphs with the goal of providing the user with an intuitive understanding of the overall behaviour of the system.

In this colour scheme, the anomaly is represented in red, the alert in yellow, and the good working in green. This colour scheme was chosen to be similar to the traffic light system used on roads, making it simple for everyone to understand and assimilate.

This system makes it quick, simple, and easy for the operator to determine the equipment's current operational state, which can help to avoid serious malfunctions or failures (when it is yellow or red).

The company's technicians have suggested limits based on their prior knowledge and data provided by the manufacturer.

The system will be coloured Green (Normal operation), Yellow (Alert) and Red (Fault), as shown in Figure 14.

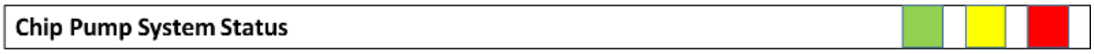

Figure 14 - Colour system for alert.

6.10 Summary

Failures in industrial facilities can result in significant losses or even put people and property in danger. In a case study involving chip pumps, it was possible to identify issues and create a model to forecast future behaviour using a dataset consisting of sensory data collected over roughly three years and factory inspections. The final chip pump in a set of three was under significant strain, according to a FMECA analysis. This effort was warranted because the load had to be transported vertically, as opposed to horizontally by the predecessor chip pumps.

The same chip pump has settlement issues that cause its vibration to increase exponentially with greater efforts. The chip pump's useful life is much shorter than necessary as a result of the vibration and increased effort that go along with it. This causes the chip pump's shaft to experience more stress than desirable.

The ability to predict sensor values up to three months in advance is very helpful for management of equipment maintenance decisions. The forecast's temporal component is completely novel because only short and medium-term forecasts were found when researching the state of the art.

Neural network predictions turned out to be accurate for this kind of issue. All variables had mean absolute percentage errors that were less than 10%.

Given the outcomes, this work gave the industry a chance to plan its scheduled maintenance stops with greater knowledge. Because it prevents unforeseen breakdowns, this decision-making assistance greatly improves asset availability and lowers costs. The method has a drawback in that it is dependent on previous sensory information. Forecasts can be more uncertain when one or more important variables change, for example as a result of component differences, environmental changes, or other changes.

By providing the neural networks with relevant training data, this methodology can be applied to different pieces of machinery, though there is no assurance that this will yield the same results in different assets. Depending on the kinds of patterns in the data, the results may be better or worse.

Chapter 7

Pulp paper press behaviour prediction

This chapter describes the work developed to predict a chip pump's behaviour. Most of this work was published in a scientific conference [147] and in a scientific journal paper [148].

7.1 Introduction

The research described in the present chapter aims to propose a model to forecast sensor values of an industrial pulp paper press for 15 days, 30 days, and 90 days.

The goal was to compare the performance of multiple prediction models, including neural networks and other machine learning methods, optimizing different features and architectures. It was defined that the forecasts of most variables should have MAPE errors below 10%.

7.2 Dataset and pre-processing

For the present analysis, a paper pulp company provided a three-year data set containing the time series of six variables: electric current (Sensor 1), oil level (Sensor 2), pressure (Sensor 3), rotation velocity (Sensor 4), temperature (Sensor 5), and torque (Sensor 6). All data were collected from sensors with a sampling frequency of one minute.

The dataset contains several repeated values as well as discrepant samples (outliers) that may be due to reading errors or production line stops. Upper outliers might have resulted from errors in sensor reading or recording, while lower outliers are most probably a result of those causes along with programmed or non-programmed downtimes.

In a predictive algorithm, the quality of the underlying data is of extreme importance. Poor quality data implies inaccurate results. For that reason, the dataset was previously processed to increase confidence in the results and facilitate convergence during the learning process. The previous chapter dealt with the study of a Chip Pump, this chapter deals with a Pulp Paper Press, which is why there is new data. Due to this reason, there is a need for further data pre-processing.

The units of the several variables are as follows: electric current is measured in Ampère (A); oil level is measured in percentage of full tank (% Tank); pressure is measured in Pascal (Pa); rotation velocity is measured in rotations per minute multiplied by 1000 (RPM×1000); temperature is measured in degrees Celsius (°C); torque is measured in Newton-meter (N×m). Figure 15 shows the six variables of the time series.

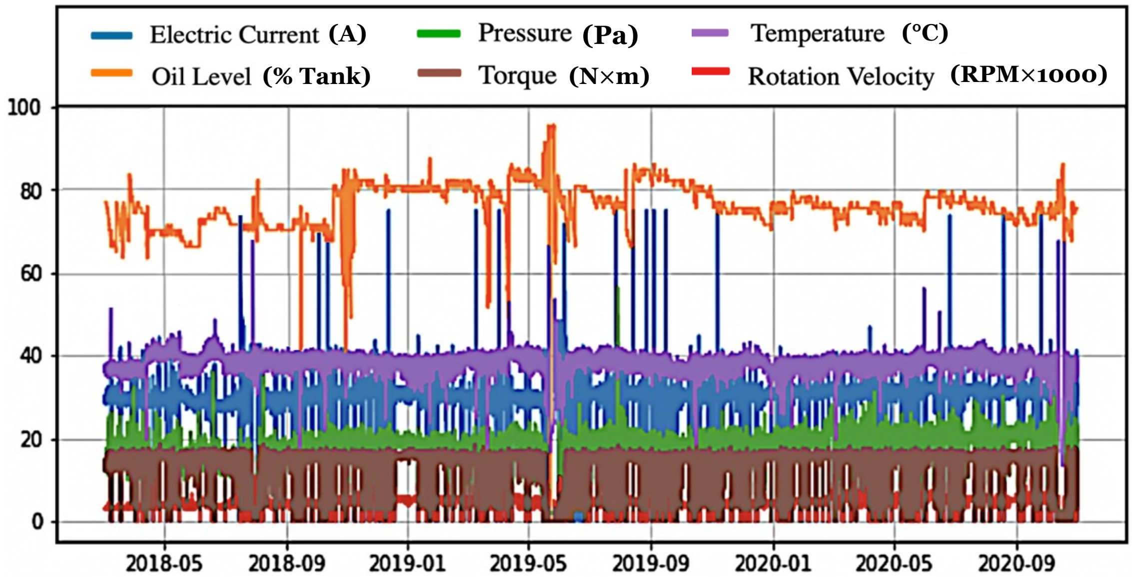

Figure 15 – Original data with a sampling period of 1 min.

Figure 15 shows that there are many outliers in the dataset (e.g., null, zeroes and repeated values); repeated values arise by sensory errors or even at the time change. The outliers are replaced by the average value of the variable in the sliding window before the outlier. This method has been described in more detail by Mateus *et al.* [149].

Therefore, the dataset was filtered using a Python algorithm developed by the authors as follows:

- Repeated values, as well as lower and upper discrepant values, were removed and replaced by the corresponding variable average value;
- Values beyond three standard deviations from the first and third quartiles on each variable were also replaced by the mean value of the variable in question.

Figure 16 shows the six-time series of the variable values collected by the sensors after being filtered by the previously described pre-processing method. As the chart shows, there are no more sudden variations, probably representing outliers, which could impair the machine learning process. Previous studies show that pre-processing discrepant data improves the learning process [43].

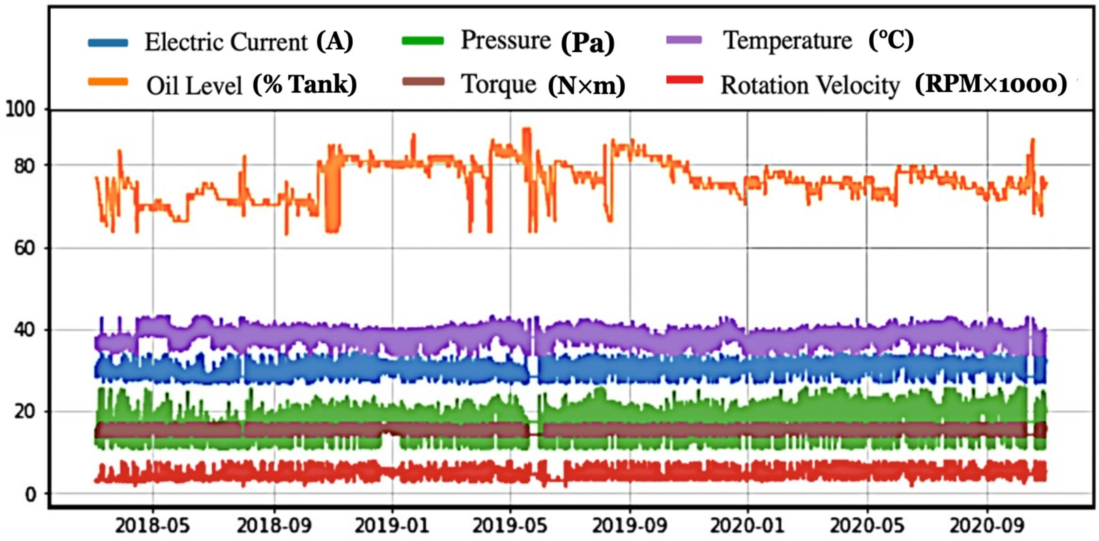

Figure 16 - Data filtered from outliers with a sampling period of 1 min.

7.3 MLP with Ratios (MLP RATIOS)

The sliding window process has already been explained in Section 6.7. A sliding window encompasses a continuous subset of a time series dataset, which slides over the latter with a certain step. The window size determines the number of data point samples from the whole dataset to be included in this subset.

The window started with the first w data points (samples) of the time series and slid to the end of the series, in steps of one for an overlapping window, or steps of w samples for a non-overlapping window.

For n variables (n = 6 in this case), data from each variable i in each sliding window with size w were grouped into 15 equal-width bins j and the corresponding absolute frequency values (vector Si,j), along with the respective average (Ai), median (Mi), standard deviation (Sdi), variance (Vi) statistics and, finally, 30 ratios ($Ri1,i2$), between each pair of variables, where $i1 \neq i2$ make up the input vector I that feeds the neural network, as represented in Equation 16.

$$I = \left(S_{1,1}, S_{1,2}, \dots, S_{n,15}, A_1, \dots, A_n, SD_1, \dots, SD_n, M_1, \dots, M_n, V_1, \dots, V_n, R_{1,2}, \dots, R_{n,n-1}\right) \quad (16)$$

For each window w: Si,j represents the value of variable i in bin j; Ai is the average value of variable i; Mi is the median value of variable i; SDi is the standard deviation; Vi is its variance; and, finally, Ri1,i2 represents the ratios between the variables collected by the sensors.

Data inputs were further standardized using the Standard Scaler library from Sklearn, before being fed into the ANN model. Standardization is a technique applied in the preparation of data, with the objective of placing them in a range of common values.

Note that each variable *i* was predicted not only from its respective past data but also from the other five variables.

In order to make it easier to identify the ANN throughout the thesis document, this ANN was called MLP RATIOS.

Time series data were separated into two groups: the first 80% from 1 January 2018 to 27 May 2020 were used for training the model; and the remaining 20% for carrying out the tests.

The architecture type chosen for the neural network is the multilayer perceptron, one of the most popular feedforward architectures, implemented using the Python Sklearn library named MLPRegressor.

The MLPRegressor uses multiple hyper parameters to optimize the generalization of the network model for prediction. Several architecture combinations were tested to find the best possible network configuration.

Adam solver was chosen as the algorithm for optimizing ANN weights, since it is a graph-based optimization algorithm recommended for large datasets, using a logistic sigmoid as the activation function, as represented in Equation 17, where x is the independent variable.

$$f(x) = \frac{1}{(1+exp(-x))} \tag{17}$$

Creation of the vector and tests to find the best value for each alternative ANN configuration took about two days to perform, due to the complexity and size of the dataset, in a shared GPU server AMD EPYC 7552, with 16 Core + Nvidia Tesla T4/V100S.

The authors tested alternative networks with one, two, three, and four hidden layers. Using one layer only yielded quite bad results, while using four layers was quite time consuming. Results from using two and three layers were quite similar so, the authors chose two layers only as the training time was faster without loss of accuracy. Alternative ANN configurations further varied the number of neurons in each layer.

A network with two hidden layers (150 and 75 neurons) was chosen, as it showed results very similar to the three hidden layers' architecture but was much faster. Figure 17 shows the chosen ANN architecture.

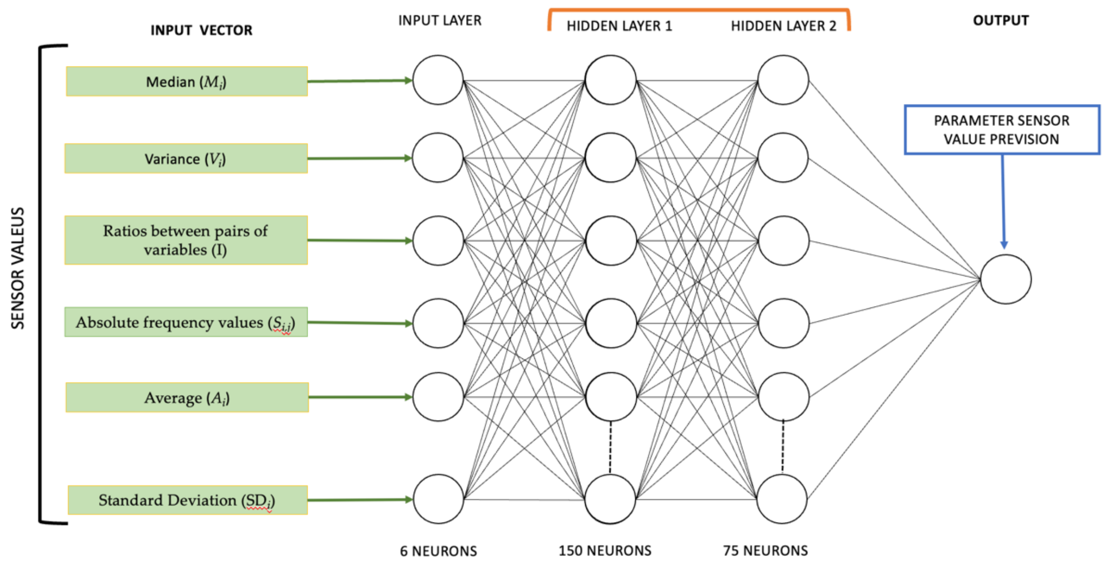

Figure 17 - Architecture of the artificial neural network.

7.4 MLP without ratios (MLP TEPEN)

This architecture is very similar to the previous one, the only difference is that this one does not have data of the ratios between pairs of variables.

This change was made as a kind of ablation study, to analyse the impact that ratios between variables have on prediction results.

This neural network was presented at a congress, in China, “The Efficiency and Performance Engineering Network 2021” (TEPEN 2021) and the “Sixth International Conference on Maintenance Engineering”.

Because of this, for future comparisons, this network was designated as TEPEN ANN. Figure 18 shows the neural network without ratios between pairs of variables.

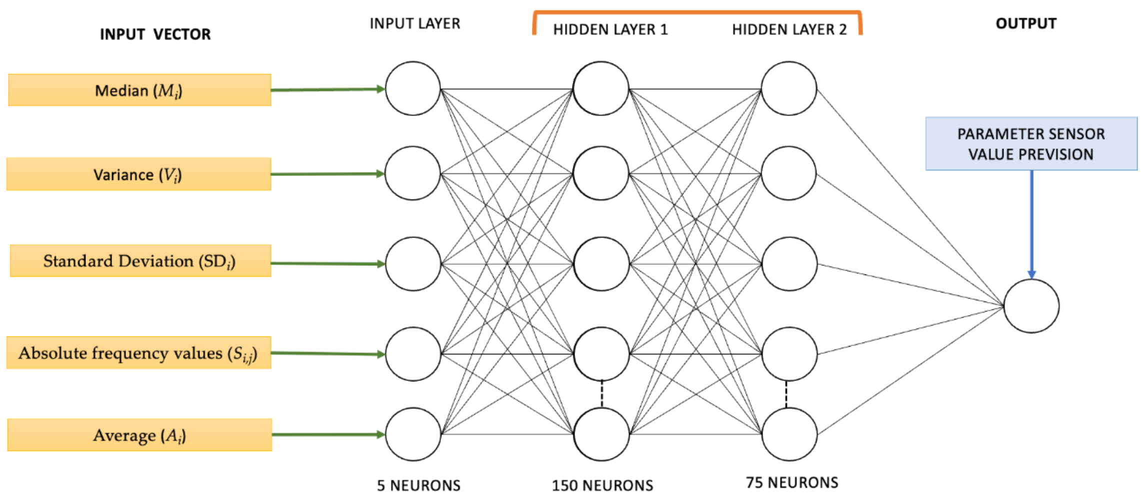

Figure 18 - Architecture of the TEPEN MLP ANN.

7.5 Experiments and results of MLP TEPEN

This sub-chapter describes the experiments carried out during the development of this research study along with some of the best results obtained through several tests. One of the main goals in creating the MLP TEPEN was to study the impact of the overlapping sliding windows on the network learning time and on the prediction results. Table 13 shows the results of the tests developed using the overlapping window.

TABLE 13 - RESULTS OF PREDICTION TESTS WITH OVERLAPPING SLIDING WINDOW

Window Size (Samples)	Oil		Current		Temperature		Torque		Pressure		Velocity	
	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS
2440	10.51	0.70	1.73	0.13	3.26	0.49	0.62	0.06	9.53	1.69	2.14	0.25
1440	7.23	0.48	1.67	0.11	3.67	0.29	0.80	0.06	7.57	1.18	2.84	0.19
720	5.94	0.19	1.56	0.09	3.55	0.19	0.54	0.06	4.74	0.78	1.84	0.14
360	5.22	0.07	1.45	0.07	3.28	0.13	0.56	0.05	5.39	0.51	1.68	0.04
180	5.44	0.13	1.15	0.06	2.59	0.10	0.55	0.04	4.94	0.36	1.43	0.08
144	5.69	0.12	1.15	0.06	2.46	0.09	0.54	0.04	4.62	0.34	1.51	0.08
90	5.02	0.11	1.08	0.05	2.46	0.09	0.56	0.04	4.53	0.31	1.43	0.07
45	5.24	0.20	1.08	0.06	2.23	0.11	0.56	0.03	4.97	0.37	1.46	0.07
24	5.60	0.41	1.14	0.07	2.29	0.19	0.52	0.03	4.96	0.69	1.49	0.08
12	5.76	0.86	1.10	0.11	2.16	0.34	0.53	0.03	4.70	1.25	1.61	0.13

In Table 13, it can be seen that the window size of 2440 samples is the one that obtains the worst loss and MSE results for all variables. In general, the window can be reduced to six hours (360 samples) or, in some cases, up to 12 minutes (12 samples), since they all present good prediction results.

Figure 19 shows the rotation velocity learning history, using a 12-sample window with overlap. The curve shows that the model learns in the first epochs. Through the curve it is concluded that the learning of the Neural Network is quite fast, as it learns most of the patterns and the loss decreases to almost zero in only three epochs. This model presents a Mean Squared Error of 1.10 (Table 13).

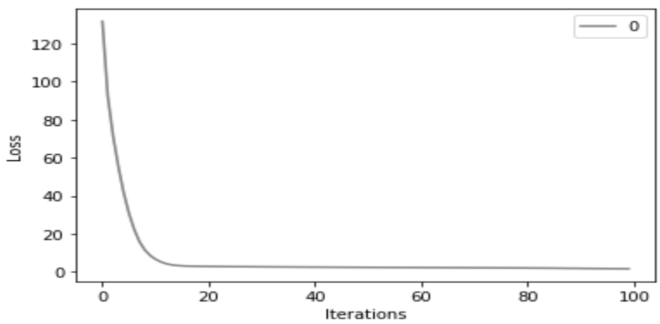

Figure 19 - Rotation velocity learning curve, using window with overlap.

TABLE 14 - RESULTS OF PREDICTION TESTS USING NON-OVERLAPPING SLIDING WINDOW

Window Size (Samples)	Oil		Current		Temperature		Torque		Pressure		Velocity	
	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS	MSE	LOSS
2440	66.93	2016.18	20.61	156.63	26.90	348.65	6.52	9.13	8.12	17.95	1.00	0.19
1440	66.19	1985.44	19.06	153.46	25.40	332.34	5.34	9.03	8.31	18.00	0.65	0.36
720	65.94	1945.29	18.79	153.35	25.32	329.65	4.99	9.08	8.21	20.09	1.19	0.59
360	54.69	1327.35	10.13	44.62	15.58	135.72	1.03	0.34	4.50	10.41	1.13	0.75
180	43.93	836.25	4.64	7.71	8.25	42.00	0.53	0.14	4.25	8.76	1.07	0.73
144	43.95	837.15	4.69	7.81	8.28	41.90	0.52	0.14	4.16	8.04	1.14	0.83
90	25.99	267.54	1.30	0.65	1.61	3.39	0.46	0.12	4.04	7.98	1.18	0.76
45	8.60	26.05	1.02	0.53	1.59	1.92	0.47	0.13	3.80	7.21	1.15	0.57
24	4.64	11.07	0.74	0.44	1.58	1.68	0.49	0.11	4.15	3.58	1.41	0.17
12	5.45	3.61	0.78	0.26	1.77	0.78	0.48	0.10	4.72	1.13	1.56	0.09

Table 14 presents the corresponding results for the tests carried out with a non-overlapping sliding window. The same table shows the learning problems resulting from Neural Network when using large size windows without overlapping. However, window sizes with 24 and 12 samples generate results that are already satisfactory.

In the absence of the overlapping technique, the Neural Network receives much less input data per epoch, which delays the learning process. This problem is, in part, overcome by reducing the window size, making the network receive less input data samples per epoch. However, when using smaller windows, it can be more difficult to catch larger patterns.

Although the learning is slow, the network can learn, presenting a loss value of 0.09 in the window of 12 samples for the variable in question.

Figure 20 illustrates the learning history of the network to predict Rotation Velocity using a window size of 12 non-overlapping samples. A slow slope is evident, demonstrating a slower learning rate when compared to an overlapping window.

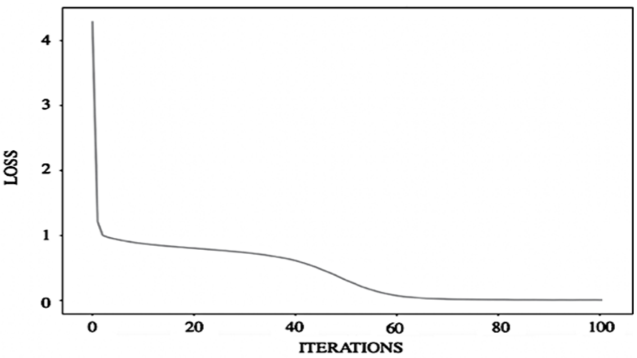

Figure 20 – Rotation velocity learning curve using a non-overlapping window.

7.6 Tests and result of ANN ratios

7.6.1 Results for overlapping sliding windows

The algorithm developed was tested for forecasts of 15, 30, and 90 days in advance. The training took up to 1000 learning epochs in each of the tests, with overlapping sliding window sizes w of 720, 1440, 2880, and 4320 samples.

Best results were achieved for window sizes with either 12 or 24 h (720 or 1440 samples). Hence, detailed results will be presented only for these two window sizes.

Tables 15–17 show the results achieved for the six variables (sensors) in terms of MAPE, MSE, and number of iterations (ITER), which the training requires to be completed.

TABLE 15 - COMPARATIVE MAPE RESULTS.

		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
720 Samples	***90 Days***	3.441	6.231	16.286	4.977	2.696	4.053
	30 Days	2.295	4.642	14.643	4.006	2.332	4.112
	15 Days	2.205	4.306	12.453	3.717	1.734	3.678
		Current	*Oil Level*	**Pressure**	**Temperature**	**Torque**	**Velocity**
1440 Samples	***90 Days***	3.623	6.696	21.426	4.878	2.612	4. 451
	30 Days	2.310	5.124	14.034	4.686	2.049	4.234
	15 Days	2.541	4.319	13.633	4.363	1.864	3.896

TABLE 16 - COMPARATIVE MSE RESULTS.

		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
720 Samples	***90 Days***	1.374	5.804	3.954	2.336	0.638	1.485
	30 Days	1.094	5. 358	4.153	2.201	0.735	1.379
	15 Days	1.080	5.013	4.056	2.060	0.670	1.454
		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
1440 Samples	***90 Days***	1.413	6.476	4.518	2.271	0.621	1.584
	30 Days	1.010	5.385	4.123	2.336	0.632	1.668
	15 Days	1.173	4.739	3.808	2.374	0.643	1.517

TABLE 17 - COMPARATIVE NUMBER OF ITERATIONS.

		Current	Oil Level	Pressure	Temperature	Torque	Velocity
720 Samples	**90 Days**	122	161	363	201	58	164
	30 Days	120	173	353	230	56	159
	15 Days	104	132	425	188	63	190
1440 Samples		***Current***	***Oil Level***	***Pressure***	***Temperature***	***Torque***	***Velocity***
	90 Days	136	164	415	217	64	179
	30 Days	126	144	435	249	64	213
	15 Days	105	115	432	208	71	168

Evaluation results show that it is possible to predict variable (sensor) values with 3 months, 1 month, and 15 days in advance with a reasonable degree of accuracy. Most variables show MAPE errors below 10%.

In general, a window size of 720 samples (12 h) over 1440 samples (24 h), not only has a shorter learning time, but it also yields better accuracy results in terms of MAPE and MSE. Hence, a window size of 720 samples was selected as a good sampling size.

Almost all variables show large fluctuations, including striking peaks such as those shown in Figure 21 and Figure 22. Hence, to stabilize the output and to visualize better actual and predicted time series values on each variable, they were smoothed using a rolling average filter of 1 day.

Figures 21 and 22 show two examples of real time series in blue and 90-day forecasts in orange, after smoothing of 2 minutes using overlapping. Pressure is the most difficult variable and torque is the easiest variable to predict, as shown in Table 13. That is the reason why they were chosen as examples. According to Table 13, *Pressure* was the variable that had the highest MAPE error and torque was the one that overall had the lowest MAPE error.

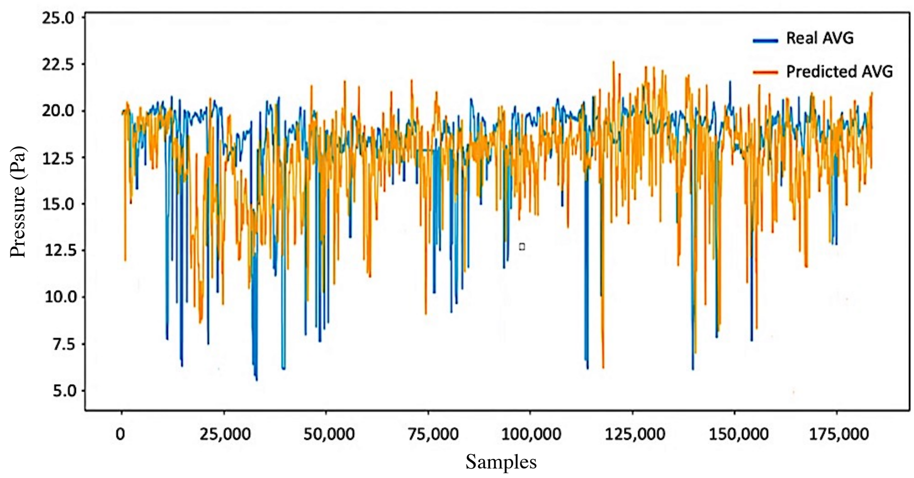

Figure 21 - Graph of Real Pressure signal and 90-day forecast.

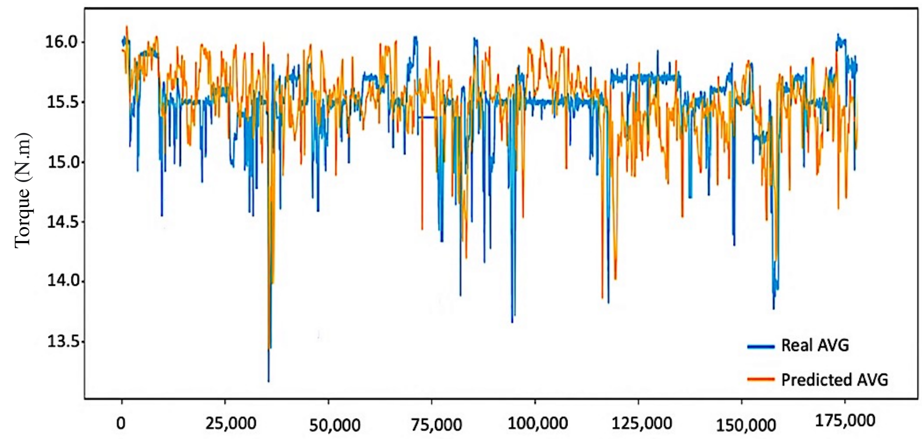

Figure 22 - Graph of Real Torque signal and 90-day Forecast.

7.6.2 Results with non-overlapping sliding windows

Using overlapping windows showed good prediction accuracies for all variables. However, their training times are quite large, taking on average more than two days for each variable (using a Tesla V100S-PCIE-32GB). Hence, non-overlapping windows were assessed to reduce learning time.

Using non-overlapping windows, the input vector in the neural network contains fewer data points, thus making its processing much faster. On average, this method allowed us to reduce the learning time to only seven minutes (using a MacBook Pro M1 from 2020 with 8 GB of RAM with MacOS Monterey).

The use of non-overlapping windows produces worse long-term forecasts (90 days) than the overlapping windows method. However, the short-term (15 days) results are good. It should be noted that the neural network is the same, regardless of whether it is for short/medium or long-term predictions. It is only the data included in the input vector that change.

Figure 23 presents the real signal of the electric current and the 15-days forecast after smoothing, using non-overlapping windows.

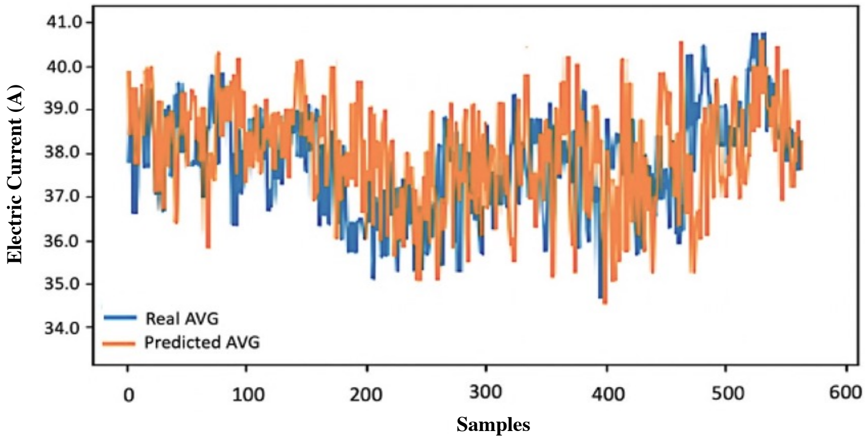

Figure 23 – Results of 15-day forecast using non-overlapping windows for current.

7.7 Analysis of the results

7.7.1 Comparison between MLP TEPEN and MLP RATIOS

The present research corresponds to an optimization of the features of the neural network vector already developed by the authors. The new vector contains new ratios among variables [47]. This is the difference between the two vectors, as is presented in Table 18. This comparison is made for 90-day forecast results.

TABLE 18 - COMPARATIVE MSE RESULTS BETWEEN OLD VECTOR AND NEW VECTOR

		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
720 Samples	*Old vector*	1.564	5.943	4.741	3.553	0.544	1.842
	New vector	1.374	5.804	3.954	2.336	0.638	1.485
		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
1440 Samples	*Old vector*	1.672	7.231	3.670	7.571	0.802	2.843
	New vector	1.413	6.476	4.518	2.271	0.621	1.584

Analysing Table 18, the prediction results of the new feature vector are generally much better than the old vector, except for the torque parameter that maintains identical values.

7.7.2 Comparison between LSTM, GRU and MLP RATIOS

Table 19 shows the comparison of the MAPE results of the prediction models using LSTM, GRU and the feedforward model that were published in papers [83], [87].

The comparison is made by analysing the MAPE errors of 30 days in advance forecast for each variable. The 30-day forecast was selected because it was the time gap defined as the objective for the project.

TABLE 19 - COMPARISON OF MAPE RESULTS

	Current	***Oil Level***	***Pressure***	***Temperature***	***Torque***	***Velocity***
GRU-ReLU	2.52	2.94	9.91	2.84	3.03	15.05
GRU-Sigmoid	**2.22**	**2.72**	**9.29**	2.74	2.88	12.42
LSTM-ReLU	2.42	2.92	10.36	**2.30**	3.72	17.19
MLP RATIOS (720 Samples)	2.30	4.64	14.64	4.01	2.33	4.11
MLP RATIOS (1440 Samples)	2.54	4.32	13.63	4.36	**1.86**	**3.90**

The first row and second row of the column present the results of the GRU prediction models using the ReLU and Sigmoid activation functions. The third row of the column presents the results of a traditional LSTM model using the ReLU activation function. Finally, the last two lines present the results of the MLP neural network developed, presented and explained in the previous chapters.

Analysing Table 19, it is concluded that the current parameter has very similar prediction results in all models. The GRU and LSTM models have similar prediction results, however the GRU-SIGMOID has a slightly lower MAPE error. In the case of the pressure parameter, the GRU models have the best prediction results, and the GRU-SIGMOID is the one that achieves the best results.

In the temperature parameter, although the results do not show a significant difference, it is the LSTM-ReLU model that presents the smallest prediction error. Regarding torque, the MLP models present the best prediction results, with the best model the MLP-1440 SAMPLES. The biggest difference in results occurs in the velocity parameter, where the MLP models present much better prediction results than the other models, obtaining much lower MAPE errors.

MLP networks are simpler than GRU models and, in turn, the GRU network is simpler than LSTM. Observing the results, it is concluded that both prediction models can predict the future values of an industrial paper press, 30 days in advance, with MAPE, in general less than 10%. The difference in results of the velocity variable is noted, where only feedforward networks, despite their simplicity, achieved a MAPE error of less than 5%. In short, there is no better overall model, because each variable has the forecast model that best suits its data, as shown in Table 19. Therefore, for optimal prediction, it is important to plan and optimize the machine learning models, with the best model that achieves the minimum difference error for each variable.

7.7.3 Comparison between XGBoost and feedforward network

In this section, we compare the ML feedforward network with another alternative model, namely XGBoost.

Random forest models have, for many problems, a low error. However, for the present problem, they do not have the ability to follow the trend of the parameters to be predicted. They were tested, but the results were not acceptable, which is why they are not included in this document. The first row displays the results of the XGBoost forecast model.

TABLE 20 - COMPARISON OF MSE BETWEEN XGBOOST AND FEEDFORWARD NETWORK

	Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
XGBOOST	1.075	4.953	0.618	1.304	1.576	3.356
MLP RATIOS (720 Samples)	1.374	5.804	3.954	2.336	0.638	1.485
MLP RATIOS (1440 Samples)	1.413	6.476	4.518	2.271	0.621	1.584

The XGBoost model, in addition having a very fast training, presents good results in the prediction of electric current, pressure and temperature, with the only disadvantage of the difficulty of identifying peaks of values. It is noted that only MLP models can follow oil level parameter trends. The XGBoost algorithm needs four minutes to create the vector and train the model. It should be noted that the input vector is with non-overlapping sliding windows and the machine used is a MacBook Pro M1 from 2020 with 8 GB of RAM with MacOS Monterey. Figure 24 shows an example of a forecast of 30-day values (after smoothing) for velocity using XGBoost.

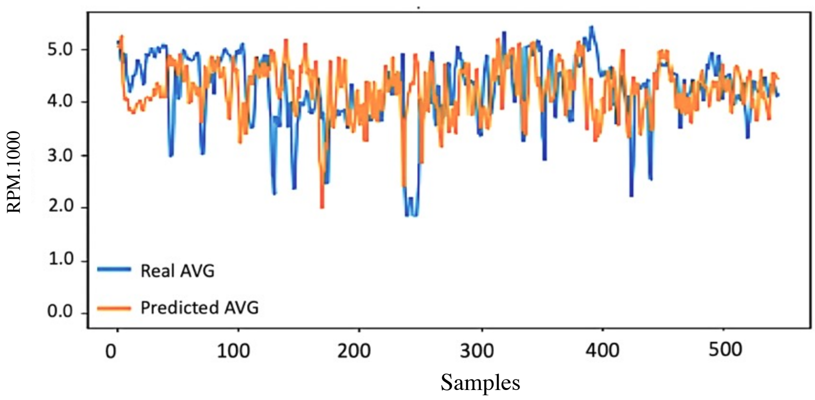

Figure 24 – Results of 30-day forecast for velocity using XGBoost.

7.8 Discussion

XGBoost is a scalable and highly accurate implementation of gradient boosting that pushes the limits of computing power for boosted tree algorithms, being built largely for energizing machine learning model performance and computational speed.

One of the advantages is that the proposed method can perform short, medium, and long-term predictions on other equipment, providing that there are the necessary data and processing power available. The features used and developed to feed the machine learning models should be available in a wide range of industrial equipment. Nonetheless, for each specific situation, the data pre-processing or the neural network architecture may need to be modified and there are not, *a priori,* guarantees of similar results.

This study focuses on predicting the future behaviour of the machine in the short, medium and long-term. Sensor failures will be detected as a malfunction, which needs further analysis and diagnosis.

The results of these predictions can then be processed, for example using a neural network classifier, to classify the condition of the equipment into one of the following states: failure, warning, or good operation, as presented in Section 6.9.

7.9 Summary

Forecasting is very important to make better decisions in maintenance and other areas. Predicting the probable future behaviour of a physical asset brings numerous benefits. For instance, based on accurate predictions, and knowing the respective nominal operating values recommended by the equipment's manufacturer, it is possible to identify anomalies in advance for the equipment, namely in the short, medium and long-term.

The proposed algorithms make it possible to predict the behaviour of a pulp paper press in the long-term, supported by the time series acquired from sensors installed on it. This way, it is possible to optimize long-term programmed stops to avoid production downtimes. These prediction models will be enhanced by the addition of a classification network that will classify the machine into one of three states: normal operation, alert, or malfunction.

This chapter presented a valuable comparison between the input vector of the neural network using overlapping and non-overlapping sliding windows, showing the results of the tests performed, with unequivocal conclusions about the advantages and limitations of each technique used.

Data resampling can make the prediction process much faster since it reduces the dataset considerably.

The use of sliding windows over time series is necessary for training. Overlapping windows offer learning in less epochs. Larger windows make it easier to catch long trends, but the optimal window size needs to be determined experimentally. Overlapping windows offer more input data to the Neural Network in each epoch and thus generating faster learning rates and better prediction results.

Another drawback of non-overlapping windows is the limitation imposed by small datasets. Another major disadvantage of this type of sliding windows is the slower learning rate, especially for small window sizes. Its great advantage is the speed of processing when the input vector is created.

The number of data points present in the neural network input vector, as well as the prediction gap, have a direct impact on the prediction accuracy. On the one hand, a larger sliding window increases the prediction errors, but a smaller window has difficulty in predicting peaks. On the other hand, the larger the prediction gap, the more difficult the prediction becomes.

The results achieved for the short-term, midterm and long-term were comparable to or better than the state of the art. Long-term forecasts using overlapping windows showed very good accuracies, because the predictions of most parameters produce MAPE errors below 10%, that is the objective of the research presented, as shown in Section 9.2. However, they take a large processing time. Short-term forecasts using non-overlapping windows can significantly reduce this shortcoming.

The XGBoost Model produces fast and good results in the prediction of electric current, pressure, and temperature.

Future work includes applying this method to other variables and comparing it against alternative machine learning models for prediction. Additionally, other machine learning methods, such as unsupervised clustering, will be studied to classify the future condition state of the asset based on the forecasts resulting from the proposed ANN.

The number of input features can also be optimized using techniques such as Principal Component Analysis (PCA), or Probabilistic Principal Component Analysis (PPCA). Other approaches, namely Hidden Markov Models (HMM) will also be explored.

Chapter 8

Pulp Paper Press Classification

This chapter describes work to predict a chip pump's probable future state. Most of this work was published in a scientific international journal [150].

8.1 Introduction

Clustering is applied in several areas, especially pattern recognition, data mining, and decision support. One of the most well-known non-hierarchical data grouping techniques is K-means [151],[152].

The present chapter describes a case study where data grouping methods are applied to determine distinct states of operation of a paper press. The method used for this task was K-means. The data set for this purpose is composed of the equipment variables identified in Chapter 7.

Due to the large correlation between the state of an equipment and its lubrication quality, it was decided to develop a classification algorithm for the lubricants of the press. In this study we analyse the following oil parameters: Viscosity at 100 °C, PQ Index, TAN (Total Acid Number), Al, Cr, Cu, Fe, Na, Ni, Pb, Si, Sn.

The present research aims to propose a model to determine the probable future states of an equipment, based on the data set presented in Section 7.4. State classification is done using neural networks, which have an input vector that uses threshold lines in the various sensors according to the manufacturer's recommendations to classify the asset's condition.

The equipment's states of operation must be determined using a clustering algorithm, such as k-means.

One of the objectives of this study is to develop a lubricant classifier algorithm based on neural networks with errors below 5% compared to the results of human experts.

8.2 Equipment nominal operation zones

A warning system similar to the one described in Section 6.9 was created, but this time on a paper press.

According to the manufacturer's user manual, the equipment is recommended to work between a predefined range for all six variables. Figure 25 depicts yellow lines representing the lower and upper temperature thresholds of the equipment when working in its normal functioning zone. Likewise, when the equipment works beyond the red lines, it generates a red alert indicating that it is working in its failure zone and hence needs urgent attention. When the equipment is working between the yellow and red lines, it generates a yellow alert informing that the equipment needs to be checked for possible overload or anomaly.

The dataset was enriched with this classification, namely indicating whether each variable value was within the range of normal, alert, or high risk of failure operation values provided by the equipment's manufacturer.

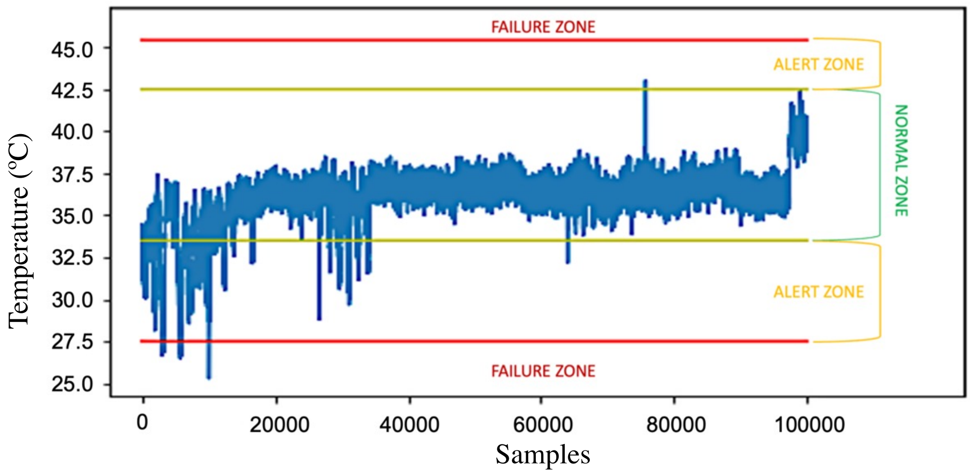

Figure 25 - Normal, alert, and failure zones for the variable temperature.

8.2.1 Sensor values predicted at 30 days

The aim of this study was to classify the state of the equipment 30 days in advance. The dataset was enriched with 30-day forecasts of each variable from a previously developed and validated neural network that can predict sensor variable values at 30 days with a MAPE error of less than 10% [147], [148]. Figure 26 illustrates the respective predicted time series for the six variables (sensors). This forecast was described in detail in the previous chapter.

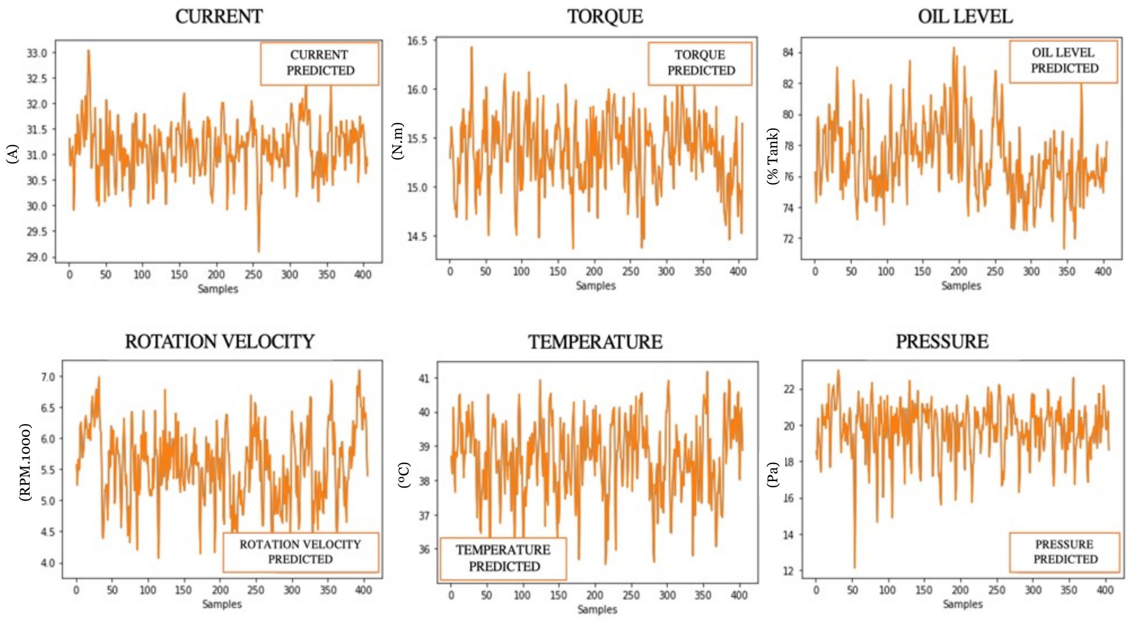

Figure 26 - Predicted time series of sensor value.

8.2.2 Oil Database

As shown in Chapter 5, oil analysis is an extremely important tool in predictive maintenance. Based on this, it is possible to evaluate the condition of the fluids and, by consequence, the equipment.

Therefore, to assess the reliability of assets and increase their availability, in addition to clarify the state of the equipment considering its sensors, it is extremely important to know the state of its lubricants.

It should be noted that, through oil analysis, it is also possible to identify problems before they become failures. Due to the above-mentioned reasons, the authors also decided to perform an algorithm for fluid classification.

The present classification has a database that contains 179 oil samples, and each oil sample contains the analysis of 12 parameters of the lubricant. The parameters under analysis are Viscosity at 100°C PQ Index, TAN (Total Acid Number), Al, Cr, Cu, Fe, Na, Ni, Pb, Si, Sn. Table 21 presents a summary of statistical parameters of the variables used in this study.

TABLE 21 - METRICS OF OIL ANALYSIS

	Units	Mean	Min	Max	Var	Std
PQIndex	ppm	131.78	0.00	6732.00	396718.55	631.63
TAN (Total Acid Number)	mgKOH/g	1.26	0.18	2.85	0.26	0.52
Al Content	ppm	1.30	0.00	15.00	8.00	2.84
Cr Content	ppm	5.59	0.00	2.,00	34.02	5.85
Cu Content	ppm	9.16	0.00	243.00	815.87	28.65
Fe Content	ppm	260.17	2.00	1231.00	91004.30	302.55
Na Content	ppm	5.21	0.00	38.00	25.82	5.10
Ni Content	ppm	4.20	0.00	26.00	17.16	4.16
Pb Content	ppm	0.51	0.00	30.00	6.25	2.51
Si Content	ppm	2.39	0.00	22.00	8.10	2.85
Sn Content	ppm	1.07	0.00	8.00	2.62	1.62
Viscosity at 100°C	m²/s	3035.76	954.40	4146.20	168647.64	436.90

Analysing Table 21, key variables such as Index PQ, TAN, Fe and Viscosity at 100 ºC have a large volatility. This volatility is a consequence of the differences in the state of the oil in the various analysis. Figure 27 graphically shows the variability of those parameters. The X axis represents the oil analysis samples, and the Y axis represents the results of each parameter normalized between zero and one.

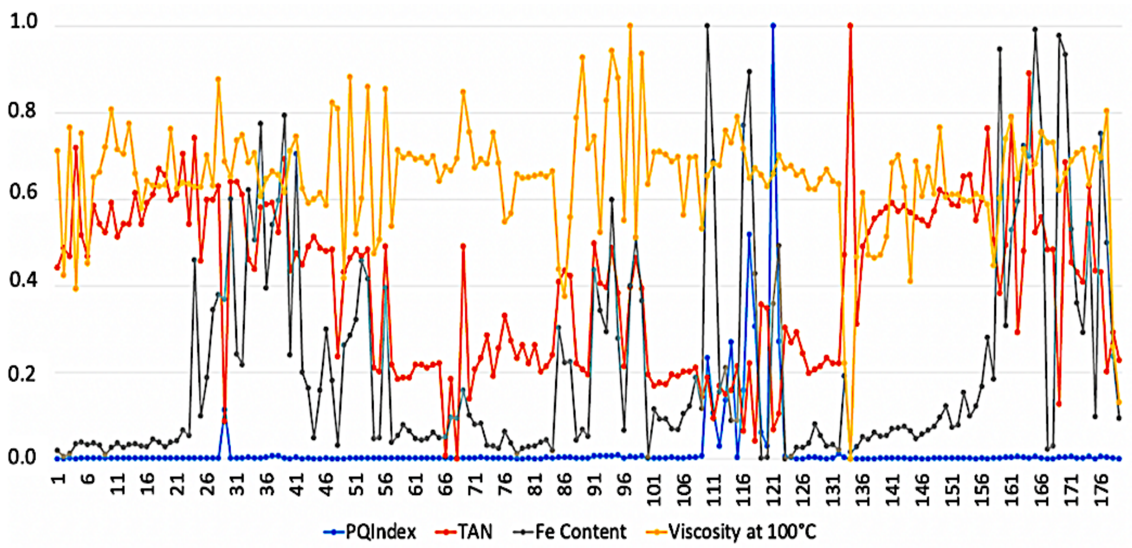

Figure 27 - Graph of all oil analysis (Index PQ, TAN, Fe and Viscosity at 100 ºC).

8.3 Methodology

This section describes the methodology used to carry out this study. Two flowcharts explain the methodology. One shows the methodology used to classify the state of the press (Figure 28) and the other one presents the methodology used to classify the lubricant in the equipment (Figure 29).

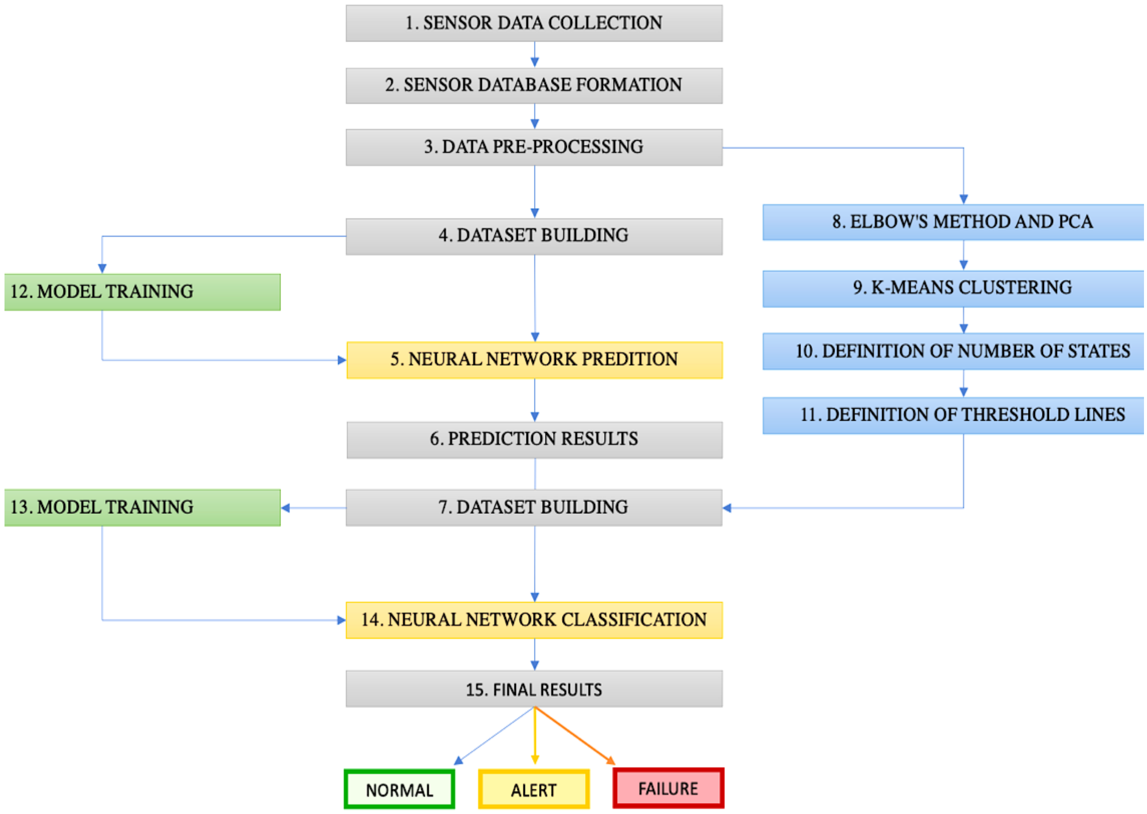

Figure 28 - Methodology used to classify press condition.

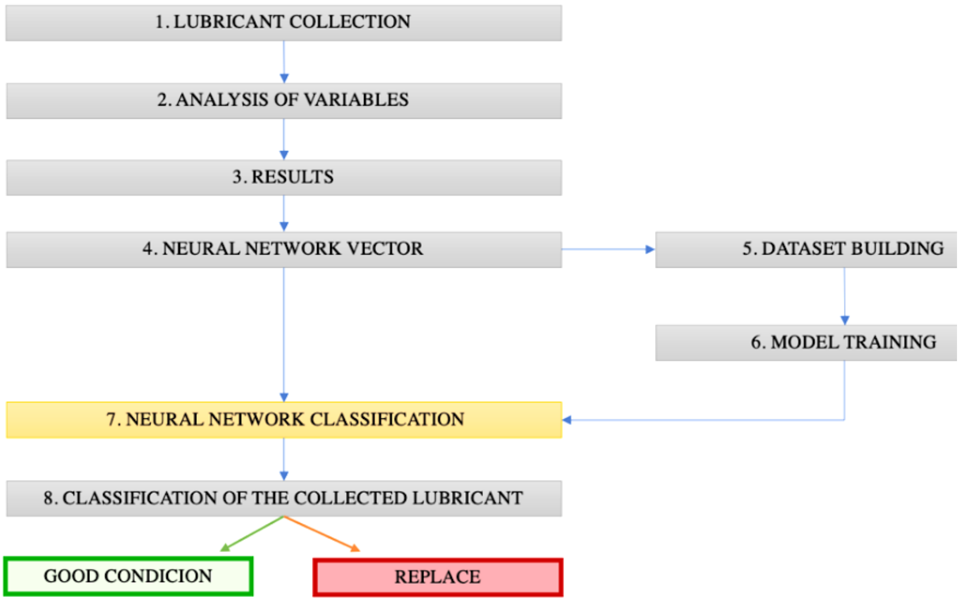

Figure 29 - Methodology used to classify the lubricating oil of the press.

8.4 Clustering operating states

The original filtered paper press data explained in section 7.2 was used in this cluster analysis. After applying the k-means method for K values between 1 and 10, the ideal number of clusters was determined using the Elbow method. For this purpose, the relationship between the number K [1, 10] of clusters and the sum of the squares of the distances between each point and the centroid of the cluster, where it was grouped, was graphically represented. Then, the ideal number of clusters is where the value of the dependent variable begins to stabilize, which visually resembles the shape of an elbow, hence the name of the method. From this value of K, the function starts to move almost parallel to the abscissa axis.

The K value corresponding to this point, where the error starts to stabilize, is the optimal K value, that is, it represents the ideal number of clusters (Figure 30).

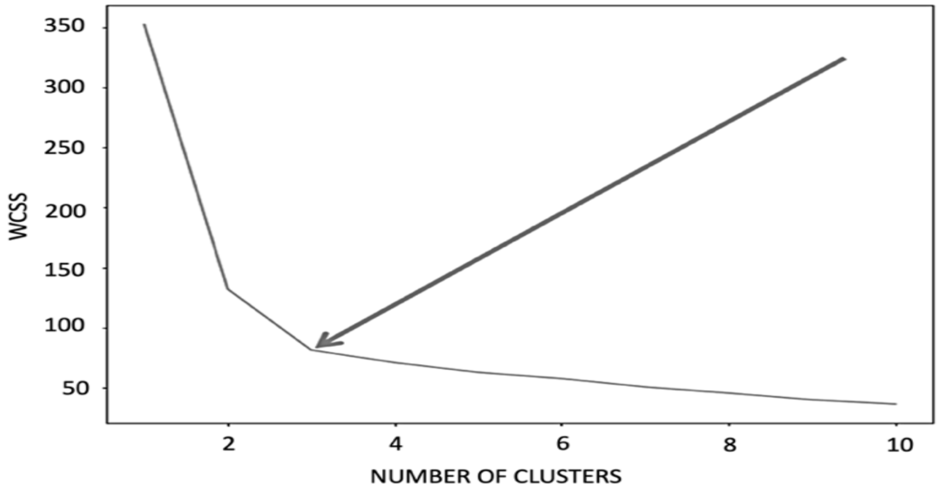

Figure 30 - Application of the elbow method to determine the number of clusters.

From the analysis of the graph presented in Figure 30, it was concluded that the ideal number of clusters was K=3. The next step was to convert the multidimensional dataset (6 variables) into two dimensions (variables), just to be able to visualize the distribution of the dataset more easily by the three defined clusters. To this end, principal component analysis was applied to the initial data set to represent the data in the two most representative principal components (PC1 and PC2) in terms of the explained variance (Figure 31).

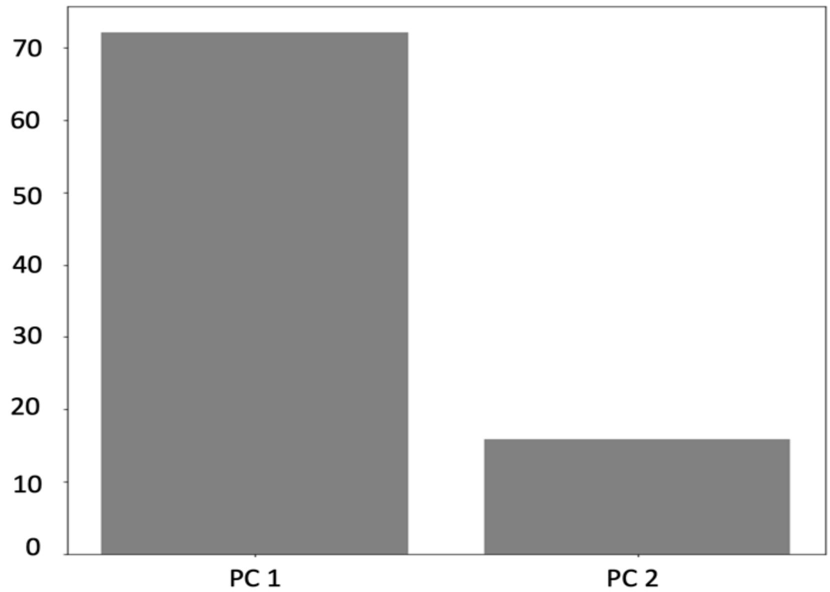

Figure 31 - Variance in PC1 and PC2.

The variance proportion is how much "variance" is explained by the principal component.

PC1 has a variance proportion of approximately 0.71, that is, PC1 explains about 71% of the variance occurring in the data set. The sum of PC1 and PC2 explains about 87%.

Due to this, it can be said that there is a minimal loss of information when using the two main components to form the three clusters as indicated by the Elbow Method.

Figure 32 presents the dataset classified according to the three defined clusters, as well as their centroids. The X axis represents Principal Component 1 and the Y axis represents Principal Component 2.

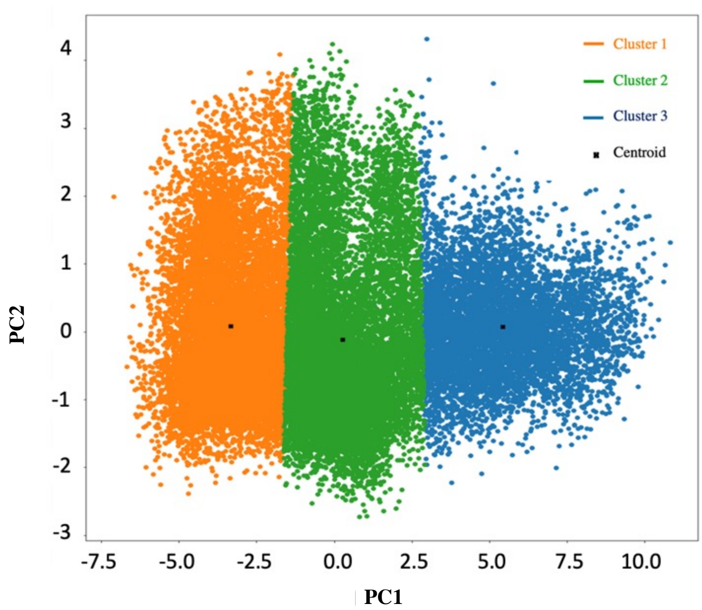

Figure 32 - Representation of the data in the three clusters.

Finally, the data were represented in Figure 33 according to several scatter plots that allow a visual analysis of the degree of association between the six-variables under analysis.

The last row and column in Figure 33 show the cluster rating of each data point. Analysing the various dispersion plots, three distinct states of asset functioning are unequivocally identified according to the different variables under analysis. This reinforces the idea that the equipment has three distinct operating states.

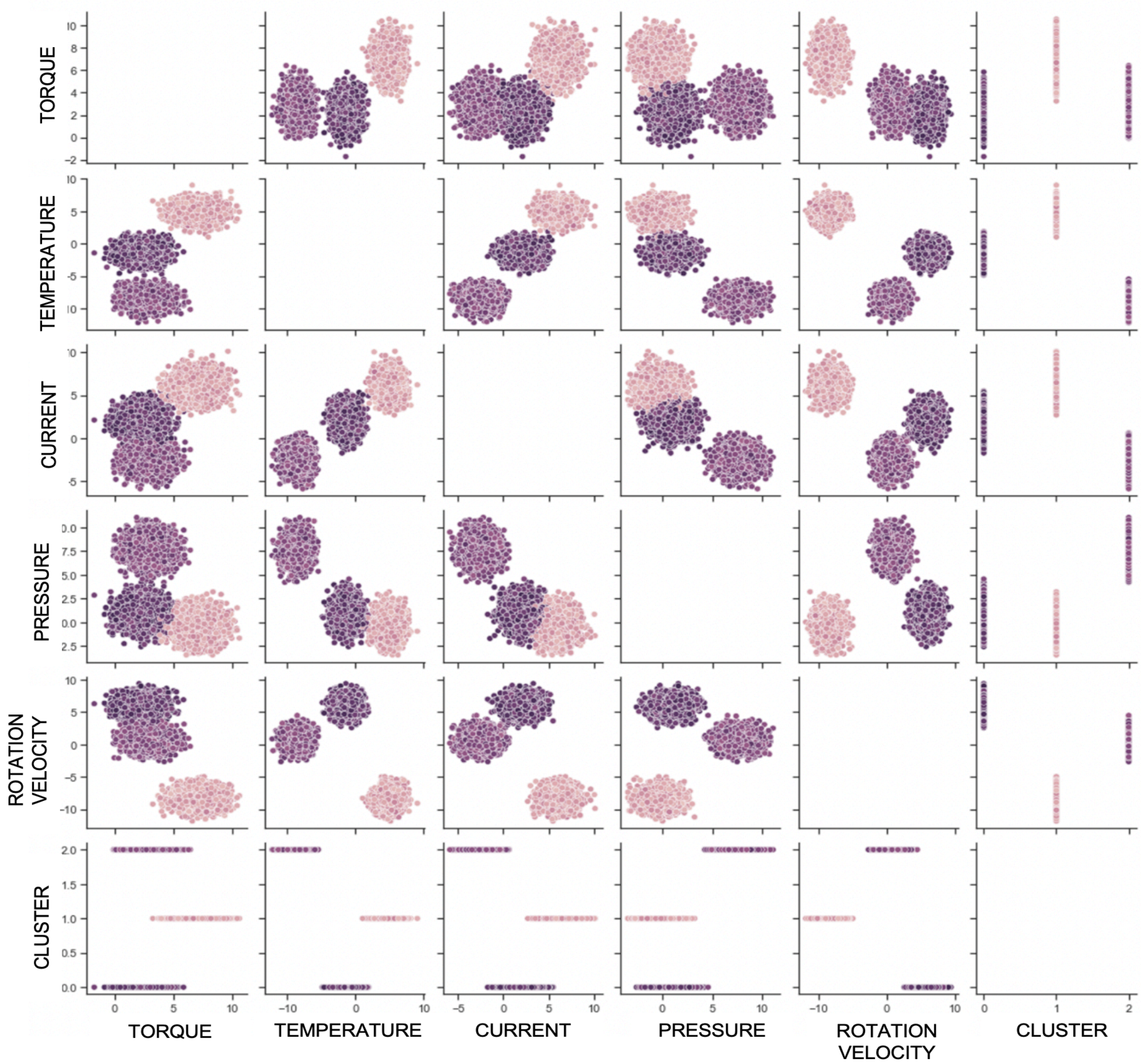

Figure 33 - Matrix of the scattering plots of five variables and the cluster.

Analysing Figure 33, it is easy to identify three distinct groups of data in all the ratios between the variables. Temperature has higher values when current or torque are higher. Temperature is lower when the current is lower. It can be said that temperature and electric current are directly proportional. The electric current is higher when the torque is higher.

8.5 Neural networks architecture

8.5.1 Network classification for the state of the paper press

A neural model was developed to automatically classify each data sample into one of the three operational states. To achieve this classification, a 30-day data prediction performed by a neural network using MLPRegressor was used. This prediction database was separated into two parts: the first 80% used for training the model; and the remaining 20% for carrying out the performance tests.

In this classification, we chose to work with feedforward architectures (Multilayer Perceptron), using the Sklearn Python library model called MLPClassifier. Knowing that the dataset is very large, we chose to use a graph-based optimization algorithm named "adam", using a logistic sigmoid as activation function [153].

Several architecture combinations were tested to find the best possible network configuration in terms of accuracy, resulting in a final architecture with an accuracy above 96%.

Knowing that three clusters cover almost all the variance of the data, as indicated in the previous section, the authors defined that the neural network would classify the machine in one of three states.

The network is composed of a first layer with 6 neurons that receive information from the press sensors, then the information is processed by 2 layers of hidden neurons (100,10) and the output of this neural network is the following: Good, Alert, Failure. Figure 34 depicts the ANN architecture chosen for classification of the state of the paper press.

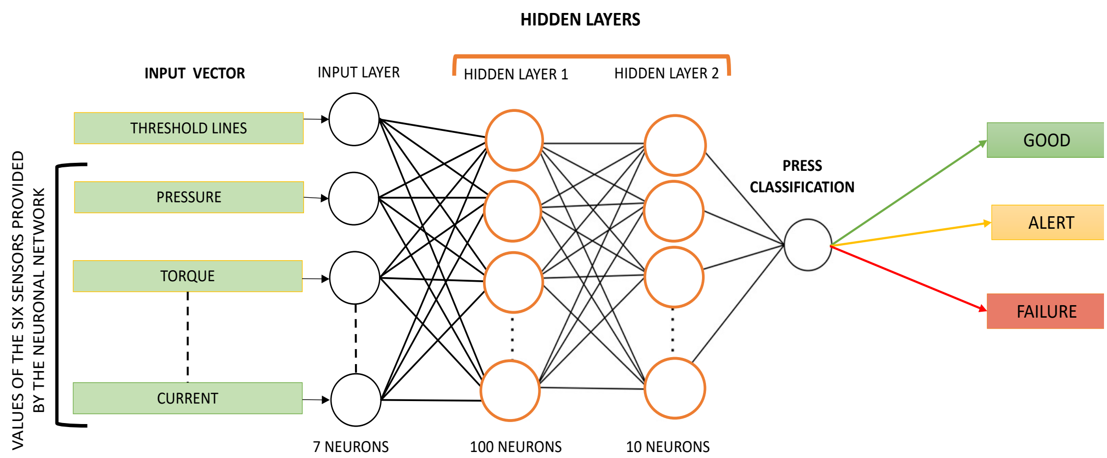

Figure 34 - Architecture of the ANN for classification of the state of the paper press.

Training this classification network takes approximately 4 minutes using Apple's M1 processor. The neural network required 150 iterations.

8.5.2 Neural network for press lubricant classification

The network is composed of a first layer with 12 neurons that receive information from the press sensors. The information is processed by three layers of hidden neurons (500, 100,10) and the output of this neural network is the following: Oil in good

Condition or Replace the oil. Figure 35 depicts the ANN architecture chosen for oil classification.

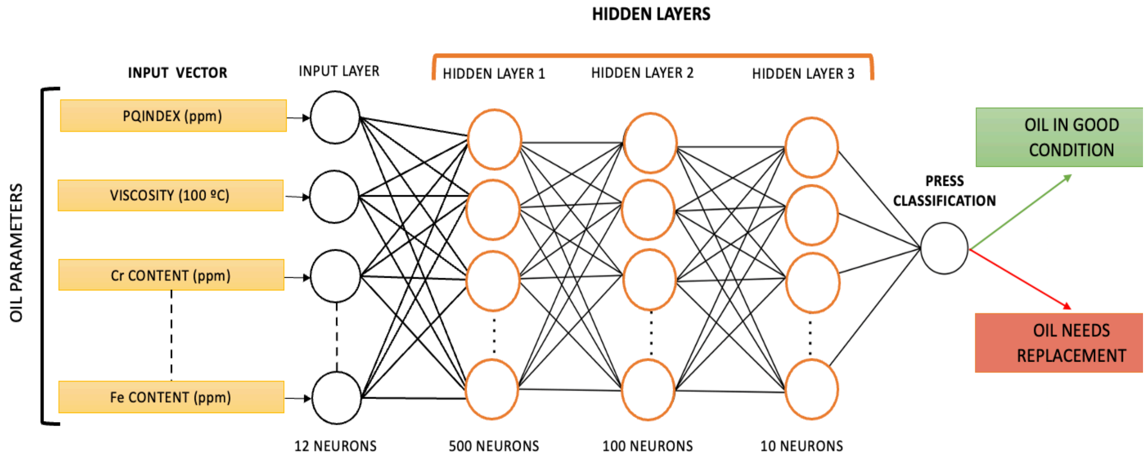

Figure 35 - Architecture of the ANN for classification of oil.

For this architecture, we chose the “lbfgs” solvers, using the “Relu” as the activation function. The network needed 455 iterations for training. Training this classification network takes approximately 2 minutes using Apple's M1 processor.

8.6 Press state classification results

The classifier algorithm showed very good results, as the error was below the defined p-value of 5%. The results of the press state of the classification network were compared with the results of the classification network. The network's accuracy compared to human experts was 96%.

Table 22 shows the different metrics and accuracy in the different states of the press. It was in the failure classification that the network had its highest success rate.

Table 23 shows the excellent result that the neural network had in classifying the failure state, as this is the most important state to classify, since identifying it in advance, it can prevent failures. It is also mentioned that the classification accuracy of the remaining states is above 95%, which was the goal.

TABLE 22 - PRESS CLASSIFICATION RESULTS.

Classification	Precision	Recall	F1-Score
Normal	0.96	1.00	0.98
Alert	0.98	0.85	0.91
Failure	1.00	0.90	0.94
Accuracy			0.96
Macro AVG	0.98	0.91	0.94
Weighted AVG	0.96	0.96	0.96

8.7 Lubricating oil classification results

The condition of the press depends heavily on the state of the lubricant. Therefore, it is essential to analyse the oil and to know its state of degradation. The algorithm created had an error well below the defined p-value of 5%, because the accuracy of this classifier was 98%. Thus, the algorithm proved to be quite reliable to classify the oil of any equipment, if they have a robust oil analysis database. Table 23 presents the performance metrics of the classification of the lubricating oil results. It should be noted that it obtained a 100% prediction in the degraded oils. Hence, in this case no damage to the industrial equipment was caused.

TABLE 23 - PERFORMANCE OF THE CLASSIFICATION OF THE LUBRICATING OIL

Classification	Precision	Recall	F1-Score	Support
Oil in good Condition	0.96	1.00	0.98	27
Replace the oil	1.00	0.94	0.97	18
Accuracy			0.98	45
Macro AVG	0.98	0.97	0.98	45
Weighted AVG	0.98	0.98	0.98	45

Table 24 presents a confusion matrix where it can be observed that the classification network only missed a classification, comparing its results with the results of human experts.

The classifier network only missed one classification out of 45 analyses. It classified an oil as degraded when the expert classified it as good.

However, it should be noted that for the neural network to classify this lubricating oil as degraded, it is because it is close to its degradation limit. It should be noted that the human expert may also have made a classification error.

TABLE *24* - LUBRICATING OIL RESULTS (CONFUSION MATRIX)

		Predictive Value	
		Oil in good Condition	Replace the oil
Real	Oil in good Condition	27	0
	Replace the oil	1	17

8.8 Limitations

A limitation of this project is that it is based on machine learning, which uses inductive reasoning. There are no guarantees about the certainty of fault detection. However, the quality of the predictive data that is fed into the press state classification network is quite good. The classification network learned according to the threshold lines that were stipulated based on the equipment's history and the technicians' recommendations.

Finding the hyperparameters to be used in the lubricant classification network and machine status classification network, aiming that the results meet the predefined objectives and errors, was the biggest challenge to overcome in this work. Obtaining a reliable and good-quality long-term forecast was the other difficulty encountered in this study. To achieve small error margins, it is necessary to use deep knowledge of the machine being modelled, as well as machine learning methods. The parameters and methods are only valid for the machine being studied, even though similar procedures may be followed to pursue similar or better results for other machines.

8.9 Summary

Through clustering using K-means it was shown that it is possible to identify equipment operating zones and to define how many states the classifier network should have at its output.

The model developed can predict the probability of future failures in a paper press considering a long-term forecast.

The contribution of a 30-day classification is innovative and provides a great advantage in industrial planning, as it allows to schedule stops one month in advance.

This methodology is very important in this area, as it can be applied to monitor this and other equipment automatically, if these assets have a robust database of sensor history.

It should be noted that, in this case we chose to perform a 30-day classification so, we introduced in the classification network the results of the 30-day sensor forecasts. To make a forecast with another time frame, for example 90 days, it is only necessary to introduce in the classification network the results of that same forecast. The prediction process is explained in Chapter 7.

The results of the lubricant classifier developed with neural networks were compared with the human expert classifications and converged at 96%.

These contributions can have a huge impact on the quality of operation and availability of assets, thus reducing maintenance costs and improving reliability.

Chapter 9

Main contributions

9.1 Contributions

This doctoral thesis offers several contributions in the field of Predictive Maintenance.

The answer to research question number one appears in Chapter 5 and has several contributions to the state of the art, which are highlighted as follows:

- Development of a Diesel engine lube oil deterioration level formula based on PCA;
- Creation of a model to assist in the decision of when to replace a Diesel engine lubricating oil. This model was validated by comparing its results with human expert ratings and the ratings of an ANN network.

As indicated in Section 1.4, the remaining research questions were answered through studies of paper industry equipment. Research questions 2 and 3 concern an industrial chip pump. The answer to these questions appears in Chapter 6 and offers the following contributions:

- A comprehensive diagnosis of the potential causes of early cracks in the chip pump shaft number 3 was carried out using the Ishikawa diagram;
- Identification of the causes of frequent cracks in the shaft of chip pump number 3. The cause of these frequent failures was a much higher stress than its predecessors, which made it work above the nominal specifications. This stress is justified by the fact that it transports its load vertically, while the previous chip pumps did so horizontally. The same chip pump has deficiencies in its seating that increase its vibration. The greater vibration associated with greater tension stress causes the chip pump shaft to suffer more stress than recommended, which shortens its life cycle;
- Using FMECA analysis, we evaluated failure modes, their impact and how to prevent those same failures of chip pump number 3 from happening. The FMECA analysis also defined criteria and verifications that should be included in the equipment preventive maintenance plan;
- To predict the operating conditions of chip pump number 3, short and long-term prediction models have been developed. The short-term prediction is based on a

self-adaptive exponential smoothing. The long-term forecasting model offers a 90-day ahead forecast and uses shallow networks, specifically MLP;

- The 90-day forecast for the long-term model is noteworthy, since no sensory forecasts of equipment with such a temporal antecedence were found in the state-of-the-art;
- The results of the network for long-term prediction of chip pump sensor values generally had a MAPE error of less than 10%;
- Work on chip pumps gives the industry a chance to plan scheduled maintenance stops with greater knowledge;
- Implementation of a fault alert system using colours. This system presents three states of the equipment for user information: Red for the malfunction, yellow for the alert, and green for the good operation. This system makes it quick, easy, and simple for the operator to understand the equipment's current state of operation, which can help prevent serious faults or breakdowns.

The remaining research questions focused their study on a pulp paper press. The answers to research question number 4 are presented in Chapter 7 and offered the following contributions:

- Algorithms were created to predict the behaviour of a pulp paper press in the short, medium and long-term using feedforward and XGBoost neural networks;
- Different input vectors were analysed in the neural network and various ANN features and architectures were tested;
- The predictions were performed using input vectors with overlapping and non-overlapping sliding windows, concluding that overlapping influences the results and the network learning process;
- Comparison of the prediction results of sensors of a pulp paper press with vectors with and without relation among the variables;
- Comparisons between feedforward neural network and XGBoost algorithm were made, in order to determine which is the best prediction method for each specific parameter;
- LSTM and GRU nets were compared with the results of shallow feedforward and XGBoost nets.

From the resolution of question number 5, present in Chapter 8, came the following contributions:

- Development of an algorithm capable of classifying the future state of the pulp press using feedforward neural networks. It is pointed out that this classifier algorithm uses the data predicted from the previously announced algorithm;

- The optimal number of states to classify a press using the k-means algorithm and the elbow method was found.

The resolution of question number 5 also appears in Chapter 8 and gave the last contributions of this doctoral thesis, highlighting two important issues:

- An alert system was created, finding the pulp paper press operating zones using threshold lines;
- Development of a lubricant classifier based on feedforward neural networks. This classifier had a success rate over 96% when comparing its classifications with those of human experts.

The contributions identified here can have a positive impact on equipment availability and, consequently reduce the company's production costs while increasing reliability [15], [114]-[116].

It is emphasised that all research questions were successfully answered, thus fulfilling all the objectives proposed for this doctoral thesis.

9.2 Publications

These state-of-the-art contributions were published in scientific journals and presented at conferences.

9.2.1 Papers in international journals

Rodrigues, J.A., Costa, I., Farinha J., Mendes M., Margalho L.. 2020. Predicting Motor Oil Condition Using Artificial Neural Networks and Principal Component Analysis. Eksploatacja i Niezawodnosc - Maintenance and Reliability 22 (3): 440–48. https://doi.org/10.17531/ein.2020.3.6. **Appendix A**

Rodrigues J.A., Farinha J., Cardoso A.J.M.Predictive Maintenance Tools – A Global Survey. WSEAS TRANSACTIONS ON SYSTEMS AND CONTROL 2021; 16: 96–109, https://doi.org/10.37394/23203.2021.16.7. **Appendix B**

Rodrigues J.A., Farinha J., Mendes M., Cardoso A.J.M. Mateus R. Short and long forecast to implement predictive maintenance in a pulp industry. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2021; 24: 33–41, https://doi.org/10.17531/ein.2022.1.5. **Appendix C**

Rodrigues J.A., Farinha J., Mendes M., Cardoso A.J.M., Mateus R. Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition. Energies 2022; 15(17): 6308, https://doi.org/10.3390/en15176308. **Appendix D**

Rodrigues J., Martins A., Farinha J. et al. Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning. Energies 2022. doi:10.3390/en15249387, https://doi.org/10.3390/en15249387. **Appendix E**

9.2.2 Papers in international conferences

Rodrigues, J. A., Farinha, J. T., Costa, I., Mendes, M., & Margalho, L. (2019). "Modelling Diesel Engine Oil Condition Using Artificial Neural Networks". eMaint19, Stockholm.

Rodrigues J. A., Farinha J. T., Cardoso A.J.M. et al. Prediction of Sensor Values in Paper Pulp Industry Using Neural Networks. In Zhang H, Feng G, Wang H et al. (eds): Proceedings of IncoME-VI and TEPEN 2021, Cham, Springer International Publishing: 2023: 281–291, https://doi.org/10.1007/978-3-030-99075-6_24. **Appendix F**

9.2.3 Papers in national conferences

Rodrigues J.A., Farinha J., Martins A., Margalho L. (2019). Análise Multivariada de Óleos Lubrificantes. Atas do 15.º Congresso Nacional de Manutenção. 21 e 22 de Novembro de 2019 - Altice Fórum Braga, Braga, Portugal. ISBN: 978-989-8200-20-4

Rodrigues J.A., Farinha J., Mendes M., Martins A., Mateus B. (2021). Previsão e Diagnóstico de Avarias de uma Prensa Industrial, com Análise FMECA e Redes Neuronais. Actas do 16.º Congresso Nacional de Manutenção. 23 e 24 de Novembro de 2021 - Centro de Congressos de Aveiro, Aveiro, Portugal.

Rodrigues J.A., Farinha J., Mendes M., Mateus R., Cardoso A.J.M., "Identificação dos Estados de Funcionamento de um Equipamento Industrial através do Método de Agrupamento de Dados K-means", 1º Congresso Nacional em Engenharia e Gestão de Ativos, Coimbra, Portugal, 26-27 de maio 2022.

9.3 Awards

The contributions of this thesis received some awards, namely:

- Award for best presentation at international conference – TEPEN 2021& IncoME-VI, China.

- 2nd prize in the Young Engineer Innovation Award - PIJE 2021 (Ordem dos Engenheiros).

Chapter 10

Conclusions

10.1 Final Remarks

This doctoral thesis was divided into three sets of distinct assets. First, the study focused on the lubricating oil of Diesel engine buses. Then the study advanced to the identification of a frequent failure of the axis of a chip pump in a paper industry, and then a model was created to predict its behaviour and alert about possible future failures. The last study used data from paper presses to predict their performance and determine their probable status in the future. The following conclusions arise from the work done.

Proper lubrication of an engine or a machine is extremely important, since a lubricating oil has the role of minimising the wear produced by friction between moving surfaces.

Besides lubricating, the oil also acts as a coolant, assists in the internal sealing process. It prevents leaks in minimum spaces, protects the asset from external contaminants and promotes protection against corrosion. These reasons show that it is essential to monitor and know when to replace a lubricating oil. It is also important to know which variables are most important depending on the asset.

This doctoral thesis presents a formula for calculating the percentage of degradation of an engine lubricating oil using PCA, which allows to classify the state of oils with high precision. The formula developed was validated through comparison of the results with the classifications of neural models and human experts. The results converged by more than 90%, which shows that this formula is viable, reliable.

This formula can help companies in the decision-making process of changing the lubricants of their assets, avoiding early changes, or avoiding continuing with oils that no longer fulfil their function with quality. The proper use of this formula can have a consequent cost reduction and a positive impact on equipment availability.

This study, based on a multivariate analysis, carried out in two public transport companies, can and should be extended to other types of companies and to other types of equipment.

Failures and breakdowns in an industrial environment are one of the main challenges of a company's success, as they harm the company's finances, its market image and can, sometimes, put people's lives at risk.

The final chip pump, in a set of three under study, had a chronic failure on its shaft. The cause of this chronic malfunction was that this chip pump was under significant stress, according to an FMECA analysis. This effort was justified because the load had to be transported vertically, rather than horizontally, by the predecessor chip pumps. All chip pumps are of the same model and the same power. The same chip pump has fixation problems that cause its vibration to increase exponentially with the greater efforts. The life of the chip pump shaft is much shorter than expected due to excessive vibration, which damages its shaft.

A model was created to predict the future behaviour of the chip pump using a dataset composed of sensory data collected over approximately three years and factory inspections. The neural network predictions proved to be accurate for this type of problem. All variables presented mean absolute percentage errors of less than 10%.

This study offered the opportunity to plan maintenance shutdowns with greater certainty. This way, there will be a much lower probability of failure.

By using neural networks with relevant, quality training data, this prediction and alerting methodology can be applied to different machines, although there is no guarantee that this will produce the same results on different assets. Depending on the types of patterns in the data, the results can be better or worse. This methodology has the main disadvantage of relying on previous sensory information. Forecasts can be more uncertain when the data are not of good quality.

Regarding the pulp paper press, the suggested prediction model enables knowledge of a cellulosic paper press' long-term behaviour, supported by time series collected from sensors installed on it. Long-term scheduled downtime can be optimised in this way, and production downtime can be avoided. A classification network will classify the machine into one of three states: normal operation, alert or malfunction.

The pulp paper press prediction work offers a priceless comparison of the neural network input vector using overlapping and non-overlapping sliding windows. It presents the results of the tests conducted and draws clear conclusions about the benefits and drawbacks of each technique. Resampling data can make the forecasting process much faster, as it greatly reduces the dataset.

Training requires the use of sliding windows along time series. Learning is possible with overlapping windows in fewer epochs. Long trends are easier to capture when the window is larger, but the ideal window size must be determined experimentally. Overlapping windows give the neural network more input at each epoch, which leads to faster learning rates and more accurate prediction outcomes.

Non-overlapping sliding windows show processing speed but overall have worse long-term results.

The neural network's input vector's size directly affects how accurately predictions are made. While a smaller sliding window is, more difficult to predict peaks, a larger sliding window increases prediction errors. On the other hand, the forecast gets harder the wider the forecast gap.

The short and medium-term results are on par with or better than the state-of-the-art.

The forecasts of the majority of parameters have MAPE errors below 10%, as was the goal of the research presented. The long-term forecasts using overlapping windows demonstrated very good accuracy. They take a long time to process. Non-overlapping windows can be used to expedite the process for short-term forecasts.

The XGBoost Model presents fast and good results in the prediction of electric current, pressure and temperature.

The ability to predict sensor values up to three months in advance is very useful for managing equipment maintenance decisions. The time component of the forecast is completely innovative and offers the possibility of classifying the machine in the future.

The operating zones of the equipment can be determined using K-means, and this allows one to specify how many states the classification network's output should contain.

A database of long-term forecast values can be used to classify future failures in a paper press using the currently developed model.

The contribution of a long-term classification is innovative and brings a great advantage in industrial planning, as it allows to schedule stops several weeks in advance.

This methodology is very important in this area, as it can be applied to monitor these and other equipment automatically, if the assets have a robust database of previous sensory readings.

A lubricant classifier for pulp paper press was also developed using neural networks. The results of this classifier were compared with human expert classifications and were found to be more than 96% accurate.

Within the scope of this doctoral thesis, it was found that the greatest difficulty encountered was that some databases had missing or repeated values. Hence, there was a more complex data processing, but not enough to affect the reliability of the results achieved.

The methods used and research done throughout this doctoral thesis will have very beneficial practical repercussions for the industrial sector, particularly in the areas of maintenance, safety, quality, sustainability, and efficiency, as it will maximise the availability of assets, working in tandem with the Systems of Quality and Maintenance Management.

These positive developments across several industries will primarily lead to cost reductions, which will give the industries an advantage in the marketplace, because they will be able to offer more affordable prices and quality of service.

10.2 Future Work

Future research includes testing other lubricant classification techniques, as well as extracting information from their analysis that indicate a specific malfunction.

Regarding predictive models, it is intended to evolve to include more variables and different machine learning models.

Future work includes additional experiments to improve the input vectors and optimize other neural network hyperparameters.

In addition to including more variables, this methodology should be tried on more assets in order to create the most robust, reliable and generic prediction and classification model possible.

References

1. Martins, A.B.; Torres Farinha, J.; Marques Cardoso, A. Calibration and Certification of Industrial Sensors – a Global Review. *WSEAS TRANSACTIONS ON SYSTEMS AND CONTROL* **2020**, *15*, 394–416, doi:10.37394/23203.2020.15.41.

2. Martins, A.; Fonseca, I.; Farinha, J.; Reis, J.; Cardoso, A.J.M. Maintenance Prediction through Sensing Using Hidden Markov Models-A Case Study. *Applied Sciences* **2021**, *11*, 7685, doi:10.3390/app11167685.

3. Uygun, Y. *Industry 4.0: Principles, Effects and Challenges – Nova Science Publishers*; Nova Science Publishers, **2020**; ISBN 1-5361-8423-3.

4. Cachada, A.; Barbosa, J.; Leitño, P.; Gcraldcs, C.A.S.; Deusdado, L.; Costa, J.; Teixeira, C.; Teixeira, J.; Moreira, A.H.J.; Moreira, P.M.; et al. Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. In Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA); September **2018**; Vol. 1, pp. 139–146.

5. Barata, J. A Systematic Approach to Design Product Traceability in Industry 4.0: Insights from the Ceramic Industry. *26TH INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS DEVELOPMENT (ISD2017 CYPRUS)*, **2017**.

6. Cruz, S.; Paulino, A.; Duraes, J.; Mendes, M. Real-Time Quality Control of Heat Sealed Bottles Using Thermal Images and Artificial Neural Network. Journal of Imaging 2021, 7, 24, doi:10.3390/jimaging7020024.

7. Cachada, A.; Barbosa, J.; Leitño, P.; Gcraldcs, C.A.S.; Deusdado, L.; Costa, J.; Teixeira, C.; Teixeira, J.; Moreira, A.H.J.; Moreira, P.M.; et al. Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. In Proceedings of the **2018** IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA); September 2018; Vol. 1, pp. 139–146.

8. EN 13306:2017 - Maintenance - Maintenance Terminology Available online: https://standards.iteh.ai/catalog/standards/cen/5af77559-ca38-483a-9310-823e8c517ee7/en-13306-2017 (accessed on 25 May 2021).

9. Kumar, U.; Galar, D.; Parida, A.; Stenström, C.; Berges, L. Maintenance Performance Metrics: A State-of-the-art Review. *Journal of Quality in Maintenance Engineering* **2013**, *19*, 233–277, doi:10.1108/JQME-05-2013-0029.

10. Rao, B.K.N. *Handbook of Condition Monitoring*; Elsevier, **1996**; ISBN 978-1-85617-234-9.

11. Carnero, M.C. Selection of Diagnostic Techniques and Instrumentation in a Predictive Maintenance Program. A Case Study. *Decision Support Systems* **2005**, *38*, 539–555, doi:10.1016/j.dss.2003.09.003.

12. Selcuk, S. Predictive Maintenance, Its Implementation and Latest Trends. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture* **2017**, *231*, 1670–1679, doi:10.1177/0954405415601640.

13. Patwardhan, A.; Verma, A.K.; Kumar, U. A Survey on Predictive Maintenance Through Big Data. In Proceedings of the Current Trends in Reliability, Availability, Maintainability and Safety; Kumar, U., Ahmadi, A., Verma, A.K., Varde, P., Eds.; Springer International Publishing: Cham, **2016**; pp. 437–445.

14. Bousdekis, A.; Lepenioti, K.; Apostolou, D.; Mentzas, G. A Review of Data-Driven Decision-Making Methods for Industry 4.0 Maintenance Applications. *Electronics* **2021**, *10*, 828, doi:10.3390/electronics10070828.

15. Rodrigues, J.; Torres Farinha, J.; Marques Cardoso, A. Predictive Maintenance Tools – A Global Survey. *WSEAS TRANSACTIONS ON SYSTEMS AND CONTROL* **2021**, *16*, 96–109, doi:10.37394/23203.2021.16.7.

16. John Moubray *Reliability-Centered Maintenance*; Second Edi.; Industrial Press Inc.: New York, **1997**; ISBN 0-8311-3078-4.

17. Monchy, F. *A Função Manutenção*; **1989**; ISBN 8487022111.

18. Farinha, J.T. *A Terologia e as Novas Ferramentas de Gestão*; Monitor.; **2011**; ISBN 9789729413827.

19. Husband, T.; Basker, B.A. Maintenance Engineering: The Current State of the Art. *Production Engineer* **1976**, *55*, 77–82, doi:10.1049/tpe.1976.0029.

20. Shrivastav, O.P. Industrial Maintenance: A Discipline in Its Own Right. *World Transactions on Engineering and Technology Education* **2005**, *4*, 4.

21. Poór, P.; Basl, J.; Zenisek, D. Predictive Maintenance 4.0 as next Evolution Step in Industrial Maintenance Development. In Proceedings of the 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE); March **2019**; pp. 245–253.

22. Monchy, F.; Mirochnikoff, Y. La Fonction Maintenance : Formation à La Gestion de La Maintenance Industrielle. *undefined* **1987**.

23. Ran, Y.; Zhou, X.; Lin, P.; Wen, Y.; Deng, R. A Survey of Predictive Maintenance: Systems, Purposes and Approaches. *arXiv:1912.07383 [cs, eess]* **2019**.

24. Farinha, J.M.T. *Asset Maintenance Engineering Methodologies*; 1st edition.; CRC Press: Boca Raton, FL, **2018**; ISBN 978-1-138-03589-8.

25. Rahmati, S.H.A.; Ahmadi, A.; Govindan, K. A Novel Integrated Condition-Based Maintenance and Stochastic Flexible Job Shop Scheduling Problem: Simulation-Based Optimization Approach. *Ann Oper Res* **2018**, *269*, 583–621, doi:10.1007/s10479-017-2594-0.

26. Sheu, D.D.; Kuo, J.Y. A Model for Preventive Maintenance Operations and Forecasting. *J Intell Manuf* **2006**, *17*, 441–451, doi:10.1007/s10845-005-0017-6.

27. Parida, A.; Kumar, U. Maintenance Productivity and Performance Measurement. In *Handbook of Maintenance Management and Engineering*; Ben-Daya, M., Duffuaa, S.O., Raouf, A., Knezevic, J., Ait-Kadi, D., Eds.; Springer: London, **2009**; pp. 17–41 ISBN 978-1-84882-472-0.

28. Balduíno, M.; Torres Farinha, J.; Marques Cardoso, A. Production Optimization versus Asset Availability – a Review. *WSEAS TRANSACTIONS ON SYSTEMS AND CONTROL* **2020**, *15*, 320–332, doi:10.37394/23203.2020.15.33.

29. Paolanti, M.; Romeo, L.; Felicetti, A.; Mancini, A.; Frontoni, E.; Loncarski, J. Machine Learning Approach for Predictive Maintenance in Industry 4.0. In Proceedings of the 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA); July **2018**; pp. 1–6.

30. Jun, L.; Huibin, X. Reliability Analysis of Aircraft Equipment Based on FMECA Method. *Physics Procedia* **2012**, *25*, 1816–1822, doi:10.1016/j.phpro.2012.03.316.

31. Compare, M.; Baraldi, P.; Zio, E. Challenges to IoT-Enabled Predictive Maintenance for Industry 4.0. *IEEE Internet of Things Journal* **2020**, *7*, 4585–4597, doi:10.1109/JIOT.2019.2957029.

32. Gbadamosi, A.-Q.; Oyedele, L.O.; Delgado, J.M.D.; Kusimo, H.; Akanbi, L.; Olawale, O.; Muhammed-yakubu, N. IoT for Predictive Assets Monitoring and Maintenance: An Implementation Strategy for the UK Rail Industry. *Automation in Construction* **2021**, *122*, 103486, doi:10.1016/j.autcon.2020.103486.

33. Yeh, R.H.; Kao, K.-C.; Chang, W.L. Preventive-Maintenance Policy for Leased Products under Various Maintenance Costs. *Expert Systems with Applications* **2011**, *38*, 3558–3562, doi:10.1016/j.eswa.2010.08.144.

34. Yan, J.; Meng, Y.; Lu, L.; Li, L. Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. *IEEE Access* **2017**, *5*, 23484–23491, doi:10.1109/ACCESS.2017.2765544.

35. Kiangala, K.S.; Wang, Z. Initiating Predictive Maintenance for a Conveyor Motor in a Bottling Plant Using Industry 4.0 Concepts. *Int J Adv Manuf Technol* **2018**, *97*, 3251–3271, doi:10.1007/s00170-018-2093-8.

36. Sajid, S.; Haleem, A.; Bahl, S.; Javaid, M.; Goyal, T.; Mittal, M. Data Science Applications for Predictive Maintenance and Materials Science in Context to Industry 4.0. *Materials Today: Proceedings* **2021**, *45*, 4898–4905, doi:10.1016/j.matpr.2021.01.357.

37. Daily, J.; Peterson, J. Predictive Maintenance: How Big Data Analysis Can Improve Maintenance. In *Supply Chain Integration Challenges in Commercial Aerospace*; Richter, K., Walther, J., Eds.; Springer International Publishing: Cham, **2017**; pp. 267–278 ISBN 978-3-319-46154-0.

38. Hegedűs, C.; Kosztyán, Z.T. 46 THE CONSIDERATION OF MEASUREMENT UNCERTAINTY IN FORECAST AND MAINTENANCE RELATED DECISIONS. **2011**, *1*, 14.

39. Hashemian, H.M. State-of-the-Art Predictive Maintenance Techniques. *IEEE Transactions on Instrumentation and Measurement* **2011**, *60*, 226–236, doi:10.1109/TIM.2010.2047662.

40. Carvalho, T.P.; Soares, F.A.A.M.N.; Vita, R.; Francisco, R. da P.; Basto, J.P.; Alcalá, S.G.S. A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance. *Computers & Industrial Engineering* **2019**, *137*, 106024, doi:10.1016/j.cie.2019.106024.

41. Galar, D.; Stenström, C.; Parida, A.; Kumar, R.; Berges, L. Human Factor in Maintenance Performance Measurement. In Proceedings of the 2011 IEEE International Conference on Industrial Engineering and Engineering Management; December **2011**; pp. 1569–1576.

42. Sahal, R.; Ali, M.I.; Breslin, J. Big Data and Stream Processing Platforms for Industry 4.0 Requirements Mapping for a Predictive Maintenance Use Case. *Journal of Manufacturing Systems* **2020**, *54*, 138–151, doi:10.1016/j.jmsy.2019.11.004.

43. Ferreira, L.A. *Uma Introdução à Manutenção*; 1º Edição.; Publindústria, Edições Técnicas, **1998**; ISBN 972-95794-4-X.

44. English, J.R.; Yan, L.; Landers, T.L. A Modified Bathtub Curve with Latent Failures. In Proceedings of the Annual Reliability and Maintainability Symposium 1995 Proceedings; January **1995**; pp. 217–222.

45. Sellitto, M.A. Formulação estratégica da manutenção industrial com base na confiabilidade dos equipamentos. *Prod.* **2005**, *15*, 44–59, doi:10.1590/S0103-65132005000100005.

46. Raposo, H.; Farinha, J.T.; Fonseca, I.; Ferreira, L.A. Condition Monitoring with Prediction Based on Diesel Engine Oil Analysis: A Case Study for Urban Buses. *Actuators* **2019**, *8*, 14, doi:10.3390/act8010014.

47. Gajewski, J.; Vališ, D. The Determination of Combustion Engine Condition and Reliability Using Oil Analysis by MLP and RBF Neural Networks. *Tribology International* **2017**, *115*, 557–572, doi:10.1016/j.triboint.2017.06.032.

48. Parlak, A.; Islamoglu, Y.; Yasar, H.; Egrisogut, A. Application of Artificial Neural Network to Predict Specific Fuel Consumption and Exhaust Temperature for a Diesel Engine. *Applied Thermal Engineering* **2006**, *26*, 824–828, doi:10.1016/j.applthermaleng.2005.10.006.

49. Le, V.T.; Lim, C.P.; Mohamed, S.; Nahavandi, S.; Yen, L.; Gallasch, G.E.; Baker, S.; Ludovici, D.; Draper, N.; Wickramanayake, V. Condition Monitoring of Engine Lubrication Oil of Military Vehicles: A Machine Learning Approach. *PEER REVIEW* **2017**, 8.

50. Zhu, J.; Yoon, J.; He, D.; Bechhoefer, E. Online Particle-Contaminated Lubrication Oil Condition Monitoring and Remaining Useful Life Prediction for Wind Turbines: Particle Contamination Models. *Wind Energy* **2014**, *18*, doi:10.1002/we.1746.

51. Kumar, S.; Mukherjee, P.S.; Mishra, N.M. Online Condition Monitoring of Engine Oil. *Industrial Lubrication and Tribology* **2005**, *57*, 260–267, doi:10.1108/00368790510622362.

52. Yonghui, Y.; Weihua, W.; Xinpin, Y.; Hanliang, X.; Chengtao, W. An Integrated On-Line Oil Analysis Method for Condition Monitoring. *Meas. Sci. Technol.* **2003**, *14*, 1973–1977, doi:10.1088/0957-0233/14/11/015.

53. El-Hag, A.H.; Saker, Y.A.; Shurrab, I.Y. Online Oil Condition Monitoring Using a Partial- Discharge Signal. *IEEE Transactions on Power Delivery* **2011**, *26*, 1288–1289, doi:10.1109/TPWRD.2010.2073551.

54. Zhu, X.; Zhong, C.; Zhe, J. A High Sensitivity Wear Debris Sensor Using Ferrite Cores for Online Oil Condition Monitoring. *Meas. Sci. Technol.* **2017**, *28*, 075102, doi:10.1088/1361-6501/aa6adb.

55. Du, L.; Zhe, J. A High Throughput Inductive Pulse Sensor for Online Oil Debris Monitoring. *Tribology International* **2011**, *44*, 175–179, doi:10.1016/j.triboint.2010.10.022.

56. Westerholm, Roger.; Li, Hang. A Multivariate Statistical Analysis of Fuel-Related Polycyclic Aromatic Hydrocarbon Emissions from Heavy-Duty Diesel Vehicles. *Environ. Sci. Technol.* **1994**, *28*, 965–972, doi:10.1021/es00054a032.

57. Capone, S.; Zuppa, M.; Presicce, D.S.; Francioso, L.; Casino, F.; Siciliano, P. Metal Oxide Gas Sensor Array for the Detection of Diesel Fuel in Engine Oil. *Sensors and Actuators B: Chemical* **2008**, *131*, 125–133, doi:10.1016/j.snb.2007.12.029.

58. Mukherjee, A.; Kundu, P.K.; Das, A. Transmission Line Fault Classification under High Noise in Signal: A Direct PCA-Threshold-Based Approach. *J. Inst. Eng. India Ser. B* **2022**, *103*, 197–211, doi:10.1007/s40031-021-00601-w.

59. Shaban, K.; El-Hag, A.; Matveev, A. A Cascade of Artificial Neural Networks to Predict Transformers Oil Parameters. *IEEE Transactions on Dielectrics and Electrical Insulation* **2009**, *16*, 516–523, doi:10.1109/TDEI.2009.4815187.

60. Niu, X.; Yang, C.; Wang, H.; Wang, Y. Investigation of ANN and SVM Based on Limited Samples for Performance and Emissions Prediction of a CRDI-Assisted Marine Diesel Engine. *Applied Thermal Engineering* **2017**, *111*, 1353–1364, doi:10.1016/j.applthermaleng.2016.10.042.

61. Ghobadian, B.; Rahimi, H.; Nikbakht, A.M.; Najafi, G.; Yusaf, T.F. Diesel Engine Performance and Exhaust Emission Analysis Using Waste Cooking Biodiesel Fuel with an Artificial Neural Network. *Renewable Energy* **2009**, *34*, 976–982, doi:10.1016/j.renene.2008.08.008.

62. Li, X.; Li, J.; He, D.; Qu, Y. Gear Pitting Fault Diagnosis Using Raw Acoustic Emission Signal Based on Deep Learning. *EiN* **2019**, *21*, 403–410, doi:10.17531/ein.2019.3.6.

63. Li, Y.; Wang, K. Modified Convolutional Neural Network with Global Average Pooling for Intelligent Fault Diagnosis of Industrial Gearbox. *EiN* **2019**, *22*, 63–72, doi:10.17531/ein.2020.1.8.

64. Kanawaday, A.; Sane, A. Machine Learning for Predictive Maintenance of Industrial Machines Using IoT Sensor Data. In Proceedings of the 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS); November **2017**; pp. 87–90.

65. Kolokas, N.; Vafeiadis, T.; Ioannidis, D.; Tzovaras, D. Forecasting Faults of Industrial Equipment Using Machine Learning Classifiers. In Proceedings of the 2018 Innovations in Intelligent Systems and Applications (INISTA); July 2018; pp. 1–6.

66. Tian, Z. An Artificial Neural Network Method for Remaining Useful Life Prediction of Equipment Subject to Condition Monitoring. *J Intell Manuf* **2012**, *23*, 227–237, doi:10.1007/s10845-009-0356-9.

67. Rafiee, J.; Arvani, F.; Harifi, A.; Sadeghi, M.H. Intelligent Condition Monitoring of a Gearbox Using Artificial Neural Network. *Mechanical Systems and Signal Processing* **2007**, *21*, 1746–1754, doi:10.1016/j.ymssp.2006.08.005.

68. Heidarbeigi, K.; Ahmadi, H.; Omid, M.; Tabatabaeefar, A. Evolving an Artificial Neural Network Classifier for Condition Monitoring of Massy Ferguson Tractor Gearbox. *International Journal of Applied Engineering Research* **2010**, *5*, 2097–2107.

69. Karpenko, M.; Sepehri, N. Neural Network Classifiers Applied to Condition Monitoring of a Pneumatic Process Valve Actuator. *Engineering Applications of Artificial Intelligence* **2002**, *15*, 273–283, doi:10.1016/S0952-1976(02)00068-4.

70. Wang, J.; Li, C.; Han, S.; Sarkar, S.; Zhou, X. Predictive Maintenance Based on Event-Log Analysis: A Case Study. *IBM Journal of Research and Development* **2017**, *61*, 11:121-11:132, doi:10.1147/JRD.2017.2648298.

71. Allah Bukhsh, Z.; Saeed, A.; Stipanovic, I.; Doree, A.G. Predictive Maintenance Using Tree-Based Classification Techniques: A Case of Railway Switches. *Transportation Research Part C: Emerging Technologies* **2019**, *101*, 35–54, doi:10.1016/j.trc.2019.02.001.

72. Elhag, T.M.S.; Wang, Y.-M. Risk Assessment for Bridge Maintenance Projects: Neural Networks versus Regression Techniques. *Journal of Computing in Civil Engineering* **2007**, *21*, 402–409, doi:10.1061/(ASCE)0887-3801(2007)21:6(402).

73. Balluff, S.; Bendfeld, J.; Krauter, S. Short Term Wind and Energy Prediction for Offshore Wind Farms Using Neural Networks. In Proceedings of the 2015 International Conference on Renewable Energy Research and Applications (ICRERA); November **2015**; pp. 379–382.

74. T, D.; P, P. Power Consumption Prediction in Cloud Data Center Using Machine Learning. *International Journal of Electrical and Computer Engineering (IJECE)* **2020**, *10*, 1524–1532, doi:10.11591/ijece.v10i2.pp1524-1532.

75. Hongxiang, T.; Yuntao, L.; Xiangjun, W. Application of Neural Network to Diesel Engine SOA. In Proceedings of the 2011 Third International Conference on Measuring Technology and Mechatronics Automation; January **2011**; Vol. 1, pp. 555–558.

76. Okoh, C.; Roy, R.; Mehnen, J. Predictive Maintenance Modelling for Through-Life Engineering Services. *Procedia CIRP* **2017**, *59*, 196–201, doi:10.1016/j.procir.2016.09.033.

77. Makridis, G.; Kyriazis, D.; Plitsos, S. Predictive Maintenance Leveraging Machine Learning for Time-Series Forecasting in the Maritime Industry. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC); September **2020**; pp. 1–8.

78. Zhagparov; Buribayev, Zh.; Joldasbayev, S.; Yerkosova, A.; Zhassuzak, M. Building a System for Predicting the Yield of Grain Crops Based On Machine Learning Using the XGBRegressor Algorithm. In Proceedings of the 2021 IEEE International Conference on Smart Information Systems and Technologies (SIST); April **2021**; pp. 1–5.

79. Dong, W.; Huang, Y.; Lehane, B.; Ma, G. XGBoost Algorithm-Based Prediction of Concrete Electrical Resistivity for Structural Health Monitoring. *Automation in Construction* **2020**, *114*, 103155, doi:10.1016/j.autcon.2020.103155.

80. Zhang, Z.-Y.; Wang, K.-S. Wind Turbine Fault Detection Based on SCADA Data Analysis Using ANN. *Adv. Manuf.* **2014**, *2*, 70–78, doi:10.1007/s40436-014-0061-6.

81. Mateus, B.; Mendes, M.; Farinha, J.; Assis, R.; Cardoso, A.J.M. Comparing LSTM and GRU Models to Predict the Condition of a Pulp Paper Press. *Energies* **2021**, doi:10.3390/en14216958.

82. Ayvaz, S.; Alpay, K. Predictive Maintenance System for Production Lines in Manufacturing: A Machine Learning Approach Using IoT Data in Real-Time. *Expert Systems with Applications* **2021**, *173*, 114598, doi:10.1016/j.eswa.2021.114598.

83. Fonseca, D.J.; Navaresse, D.O.; Moynihan, G.P. Simulation Metamodeling through Artificial Neural Networks. *Engineering Applications of Artificial Intelligence* **2003**, *16*, 177–183, doi:10.1016/S0952-1976(03)00043-5.

84. Partovi, F.Y.; Anandarajan, M. Classifying Inventory Using an Artificial Neural Network Approach. *Computers & Industrial Engineering* **2002**, *41*, 389–404, doi:10.1016/S0360-8352(01)00064-X.

85. Mateus, B.C.; Mendes, M.; Farinha, J.T.; Cardoso, A.M. Anticipating Future Behavior of an Industrial Press Using LSTM Networks. *Applied Sciences* **2021**, *11*, 6101, doi:10.3390/app11136101.

86. Biochemical Pulping of Wheat Straw, Solutions! & TAPPI JOURNAL, August **2004**, Vol. 3(8) (136KB) Available online: https://imisrise.tappi.org/TAPPI/Products/04/AUG/04AUG03.aspx (accessed on 16 December 2022).

87. Szabó, L.; Soria, A.; Forsström, J.; Keränen, J.T.; Hytönen, E. A World Model of the Pulp and Paper Industry: Demand, Energy Consumption and Emission Scenarios to 2030. *Environmental Science & Policy* **2009**, *12*, 257–269, doi:10.1016/j.envsci.2009.01.011.

88. Suomela, J. Condition Monitoring of Paper Machine with Thermal Imaging.; Maldague, X.P., Rozlosnik, A.E., Eds.; Orlando, FL, March **2002**; pp. 143–150.

89. Bissessur, Y.; Martin, E.B.; Morris, A.J. Machine Condition Monitoring for Consistent Paper Production. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* **1999**, *213*, 141–151, doi:10.1243/0954408991529898.

90. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis - Forecasting and Control*; **1994**.

91. Chatfield, C.; Xing, H. *The Analysis of Time Series: An Introduction with R*; **2019**; Vol. 53; ISBN 9788578110796.

92. Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis Forecasting and Control.*; John Wiley.; New Jersey, **2016**.

93. Rob J. Hyndman; Yeasmin Khandakar Automatic Time Series Forecasting: The Forecast Package for R. *Journal of Statistical Software*, **2008**, *27*, 22.

94. Chung, W.W.C.; Leung, S.W.F. Collaborative Planning, Forecasting and Replenishment: A Case Study in Copper Clad Laminate Industry. *Production Planning & Control* **2005**, *16*, 563–574, doi:10.1080/09537280500112488.

95. Granger, C.; Newbold, P. *Forecasting Economic Time Series*; New York, **1986**.

96. Brockwell, P.J.; Davis, R.A. *Introduction to Time Series and Forecasting*; Springer: New York, 2002; ISBN 0-387-95351-5.

97. Zhang, G.P. Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Neurocomputing* **2003**, *50*, 159–175, doi:10.1016/S0925-2312(01)00702-0.

98. Spearman, C. The Abilities of Man Their Nature and Measurement. *Nature* **1927**, *120*, 181–183, doi:10.1038/120181a0.

99. Sternberg, R.J. Intelligence: State of the Art. *Dialogues in Clinical Neuroscience* **2012**, *14*, 19–27.

100. Mondal, B. Artificial Intelligence: State of the Art. In *Intelligent Systems Reference Library*; **2019**; Vol. 172, pp. 389–425.

101. Brodić, D.; Amelio, A. Artificial Intelligence and Turing Test. In; **2020**; pp. 1–5.

102. Hopcroft, J.E. Turing Machines. *Scientific American* **1984**, *250*, 86-E9.

103. French, R.M. The Turing Test: The First 50 Years. *Trends in Cognitive Sciences* **2000**, *4*, 115–122, doi:10.1016/S1364-6613(00)01453-4.

104. Copeland, B.J. The Turing Test*. Minds and Machines **2000**, 10, 519–539, doi:10.1023/A:1011285919106.

105. Abraham, T.H. (Physio)Logical Circuits: The Intellectual Origins of the McCulloch–Pitts Neural Networks. *Journal of the History of the Behavioral Sciences* **2002**, *38*, 3–25, doi:10.1002/jhbs.1094.

106. McCulloch, W.S.; Pitts, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics* **1943**, *5*, 115–133, doi:10.1007/BF02478259.

107. Kovács, Z.L. *Redes Neurais Artificiais*; Editora Livraria da Fisica, **2002**; ISBN 978-85-88325-14-2.

108. Strawn, G.; Strawn, C. Masterminds of Artificial Intelligence: Marvin Minsky and Seymour Papert. *IT Professional* **2016**, *18*, 62–64, doi:10.1109/MITP.2016.116.

109. Oymak, S.; Soltanolkotabi, M. Toward Moderate Overparameterization: Global Convergence Guarantees for Training Shallow Neural Networks. *IEEE Journal on Selected Areas in Information Theory* **2020**, *1*, 84–105, doi:10.1109/JSAIT.2020.2991332.

110. Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding Deep Learning (Still) Requires Rethinking Generalization. *Commun. ACM* **2021**, *64*, 107–115, doi:10.1145/3446776.

111. Piotrowski, A.P.; Napiorkowski, J.J.; Piotrowska, A.E. Impact of Deep Learning-Based Dropout on Shallow Neural Networks Applied to Stream Temperature Modelling. *Earth-Science Reviews* **2020**, *201*, 103076, doi:10.1016/j.earscirev.2019.103076.

112. Jonsson, P. Company-Wide Integration of Strategic Maintenance: An Empirical Analysis. *International Journal of Production Economics* **1999**, *60–61*, 155–164, doi:10.1016/S0925-5273(98)00148-0.

113. Carnero, Mc. An Evaluation System of the Setting up of Predictive Maintenance Programmes. *Reliability Engineering & System Safety* **2006**, *91*, 945–963, doi:10.1016/j.ress.2005.09.003.

114. Yamashina, H. Japanese Manufacturing Strategy Competing with the Tigers. *Business Strategy Review* **1996**, *7*, 23–36, doi:10.1111/j.1467-8616.1996.tb00120.x.

115. Sazli, M. A Brief Review of Feed-Forward Neural Networks. *Communications, Faculty Of Science, University of Ankara* **2006**, *50*, 11–17, doi:10.1501/0003168.

116. Tahmasebi, P.; Hezarkhani, A. Application of a Modular Feedforward Neural Network for Grade Estimation. *Nat Resour Res* **2011**, *20*, 25–32, doi:10.1007/s11053-011-9135-3.

117. Adeli, H.; Yeh, C. Perceptron Learning in Engineering Design. *Computer-Aided Civil and Infrastructure Engineering* **1989**, *4*, 247–256, doi:10.1111/j.1467-8667.1989.tb00026.x.

118. Noriega, L. Multilayer perceptron tutorial. School of Computing. Staffordshire University, **2005**, 4.5: 444.

119. Elavarasan, D.; Vincent, D.R. Reinforced XGBoost Machine Learning Model for Sustainable Intelligent Agrarian Applications. *Journal of Intelligent & Fuzzy Systems* **2020**, *39*, 7605–7620, doi:10.3233/JIFS-200862.

120. Lara-Benítez, P.; Carranza-García, M.; Luna-Romera, J.M.; Riquelme, J.C. Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting. *Applied Sciences* **2020**, *10*, 2322, doi:10.3390/app10072322.

121. Cho, M.; Kim, C.; Jung, K.; Jung, H. Water Level Prediction Model Applying a Long Short-Term Memory (LSTM)–Gated Recurrent Unit (GRU) Method for Flood Prediction. *Water* **2022**, *14*, 2221, doi:10.3390/w14142221.

122. Yu, Z.; Moirangthem, D.S.; Lee, M. Continuous Timescale Long-Short Term Memory Neural Network for Human Intent Understanding. Frontiers in Neurorobotics **2017**, 11.

123. Santra, A.S.; Lin, J.-L. Integrating Long Short-Term Memory and Genetic Algorithm for Short-Term Load Forecasting. *Energies* **2019**, *12*, 2040, doi:10.3390/en12112040.

124. Ertunc, H.M.; Ocak, H.; Aliustaoglu, C. ANN- and ANFIS-Based Multi-Staged Decision Algorithm for the Detection and Diagnosis of Bearing Faults. *Neural Comput & Applic* **2013**, *22*, 435–446, doi:10.1007/s00521-012-0912-7.

125. Li, D.; Sun, G.; Miao, S.; Gu, Y.; Zhang, Y.; He, S. A Short-Term Electric Load Forecast Method Based on Improved Sequence-to-Sequence GRU with Adaptive Temporal Dependence. *International Journal of Electrical Power & Energy Systems* **2022**, *137*, 107627, doi:10.1016/j.ijepes.2021.107627.

126. Qi, Y. Random Forest for Bioinformatics. In *Ensemble Machine Learning: Methods and Applications*; Zhang, C., Ma, Y., Eds.; Springer US: Boston, MA, **2012**; pp. 307–323 ISBN 978-1-4419-9326-7.

127. Amit, Y.; Geman, D. Shape Quantization and Recognition with Randomized Trees. *Neural Computation* **1997**, *9*, 1545–1588, doi:10.1162/neco.1997.9.7.1545.

128. Biau, G. Analysis of a Random Forests Model. *Journal of Machine Learning Research* **2012**, *13*, 1063–1095.

129. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5–32, doi:10.1017/CBO9781107415324.004.

130. Amruthnath, N.; Gupta, T. Factor Analysis in Fault Diagnostics Using Random Forest. *Ind Eng Manage* **2019**, *8*, 1–10, doi:10.4172/2169-0316.1000278.

131. Liaw, A.; Wiener, M. Classification and Regression by randomForest. *R News* **2002**, *2*, 18–22.

132. Sinaga, K.P.; Yang, M.-S. Unsupervised K-Means Clustering Algorithm. *IEEE Access* **2020**, *8*, 80716–80727, doi:10.1109/ACCESS.2020.2988796.

133. Hongyu, K.; Sandanielo, V.; Junior, G. Análise de Componentes Principais: Resumo Teórico, Aplicação e Interpretação. *E&S Engineering and Science* **2016**, *5*, 83, doi:10.18607/ES201653398.

134. Hotelling, H. Simplified Calculation of Principal Components. *Psychometrika* **1936**, *1*, 27–35, doi:10.1007/BF02287921.

135. Niraj, M.; Kumar, P. FMECA Analysis (A Heuristic Approach) For Frequency of Maintenance and Type of Maintenance. *International Journal of Scientific & Engineering Research* **2012**, *3*, 1–5.

136. Wong, K.C.; Woo, K.Z.; Woo, K.H. Ishikawa Diagram. In *Quality Improvement in Behavioral Health*; O'Donohue, W., Maragakis, A., Eds.; Springer International Publishing: Cham, 2016; pp. 119–132 ISBN 978-3-319-26209-3.

137. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623, doi:10.7717/peerj-cs.623.

138. Rodrigues, J.A.; Gonçalves, I.; Farinha, J.T.; Mendes, M.; Margalho, L. Predicting Motor Oil Condition Using Artificial Neural Networks and Principal Component Analysis. *Eksploatacja i Niezawodnosc - Maintenance and Reliability* **2020**, *22*, 440–448, doi:10.17531/ein.2020.3.6.

139. Rodrigues, J.A.; Farinha, J.T.; Mendes, M.; Gonçalves, I.; Margalho, L. Modelling Diesel Engine Oil Condition Using Artificial Neural. eMaintenance 2019, May **2019**, Stockholm, Sweeden https://www.researchgate.net/publication/333176391_Modelling_Diesel_Engine_Oil_Condition_Using_Artificial_Neural (accessed on 16 May 2020).

140. Cerny, B.A.; Kaiser, H.F. A Study Of A Measure Of Sampling Adequacy For Factor-Analytic Correlation Matrices. *Multivariate Behavioral Research* **1977**, *12*, 43–47, doi:10.1207/s15327906mbr1201_3.

141. Rodrigues, J.A.; Farinha, J.; Mendes, M.; Mateus, R.; Cardoso, A.J.M.; Rodrigues, J. Short and Long Forecast to Implement Predictive Maintenance in a Pulp Industry. *Eksploatacja i Niezawodnosc - Maintenance and Reliability* **2021**, *24*, 33–41, doi:10.17531/ein.2022.1.5.

142. Lu, B.; Durocher, D.B.; Stemper, P. Predictive Maintenance Techniques. *IEEE*

Industry Applications Magazine **2009**, *15*, 52–60, doi:10.1109/MIAS.2009.934444.

143. Cahyana, R. A Preliminary Investigation of Information System Using Ishikawa Diagram and Sectoral Statistics. *IOP Conf. Ser.: Mater. Sci. Eng.* **2018**, *434*, 012050, doi:10.1088/1757-899X/434/1/012050.

144. Organisation internationale de normalisation; Commission électrotechnique internationale; Association française de normalisation *Internacional Standart - Failure Modes and Effects Analysis (FMEA and FMECA) - EC 60812:2018*; **2018**; ISBN 978-2-8322-5915-3.

145. Berger, R.; Benbow, D.; Ahmad Elshennawy; Walker, H. The Certified Quality Engineer Handbook, Second Edition. *Faculty and Staff Books* **2006**.

146. Lipol, L.S.; Haq, J. Risk Analysis Method: FMEA/FMECA in the Organizations. **2011**, *11*, 9.

147. Rodrigues J. A., Farinha J. T., Cardoso A.J.M. et al. Prediction of Sensor Values in Paper Pulp Industry Using Neural Networks. In Zhang H, Feng G, Wang H et al. (eds): Proceedings of IncoME-VI and TEPEN **2021**, Cham, Springer International Publishing: 2023: 281–291, https://doi.org/10.1007/978-3-030-99075-6_24.

148. Rodrigues J.A., Farinha J., Mendes M., Cardoso A.J.M., Mateus R. Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition. Energies **2022**; 15(17): 6308, https://doi.org/10.3390/en15176308.

149. Mateus, B.; Farinha, J.T.; Mendes, M.; Martins, A.B.; Cardoso, A.M. Data Analysis for Predictive Maintenance Using Time Series and Deep Learning Models – A Case Study in a Pulp Paper Industry. **2021**.

150. Rodrigues, J.; Martins, A.; Farinha, J.; Mendes, M.; Mateus, R.; Cardoso, A.J.M. Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning. *Energies* **2022**, *15*, doi:10.3390/en15249387.

151. Jain, A.K.; Murty, M.N.; Flynn, P.J. Data Clustering: A Review. *ACM Comput. Surv.* **1999**, *31*, 264–323, doi:10.1145/331499.331504.

152. Xie, J.; Jiang, S.; Xie, W.; Gao, X. An Efficient Global K-Means Clustering Algorithm. *JCP* **2011**, *6*, 271–279, doi:10.4304/jcp.6.2.271-279.

153. Sklearn.Neural_network.MLPClassifier Available online: https://scikit-

learn/stable/modules/generated/sklearn.neural_network.MLPClassifier.html
(accessed on 27 July 2022).

Appendices

Appendix A

Article citation info:
RODRIGUES J, COSTA I, TORRES FARINHA J, MENDES M, MARGALHO L. Predicting motor oil condition using artificial neural networks and principal component analysis. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2020; 22 (3): 440–448, http://dx.doi.org/10.17531/ein.2020.3.6.

João RODRIGUES
Inês COSTA
J. TORRES FARINHA
Mateus MENDES
Luís MARGALHO

PREDICTING MOTOR OIL CONDITION USING ARTIFICIAL NEURAL NETWORKS AND PRINCIPAL COMPONENT ANALYSIS

PROGNOZOWANIE STANU OLEJU SILNIKOWEGO ZA POMOCĄ SZTUCZNYCH SIECI NEURONOWYCH I ANALIZY SKŁADOWYCH GŁÓWNYCH

The safety and performance of engines such as Diesel, gas or even wind turbines depends on the quality and condition of the lubricant oil. Assessment of engine oil condition is done based on more than twenty variables that have, individually, variations that depend on the engines' behaviour, type and other factors. The present paper describes a model to automatically classify the oil condition, using Artificial Neural Networks and Principal Component Analysis. The study was done using data obtained from two passenger bus companies in a country of Southern Europe. The results show the importance of each variable monitored for determining the ideal time to change oil. In many cases, it may be possible to enlarge intervals between maintenance interventions, while in other cases the oil passed the ideal change point.

Keywords: *condition monitoring, oil analysis, multivariate analysis, predictive maintenance.*

Bezpieczeństwo i wydajność silników takich, jak silniki Diesla czy gazowe, a nawet turbiny wiatrowe, zależą od jakości i stanu oleju smarowego. Stanu oleju silnikowego ocenia się na podstawie ponad dwudziestu zmiennych, z których każda ulega wahaniom w zależności od typu i zachowania silnika oraz innych czynników. W niniejszym artykule opisano model, który pozwala na automatyczną klasyfikację stanu oleju, z wykorzystaniem sztucznych sieci neuronowych i analizy składowych głównych. Badania przeprowadzono na podstawie danych uzyskanych od dwóch przewoźników pasażerskich działających na terenie jednego z krajów położonych na południu Europy. Wyniki pokazują, że każda z monitorowanych zmiennych ma znaczenie dla określenia idealnego czasu na wymianę oleju. Podczas gdy w wielu przypadkach w badanych przedsiębiorstwach możliwe było zwiększenie odstępów czasowych między działaniami konserwacyjnymi, w innych, idealny moment wymiany oleju został przekroczony.

Słowa kluczowe: *monitorowanie stanu, analiza oleju, analiza wielowymiarowa, konserwacja predykcyjna.*

1. Introduction

Condition monitoring of engines' oil is a strategic area in the maintenance management field. Replacing the oil too early represents unnecessary unavailability, as well a financial and environmental costs which could be spared. Replacing it too late can impair the oil's ability to protect the engine, therefore increasing the chances of damage and premature ageing of the engine, or even the risk of causing accidents which can endanger people, equipments or vehicles in urban environments. The use of modern tools from data mining and Artificial Intelligence (AI) can contribute to help make the right decision at the right time, thus protecting the environment, the companies' profits and the safety of people and property.

The present paper discusses a methodology to create models to facilitate the process of oil analysis, tested with a dataset for oil of Diesel engines, from urban passenger buses. Preliminary work was already done [14], using Artificial Neural Networks (ANN). In the present research, the neural models were improved and the results are compared with analysis using multivariate systems, namely Principal Component Analysis (PCA). PCA showed the relevance of each variable is different, and some of the variables may even have a negative impact on the predictive power of the ANN.

Data used for the experiments come from two passenger bus companies. Each company provided a dataset, containing results of laboratory analysis of the oils and their classification, according to human experts of a specialized oil analysis company. Data were mined and neural models were created, for both datasets separated and combined.

The remainder of the paper is organized as follows. Section 2 presents a summary of the state of the art. Section 3 describes the datasets used. Section 4 describes the neural networks. Section 5 describes the analysis performed using multivariate systems. Section 7 presents a comparative and critical analysis of the results obtained. Section 8 highlights the main contributions of the present research. Section 9 presents some conclusions and outlines future work.

2. Literature Review

2.1. Condition monitoring of Diesel engines' oil

Condition monitoring of Diesel engines' oil has been subject to study using different approaches, including machine learning methods. Raposo et al. present a study about condition monitoring based on oil in the Diesel engines of a fleet of urban buses. The study shows

(*) Tekst artykułu w polskiej wersji językowej dostępny w elektronicznym wydaniu kwartalnika na stronie ***www.ein.org.pl*** the evolution of oil degradation and develops a predictive maintenance policy for oil replacement [13]. The methodology presented by the authors considers only some variables of the oils, showing very interesting results about the P-F curve accompaniment. The P-F curve is the interval between the detection of a Potential failure and the actual Failure— that is, the interval where a maintenance team should intervene to prevent a potential failure from happening.

Gajewski & Valis present a study that focuses on heavy transport systems. The types of oils were obtained from several dozens engines of heavy crawlers. The study uses these data with neural networks, in order to identify the patterns that model the system deterioration [5]. Hongxiang et al. [7] use a feed forward neural network to classify different types of oil and their running/not running condition. Parlak et al. [12] use an ANN to predict specific fuel consumption and Diesel engine temperature.

2.2. Online condition monitoring of engine oil

Monitoring the condition or the engine's oil in real time is also a long sought goal. Oil degradation depends on the working time, kilometers driven, the driving speed and habits, type of motor, age of the motor and many other variables which may cause faster or slower degradation. Therefore, for replacement of the oil at the best time, it is necessary to make analysis to determine the state of degradation of the oil. The main problems with laboratory analysis is that they are a laborious process, which requires human intervention. Even though it is only necessary each several km, the idea is to automate the process, thus lowering the costs and reducing the probability of human error.

Accurate online monitoring has two main advantages: i) it reduces downtime necessary to inspect the oil; and ii) it increases chances that the oil will be changed at the best time, not too early and not too late. On the downside, it requires adequate sensors that can be put in the engines, in contact with the oil. The sensors must be robust enough to endure the operating conditions without failing and they must be precise enough to give accurate readings. J. Zhu et al. present a fair review of state of the art sensors and methods for online condition monitoring of engines' oil [18]. The authors classify the sensors in four different groups: electromagnetic, physical, chemical and optical. Electromagnetic sensors measure the dielectric constant of the oil. A second type measures the oil's conductivity. A third type measures magnetic susceptibility. A fourth type measures oil viscosity. As for physical methods, Zhu et al. mention the viscometer, ultra sound, thermal conductivity sensor and ferrography. As for chemical methods, the techniques reviewed include pH measurement and thin-film contaminant monitor. As for optical techniques, they are reflectometry and infrared absorption.

S. Kumar et al. propose a method for condition monitoring oil engine online, based on an optical sensor that transduces oil darkness into electrical resistance [8]. The colour change of the oil is one of the variables that directly correlates to the quality of the oil. Therefore, the authors argue that monitoring the oil colour it is possible to determine the degradation of the oil and change it at the most appropriate time. The method is in part similar to [17], where Yonghui et al., who combine the use of a fibre optic transducer and an inductive sensor. The inductive sensor detects large ferrous and some non-ferrous wear debris, while the optical sensor detects small particles contamining the oil.

El-Hag et al. use features extracted from acousting and radio frequency partial discharge signals to monitor oil condition in power transformers [4]. Pulse width, rise time and frequency components are used to train a neural network to assess the level of degradation of the oil. J. Zhu et al. propose a method for condition monitoring wind turbine oil using commercially available sensors to measure oil viscosity and dielectric constant sensors [19]. The sensor readings are calibrated based on the relationship between particle concentration and oil degradation. X. Zhu et al. propose a method to condition monitor oil using a sensor that detects wear debris by measuring the inductance change of two planar coils wound around a pair of ferrite cores [20]. The method is in part similar to Du et al.'s approach, which also uses inductive sensors to measure metallic, ferrous and non-ferrous, particles in lubricant oil [3].

2.3. Use of PCA and Artificial Neural Networks in condition monitoring

Principal Component Analysis is one method of multivariate analysis very popular for data mining. It is a statistical procedure to transform data, extract features and determine the most important variables of a dataset. Through PCA analysis, it is therefore possible to predict which variables deserve to be monitored and which variables are candidate to be removed without loosing predictive power. Westerholm and Li use PCA to determine the relationship between fuel parameters and the amount of particles in Diesel motor emissions [16]. Capone et al. use PCA to determine the amount of unburned fuel in lubricating oil [1].

Different Artificial Neural Network architectures have also been used for learning and predicting oil condition. Shaban et al. use a cascade of artificial neural networks to predict transformer oil parameters [15]. Niu et al. compare the performance of ANN and Support Vector Machines for predicting motor emissions [11]. Ghobadian et al. use an ANN to model the performance of a diesel engine using waste oil [6]. Li X et al. and Li Y et al. use convolution neural networks to detect gear faults based on different signals, namely sounds produced by the gears [9] and operational parameters [10].

3. Datasets Used

The present research was performed using two datasets, obtained from two different public transportation bus companies, named A and B for that purpose. The datasets contain the results of 21 parameters of the laboratory analysis of the oil, taken from the buses at different stages. Each sample also contains the bus and the oil mileage. That is a total of 23 parameters which are used as input variables for the present analysis. The 23 parameters are: mileage of the bus, mileage of the oil, amount of antifreeze found in the oil, percentage of fuel, Finacheck water content, sooth, nitration, oxidation, sulfation, TBN, viscosity at 100°C, Al, Cr, Cu, Fe, Mo, Na, Ni, Pb, Si, Sn, V and PQ. The variables were normalized and used as inputs to the neural networks and PCA analysis. The datasets also contain the decision of the specialized company, marked as 1, 2 or 3. Decision 1 means that the company decided the oil is in good condition and can be maintained for normal bus operation. Decision 2 means the oil is reaching the point where it needs to be replaced. Decision 3 means the oil has passed the point when it should have been replaced and the bus must be immediately stopped for safety reasons. Dataset A contains a total of 47 samples, obtained from a number of different buses of company A. Dataset B contains a total of 88 samples, obtained from twenty two different buses, four samples from each bus. For the present study it was not possible to obtain larger datasets—a limitation which could not be overcome. Nonetheless, the results obtained for the neural models and PCA were consistent. They were repeated a number of times and they are repeatable in similar circumstances. Many neural models showed good performance and small error in the train and test sets. The ones preferred for analysis were those with better performance in the test set, thus showing the model is general. That shows the results are valid and the method could be scaled up to larger datasets. Since PCA is a factor-analysis method, the adequacy of the datasets for PCA was tested using Kaiser-Meyer-Olkin (KMO) test [2]. KMO test gives a score between 0 and 1, where in general the higher scores mean the dataset contains enough diversity of samples to apply factor analysis. A low score, on the other hand, means there are high correlations between the variables and the results of the factoring process are unreliable. The KMO test gives a score of 0.35 for dataset A, which is very low, and a score of 0.636 for dataset B, which is acceptable, meaning there is more confidence in the factor analysis results for dataset B.

4. Neural Models

In the present research the neural models used were shallow Feed Forward Neural Networks, with one hidden layer of variable width (number of neurons), and one output layer. The hidden neurons used a sigmoid transfer function, which can be a universal approximation, maintaining the output in the range [0, 1]. The output neuron used a linear transfer function (relu), to allow for a wider amplitude of the output and facilitate the learning process. The models were created and tested in MatlabTM1. Training was performed using the Levenberg-Marquartd method. The Mean Squared Error (MSE) and correlation factor R were used for performance assessment. Training was performed with 70 % of the samples, validation with 15 % and test with the remainder 15 %. The results of the training process are variable for each experiment. That happens when the initial weights and bias of the neurons are not set for a specific value, or if the samples for the training and validation sets are chosen randomly at each experiment. Therefore, the results of the experiments presented below are selected from a number of runs. During training, the training process was stopped when the error increased in the validation set for two consecutive epochs, and the best model was retained. The output obtained from the neural models is a floating point number. That is desirable, so that it is interpreted as a measure of the quality of the oil: the lowest the value, the better the quality of the oil. On the other hand, it is also important to map the output to 1, 2 or 3, in order to obtain a model that can be compared to the classification of the human experts, as described in Section 3. So it was mapped in the discrete interval [1, 3] using the following rules: Anything below 1.50 was mapped to 1; Numbers in the interval [1.50, 2.50[were mapped to 2 and everything greater or equal to 2.50 was mapped to 3.

Table 1. R and MSE obtained for different network sizes, with dataset A

Hidden layer size	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
1	0.98	0.83	0.78	0.91	0.485
2	0.98	0.76	0.80	0.93	0.234
3	0.99	0.96	0.96	0.98	0.051
4	0.99	0.87	0.88	0.95	0.258
5	0.98	0.91	0.70	0.92	0.229
6	0.95	0.90	0.89	0.92	0.216
7	0.98	0.83	0.98	0.96	0.178
8	0.93	0.83	0.88	0.89	0.291
9	0.92	0.88	0.98	0.91	0.178
10	0.90	0.99	0.91	0.89	0.015

Table 2. Confusion matrices of the errors of the model trained with data from Company A. The model shows two prediction errors for Company A, but 39 errors for Company B

Predicted	Company A			Company B		
3	0	0	**14**	0	0	**4**
2	0	**11**	1	0	**0**	2
1	**20**	1	0	**45**	15	22
Actual	1	2	3	1	2	3

4.1. Model for Company A

In order to get the best possible results, it is important to determine the optimal size of the neural network, so that it is able to abstract and retain as much information as possible without overfitting the training data.

Table 1 shows the R and MSE obtained for a number of neurons between 1 and 10, for dataset A. Many models show good R and MSE. The model with three neurons is one of the best, since it shows a good R for all the dataset and a small MSE. It also shows a good R for the test set, which means it is a good general model, performing well even for data that it has never seen. The number of neurons is small for the number of inputs, but it is probably in line with the size of the dataset, which is also small. The model was trained in three epochs, after which it started to show signs of overfitting and, therefore, further training was rejected.

Table 2 shows the confusion matrices with a summary of the distribution of the errors of the model described above, when applied to both datasets. As the table shows, the predictions of the model for Company A are very close to the desired output. There are only two errors, when the model predicted 2 and the decision was 1 and the model predicted 3 while the company decided 2. In both cases the company was more defensive than the model. It should be mentioned that the decisions made by the company are also prone to human error, so the errors shown are not necessarily problems of the model—they can be because of outliers in the dataset.

When the same model was applied to the data obtained from Company B, there were a total of 39 errors in 88 samples. All the errors apparently happen because Company B was more defensive than the model, showing a very clear trend: the experts rated the oil worse than the model which performed very well for Company A. This proves that Company B replaces the oil, on average, before Company A. That may happen because of different maintenance policies, different motors or different oil brands.

Table 3. R and MSE obtained for different network sizes, with dataset B

Hidden layer size	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
1	0.98	0.97	0.86	0.95	0.060
2	0.95	0.93	0.88	0.94	0.110
3	0.99	0.90	0.88	0.96	0.222
4	0.95	0.95	0.83	0.92	0.078
5	0.97	0.86	0.80	0.93	0.161
6	0.86	0.84	0.86	0.86	0.249
7	0.91	0.91	0.91	0.90	0.256
8	0.97	0.89	0.93	0.94	0.144
9	**0.95**	**0.91**	**0.94**	**0.95**	**0.073**
10	0.92	0.89	0.92	0.90	0.184

Table 4. Confusion matrices of the errors of the model trained with data from Company B. The model shows six prediction errors for Company B, but thirty three errors for Company A

Predicted	Company A			Company B		
3	16	12	**14**	0	0	**26**
2	4	**0**	0	4	**15**	2
1	**0**	0	1	**41**	0	0
Actual	1	2	3	1	2	3

4.2. Model for Company B

Table 3 shows the R and MSE obtained from a number of neurons between 1 and 10, for models trained with dataset B. The network with one neuron in the hidden layer shows the lowest MSE. However, it also shows a poor R in the test set, compared to the train and validation sets, meaning the model is a bit overfitted. The model with nine neurons in the hidden layer is better, considering that it shows a high R for all dataset, a high R for the test set and a small MSE. That shows the model is more general for the specific problem being addressed. The best performance for the test and validation sets was obtained at epoch four, and after that the model starts to overfit the data.

Table 4 shows the confusion matrices for the model, applied to datasets B and A. On dataset B, there are six errors of the model, compared to the decision of the company. In four situations the model predicted 2, while the company decided 1. So, the model was more defensive, proposing the bus to replace the oil, while the company decided it was good to circulate. In two cases, the model predicted 1 and the company decided 2.

When the same model was applied to data from Company A, there was a total of thirty three errors. In one situation the company was more defensive: one time the model predicted 1 and the company, apparently, decided to stop the bus. In thirty two situations the model was more defensive: sixteen times the model predicted the bus should be taken out of circulation and the company decided the oil was good, twelve times the model predicted the bus should be taken out of circulation and the company just decided to replace the oil, and four times the model predicted the oil should be replaced and the company decided it was good. The results of this confusion matrix are according to the one shown in Table 2: the companies follow different policies, and Company B replaces the oil much earlier than Company A.

5. Principal Component Analysis

5.1. Introduction

PCA is a statistical procedure used to map a set of correlated variables into a new set of uncorrelated variables, called principal components. The principal components are calculated by decreasing order of importance. The first component is the most important, the last is the less important explanatory variable. Each principal component identified is a linear combination of all the original variables. PCA was applied to the datasets presented above, in order to understand the companies' policies, the state of the oils when the samples were collected for chemical analysis, as well as to determine which variables are more important to measure for correct assessment of the situation of the oil. The PCA experiments and analysis were performed using R Studio software.

5.2. Loadings for each variable of the dataset

Table 5 shows results of the PCA analysis for datasets A and B. As the table shows, Si, Fe, Al and Cr contents are the four most important variables, with loadings above 0.7, for dataset A.

For dataset B, Fe, Soot and Cr are the top three variables, and the only ones with score above 0.7. Five of the top ten variables are related to oil status and five are related to wear and contamination.

5.3. Company A

For Company A, the most important components are polluting metallic agents, which are generated by motor wear: slip wear, wear due to friction, wear due to metal fatigue, and wear due to cutting. This means the bus motors suffer a lot of wearing, being advisable the use of additives to reduce friction, oil leaks and even fumes. The use of the right additives might also increase motor expected life.

Table 6 shows the calculated percentage of deterioration of the oil, for each oil sample, as well as the average deterioration. The percentage of deterioration was obtained multiplying the PCA loadings by the normalized value of each variable. So a kind of weighted average is obtained and then compared to the reference values obtained from the oil datasheet.

As the table shows, four different oils are already well beyond their expected useful life, based on manufacturer recommendations. More than 41 % of the samples have passed 60 % degradation. Another relevant aspect is that there is a large variability between samples. The standard deviation in the dataset is 0.312. In general, Company A apparently has a poor maintenance policy. On one hand, the motors suffer a lot of wearing when the oils are still used beyond their reference limits. That can damage the motors and reduce their useful life. On the other hand, some oils may be changed while they are still good, causing unnecessary financial and environmental costs. In two

Table 5. Most important variables, according to PCA analysis, for datasets A and B. The ten first variables that are in both datasets are highlighted in bold

Order of Relevance	Variable	Loading	Variable	Loading
1	**Si Content**	0.872	**Fe Content**	0.889
2	**Fe Content**	0.864	**Soot**	0.835
3	**Al Content**	0.789	**Cr Content**	0.781
4	**Cr Content**	0.729	Viscosity at 100°C	0.689
5	**Sn Content**	0.668	**Sn Content**	0.682
6	**PQ Index**	0.551	**Cu Content**	0.611
7	Ni Content	0.425	Pb Content	0.571
8	**Soot**	0.441	**Sulfation**	0.507
9	**Oxidation**	0.412	Nitration	0.496
10	V Content	0.376	**Oxidation**	0.488
11	**Cu Content**	0.282	**Al Content**	0.482
12	**Sulfation**	0.266	**Si Content**	0.423
13	Mo Content	0.166	**PQ Index**	0.32
14	Pb Content	0.134	Na content	0.166
15	Fuel	0.132	Antifreeze	0.162
16	Na Content	0.118	Water Content	0.127
17	Viscosity	0.089	Mo Content	0.072
18	TBN	0.069	V Content	0.02
19	Nitration	-0.003	Ni Content	-0.008
20	Water content	-0.14	Fuel Content	-0.134
21	Antifreeze	-0.142	TBN Content	-0.395

cases the oils are apparently at just 15 % of their useful life when they were replaced, according to the results of the laboratorial analysis.

5.4. Company B

Table 6 shows the percentage of degradation of the oil samples. As the table shows, none of the samples is out of the limits proposed by the manufacturer, showing the oil is in good condition to protect the Diesel engines. In general the dataset is also more homogenous, with most of the variables near the average value. The standard deviation is 0.126, which is much lower than the standard deviation calculated for dataset A.

Table 8 shows tha average deterioration of oil for each of the twenty two buses. The average is always in the range 30 – 59 %. Bus 1814 shows the lowest average of all. Bus 2160 shows the highest average, with 58.65 %, which is already a very high value. In general, however, the buses of company B show a low level of wear, with the five variables related to wearing and contamination within the ten highest principal components because of their importance in oil deterioration.

6. Neural model with reduced dimensionality

6.1. Merging datasets

Since both datasets contain the same 23 input variables and three-levels output variable, they are fit to be merged together, aiming to produce one more general model. However, since data come from different sources, it is important to avoid skewing the results towards the policies of one of the companies, for the previous analysis have shown that the companies follow different policies and the datasets may have outliers. The problem can be seen as the typical problem of imbalanced datasets, which is solved using techniques such as oversampling the less frequent data or undersampling the most frequent data. In the present case, considering the small dimension of the dataset, data from Company A were oversampled choosing randomly additional samples from dataset A after cross validation.

The merged dataset was named AB and it contains a total of 176 samples, 88 from each dataset. The neural models used for experiments with dataset AB all contained seven neurons in the hidden layer. The number of inputs, however, varied: i) one of the experiments was run with all the 21 laboratorial variables and the mileage as inputs for the neural network; and ii) a second experiment was run using as inputs to the neural network just the ten variables highlighted in Table 5.

Table 9 shows the R and MSE obtained for different models. As the table shows, the model with 23 input variables was very good, with a very small MSE. The model with just 12 inputs is also very good, with R 0.94 for all the dataset and 0.84 for the test set. The table shows that reducing the number of variables it was still possible to generate a good neural model, with good R and a small MSE of just 0.15

Table 10 shows the confusion matrices of the errors counted when simulating the two models trained with dataset AB. The model trained with all the 23 inputs generates 13 prediction errors: five for samples of company B and 8 for company A. .

The model trained with just 12 input variables produces a total of 14 prediction errors: 7 for each company. Those results also show that the model trained with just 10+2 variables seems more general than the model trained with all the variables, for it generates the same number of prediction errors for each company. The smaller model, retaining less information, still shows good performance marks and is perhaps the most general, as discussed in Section 7.

Table 6. Percentage of deterioration of the oil for dataset A

Sample	Bus #	% Deterioration	Sample	Bus #	% Deterioration
1	122	31.7	25	267	66.7
2	122	13.6	26	270	32.5
3	203	49.9	27	270	54.2
4	203	18.8	28	270	89.1
5	214	52.1	29	282	54.6
6	214	150.6	30	282	51.7
7	214	35.6	31	283	69.0
8	214	17.6	32	283	70.7
9	219	13.4	33	289	73.1
10	246	79.1	34	289	85.3
11	246	64.0	35	290	49.3
12	247	83.6	36	294	55.8
13	248	73.6	37	294	54.3
14	249	46.5	38	297	62.3
15	251	54.1	39	209	72.5
16	252	69.2	40	209	73.3
17	254	162.1	41	301	31.0
18	259	128.0	42	301	86.8
19	260	88.3	43	301	31.7
20	265	118.3	44	304	66.4
21	266	48.3	45	304	51.6
22	266	45.6	46	304	54.8
23	266	63.7	47	304	35.7
24	267	55.2	**Average**		62.5

Table 7. Percentage of deterioration (Det.) of the oil for dataset B

Sample	Bus #	% Det.	Sample	Bus #	% Det.	Sample	Bus #	% Det.
1	2175	31.1	31	1737	29.1	61	2127	64.6
2	2175	30.4	32	1737	28.4	62	2127	56.7
3	2175	38.7	33	2148	84.2	63	2127	37.4
4	2175	30.0	34	2148	46.1	64	2127	38.8
5	1730	31.9	35	2148	41.0	65	2119	81.9
6	1730	40.5	36	2148	48.7	66	2119	40.0
7	1730	40.1	37	2131	64.1	67	2119	37.1
8	1730	33.0	38	2131	49.1	68	2119	35.3
9	1764	37.3	39	2131	40.8	69	1708	74.8
10	1764	29.3	40	2131	51.9	70	1708	31.2
11	1764	36.4	41	1814	42.8	71	1708	46.2
12	1764	34.2	42	1814	27.7	72	1708	32.2
13	1778	32.9	43	1814	24.5	73	1727	35.5
14	1778	37.7	44	1814	25.0	74	1727	33.9
15	1778	51.0	45	2169	52.8	75	1727	26.8
16	1778	36.3	46	2169	37.2	76	1727	31.0
17	1739	31.6	47	2169	42.6	77	1743	37.4
18	1739	32.5	48	2169	47.2	78	1743	33.3
19	1739	32.4	49	2128	40.3	79	1743	34.7
20	1739	30.6	50	2128	42.6	80	1743	31.0
21	2159	39.8	51	2128	31.8	81	1734	41.3
22	2159	27.1	52	2128	39.6	82	1734	31.4
23	2159	40.0	53	2150	32.7	83	1734	30.7
24	2159	32.3	54	2150	61.4	84	1734	35.9
25	2136	44.2	55	2150	35.8	85	2160	81.1
26	2136	47.1	56	2150	37.2	86	2160	58.7
27	2136	38.4	57	2152	68.1	87	2160	46.5
28	2136	42.0	58	2152	40.4	88	2160	48.3
29	1737	33.5	59	2152	49.1	**Average**		**40.6**
30	1737	32.0	60	2152	37.5			

Table 8. Average deterioration of the oil for each of the twenty two buses for Company B

Bus #	% Deterioration	Bus #	% Deterioration
2175	32.55	1730	36.38
1764	34.30	1778	39.48
1739	31.78	2159	34.80
2136	42.93	1737	30.75
2148	55.00	2131	51.48
1814	30.00	2169	44.95
2128	38.58	2150	41.78
2152	48.78	2127	49.38
2119	48.58	1708	46.10
1727	31.80	1743	34.10
1734	34.83	2160	58.65

Table 9. R and MSE obtained for the neural models trained with dataset AB, using all the input variables and then just 12 selected input variables

# inputs	R (train set)	R (validation set)	R (test set)	R (all dataset)	MSE (all dataset)
21+2	0.97	0.84	0.86	0.93	0.27
10+2	0.96	0.86	0.84	0.94	0.15

Table 10. Confusion matrices of the errors of the model trained with data from AB

# inputs	Predicted	Companies A&B			Company A			Company B		
21+2	3	1	0	**52**	0	0	**25**	1	0	**27**
	2	7	**34**	3	4	**19**	2	3	**15**	1
	1	**77**	1	1	**36**	1	1	**41**	0	0
10+2	3	0	0	**47**	0	0	**24**	0	0	**23**
	2	4	**34**	7	2	**19**	4	2	**15**	3
	1	**81**	1	2	**38**	1	0	43	0	2
	Actual	1	2	3	1	2	3	1	2	3

7. Discussion and comparison of the results obtained

Using ANN models it was possible to determine companies A and B follow different policies, with Company B being more defensive than Company A. PCA confirmed these results, showing that Company B replaces the oil in the interval from 30 to 59 % of deterioration, while Company A sometimes even passes the limit established by the manufacturer. Using ANN it was

Table 11. Classification of the different samples by the human expert, the Artificial Neural Network, and level of deterioration of the oil according to PCA analysis. Samples where the human expert and the ANN differ are marked in bold.

Human	ANN	Det. (%)	Human	ANN	Det. (%)	Human	ANN	Det. (%)	Human	ANN	Det. (%)
1	1	31.7	1	1	51.6	2	3	37.5	1	1	36.3
1	1	13.6	2	2	54.8	3	3	64.6	1	1	31.6
2	2	49.9	1	1	35.7	3	3	56.7	1	1	32.5
1	1	18.8	1	1	31.7	2	2	37.4	1	1	32.4
1	1	52.1	1	1	13.6	2	2	38.8	1	1	30.6
3	3	150.6	2	2	49.9	3	3	81.9	3	3	39.8
1	1	35.6	1	1	18.8	3	3	40	2	2	27.1
1	1	17.6	1	1	52.1	2	2	37.1	2	2	40
1	1	13.4	3	3	150.6	2	2	35.3	2	2	32.3
3	3	79.1	1	1	35.6	1	1	74.8	3	3	44.2
1	1	64	1	1	17.6	1	1	31.2	3	3	47.1
2	2	83.6	1	1	13.4	1	1	46.2	3	3	38.4
2	2	73.6	3	3	79.1	1	1	32.2	3	3	42
1	1	46.5	1	1	64	**1**	**2**	35.5	1	1	33.5
1	1	54.1	2	2	83.6	1	1	33.9	1	1	32
2	2	69.2	2	2	69.2	1	1	26.8	1	1	29.1
3	3	162.1	3	3	162.1	1	1	31	1	1	28.4
3	3	128	3	3	128	**1**	**2**	37.4	3	3	84.2
3	3	88.3	3	3	88.3	1	1	33.3	3	3	46.1
3	3	118.3	3	3	118.3	1	1	34.7	2	2	41
1	1	48.3	1	1	48.3	1	1	31	3	3	48.7
1	1	45.6	1	1	45.6	**3**	**2**	41.3	3	3	64.1
2	2	63.7	2	2	63.7	1	1	31.4	**3**	**2**	49.1
2	2	55.2	2	2	55.2	1	1	30.7	1	1	40.8
3	3	66.7	3	3	66.7	1	1	35.9	3	3	51.9
1	1	32.5	1	1	32.5	3	3	81.1	1	1	42.8
1	1	54.2	1	1	54.2	3	3	58.7	1	1	27.7
3	3	89.1	3	3	89.1	3	3	46.5	1	1	24.5
3	**2**	**54.6**	**3**	**2**	54.6	3	3	48.3	1	1	25
1	1	51.7	1	1	51.7	1	1	31.1	2	2	52.8
3	3	69	3	3	69	1	1	30.4	2	2	37.2
2	2	70.7	2	2	70.7	2	2	38.7	**1**	**3**	42.6
3	3	73.1	3	3	73.1	1	1	30	3	3	47.2
3	3	85.3	3	3	85.3	1	1	31.9	3	3	40.3
1	1	49.3	1	1	49.3	1	1	40.5	**3**	**2**	42.6
1	1	55.8	1	1	55.8	1	1	40.1	2	2	31.8
2	2	54.3	2	2	54.3	1	1	33	3	3	39.6
2	2	62.3	**1**	**2**	62.3	1	1	37.3	1	1	32.7
3	**2**	72.5	**1**	**2**	72.5	1	1	29.3	3	3	61.4
2	**3**	73.3	3	3	73.3	1	1	36.4	2	2	35.8
1	1	31	1	1	31	1	1	34.2	2	2	37.2
3	3	86.8	2	2	73.6	1	1	32.9	3	3	68.1
2	**1**	31.7	1	1	46.5	1	1	37.7	3	3	40.4
3	3	66.4	1	1	54.1	1	1	51	3	3	49.1

possible to create predictive models to classify the oils, with a very high degree of accuracy. The predictions of the models are sometimes different from the decisions of the companies, but some of the errors may be due to poor decisions of the companies. Using PCA it was onfirmed that in both companies it is possible to identify oils which were replaced very early and oils which were replaced at a much more advanced degree of deterioration.

PCA also showed that the variables measured during oil analysis have different importance to assess the quality of the oil and predict the company decision. ANN modeling confirmed this result, since it was possible to train a model with very high accuracy using just 10 of the 21 variables.

Table 11 compares the results obtained for dataset AB. It shows the classification of the oil by the human experts, the classification given by the artificial neural network trained with 12 variables and the percentage of oil degradation calculated using PCA. The samples where the classification of the neural model is different from the human expert are marked in bold. As the table shows, there is a large variability of results. But in general the ANN's classification and the PCA classification are coherent and arguably better than the human experts. As a reference, the average average deterioration for class 1 is 37.86 % for the human experts and 37.03 % for the ANN. For class 2 it was 52.07 % for humans and 52.44 % for the ANN. And for class 3, the average deterioration is 73.05 % for humans and 74.14 % for the ANN. The average deteriorations show that the ANN was successful clustering the most deteriorated oils in class 3 and the less deteriorated oils in class 1, more than the human experts.

Looking in more detail at the situations where the ANN and the human classification differ, it is possible to conclude that there is a high probability that some of the misclassifications are human rather than design or training limitations of the ANN. For example, in Table 11 the first error happens in a sample where PCA determines a level of oil degradation of 54.6. That sample was classified by the company as a 3 and by the neural model as a 2. In fact, the average oil degradation for oils classified as 2 is approximately 52. The third error happens for a sample with oil degradation 31.7 according to PCA, which is below the average degradation for class 1 (approximately 37). That sample was classified as 2 by the experts and 1 by the neural network. Many of the remainder errors are similar to the ones already described, which shows the neural model is very much according to the results obtained with PCA analysis.

8. Main contributions

The present paper proposes different novel contributions to the state of the art, which can be highlighted as follows.

- The results show that it is possible to create good artificial neural models to classify the oils. Moreover, the models can perform possibly with even less errors than human experts.
- Using PCA, the relevance of the variables monitored for oil analysis was determined, thus providing a better insight into the importance of each variable.
- The results also show that a good neural model does not need to use all the variables. In fact, a good model was created with just 12 input variables. This helps the process of determining the right time for oil change.

9. Conclusion

Condition monitoring of engines' oil is very important to prolong the engine life, avoid unnecessary pollution and also accidents due to engine overheating or other failures. The present paper describes experiments to create different artificial neural models that can help classify the state of deterioration of the oils with high accuracy. Because of the different policies followed by different companies, it may be difficult to create one single model that fits all policies. But it was possible to create models that showed good performance for two different companies. Those models may even generalize and learn a balance between the two policies. The results of the neural models were convergent with the results of PCA. PCA determines which companies follow the best policies for oil replacement and which variables are best predictors. The present analysis may be useful to help companies make the best decisions at the best time, or even decide which variables are more important to monitor. Future research includes fine tuning the models with more data and proposing a model to automate the process, as well as testing other classification or future extraction techniques

References

1. Capone S, Zuppa M, Presicce D S, Francioso L, Casino F, Siciliano P. Metal oxide gas sensor array for the detection of diesel fuel in engine oil. Sensors and Actuators B: Chemical 2008; 131(1): 125-133, https://doi.org/10.1016/j.snb.2007.12.029.
2. Cerny B A, Kaiser H F. A study of a measure of sampling adequacy for factor-analytic correlation matrices. Multivariate Behavioral Research 1977; 12(1): 43-47, https://doi.org/10.1207/s15327906mbr1201_3.
3. Du L, Zhe J. As high throughput inductive pulse sensor for online oil debris monitoring. Tribology International 2011; 44(2): 175-179, https://doi.org/10.1016/j.triboint.2010.10.022.
4. El-Hag A H, Saker Y A, Shurrab I Y. Online oil condition monitoring using a partial discharge signal. IEEE Transactions on Power Delivery 2010; 26(2): 1288-1289, https://doi.org/10.1109/TPWRD.2010.2073551.
5. Gajewski J, Valis D. The determination of combustion engine condition and reliability using oil analysis by MLP and RBF neural networks. Tribology International 2017; 115: 557- 572, https://doi.org/10.1016/j.triboint.2017.06.032.
6. Ghobadian B, Rahimi H, Nikbakht A, Najafi G, Yusaf T. Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network. Renewable energy 2009; 34(4): 976-982, https://doi.org/10.1016/j.renene.2008.08.008.
7. Hongxiang T, Yuntao L, Xiangjun W. Application of neural network to diesel engine SOA. Third International Conference on Measuring Technology and Mechatronics Automation 2011; IEEE, https://doi.org/10.1109/ICMTMA.2011.141.
8. Kumar S, Mukherjee P, Mishra N. Online condition monitoring of engine oil. Industrial lubrication and tribology 2005; 57(6): 260-267, https://doi.org/10.1108/00368790510622362.
9. Li X, Li J, He D, Qu Y. Gear pitting fault diagnosis using raw acoustic emission signal based on deep learning. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21 (3): 403-410, https://doi.org/10.17531/ein.2019.3.6.
10. Li Y, Wang K. Modified convolutional neural network with global average pooling for intelligent fault diagnosis of industrial gearbox. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22 (1): 63-72, https://doi.org/10.17531/ein.2020.1.8.

11. Niu X, Yang C, Wang H, Wang Y. Investigation of ann and svm based on limited samples for performance and emissions prediction of a crdi-assisted marine diesel engine. Applied Thermal Engineering 2017; 111: 1353-1364, https://doi.org/10.1016/j.applthermaleng.2016.10.042.
12. Parlak A, Islamoglu Y, Yasar H, Egrisogut A. Application of artificial neural network to predict specific fuel consumption and exhaust temperature for a diesel engine. Applied Thermal Engineering 2006; 26(8-9): 824-828, https://doi.org/10.1016/j.applthermaleng.2005.10.006.
13. Raposo H, Farinha J T, Fonseca I, Galar D. Predicting condition based on oil analysis - a case study. Tribology International 2019; 135: 65-74, https://doi.org/10.1016/j.triboint.2019.01.041.
14. Rodrigues J, Costa I, Farinha J T, Mendes M, Margalho L. Modelling diesel engine oil condition using artificial neural networks. eMaintenance 2019, Stockholm, Sweden.
15. Shaban K, El-Hag A, Matveev A. A cascade of artificial neural networks to predict transformers oil parameters. IEEE Transactions on Dielectrics and Electrical Insulation 2009; 16(2): 516-523, https://doi.org/10.1109/TDEI.2009.4815187.
16. Westerholm R, Li H. A multivariate statistical analysis of fuel-related polycyclic aromatic hydrocarbon emissions from heavy-duty diesel vehicles. Environmental Science & Technology 1994; 28(5): 965-972, https://doi.org/10.1021/es00054a032.
17. Yonghui Y, Weihua W, Xinpin Y, Hanliang X, Chengtao W. An integrated online oil analysis method for condition monitoring. Measurement Science and Technology 2003; 14(11): 1973-1977, https://doi.org/10.1088/0957-0233/14/11/015.
18. Zhu J, He D, Bechhoefer E. Survey of lubrication oil condition monitoring, diagnostics, and prognostics techniques and systems. Journal of Chemical Science and Technology 2013; 2(3): 100-115.
19. Zhu J, Yoon J M, He D, Bechhoefer E. Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines. Wind Energy 2015; 18(6): 1131-1149, https://doi.org/10.1002/we.1746.
20. Zhu X, Zhong C, Zhe J. A high sensitivity wear debris sensor using ferrite cores for online oil condition monitoring. Measurement Science and Technology 2017; 28(7), https://doi.org/10.1088/1361-6501/aa6adb.

João RODRIGUES
CISE, Univ. Beira Interior, Covilhã, 6201-001, Portugal
and
Industrial Eng. and Management, Univ. Lusófona, Campo Grande 376, 1749-024, Lisboa, Portugal

Inês COSTA
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal

J. Torres FARINHA
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
and
CEMMPRE, Coimbra University, DEM, Polo 2, 3030-290 Coimbra, Portugal

Mateus MENDES
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
and
ISR, Coimbra University, DEEC, Polo 2, 3030-290 Coimbra, Portugal

Luís MARGALHO
Polytechnic Institute of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal

E-mails: j.antunesr@hotmail.com, a21260426@isec.pt, tfarinha@isec.pt, mmendes@isr.uc.pt, lmelo@isec.pt

Appendix B

EKSPLOATACJA I NIEZAWODNOŚĆ
MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 24 (2022), Issue 1
journal homepage: http://www.ein.org.pl

Article citation info:
Antunes Rodrigues J, Torres Farinha J, Mendes M, Mateus R, Marques Cardoso A. Short and long forecast to implement predictive maintenance in a pulp industry. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2022; 24 (1): 33–41, http://doi.org/10.17531/ein.2022.1.5.

Indexed by: Web of Science Group

Short and long forecast to implement predictive maintenance in a pulp industry

João Antunes Rodrigues[a,b]*, José Torres Farinha[c,d], Mateus Mendes[c,e]*, Ricardo Mateus[b], António Marques Cardoso[a]

[a] CISE—Electromechatronic Systems Research Centre, University of Beira Interior, Calçada Fonte do Lameiro,62001-001 Covilhã, Portugal
[b] EIGeS - Research Centre in Industrial Engineering, Management and Sustainability, Lusófona University, Campo Grande, 376, 1749-024 Lisboa, Portugal
[c] Polytechnic of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
[d] University of Coimbra, CEMMPRE - Centre for Mechanical Engineering, Materials and Processes, 3030-788 Coimbra, Portugal
[e] University of Coimbra, ISR - Institute of Systems and Robotics, 3004-531 Coimbra, Portugal

Highlights

- This article presents a predictive model for a wood chip pump system.
- The Ishikawa diagram and the FMECA analysis were used to identify possible causes of system failures.
- Development of an algorithm for predicting the values of equipment sensors in the short and long term.
- The prediction made through Neural Networks had a mean absolute percentage error in all variables lower than 10%.

Abstract

Predictive maintenance is very important for effective prevention of failures in an industry. The present paper describes a case study where a wood chip pump system was analyzed, and a predictive model was proposed. An Ishikawa diagram and FMECA are used to identify possible causes for system failure. The Chip Wood has several sensors installed to monitor the working conditions and system state. The authors propose a variation of exponential smoothing technique for short time forecasting and an artificial neural network for long time forecasting. The algorithms were integrated into a dashboard for online condition monitoring, where the users are alerted when a variable is determined or predicted to get out of the expected range. Experimental results show prediction errors in general less than 10 %. The proposed technique may be of help in monitoring and maintenance of the asset, aiming at greater availability.

Keywords

predictive maintenance, condition based maintenance, time series, artificial neural networks, forecasting.

1. Introduction

As technology evolves, industrial processes are forced to adapt. That is currently the case with Industry 4.0, which may require process changes in all areas, including tracking products[2], monitoring and predicting production [36], quality control [37], or condition-based maintenance [4], among other uses of sensor networks and algorithms.

Due to this fact, there is a need for maintenance departments to reorganize, integrate new sensors, and process collected data for better performance. Machine learning can be beneficial in quality management and control, reducing maintenance costs, and improving the overall manufacturing process. That can make a key difference in modern industries.

This article presents a case study, where data analysis is performed and a predictive system is developed for a wood chip pump system, operating in an industrial paper company. This asset had frequent failures on an axis. The pump shaft and its entire fastening system had a much shorter life cycle than recommended by the manufacturer. The shaft opened cracks quickly. The analysis aimed to determine the cause of failure, as well as other potential failures.

To identify all possible causes of malfunction, Ishikawa Diagram, and Failure Mode, Effects, and Criticality Analysis (FMECA) were used. After identification of the actual cause, sensors were installed for monitoring key condition variables of the system's equipment to improve its reliability.

A global analysis of the data collected from the sensors installed in each equipment, including their minimum and maximum expected values, is presented. The variables' behaviour is studied, including graphical analysis for visualization, and forecast algorithms based on time series and Artificial Neural Networks (ANN) are applied.

A short term prediction model, with a gap of 5 days, was implemented, based on the common technique of exponential smoothing. A long term prediction model, with a gap of 3 months, was implemented, based on artificial neural networks. The short term gap of 5

(*) Corresponding author.

E-mail addresses: J. Antunes Rodrigues (ORCID: 0000-0002-8210-5468): p5942@ulusofona.pt, J. Torres Farinha (ORCID: 0000-0002-9694-8079): tfarinha@isec.pt, M. Mendes (ORCID: 0000-0003-4313-7966): mmendes@isec.pt, R. Mateus (ORCID: 0000-0003-3630-6426): p5768@ulusofona.pt, A. Marques Cardoso (ORCID: 0000-0001-8737-6999): ajmc@ubi.pt

days is adequate for the company to prepare small interventions. The long term gap allows the company to adequately prepare and schedule maintenance interventions, thereby avoiding loss of production and optimizing downtime. The duration of the gaps was decided so that a competitive advantage is achieved by reducing maintenance downtime and increasing production time.

A dashboard was developed, in which some alerts are displayed through semaphores, along with some quantitative and graphical information.

The system is designed to avoid unexpected failures and to reduce costs as much as possible, which are two of the main objectives of a good maintenance policy [22].

The present paper describes a case study where different diagnostic and prediction tools are combined to improve maintenance performance and maximize equipment availability. The fault diagnosis methodology as well as the prediction method proposed can be adapted and applied to other equipment. Fault diagnosis methods are suitable for any type of equipment, while the machine learning methods can be applied to any dataset with adaptations and proper training.

The paper is organized as follows. Section 2 presents related work and the theoretical framework. Section 3 describes the chip pump system and its diagnosis. Section 4 presents the system's condition monitoring variables. Section 5 is about the condition variables global analysis. Section 6 presents the approach about short time forecast. Section 7 proposes the approach of long- time forecast. Finally, section 8 draws some conclusions and proposes future work.

2. Background

2.1. Predictive Maintenance and Diagnosis

Predictive maintenance aims to maximize the system's availability, based on the identification of the weakest components of this physical asset [29].

According to the European Standard EN 13306:2017, a failure is the loss of the ability of an item to perform a required function after its failure, which may be complete or partial [38].

Predictive maintenance currently uses a lot of hardware to collect and store data and software to analyse it. Farinha (2018) presents an overview of the subject [9]. The purpose of predictive maintenance is to enable proactive scheduling of corrective work and thus avoiding unexpected equipment failures [33].

Maintenance optimization is a priority, due to the great trend in simulation-based optimization [28]. Currently, the best maintenance plans are tirelessly sought to minimize the overall cost of maintenance or to maximize the production and availability of assets [31]. Maintenance costs can reach 50% of production costs, which reinforces the importance of improving this area [1][26].

Predictive maintenance has evolved since visual inspection, which was its first method. Currently, with the advance of sensors and computer power, several advanced signal processing techniques are used based on pattern recognition, classification, clustering, and prediction algorithms [25].

According to FMECA reliability theory process, several types of failure mode, reasons, effects, and criticality of assets can be determined [16].

After detecting all possible failures through the Ishikawa Diagram and subsequent FMECA analysis, the main objective of predictive maintenance is to avoid the same failures by predicting them in advance.

2.2. Industry 4.0 in Industrial Maintenance

As hardware prices decrease and computing power increases, the Internet of Things (IoT) is increasingly more present in the industry [12][6]. That is a key factor to make processes predictable, simpler, controllable, and efficient, thus reducing equipment manufacturing and maintenance costs as much as possible [35].

Industry 4.0 is a result of the technological revolution, thus helping predictive maintenance [19][34]. In such a globalized and competitive market, it is necessary to make decisions about people and equipment all the time. Predictive maintenance decisions of this kind, in general, depend on massive amounts of data [7][30]. Predicting with low error the need to perform maintenance operations on the assets at a certain future point in the medium and long term is one of the main challenges in this field [14].

Due to the importance that IoT has acquired in recent years in industry and maintenance, a new concept applied specifically to the industrial sector has emerged, which is Industrial Internet of Things (IIoT).

To have an accurate forecast, it is imperative to have timely calibration and certification of industrial sensors. This is indispensable because, without the support of metrology based on measurement quality, there could be evaluation errors and discrepant data, which can result in prediction errors, poor forecasting, risks, large costs, and, consequently, loss of confidence from the market [23].

According to Hashemian, condition-based maintenance techniques for equipment and industrial processes are divided into three categories. The first category uses signals from existing process sensors, such as resistance temperature detectors and thermocouples, to help verify the performance of assets [13]. The second category depends on signals from test sensors that are installed on the equipment. The third category involves injecting a test signal into the equipment. The present work falls into the second type, as it depends on sensor signals that are installed in the equipment to measure the operational parameters.

2.3. Other Related Work

In this section some works are presented, whose aim is to predict the values of sensors installed in equipment, stressing the important of this research field for predictive maintenance using Artificial Intelligence (AI).

Kanawaday *et al.* took advantage of the machine data generated by various sensors by applying different data analysis algorithms to obtain information that help in making decisions [17]. The data captured by the sensors were always accompanied by the date and time, both of which are vital parameters for predictive modelling. The same authors used the Auto Regressive Integrated Moving Average (ARIMA) forecast in the sensor database of a longitudinal cutting machine [11][10][8].

Short-term forecasting work in maintenance has also been carried out by other authors. However, it should be noted that those studies are only focused on short-term forecasting, which shows a clear limitation in the area of long-term forecasting. An example of this type of study is the work presented below.

Kolocas *et al.* presented a predictive maintenance methodology to predict possible equipment failures of an industrial equipment in real time, using data from process sensors of operation periods. The alert period for the failure of the asset is forecasted in short-term, since a forecast gap was defined around 5-10 minutes before the incident occurred [20].

The following review section demonstrates a promising avenue of research in the use of neural networks in the area of predictive maintenance.

Tian [32] developed an Artificial Neural Network (ANN) based method designed to achieve more accurate remaining life prediction of equipment subject to condition monitoring. The proposed ANN method is validated using vibration monitoring data collected from pump bearings. The ANN model has as input to the network the age of the equipment and current condition measurement values and inspection performed. The network gives a percentage of the asset's life as an output.

Rafiee *et al.* [27] used a 2-layer perceptron neural network to detect gear and bearing failures and identify gearboxes using a new feature vector updated by the standard deviation of wavelet packet coefficients of vibration signals. Synchronization of vibration signals used cubic Hermite interpolation by parts.

Heidarbeigi *et al*. [15] developed a neural network built to predict gearbox failures. In this project a backpropagation learning algorithm and a multilayer network were used. The network has three classification outputs, which are: worn, broken teeth of gear, and faultless condition. The ideal Multilayer Perceptron Neural Network (MLP) selected for classification exhibited a 489-10-3 layer structure and had 87% accuracy. The model shown works based on vibration differences, so it can be used in other applications.

Karpenko [18] developed a neural network pattern classifier to diagnose and identify failures in an actuator of a Fisher-Rosemount 667 industrial process valve. The network is trained with experimental data obtained from the asset. The test results show that the resulting multilayer feedforward network can detect and identify various types of failure.

Wang [33] presents an artificial intelligence algorithm based on neural networks to identify failures in diesel engine lubrication pumps using vibration data. The algorithm has been tested on more than fifty lube pumps which have proven its effectiveness.

The studies mentioned above show that neural networks using monitoring data such as vibration and temperature can detect and even anticipate failures. That is useful in the diagnosis of faults with high reliability, as well as foreseeing potential failures and preventing them from happening. The research carried out also shows that there is gap in a long-term forecasts, specially predicting with 3 months advance. Nonetheless, this should be a research goal, because industries often need several weeks to prepare and carry out complex maintenance operations with minimum downtime.

3. Chip Pump System: Problem and Diagnosis

The chip pump system is depicted in Figure 1. It comprises three chip pumps, each one fed by one asynchronous motor through a mechanical connection. The inputs of the system are wood chips and liquor. The final product is a mixture of them.

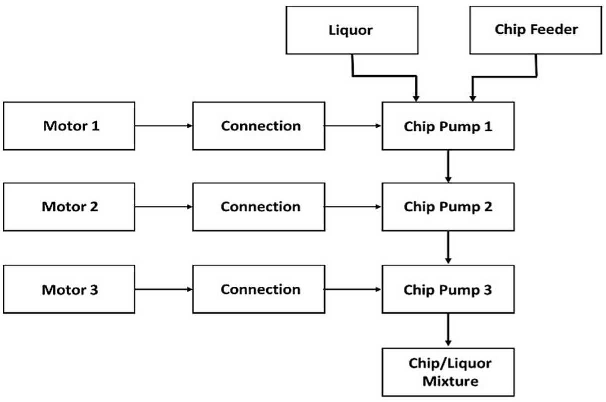

Fig. 1. Chip Pump System

The company found that the shaft of the chip pump 3 depicted in Figure 1 had shorter life services than expected. Frequent failures on that chip pump had led to cracks in the shaft, damaging its fixation cones.

Pressure is an important parameter in diagnosis, and active diagnosis is a proposal for future work to be developed after this manuscript. After several measurements, it was concluded that the pressure exerted by the mixture at the output of the chip pump increases, as shown in Figure 2.

Ishikawa diagrams allow to carry out an exhaustive diagnosis of the potential causes of equipment defects [5]. Figure 3 shows the Ishikawa diagram carried out for the fissure or breakage of the shaft and cone of the chip pump 3.

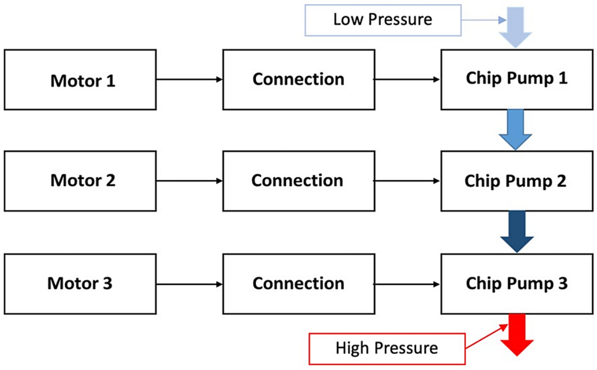

Fig. 2. Pressure increases throughout the system

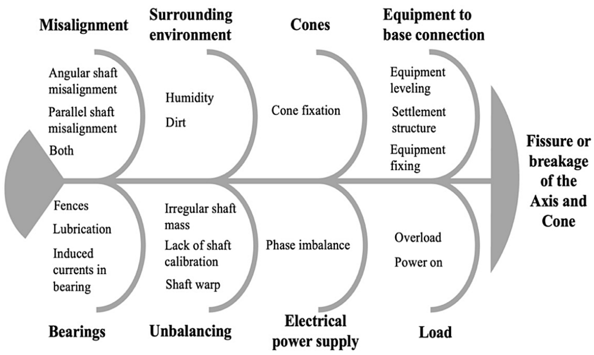

Fig. 3. Ishikawa diagram about fissure or breakage of the shaft and cone

The previous root-cause approach was complemented by a FMECA following the guidelines given by the IEC 60812:2018 [24].

FMECA allows the identification of the main possible problems in the asset. This type of analysis can be developed through a hierarchy of potential failures, complemented by a list of recommendations for avoiding them through maintenance techniques.

Through FMECA it is possible: to develop a working method; to evaluate modes of failure and their impact, to organize them; to identify the points of failure and verify the integrity of the system; to resolve failures faster; and, finally, to define criteria for tests and verifications that must be included in the preventive maintenance plan. A failure analysis can be used to understand the asset's failure mechanism. FMECA includes Failure Mode and Effect Analysis (FMEA) and the Criticality Analysis (CA) [3], [21].

The main problem was identified as the "fissure or breakage of shaft and cone", according to the FMECA matrix illustrated in Figure 4.

Based on the Ishikawa diagram and the FMECA analysis, and subsequent vibration analysis, it was possible to conclude that the actual cause of the defects was the poor seating of the chip pump machine, which was causing excessive vibration, cracking the shaft and consequently damaging the cones.

4. Chip Pump System Monitoring

Following the correction of the problem, the company decided to install a monitoring system over the key variables identified in the Ishikawa and FMECA analysis.

The system has the following sensors to monitor its condition: accelerometers; temperature sensors in roller bearings, in oil circuits, and in motor windings; load sensors; pressure sensors; flow meters; and rotation meters. Sensor readings are recorded every minute.

Figure 5 gives a global vision of the variables that are continuously monitored.

Equipament	Chip Pump				Prepared by		Team Company					
Team	Company				Date		2021					
Equipment Module	Function	Failure Mode	Failure Effect	Severity	Potential Cause of Failure	Occurrence	Preventive Action	Detection Action	Detection	RPN	Recommended Actions	Responsible and Deadline
Chip Pump (412-306)	Mechanical Traction / Mechanical Transmission	Crack or breakage of shaft and cone	Stop system for Chip Pump (412-306) and production system		**Misalignment**							
				2	Angular shaft misalignment	1	Angular alignment	Vibration Analysis	3	6	Perform alignment	
				2	Angular shaft misalignment	1	Angular alignment	Vibration Analysis	3	6	Perform alignment	x
				3	Angular and parallel shaft misalignment	2	Angular and parallel alignment	Vibration Analysis	3	18	Perform alignment	
					Imbalance							
				3	Irregular shaft mass	1	Regulate the mass	Vibration Analysis	4	12	Replacement	
				3	Lack of shaft calibration	1	Calibration	Vibration Analysis	3	9	Perform calibration	
				3	Wash Shaft	1	Replacement	Vibration Analysis	3	9	Replacement	
					Cones							
				4	Cone fixing	1	Fix the cone	Vibration Analysis	3	12	Perform predictive inspection	
					Connecting the equipment to the base							
				4	Leveling of equipment	2	Leveling the equipment	Vibration analysis and leveling check	3	24	Inspect settlement	
				4	Equipment laying structure	4	Fix the equipment	Visual displacement of the equipment	1	16	Inspect settlement	
				4	Fixing the equipment	4	Fix the equipment	Visual displacement of the equipment	1	16	Using standard screws	
					Environment							
				2	Humidity	1	Correct infiltrations	Existence of fungi	1	2	Perform isolation	
				2	Dirtiness	1	Check the cleaning of the equipment	Existence of dirt	1	2	Perform a clean-up	
					Bearings							
				2	Seals	4	Leak control or replacement	Leak checking	1	8	Perform lubrication	
				2	Lubrication	1	Lubricate bearings	Excessive friction in bearings	2	4	Perform lubrication	
				2	Induced currents in bearing	1	Improve housing insulation	Measurement of the current in the rotor	4	8	Improve housing insulation	
					Electric Power							
					Engine windings temperature		Download load	Temperature Measurement of windings			Do not exceed the recommended load	
				3	Phase imbalance	2	Balance phases	Phase measurement	1	6	Systemic phase control	
					Load							
				4	Overload	4	Respect the maximum recommended load	Analysis of Vibrations, Temperature and Electric Currents of the Motor	3	48	Follow the equipment standard	
				4	Start	4	Start at the proper speed in sequence and in conjunction with the start-up of the previous pumps	Tachometers / Voltimeters and Amperimeters	4	64	Follow the manufacturer's procedures	

Fig. 4. FMECA analysis of fissure or breakage of shaft and cone

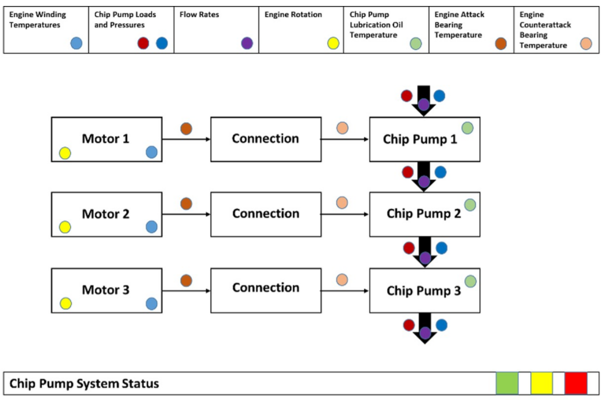

Fig. 5. Global vision of the variables that are continuously monitored

A more sophisticated algorithm to determine the relationship between predicted variables and the possibility of asset failure is out of the scope of the current project. That is a work to be developed in the future, in a separate project. The goal of the present project is just to monitor the equipment status and to predict future values. A short time prediction is performed, to anticipate future values five days in advance. A long-time prediction is performed, for three months in advance.

Relying on the forecast results, the company can anticipate malfunctions when peaks or ebbs in the predicted parameters are detected. By preventing and anticipating these failures, the company reduces its operating and maintenance costs.

5. Condition variable global analysis

The first analysis made on the condition monitoring variables was about their average and amplitude. The average, minimum and maximum values, and the time when the two latter occurred, were analyzed for all variables: vibration; temperature of attack and counterattack bearings, oil, and motor windings; load; pressure; flow; and rotation velocity.

This section presents statistics of temperature and pressure values for the three chip pumps from May 2017 to August 2019 (Tables 1-4). Pressure increases significantly throughout the system, as the mixture increases density.

Table 2 presents a comparison of engine winding temperatures from May 2017 to August 2019.

6. Short Time forecast

The short time forecast is based on an Exponential Smoothing self-adaptive, model according to Formula (1).

$$S_{t+1} = \alpha_t \times X_t + (1 - \alpha_t) S_t \tag{1}$$

Table 1. Analysis of Pressures before and after each Chip Pump

	Pressure before chip pump 1			Pressure after chip pump 1			Pressure after chip pump 2			Pressure after chip pump 3		
Year	Average value (kPa)	Max. value (kPa)	Max. value date	Average value (kPa)	Max. value (kPa)	Max. value date	Average value (kPa)	Max. value (kPa)	Max. value date	Average value (kPa)	Max. value (kPa)	Max. value date
2017	-	-	-	357.67	1031.45	2017-12-05 11:01	678.12	1492.63	2017-08-09 12:22	1007.29	1201.84	2019-06-17 16:47
2018	46.57	160.16	2018-12-12 21:51	357.19	565.89	2018-02-04 02:50	685.23	1547.28	2018-11-16 12:14	995.44	1187.94	2018-06-26 11:22
2019	48.31	162.68	2019-01-23 08:03	361.64	558.97	2019-07-18 12:32	676.26	909.47	2019-06-05 11:29	1023.69	1244.60	2019-07-22 11:54

Table 2. Chip pump lubricating oil temperature

	Chip pump 1 lubricating oil temperature			Chip pump 2 lubricating oil temperature			Chip pump 3 lubricating oil temperature		
Year	Average value (°C)	Max. value (°C)	Max. value date	Average value (°C)	Max. value (°C)	Max. value date	Average value (°C)	Max. value (°C)	Max. value date
2017	36.1	43.11	2017-11-03 21:26	37.47	51.42	2017-01-16 13:26	36.56	44.56	2017-11.03 21:26
2018	-	-	-	42.54	108.19	2018-09-28 15:44	42.71	64.87	2018-10-04 12:43
2019	54.18	61.72	2019-07-26 13:18 2019-07-06 13:19	54.07	62.5	2019-07-26 13:15	54.28	62.23	2019-03-24 12:22

Table 3. Temperature analysis of the drive pump bearing for the chip pump

	Temperature analysis of the drive pump bearing for the chip pump 1			Temperature analysis of the drive pump bearing for the chip pump 2			Temperature analysis of the drive pump bearing for the chip pump 3		
Year	Average Value (ºC)	Max. Value (ºC)	Max. Value Date	Average Value (ºC)	Max. Value (ºC)	Max. Value Date	Average Value (ºC)	Max. Value (ºC)	Max. Value Date
2017	53.58	76.87	2017-08-03 13:06	58.99	83.36	2017-10-01 16:48	68.53	95.58	2017-10-27 15:03
2018	63.75	94.62	2018-08-03 18:28	71.69	93.29	2018-08-03 18:31	72.16	105.78	2018-09-25 14:06
2019	62.02	89.37	2019-05-30 15:14	64.33	95.71	2019-05-12 17:33	68.46	105.32	2019-07-09 18:57

Table 4. Temperature analysis of the counterattack bearing to the chip pump motor

	Temperature analysis of the counterattack bearing for the chip pump 1			Temperature analysis of the counterattack bearing for the chip pump 2			Temperature analysis of the counterattack bearing for the chip pump 3		
Year	Average Value (ºC)	Max. Value (ºC)	Max. Value Date	Average Value (ºC)	Max. Value (ºC)	Max. Value Date	Average Value (ºC)	Max. Value (ºC)	Max. Value Date
2017	27.36	46.22	2017-06-20 12:12	27.71	53.90	2017-06-20 12:12	24.45	57.38	2017-06-20 14:22
2018	27.83	55.95	2018-10-03 15:02	27.93	58.86	2018-10-03 15:08	25.60	56.68	2018-10-03 15:04
2019	27.98	48.32	2019-07-11 13:49	28.78	50.77	2019-07-11 14:08	26.70	53.67	2019-07-11 14:00

where:

S_{t+1} is the expected value for time t+1

α_t is the the Auto Adaptive Smoothing Coefficient for time t ($0 \le \alpha_t \le 1$)

X_t is the variable value at time t

S_t is the expected value for time t

The Auto Adaptive Smoothing Coefficient α_t is calculated through Formula (2):

$$\alpha_{t+1} = Min(1, k_t) \tag{2}$$

where:

$$E_t = X_t - S_t \tag{3}$$

and:

$$k_t = \left| \frac{A_t}{M_t} \right| \text{, if } M_t > 0 \text{ , 0 otherwise} \tag{4}$$

$$A_t = \beta \times E_t + (1-\beta) \times A_{t-1} \text{ , } 0 \le \beta \le 1 \tag{5}$$

$$M_t = \beta \times |E_t| + (1-\beta) \times M_{t-1} \text{ , } 0 \le \beta \le 1 \tag{6}$$

E_t is the forecast error for time t. β is a parameter of the algorithm – a larger value will result in faster response of the filter.

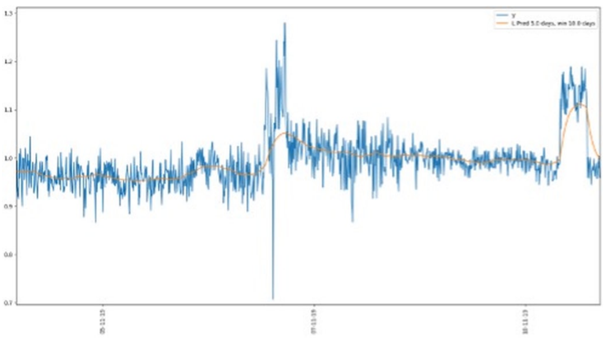

Fig. 6. Result of the short time prediction algorithm for variable Vibration

The short-term algorithm was implemented in Python. Figure 6 shows an example of the output produced by the short time prediction algorithm for vibration, with a $\beta = 0.4$. As the plot shows, the prediction follows the trends of the signal very closely. Since it is smoothed, the prediction is much more stable and immune to short spikes. For vibration, the Mean Squared Error (MSE) is 0.068 and the Mean Average Percentage Error (MAPE) is 5.61%. For pressure, the MSE is 990.64 and the MAPE is 1.36%. For the U, V, W motor winding temperatures, the MSE are 0.18, 0.21, 0.18 and the MAPE are 0.39, 0.41 and 0.36 %, respectively. For flow, MSE is 322.5, MAPE is 0.45. For the temperature of the attack roller bearing, MSE is 0.30, and MAPE is 0.59 %. For the counterattack roller bearing, the errors are 3.14 and 5.76 %. For velocity and temperature oil temperature, MSE are 254.07 and 0.15m and MAPE are 0.37% and 0.29%.

7. Long Time Forecast

To forecast the parameters, a dataset provided by the company was used. The dataset contains sensor data from 2017 to 2020, with a sampling period of 1 minute, as stated above.

The dataset was divided into two parts, 80% for training and 20% for testing. Each training iteration takes between six hours and eight hours on a computer with Intel Xeon E5-2680v2 CPU.

The code used was developed by the authors in Python, using the the ScyPy Sk-learn Library. Several mode tests were carried out and based on the results the best parameters were chosen.

It was ensured that there were no overfitting problems, as graphs were developed about the network's learning history, having presented a converging curve. The final Neural Network has two hidden layers (140-2).

7.1. Dataset, filter, smoothing and normalization

The dataset was composed of 11 variables: Vibration, Pressure, Velocity, U Winding temperature, V Winding temperature, W Winding temperature, Oil temperature, Flow, Temperature of Attack Roller Bearing, Temperature of Counterattack Bearing and Load. It should be noted that the load will not have a forecast, as it is only used as an input to the neural network.

Missing data in the dataset were filled with last known value for that variable, *i.e.* all missing or null values are replaced.

Then a median filter was applied using a sliding window with the previously defined window width (w, in samples). Finally, the data of all variables under study were normalized using the python StandardScaller library. The normalization interval used was [0, 1].

7.2. Input vector creation

To create the input vector for the neural network, a sliding window of width *wn* is applied. The following diagram illustrates the application of the window to the time series *u*.

u[n - ...]	u[n-*wn*-2]	u[n-*wn*-1]	u[n-wn]	...	u[n-1]	u[n]
			W[*wn*]	...	*W*[2]	*W*[1]

Fig. 7. A sliding window W, with size wn, is applied to the time series u, so that wn samples of the sequence u are selected to create the input to the neural network

Applying the sliding window *W* to sequence *u*, *wn* samples, from *u[n]* to *u[n-wn],* are selected to create the input vector to the neural network.

Once the *wn* samples are selected, a signature *Sn* of the window is calculated to feed as input to the neural network.

The signature *Sn* comprises the mean value of the window (m_w), the Standard Deviation (std_w), the median (med_w) of the *wn* samples, and the Power Spectrum Density (psd_w), as represented in (7). Experiments with other vectors were performed, but for succinctness the results are not presented in the paper.

$$Sn(n) = [m_w, std_w, med_w, psd_w] \tag{7}$$

Once the sequence of signatures of each window is created, a transformed dataset is constructed, with the structure represented in Figure 8.

Sn[n - ...]	Sn[n-5]	Sn[n-4]	Sn[n-3]	Sn[n-2]	Sn[n-1]	Sn[n]

Fig. 8. Representation of the transformed dataset, containing the signatures of each window wn

To train the model to predict future values, a time gap *g*, in samples, is applied to create the desired output vector. The vector is introduced, so that the predicted value *p* for time *n*+*g* is a function of *Sn[n],* as shown in (8).

$$p\,[n + g] = f(Sn[n]) \tag{8}$$

Figure 9 schematically shows the correspondence between signature *Sn* in the dataset and the predicted value *p*, where *Sn[n]* is used to predict *p[n-g]*. In the figure, *g=3*. In the experiments, *g* was the number of samples in 90 days.

wn[n - ...]	wn[n-5]	wn[n-4]	wn[n-3]	wn[n-2]	wn[n-1]	wn[n]
p[n - ...]	p[n-5]	p[n-4]	p[n-3]	p[n-2]	p[n-1]	p[n]

Fig. 9. Representation of the prediction model, where the signature of the sig - nal at time n-3 is used to predict the value at time n

The machine learning model used to make the predictions was an Artificial Neural Network, namely the MLPRegressor of the Sklearn library. The neural network after several training procedures, achieved good results. Figures 10-12 show the original signal and the prediction for different values. Those results were obtained using a multilayer neural network with two hidden layers, with 200 and 10 neurons, respectively, using the ReLU activation function. The sliding window applied on the data comprised 7 days of data.

For better stability of the values predicted, they were smoothed using median filter with window size 20.

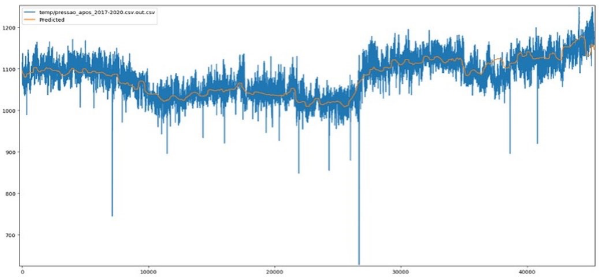

Fig. 10. Results of prediction for temperature. The signal is in blue, the prediction in orange

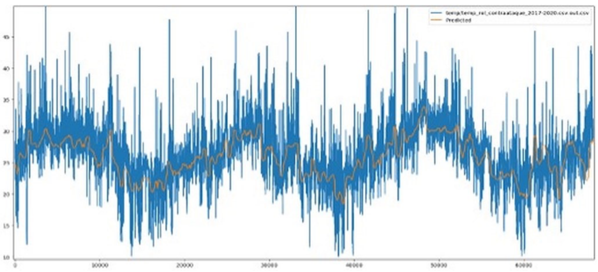

Fig. 11. Results of prediction for counterattack bearing temperature

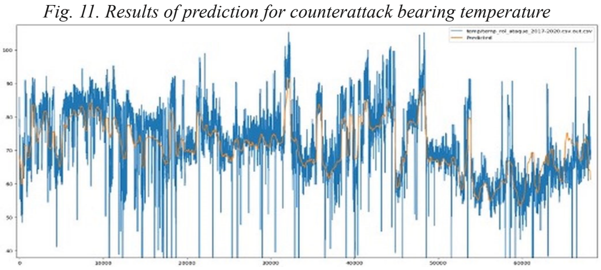

Fig. 12. Results of prediction for attack bearing temperature

To better understand the efficiency of the neural network, Table 5 shows the Mean Squared Errors for all the predicted variables.

Table 5 shows that it is possible to predict the status of the equipment in advance, with errors on average less than 10 %.

Table 5. MAPE of the 3-month forecast of all variables.

VARIABLE	MAPE
Vibration	9.47
Pressure	1.59
Velocity	1.32
Winding temperature U	4.26
Winding temperature V	4.32
Winding temperature W	4.47
Oil	5.34
Flow	3.35
Temperature Attack	6.63
Temperature Against Attack	9.72

7.3. User end interface

The end user interface was implemented through semaphores, quantitative values, and graphs, aiming to give, in an intuitive way for the user, a global vision of the system behaviour.

In this colour system, red is for the anomaly, yellow for lookout, and green for good working. This choice of colours was chosen to be like the traffic light system used on roads, making it easy to interpret and assimilate by everyone.

Through this system, it is easy, quick, and simple for the operator to know in which state of operation the equipment is, which can also contribute to prevent serious failures or malfunctions (when it is yellow or red).

The limits for green, yellow, and red were proposed by the company technicians, based on previous experience and manufacturer's information.

8. Conclusion

Failures in industrial plants can cause huge losses, or even endanger people and property. A case study of chip pumps has been described, where a dataset of approximately three years of sensory data and factory inspections were used to diagnose problems and develop a model to predict future behaviour. FMECA analysis identified that the last of three chip pumps was subjected to huge strain. Such effort was justified by the fact that it must transport its load vertically, while the predecessor chip pumps do it horizontally.

The same chip pump has deficiencies in its settlement which exponentially increase its vibration. Such vibration associated with a greater Strain effort make the shaft of the chip pump to suffer more stress than recommended, hence its useful life is doomed to be much shorter than required.

The forecast of sensor values to three months offers a great advantage for decision-making in equipment maintenance management. The temporal dimension of the forecast is totally innovative since, in the review of the state of the art, only short/medium-term forecasts were found.

Prediction made through Neural Networks proved to be valid for this type of problem. The Mean Absolute Percentage Error in all variables was below 10%.

Given the results achieved, this work offers the industry concerned the possibility of making more informed scheduled maintenance stops. This contributes very positively to increase the availability of assets as well as to reduce costs, as it reduces unexpected breakdowns. One limitation of the approach is that it relies on past sensory data. Changes in one or more key variables, for example due to differences in parts, environment, or other changes, can result in more uncertain predictions.

This methodology can be applied to other equipment by training the neural networks with appropriate data, although there is no guarantee that the same results can be achieved in another asset. The results can be better or worse, depending on the type of patterns present in the data.

This problem may be subject to future work. Other future work includes the study of more variables, as well as other machine learning models.

Funding

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowvska-Curie grant agreement 871284 project SSHARE and the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020), under Project POCI-01-0145-FEDER-029494, and by National Funds through the FCT—Portuguese Foundation for Science and Technology, under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020, and UIDP/04131/2020.

Acknowledgments

This research is sponsored by FEDER funds through the program COMPETE — Programa Operacional Factores de Competitividade—and by national funds through FCT—Fundação para a Ciência e a Tecnologia—under the project UIDB/00285/2020. This work was produced with the support of INCD funded by FCT and FEDER under the project 01/SAICT/2016 nº 022153.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ANN	Artificial Neuronal Networks
ARIMA	Auto Regressive Integrated Moving Average
FMEA	Failure Mode and Effect Analysis
FMECA	Failure Mode, Effects and Criticality Analysis
MAPE	Mean Absolute Percentage Error
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NN	Neuronal Networks
WS	Windows Size
IoT	Internet of Things
IIoT	Industrial Internet of Things

References

1. Balduíno M, Torres Farinha J, Marques Cardoso A. Production Optimization versus Asset Availability – a Review. WSEAS Transactions on Systems and Control 2020; 15: 320–332, https://doi.org/10.37394/23203.2020.15.33.
2. Barata J. A Systematic Approach to Design Product Traceability in Industry 4.0: Insights from the Ceramic Industry. [https://www.researchgate.net/publication/319998541_A_Systematic_Approach_to_Design_Product_Traceability_in_Industry_40_Insights_from_the_Ceramic_Industry]
3. Berger R, Benbow D, Ahmad Elshennawy, Walker H. The Certified Quality Engineer Handbook, Second Edition. Faculty and Staff Books 2006.
4. Cachada A, Barbosa J, Leitño P et al. Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), 2018; 1: 139–146, https://doi.org/10.1109/ETFA.2018.8502489.
5. Cahyana R. A preliminary investigation of information system using Ishikawa diagram and sectoral statistics. IOP Conference Series: Materials Science and Engineering 2018; 434: 012050, https://doi.org/10.1088/1757-899X/434/1/012050.
6. Compare M, Baraldi P, Zio E. Challenges to IoT-Enabled Predictive Maintenance for Industry 4.0. IEEE Internet of Things Journal 2020; 7(5): 4585–4597, https://doi.org/10.1109/JIOT.2019.2957029.
7. Daily J, Peterson J. Predictive Maintenance: How Big Data Analysis Can Improve Maintenance. In Richter K, Walther J (eds): Supply Chain Integration Challenges in Commercial Aerospace, Cham, Springer International Publishing: 2017: 267–278, https://doi.org/10.1007/978-3-319-46155-7_18.
8. Efthymiou K, Papakostas N, Mourtzis D, Chryssolouris G. On a Predictive Maintenance Platform for Production Systems. Procedia CIRP 2012; 3: 221–226, https://doi.org/10.1016/j.procir.2012.07.039.
9. Farinha J M T. Asset Maintenance Engineering Methodologies. 1st edition. Boca Raton, FL, CRC Press: 2018.
10. Fernandes M, Canito A, Corchado J M, Marreiros G. Fault Detection Mechanism of a Predictive Maintenance System Based on Autoregressive Integrated Moving Average Models. In Herrera F, Matsui K, Rodríguez-González S (eds): Distributed Computing and Artificial Intelligence, 16th International Conference, Cham, Springer International Publishing: 2020: 171–180, https://doi.org/10.1007/978-3-030-23887-2_20.
11. Francis F, Mohan M. ARIMA Model based Real Time Trend Analysis for Predictive Maintenance. 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), 2019: 735–739, https://doi.org/10.1109/ICECA.2019.8822191.
12. Gbadamosi A-Q, Oyedele L O, Delgado J M D et al. IoT for predictive assets monitoring and maintenance: An implementation strategy for the UK rail industry. Automation in Construction 2021; 122: 103486, https://doi.org/10.1016/j.autcon.2020.103486.
13. Hashemian H M. State-of-the-Art Predictive Maintenance Techniques. IEEE Transactions on Instrumentation and Measurement 2011; 60(1): 226–236, https://doi.org/10.1109/TIM.2010.2047662.
14. Hegedűs C, Kosztyán Z T. 46 The consideration of measurement uncertainty and maintenance related decisions. Problems of Management in the 21stcentury 2011; 1: 46-59.
15. Heidarbeigi K, Ahmadi H, Omid M, Tabatabaeefar A. Evolving an Artificial Neural Network Classifier for Condition Monitoring of Massey Ferguson Tractor Gearbox. International Journal of Applied Engineering Research 2010; 5: 2097–2107.
16. Jun L, Huibin X. Reliability Analysis of Aircraft Equipment Based on FMECA Method. Physics Procedia 2012; 25: 1816–1822, https://doi.org/10.1016/j.phpro.2012.03.316.
17. Kanawaday A, Sane A. Machine learning for predictive maintenance of industrial machines using IoT sensor data. 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2017: 87–90, https://doi.org/10.1109/ICSESS.2017.8342870.
18. Karpenko M, Sepehri N. Neural network classifiers applied to condition monitoring of a pneumatic process valve actuator. Engineering Applications of Artificial Intelligence 2002; 15(3): 273–283, https://doi.org/10.1016/S0952-1976(02)00068-4.
19. Kiangala K S, Wang Z. Initiating predictive maintenance for a conveyor motor in a bottling plant using industry 4.0 concepts. The International Journal of Advanced Manufacturing Technology 2018; 97(9): 3251–3271, https://doi.org/10.1007/s00170-018-2093-8.

20. Kolokas N, Vafeiadis T, Ioannidis D, Tzovaras D. Forecasting faults of industrial equipment using machine learning classifiers. 2018 Innovations in Intelligent Systems and Applications (INISTA), 2018: 1–6, https://doi.org/10.1109/INISTA.2018.8466309.
21. Lipol L S, Haq J. Risk analysis method: FMEA/FMECA in the organizations. International Journal of Basic & Applied Sciences IJBAS-IJENS 2011; 11(05): 9.
22. Lu B, Durocher D B, Stemper P. Predictive maintenance techniques. IEEE Industry Applications Magazine 2009; 15(6): 52–60, https://doi.org/10.1109/MIAS.2009.934444.
23. Martins A B, Torres Farinha J, Marques Cardoso A. Calibration and Certification of Industrial Sensors – a Global Review. WSEAS Transactions on Systems and Control 2020; 15: 394–416, https://doi.org/10.37394/23203.2020.15.41.
24. Organisation internationale de normalisation, Commission électrotechnique internationale, Association française de normalisation. International Standart - Failure modes and effects analysis (FMEA and FMECA) - EC 60812:2018. 2018.
25. Paolanti M, Romeo L, Felicetti A et al. Machine Learning approach for Predictive Maintenance in Industry 4.0. 2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), 2018: 1–6, https://doi.org/10.1109/MESA.2018.8449150.
26. Parida A, Kumar U. Maintenance Productivity and Performance Measurement. In Ben-Daya M, Duffuaa SO, Raouf A et al. (eds): Handbook of Maintenance Management and Engineering, London, Springer: 2009: 17–41, https://doi.org/10.1007/978-1-84882-472-0_2.
27. Rafiee J, Arvani F, Harifi A, Sadeghi M H. Intelligent condition monitoring of a gearbox using artificial neural network. Mechanical Systems and Signal Processing 2007; 21(4): 1746–1754, https://doi.org/10.1016/j.ymssp.2006.08.005.
28. Rahmati S H A, Ahmadi A, Govindan K. A novel integrated condition-based maintenance and stochastic flexible job shop scheduling problem: simulation-based optimization approach. Annals of Operations Research 2018; 269(1): 583–621, https://doi.org/10.1007/s10479-017-2594-0.
29. Ran Y, Zhou X, Lin P et al. A Survey of Predictive Maintenance: Systems, Purposes and Approaches. arXiv:1912.07383 [cs, eess] 2019.
30. Sajid S, Haleem A, Bahl S et al. Data science applications for predictive maintenance and materials science in context to Industry 4.0. Materials Today: Proceedings 2021; 45: 4898–4905, https://doi.org/10.1016/j.matpr.2021.01.357.
31. Sheu D D, Kuo J Y. A model for preventive maintenance operations and forecasting. Journal of Intelligent Manufacturing 2006; 17(4): 441–451, https://doi.org/10.1007/s10845-005-0017-6.
32. Tian Z. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. Journal of Intelligent Manufacturing 2012; 23(2): 227–237, https://doi.org/10.1007/s10845-009-0356-9.
33. Wang J, Li C, Han S et al. Predictive maintenance based on event-log analysis: A case study. IBM Journal of Research and Development 2017; 61(1): 11:121-11:132, https://doi.org/10.1147/JRD.2017.2648298.
34. Yan J, Meng Y, Lu L, Li L. Industrial Big Data in an Industry 4.0 Environment: Challenges, Schemes, and Applications for Predictive Maintenance. IEEE Access 2017; 5: 23484–23491, https://doi.org/10.1109/ACCESS.2017.2765544.
35. Yeh R H, Kao K-C, Chang W L. Preventive-maintenance policy for leased products under various maintenance costs. Expert Systems with Applications 2011; 38(4): 3558–3562, https://doi.org/10.1016/j.eswa.2010.08.144.
36. Uygun Y. Industry 4.0: Principles, Effects and Challenges – Nova Science Publishers: 2020 https://books.google.pt/books/about/Industry_4_0.html?id=pkExzgEACAAJ&redir_esc=y
37. J. Imaging | Free Full-Text | Real-Time Quality Control of Heat Sealed Bottles Using Thermal Images and Artificial Neural Network | HTML. [https://www.mdpi.com/2313-433X/7/2/24/htm].
38. EN 13306:2017 - Maintenance - Maintenance terminology. [https://standards.iteh.ai/catalog/standards/cen/5af77559-ca38-483a-9310-823e8c517ee7/en-13306-2017].

Appendix C

energies

IMPACT FACTOR 3.252

CITESCORE 5.0

Article

Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition

João Antunes Rodrigues, José Torres Farinha, Mateus Mendes, Ricardo J. G. Mateus and António J. Marques Cardoso

Special Issue
Modeling and Optimization of Electrical Systems

Edited by
Prof. Dr. José Torres Farinha

MDPI

https://doi.org/10.3390/en15176308

energies

Article

Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition

João Antunes Rodrigues [1,2,*], José Torres Farinha [3,4], Mateus Mendes [3,5,*], Ricardo J. G. Mateus [2] and António J. Marques Cardoso [1]

1 CISE—Electromechatronic Systems Research Centre, University of Beira Interior, Calçada Fonte do Lameiro, 6200-358 Covilhã, Portugal
2 EIGeS—Research Centre in Industrial Engineering, Management and Sustainability, Universidade Lusófona, Campo Grande 376, 1749-024 Lisboa, Portugal
3 Polytechnic of Coimbra—ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
4 Department of Mechanical Engineering, Centre for Mechanical Engineering, Materials and Processes, University of Coimbra, 3030-290 Coimbra, Portugal
5 Department of Electrical and Computer Engineering, Institute of Systems and Robotics, University of Coimbra, 3030-194 Coimbra, Portugal
* Correspondence: p5942@ulusofona.pt (J.A.R.); mmendes@isec.pt (M.M.)

Abstract: Forecasting has extreme importance in industry due to the numerous competitive advantages that it provides, allowing to foresee what might happen and adjust management decisions accordingly. Industries increasingly use sensors, which allow for large-scale data collection. Big datasets enable training, testing and application of complex predictive algorithms based on machine learning models. The present paper focuses on predicting values from sensors installed on a pulp paper press, using data collected over three years. The variables analyzed are electric current, pressure, temperature, torque, oil level and velocity. The results of XGBoost and artificial neural networks, with different feature vectors, are compared. They show that it is possible to predict sensor data in the long term and thus predict the asset's behaviour several days in advance.

Keywords: maintenance; neural networks; XGBoost; forecast; sensor prediction

check for updates

Citation: Rodrigues, J.A.; Farinha, J.T.; Mendes, M.; Mateus, R.J.G.; Cardoso, A.J.M. Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition. *Energies* **2022**, *15*, 6308. https://doi.org/10.3390/en15176308

Academic Editor: Valentina Colla

Received: 26 July 2022
Accepted: 25 August 2022
Published: 29 August 2022

1. Introduction

Advanced sensing technology, combined with high performance computing, help industries run with increasing reliability and competitiveness.

Industries are striving to constantly improve industrial processes and equipment. Maintenance plays a fundamental role in this field, being very important to prevent disruptions in production chains.

1.1. The Importance of Maintenance

Maintenance is a combination of technical and administrative activities required to maintain equipment, facilities, and other physical assets. The goal is to maintain those assets in the desired operational condition, or restore them so that they can fulfil their function with quality [1–3]. The main objectives of a good maintenance policy are: safety, quality, cost reduction, and availability [4]. The optimization of those four objectives at the same time is challenging, since they often conflict with each other. In those cases, it is the maintenance management's responsibility to find the best compromise solution based on the company's strategic objectives.

Predictive maintenance is one of the fastest growing types of maintenance in the industry nowadays [5]. It aims to predict the occurrence of failures before they happen, using data from sensors and state-of-the-art augmented intelligence algorithms. The algorithms are trained based on historical data, the operating condition of the assets is monitored, and the trends are predicted in near real time.

Industrial systems currently use tens, hundreds, or thousands of sensors to collect data to be used primarily to monitor processes and equipment condition [6,7].

Due to developments in data processing, along with storage algorithms and hardware, it is currently possible to store and process large quantities of data to predict the future behaviour of equipment, thus making it possible to forecast failures in advance [8].

The asset's behaviour, after being observed and analyzed, can be predicted with state-of-the-art algorithms. Such techniques have a positive impact on production reliability, security, availability and quality [9]. It should also be noted that predictive maintenance promotes environmental sustainability, as it contributes to reduce industrial downtimes, unnecessary maintenance interventions, production surpluses, and non-conforming products [10].

1.2. Industry 4.0 in Maintenance

Industry 4.0 is a consequence of scientific and technological advances, including predictive maintenance.

The amount of data extracted from industrial processes has exponentially increased due to the rise of non-invasive sensing technologies and decreasing hardware costs. However, it is essential to calibrate the sensors correctly, so that the acquired data are reliable [7,11]. Poor or incorrect data do not add value and can lead to prediction errors [12,13].

Analysis of reliable data with predictive computational techniques can avoid unnecessary equipment changes, save costs and improve safety, availability, and efficiency of processes [14].

1.3. Predictive Maintenance from an Economic Point of View

Maintenance was often seen as a source of unnecessary cost by industry, so it was often overlooked by companies. Nowadays, the role of maintenance is better understood. It is considered a key factor for the success of companies, helping them to reduce production costs and, consequently, increase profits [15].

Although applying predictive maintenance policies may involve significant costs, those costs are often less than the benefits generated from a well-planned system [16].

Most devices involve an expensive hardware network, formed by many sensors for data collection and storage. In addition to hardware, predictive maintenance requires additional costs for training staff, as well as analysing data and developing and training prediction and classification methods.

By enabling more efficient, sustainable, and higher quality production, the application of predictive maintenance also affects the company's image in the market and contributes to increase its value.

Predictive maintenance can be applied to almost all industrial equipment. However, due to its high implementation costs, technical and economic analyses must be performed before proceeding to modelling and deployment, namely determining the criticality of the equipment in case of failure or anomaly, and the potential economic losses for the company.

According to François Monchy, the more expensive the unavailability of an equipment, the more important its maintenance [17]. In other words, direct and indirect costs of equipment unavailability along with the value generated by the equipment are the most important factors to consider when choosing a maintenance policy.

The greatest advantage of predictive maintenance is that it can assess the current condition of any machine and predict when it needs maintenance before a fault happens. With a properly implemented and updated maintenance policy, it is possible to schedule equipment maintenance for times that will have the least impact in production schedule and deadlines, minimizing disruptions in production lines and improving the quality of the items produced by the factory, contributing to the profitability and sustainability of any company's business.

1.4. Artificial Neural Networks

Artificial neural networks are machine learning models with interconnected nodes distributed over several layers. The networks can be trained to recognize hidden patterns, to classify input samples into a few classes and to perform predictions. This type of model was inspired by the human brain [18,19].

The neuron is the atomic unit of a neural network. When an input vector is given, the neuron provides an output which is a function of the weighted average of the input vector's coordinates. The neurons' outputs can then be fed as inputs to other neurons in the subsequent layers.

Optimization of neural networks is a challenging problem, and it has been the topic of many works [20,21].

Feed-forward (FF) neural networks are a type of neural network in which the data flow in a single direction, from input to output, without any feedback. On the contrary, outputs in recurrent neural networks (RNN) can be fed back into the network, allowing the network to remember past events and operate in non-episodic environments.

Multi-layer perceptron (MLP) is a type of FF neural network. It comprises three types of layers: one input layer, several hidden layers, and one output layer. The main applications of MLP networks are pattern classification, recognition, and prediction [22].

As the computing power and big data increase, deep learning models are becoming more popular in several fields of science. Deep models are characterized by containing several layers, while shallow models rarely have more than three layers. For instance, deep networks are the preferred architecture in object detection or classification problems. Shallow neural networks are more adequate for prediction problems. Despite many clear distinctions between deep and shallow neural networks, some techniques developed for deep learning can help improving shallow models, and vice versa [23].

The importance of the present work is reinforced by several authors that have emphasized the necessity to change the focus from short-term (15 days) maintenance policies to long-term ones (90 days). The importance of these contributions corresponds to the increase of equipment's availability, which permits increased productivity and, at last, the success of the company [24–26].

1.5. XGboost and Random Forest

XGBoost is a scalable and highly accurate implementation of gradient boosting that pushes the limits of computing power for boosted tree algorithms, being built largely for energizing machine learning model performance and computational speed [27].

Random forest is also a popular and effective ensemble machine learning algorithm. It is widely used for classification and regression predictive modeling problems [28].

1.6. Objectives

The present research aims to propose a model to forecast sensor values of an industrial pulp paper press for 15 days, 30 days, and 90 days.

The goal was to compare the performance of multiple prediction models, including neural networks and other machine learning methods, optimizing different features and architectures.

The team defined that forecasts of most variables must have MAPE errors of less than 10%.

1.7. Contributions

Predicting in advance the values of the sensors allows us to anticipate the future state of the monitored equipment and to predict its expected operating conditions. The main contributions are:

- The approach proposed for the predictions in the present research compares and determines the best features, time windows and architectures for feed-forward shallow networks and XGBoost.

- Results are compared to LSTM and GRU.

To the best of the authors' knowledge, these are novel contributions for the area of equipment maintenance, allowing us to maximize the useful life of equipment while still minimizing risk of failure.

Similar works on industrial sensor prediction use other machine learning models, including deep networks, which require larger computational networks, which demonstrates the contribution of this study using shallow networks.

1.8. Paper Structure

The structure of the paper is as follows. Section 2 addresses the work related to this field of research. Section 3 presents the data and explains how they were treated and filtered. Section 4 shows the architecture of the underlying neural network. Section 5 presents the metrics for evaluating the neural model. Section 6 displays the tests and results of this study. Sections 7 and 8 present time series for overlapping and non-overlapping sliding windows, respectively. Section 9 shows a comparison of different feature vectors and forecast models. Finally, Section 10 presents the conclusions.

2. Related Work

2.1. Neural Networks for Prediction and Classification

This section reviews relevant works using neural networks for prediction and classification, namely in the field of predictive maintenance.

Rodrigues et al. used a neural network to predict and classify the degradation state of diesel engine oils from laboratory analysis data on 21 oil parameters, achieving an accuracy over 90% [29].

Effective maintenance is essential to keep assets at maximum availability and accident free. For these reasons, Bukhsh et al. developed a model to predict the need for railway maintenance [30].

Elhag and Wang presented an application of artificial neural networks to assess bridge risk by computing their risk scores and categories [31].

Balluff and his team developed a model to predict wind speed and pressure through recurrent neural networks [32].

Deepika and Prakash predicted the power consumption of a virtual machine with the help of backwards predictive analytics using a multi-layer perceptron, achieving a 91% accuracy [33].

Hongxiang et al. developed an algorithm using artificial neural networks (ANNs) to analyze spectroscopy data from lubricant oils. Results proved that ANNs can be used to classify distinct types of lubricants and to distinguish routine conditions of a diesel engine from operating conditions [34].

An algorithm based on a multi-layer feed-forward neural network model was developed to control a steel pickling process in several simulation cases [35].

Okoh et al. presented an approach to determine when a system needs to undergo maintenance, repair, and overhaul, before a failure occurs. One of the main innovations of this project is that forecasts were made in the long-term [36].

One of the main challenges of maintenance is to increase the availability of equipment and, hence, it is important to prognose failures before they happen. Makridis et al. presented a machine learning approach for detecting anomalies from data collected through sensors installed on vessels, predicting the condition of specific parts of the vessels' main engine [37].

In 2021, Zhagparov et al. proposed a solution to automate the prediction of grain yield based on machine learning using the XGBRegressor algorithm on the territory of the Republic of Kazakhstan. Comparisons were made with linear regression and decision tree regressor algorithms [38].

Dong et al., in 2020, developed a prediction model based on the XGBoost algorithm that considers all potential influential factors simultaneously; the objective of this model was to predict the electrical resistivity based on an experimental database [27].

In summary, according to the authors referred, among others, neural networks have high prediction accuracy and can improve support in decision making [39,40].

2.2. Condition Monitoring in Paper Press

Condition monitoring plays a central role in the maintenance of paper machines; the main objective is to maximize the availability and reduce the costs of these manufacturing units and to prevent unexpected damage or mechanical breakdowns.

The results of the tests by Suomela et al. in 2002 make it clear that thermal imaging combined with adaptive drive has great potential for monitoring paper machine components [41].

The work by Bissessur et al. features the ability to detect faults and provide early warning of impending problems based on collected vibration data and pre-processing spectra. These data processed by a neural network provide an instant decision about the state of the felt that is monitored. This method can be extended to diagnose faults in a wide range of mechanical and rotating equipment in industries [42].

Mateus et al. developed predictive models based on long-term deep memory neural networks applied to a dataset of sensor readings. The results show that it is possible to predict future behaviour up to a month in advance with reasonable confidence (errors in general inferior to 10%) using long short-term memory and gated recurrent unit deep neural networks [43,44].

3. Dataset and Pre-Processing

For the present analysis, a paper pulp company provided a three-year data set containing the time series of six variables: electric current (Sensor 1), oil level (Sensor 2), pressure (Sensor 3), rotation velocity (Sensor 4), temperature (Sensor 5), and torque (Sensor 6). All data were collected from sensors with a sampling frequency of one minute.

The dataset contains several repeated values as well as discrepant samples (outliers) that may be due to reading errors or production line stops. Upper outliers might have resulted from errors in sensor reading or recording, while lower outliers are most probably a result of those causes along with programmed or non-programmed downtimes.

In a predictive algorithm, the quality of the underlying data is of extreme importance. Poor quality data implies inaccurate results. For that reason, the dataset was previously processed to increase confidence in the results and facilitate convergence during the learning process.

The units of the several variables are as follows: electric current is measured in amperes (A); oil level is measured in percent of full tank (%Tank); pressure is measured in pascals (Pa); rotation velocity is measured in rotations per minute multiplied by 1000 (RPM $\times$ 1000); temperature is measured in degrees Celsius (°C); torque is measured in Newton-meter (N $\times$ m).

Figure 1 presents the time series collected by each sensor on the six variables.

Figure 1 shows that there are many outliers in the dataset (e.g., null, zeroes and repeated values); repeated values arise by sensory errors or even at the time change. The outliers are replaced by the average value of the variable in the sliding window before the outlier. This method has been described in more detail by Mateus et al. [45].

Therefore, the dataset was filtered using a Python algorithm developed by the authors as follows:

- Repeated values as well as lower and upper discrepant values were removed and replaced by the corresponding variable average value;
- Values beyond three standard deviations from the first and third quartiles on each variable were also replaced by the mean value of the variable in question.

Figure 2 shows the six time series of the variable values collected by the sensors after being filtered by the previously described pre-processing method. As the chart shows, there are no more sudden variations, probably representing outliers, which could impair the machine learning process. Previous studies show that pre-processing discrepant data improves the learning process [43].

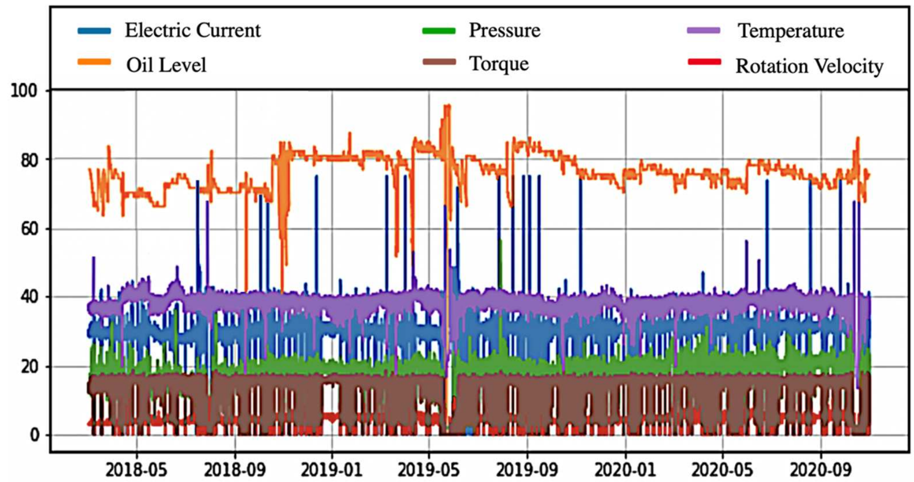

Figure 1. Plot of sensor data collected from 2018 to 2020 with a sampling period of 1 min. Raw data as provided by the company.

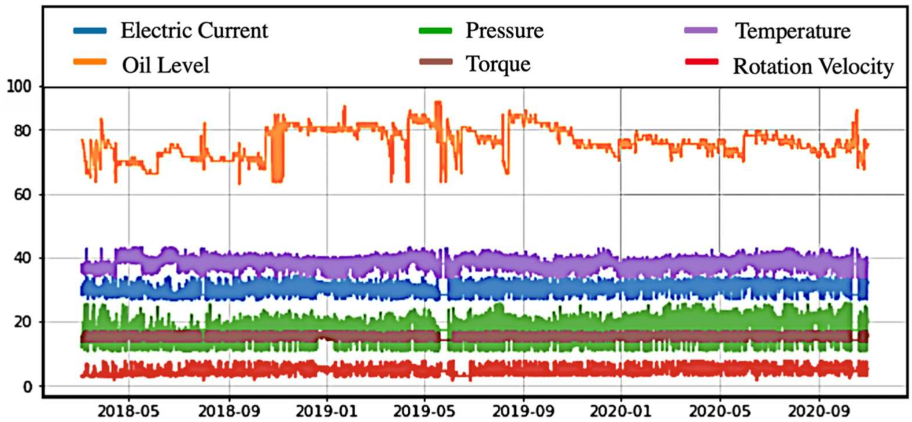

Figure 2. Data filtered from outliers with a sampling period of 1 min.

4. Artificial Neuronal Network

A sliding window encompasses a continuous subset of a time series dataset, which slides over the latter with a certain step. The window size determines the number of data point samples from the whole dataset to be included in this subset.

The window started with the first w data points (samples) of the time series and slid to the end of the series, in steps of one for an overlapping window, or steps of w samples for a non-overlapping window.

For n variables (n = 6 in this case), data from each variable i in each sliding window with size w were grouped into 15 equal-width bins j and the corresponding absolute frequency values (vector $S_{i,j}$), along with the respective average (A_i), median (M_i), standard deviation (SD$_i$),variance (V_i) statistics and, finally, 30 ratios ($R_{i1,i2}$), between each pair of variables I, where $i_1 \neq i_2$ make up the input vector I that feeds the neural network, as represented in Equation (1).

$$I = (S_{1,1}, S_{1,2}, \ldots, S_{n,15}, A_1, \ldots, A_n, SD_1, \ldots, SD_n, M_1, \ldots, M_n, V_1, \ldots, V_n, R_{1,2}, \ldots, R_{n,n-1}) \quad (1)$$

For each window w: $S_{i,j}$ represents the value of variable i in bin j; A_i is the average value of variable i; M_i is the median value of variable i; V_i is its variance; and, finally, $R_{i1,i2}$ represents the ratios between the variables collected by the sensors.

Data inputs were further standardized using the Standard Scaler library from Sklearn before being fed into the ANN model. Standardization is a technique applied in the preparation of data, with the objective of placing them in a range of common values.

Note that each variable i was predicted not only from its respective past data but also from the other five variables.

Time series data were separated into two groups: the first 80% from 1 January 2018 to 27 May 2020 were used for training the model; and the remaining 20% for carrying out the tests.

Tests were carried out with the application of various sizes w of sliding windows. Time windows w of 12, 24, 48, and 72 h were tested (720, 1440, 2880, and 4320 data point samples, respectively).

Neural Network Architecture

The architecture type chosen for the neural network is the multi-layer perceptron, one of the most popular feed-forward architectures, implemented using the Python Sklearn library named MLPRegressor.

The MLPRegressor uses multiple hyper parameters to optimize the generalization of the network model for prediction. Several architecture combinations were tested to find the best possible network configuration.

Adam solver was chosen as the algorithm for optimizing ANN weights, since it is a graph-based optimization algorithm recommended for large datasets, using a logistic sigmoid as the activation function, as represented in Equation (2), for x being the independent variable.

$$f(x) = \frac{1}{(1 + exp(-x))} \quad (2)$$

Creation vector and tests to find the best value for each alternative ANN configuration took about two days to perform, due to the complexity and size of the dataset, in a shared GPU server AMD EPYC 7552, with 16 Core + Nvidia Tesla T4/V100S.

The authors tested alternative networks with one, two, three, and four hidden layers. Using one layer only yielded quite bad results, while using four layers was quite time consuming. Results from using two and three layers were quite similar, so the authors chose two layers only as the training time was faster without loss of accuracy. Alternative ANN configurations further varied the number of neurons in each layer.

Hence, a network with two hidden layers (150 and 75 neurons, respectively) was chosen, as it showed results very similar to the three hidden layers' architecture but was much faster. Figure 3 depicts the chosen ANN architecture.

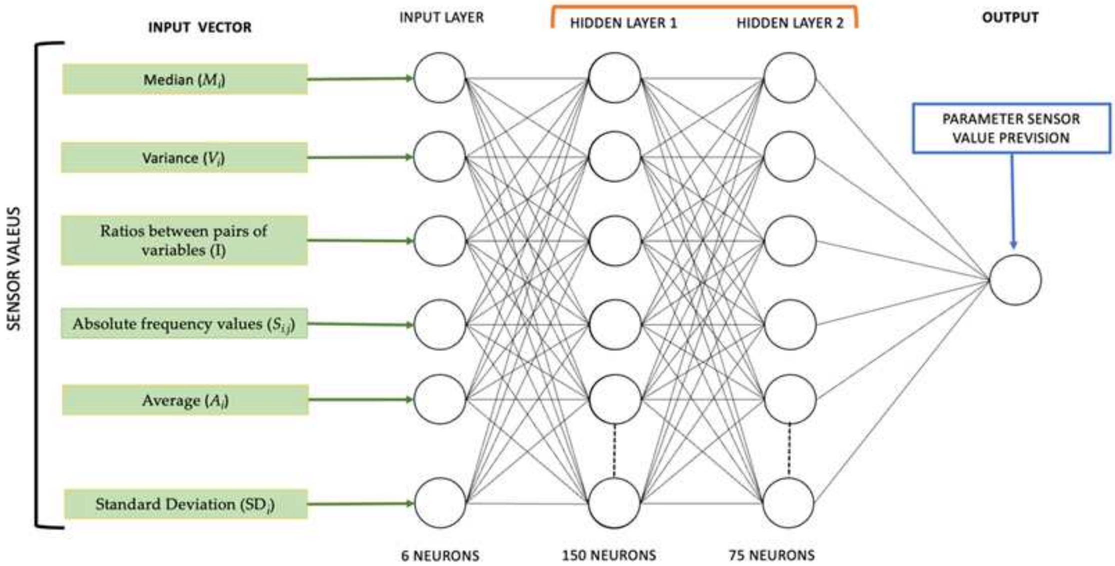

Figure 3. Architecture of the artificial neural network.

5. Model Evaluation

To assess the accuracy of the forecast model developed, three popular metrics were used: mean squared error (MSE) presented in Equation (3), and the mean absolute percent error (MAPE) presented in Equation (4) [46].

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2 \tag{3}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|Y_t - \hat{Y}_t\right|}{|Y_t|} \tag{4}$$

where Y_t represents the actual value, $\hat{Y}_t$ the predicted value, t is discrete instant time that varies between 1 and n, and n is the total number of data point samples.

6. Tests and Results for Overlapping Sliding Windows

The developed algorithm was tested for forecasts of 15, 30, and 90 days in advance. The training took up to 1000 learning epochs in each of the tests, with overlapping sliding window sizes w of 720, 1440, 2880, and 4320 samples.

Best results were achieved for windows sizes with either 12 or 24 h (720 or 1440 samples). Hence, detailed results will be presented only for these two window sizes.

Tables 1–3 show the results achieved for the six variables (sensors) in terms of MAPE, MSE, and number of iterations (ITER), which the training requires to be completed.

Evaluation results show that it is possible to predict variable (sensor) values with 3 months, 1 month, and 15 days in advance with a reasonable degree of accuracy. Most variables show MAPE errors below 10%.

In general, a window size of 720 samples (12 h) over 1440 samples (24 h), not only has a shorter learning time, but it also yields better accuracy results in terms of MAPE and MSE. Hence, a window size of 720 samples was selected as a good sampling size.

Table 1. Comparative MAPE (%) results.

		Current	Oil Level	Pressure	Temperature	Torque	Velocity
720 Samples	90 Days	3.441	6.231	16.286	4.977	2.696	4.053
	30 Days	2.295	4.642	14.643	4.006	2.332	4.112
	15 Days	2.205	4.306	12.453	3.717	1.734	3.678
1440 Samples		Current	Oil Level	Pressure	Temperature	Torque	Velocity
	90 Days	3.623	6.696	21.426	4.878	2.612	4. 451
	30 Days	2.310	5.124	14.034	4.686	2.049	4.234
	15 Days	2.541	4.319	13.633	4.363	1.864	3.896

Table 2. Comparative MSE results.

		Current	Oil Level	Pressure	Temperature	Torque	Velocity
720 Samples	90 Days	1.374	5.804	3.954	2.336	0.638	1.485
	30 Days	1.094	5. 358	4.153	2.201	0.735	1.379
	15 Days	1.080	5.013	4.056	2.060	0.670	1.454
1440 Samples		Current	Oil Level	Pressure	Temperature	Torque	Velocity
	90 Days	1.413	6.476	4.518	2.271	0.621	1.584
	30 Days	1.010	5.385	4.123	2.336	0.632	1.668
	15 Days	1.173	4.739	3.808	2.374	0.643	1.517

Table 3. Comparative number of iterations.

		Current	Oil Level	Pressure	Temperature	Torque	Velocity
720 Samples	90 Days	122	161	363	201	58	164
	30 Days	120	173	353	230	56	159
	15 Days	104	132	425	188	63	190
1440 Samples		Current	Oil Level	Pressure	Temperature	Torque	Velocity
	90 Days	136	164	415	217	64	179
	30 Days	126	144	435	249	64	213
	15 Days	105	115	432	208	71	168

7. Results with Overlapping Sliding Windows

Almost all variables show large fluctuations, including striking peaks (see Figures 1 and 2). Hence, to stabilize the output and to visualize better actual and predicted time series values on each variable, they were smoothed using a rolling average filter of 1 day.

Figures 4 and 5 present two examples of actual time series in blue and 90-day forecasts in orange, after smoothing has been applied. Pressure is the most difficult variable and torque is the easiest variable to predict, as shown in Table 1, so they were chosen as examples. According to Table 1, pressure was the variable that had the highest MAPE error and torque was the one that overall had the lowest MAPE error.

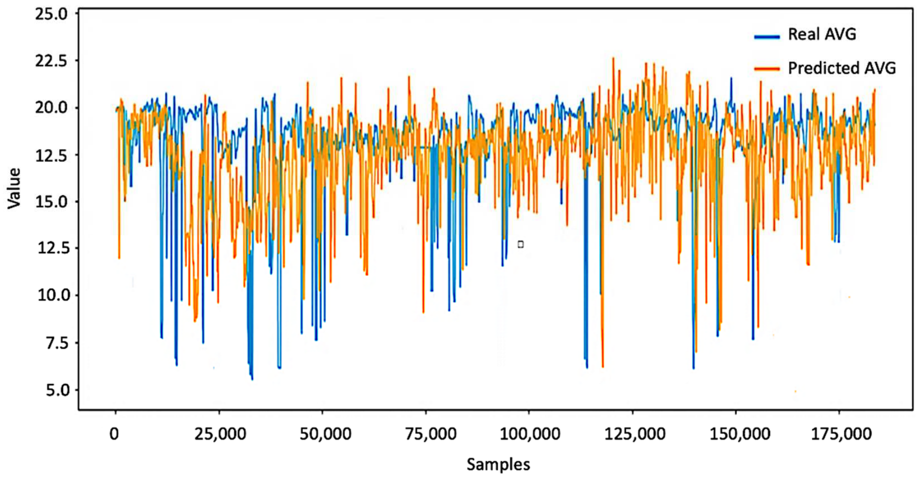

Figure 4. Real and 90-day forecast values (after smoothing) for pressure.

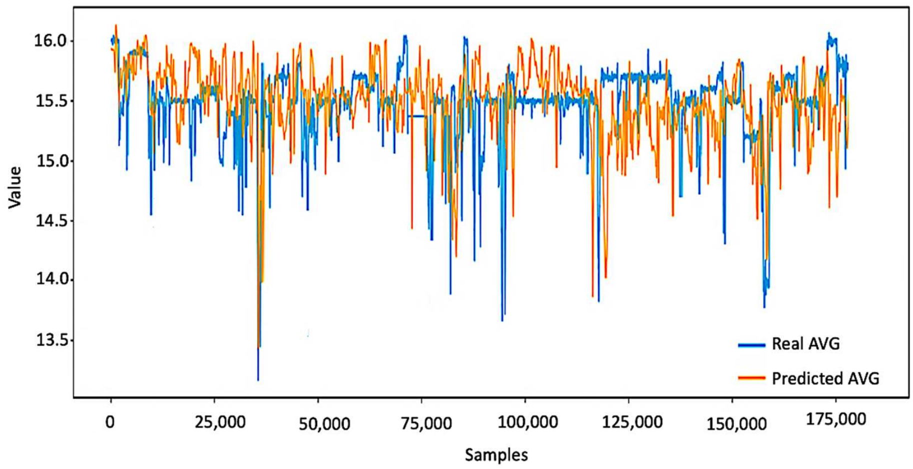

Figure 5. Real and 90-day forecast values (after smoothing) for torque.

8. Results with Non-Overlapping Sliding Windows

Using overlapping windows showed good prediction accuracies for all variables. However, their training times are quite large, taking on average more than two days for each variable (using a Cirrus workstation). Hence, non-overlapping windows were assessed to reduce learning time.

Using non-overlapping windows, the input vector in the neural network contains fewer data points, thus making its processing much faster. On average, this method allowed us to reduce the learning time to only seven minutes (using a MacBook Pro M1 from 2020 with 8 GB of RAM with MacOS Monterey).

Using non-overlapping windows yields worse long-term (90 days) forecasts than the previous overlapping window method. However, the short-term (15 days) results are good (see Figure 6). It should be noted though that the neural network is the same, regardless of whether it is for short/medium or long-term predictions. It is only the data included in the input vector that change.

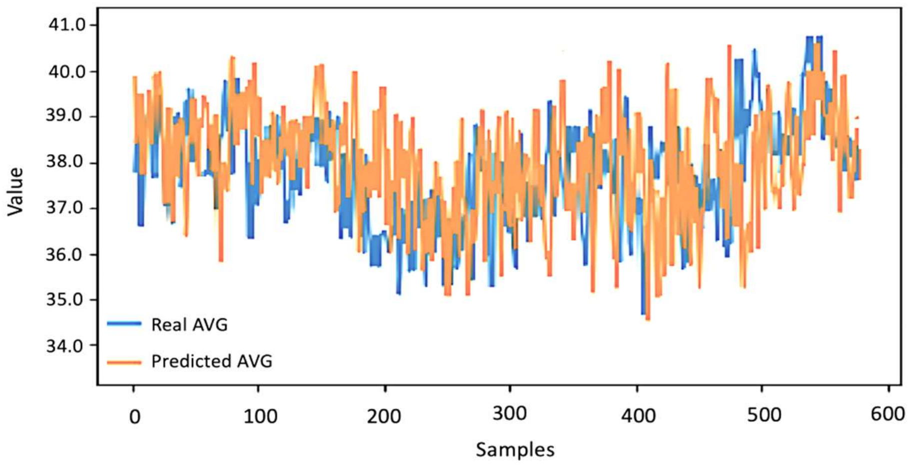

Figure 6. Real and 15-days forecast values (after smoothing) using non-overlapping windows for current.

9. Discussion

9.1. Comparation with TEPEN Vector

The present research corresponds to an optimization of the features of the neural network vector already developed by the authors. The new vector contains new ratios among variables [47]. This is the difference between the two vectors, as is presented in Table 4. This comparison is made for 90-day forecast results.

Table 4. Comparative MSE results between old vector and new vector.

		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
720 Samples	***Old vector***	1.564	5.943	4.741	3.553	0.544	1.842
	New vector	1.374	5.804	3.954	2.336	0.638	1.485
		Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
1440 Samples	***Old vector***	1.672	7.231	3.670	7.571	0.802	2.843
	New vector	1.413	6.476	4.518	2.271	0.621	1.584

Analyzing Table 4, the prediction results of the new vector are generally much better than the old vector, except for the torque parameter that maintains identical values.

9.2. Comparison between LSTM, GRU and Feed-Forward Network

Long short-term memory network is an advanced RNN, a sequential network, that allows information to persist. It can handle the vanishing gradient problem faced by RNN. Long short-term memory network (LSTM) extracts patterns from sequential data and stores these patterns in internal state variables. Each cell can retain important information for a longer period when it is used. Such information properties allow the LSTM to perform well in predicting dynamic sequences [46–48].

The gated recurrent unit (GRU) was designed by Cho et al. [49]. The closed recurrent unit is a special type of optimized recurrent neural network based on LSTM [50,51]. The difference is that the GRU combines the input port and the forgetting port in the LSTM into a single update port [52,53].

Table 5 shows a comparison of prediction models using LSTM, GRU [43,44]. A Short-Term Electric Load Forecast Method Based on Improved Sequence-to-Sequence GRU with

Adaptive Temporal Dependence. and the feed-forward model presented in this paper. The comparison is made by analysing the MAPE errors of forecast 30 days in advance for each variable. The 30-day forecast was selected because it was the time gap defined as the objective for the project.

Table 5. Comparison of MAPE results of the prediction models using LSTM, GRU and the feed-forward model presented in this paper.

	Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
GRU-ReLU	2.52	2.94	9.91	2.84	3.03	15.05
GRU-Sigmoid	**2.22**	**2.72**	**9.29**	2.74	2.88	12.42
LSTM-ReLU	2.42	2.92	10.36	**2.30**	3.72	17.19
MLP (720 Samples)	2.30	4.64	14.64	4.01	2.33	4.11
MLP (1440 Samples)	2.54	4.32	13.63	4.36	**1.86**	**3.90**

The first row and second row of the column present the results of the GRU prediction models using the ReLU and Sigmoid activation functions. The third row of the column presents the results of a traditional LSTM model using the ReLU activation function. Finally, the last two lines present the results of the MLP neural network developed, presented, and explained in the previous chapters. Table 5 presents the MAPE results, which were used to evaluate the performance of the algorithm.

Analyzing Table 5, it is concluded that the current parameter has very similar prediction results in all models. The GRU and LSTM models have similar prediction results; however, the GRU-SIGMOID has a slightly lower MAPE error. In the case of the pressure parameter, the GRU models have the best prediction results, with the GRU-SIGMOID the one that achieves the best results. In the temperature parameter, although the results do not show a significant difference, it is the LSTM-ReLU model that presents the smallest prediction error. Regarding torque, the MLP models present the best prediction results, with the best model the MLP-1440 SAMPLES. The biggest difference in results occurs in the velocity parameter, where the MLP models present much better prediction results than the other models, obtaining much lower MAPE errors.

MLP networks are simpler than GRU models and, in turn, the GRU network is simpler than LSTM. Observing the results, it is concluded that both prediction models can predict the future values of an industrial paper press, 30 days in advance, with MAPE, in general less than 10%. The difference in results in the velocity variable is noted, where only feed-forward networks, despite their simplicity, achieved a MAPE error of less than 5%. In short, there is no better overall model, because each variable has the forecast model that best suits its data, as shown in Table 5. Therefore, for optimal prediction, it is important to plan and optimize the machine learning models, with the best model that which achieves the minimum difference for each variable.

9.3. Comparison between XGBoost and Feed-Forward Network

In this section, we compare the ML feed-forward network with another alternative model, namely XGBoost.

Random forest models have, for many problems, a low error. However, for the present problem, they do not have the ability to follow the trend of the parameters to be predicted. They were tested, but the results were not acceptable, and they are not included in Table 6. The first row displays the results of the XGBoost forecast model.

The XGBoost model, in addition to having a very fast training, presents good results in the prediction of electric current, pressure and temperature, with the only disadvantage being the difficulty of identifying peaks of values. It is noted that only MLP models can follow oil level parameter trends. Figure 7 shows an example of a prediction using XGBoost. The XGBoost algorithm needs four minutes to create the vector and train the model. It should be noted that the input vector is with non-overlapping sliding windows and the machine used is a MacBook Pro M1 of 2020 with 8GB of RAM with MacOS Monterey.

Table 6. Comparison between XGBoost and feed-forward network using MSE.

	Current	*Oil Level*	*Pressure*	*Temperature*	*Torque*	*Velocity*
XGBOOST	1.075	4.953	0.618	1.304	1.576	3.356
MLP (720 Samples)	1.374	5.804	3.954	2.336	0.638	1.485
MLP (1440 Samples)	1.413	6.476	4.518	2.271	0.621	1.584

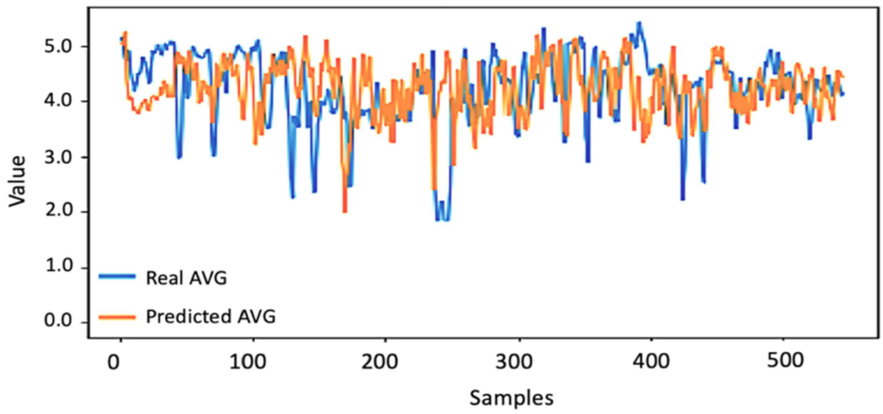

Figure 7. Real and 30-day forecast values (after smoothing) for velocity using XGBoost.

9.4. Discussion

XGBoost is a scalable and highly accurate implementation of gradient boosting that pushes the limits of computing power for boosted tree algorithms, being built largely for energizing machine learning model performance and computational speed.

One of advantages is that the proposed method can perform short-, medium-, and long-term prediction on other equipment, providing there are the necessary data and processing power available. The features used and developed to feed the machine learning models should be available in a wide range of industrial equipment. Nonetheless, for each specific situation the data pre-processing or the neuronal network architecture may need to be modified and there are no *a priori* guarantees of similar results.

This study is not intended to predict sensor failures. This study focuses on predicting the future behaviour of the machine in the short, medium, and long term. Sensor failures will be detected as a malfunction, which needs further analysis and diagnosis. That has also been clarified in the paper.

The results of those predictions can then be processed, for instance, using a classifier neural network, to classify the asset's condition into one of the following states: failure, alert, or good functioning.

10. Conclusions

Forecasting is very important to make better decisions in maintenance and other areas. Predicting the probable future behaviour of an asset brings numerous benefits. For instance, based on accurate predictions, and knowing the respective nominal operating values recommended by the asset's manufacturer, it is possible to identify anomalies in advance for the equipment in the short, medium, and long term.

The proposed algorithm makes it possible to know the behaviour of a pulp paper press in the long term, supported by the time series acquired from sensors installed on it. This way it is possible to optimize long-term programmed stops and to avoid production downtimes. This prediction model will be enhanced by the addition of a classification network that will classify the machine into one of three states: failure, alert, or good functioning.

This paper presents a valuable comparison between the input vector of the neural network using overlapping and non-overlapping sliding windows, presenting the results of the tests performed, with unequivocal conclusions about the advantages and limitations of each technique used.

The number of data points present in the neural network input vector, as well as the prediction gap, have a direct impact on the prediction accuracy. On the one hand, a larger sliding window increases the prediction errors, but a smaller window has difficulty in predicting peaks. On the other hand, the larger the prediction gap, the more difficult the prediction becomes.

The results achieved for the short term, midterm and long term were comparable to or better than the state of the art. Long-term forecasts using overlapping windows showed very good accuracies, because the predictions of most parameters present MAPE errors below 10%, that is the objective of the research presented in this paper, as shown in Section 9.2. However, they take a large processing time. Short-term forecasts using non-overlapping windows can significantly reduce this shortcoming.

The XGBoost Model presents fast and good results in the prediction of electric current, pressure, and temperature. XGBoost had errors of MSE less than two.

Future work includes applying this method to other variables and comparing it against alternative machine learning models for prediction. Additionally, other machine learning methods, such as unsupervised clustering, will be studied to classify the future condition state of the asset based on the forecasts resulting from the presented ANN.

The number of input features can also be optimized using techniques such as principal component analysis (PCA), or probabilistic principal component analysis (PPCA). Other approaches, namely hidden Markov models will also be explored.

Author Contributions: Conceptualization, J.A.R.; Formal analysis, J.A.R.; Investigation, J.A.R.; Methodology, M.M.; Project administration, J.T.F. and M.M.; Software, J.A.R.; Supervision, J.T.F. and M.M.; Validation, M.M. and R.J.G.M.; Writing–original draft, J.A.R.; Writing–review & editing, J.T.F., M.M., R.J.G.M. and A.J.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowvska-Curie grant agreement 871284 project SSHARE and the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020), under ProjectPOCI-01-0145-FEDER-029494, and by National Funds through the FCT—Portuguese Foundationfor Science and Technology, under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020, and UIDP/04131/2020. This research is sponsored by FEDER funds through the program COMPETE—Programa Operacional Factores de Competitividade—and by national funds through FCT—Fundação para a Ciência e a Tecnologia—under the project UIDB/00285/2020. This work was produced with the support of INCD funded by FCT and FEDER under the project 01/SAICT/2016 nº 022153.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
FF	Feed Forward
ITER	Iterations
MAPE	Mean Absolute Percentage Error
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
PCA	Principal Component Analysis

PPCA	Probabilistic Principal Component Analysis
RF	Random Forest.
RNN	Recurrent Neural Network
RPM	Rotations Per Minute

References

1. Kumar, U.; Galar, D.; Parida, A.; Stenström, C.; Berges, L. Maintenance Performance Metrics: A State-of-the-art Review. *J. Qual. Maint. Eng.* **2013**, *19*, 233–277. [CrossRef]
2. Standards, E. BS EN 13306:2017 Maintenance. Maintenance Terminology. Available online: http://hadidavari.com/wp-content/uploads/2018/12/BS-EN-13306-2017.pdf (accessed on 9 November 2021).
3. Rao, B.K.N. *Handbook of Condition Monitoring*; Elsevier: Amsterdam, The Netherlands, 1996. ISBN 978-1-85617-234-9.
4. Carnero, M.C. Selection of Diagnostic Techniques and Instrumentation in a Predictive Maintenance Program. A Case Study. *Decis. Support Syst.* **2005**, *38*, 539–555. [CrossRef]
5. Selcuk, S. Predictive Maintenance, Its Implementation and Latest Trends. *Proc. Inst. Mech. Eng. Part B: J. Eng. Manuf.* **2017**, *231*, 1670–1679. [CrossRef]
6. Patwardhan, A.; Verma, A.K.; Kumar, U. A Survey on Predictive Maintenance Through Big Data. In Proceedings of the Current Trends in Reliability, Availability, Maintainability and Safety; Kumar, U., Ahmadi, A., Verma, A.K., Varde, P., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 437–445.
7. Martins, A.B.; Torres Farinha, J.; Marques Cardoso, A. Calibration and Certification of Industrial Sensors—A Global Review. *Wseas Trans. Syst. Control* **2020**, *15*, 394–416. [CrossRef]
8. Bousdekis, A.; Lepenioti, K.; Apostolou, D.; Mentzas, G. A Review of Data-Driven Decision-Making Methods for Industry 4.0 Maintenance Applications. *Electronics* **2021**, *10*, 828. [CrossRef]
9. Rodrigues, J.; Torres Farinha, J.; Marques Cardoso, A. Predictive Maintenance Tools—A Global Survey. *Wseas Trans. Syst. Control* **2021**, *16*, 96–109. [CrossRef]
10. Rodrigues, J.; Farinha, J.; Mendes, M.; Mateus, R.; Cardoso, A.J.M.; Rodrigues, J. Short and Long Forecast to Implement Predictive Maintenance in a Pulp Industry. *Eksploatacja i Niezawodnosc—Maint. Reliab.* **2021**, *24*, 33–41. [CrossRef]
11. Carvalho, T.P.; Soares, F.A.A.M.N.; Vita, R.; da Francisco, R.P.; Basto, J.P.; Alcalá, S.G.S. A Systematic Literature Review of Machine Learning Methods Applied to Predictive Maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [CrossRef]
12. Galar, D.; Stenström, C.; Parida, A.; Kumar, R.; Berges, L. Human Factor in Maintenance Performance Measurement. In Proceedings of the 2011 IEEE International Conference on Industrial Engineering and Engineering Management, Singapore, 6–9 December 2011; pp. 1569–1576.
13. Sahal, R.; Ali, M.I.; Breslin, J. Big Data and Stream Processing Platforms for Industry 4.0 Requirements Mapping for a Predictive Maintenance Use Case. *J. Manuf. Syst.* **2020**, *54*, 138–151. [CrossRef]
14. Hashemian, H.M. State-of-the-Art Predictive Maintenance Techniques. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 226–236. [CrossRef]
15. Shrivastav, O.P. Industrial Maintenance: A Discipline in Its Own Right. *World Trans. Eng. Technol. Educ.* **2005**, *4*, 4.
16. Poór, P.; Basl, J.; Zenisek, D. Predictive Maintenance 4.0 as next Evolution Step in Industrial Maintenance Development. In Proceedings of the 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE), Kelaniya, Sri Lanka, 28 March 2019; pp. 245–253.
17. Monchy, F.; Mirochnikoff, Y. La Fonction Maintenance: Formation à La Gestion de La Maintenance Industrielle. *Engineering* **1987**.
18. Wang, S.-C. Artificial Neural Network. In *Interdisciplinary Computing in Java Programming*; Wang, S.-C., Ed.; The Springer International Series in Engineering and Computer Science; Springer: Boston, MA, USA, 2003; pp. 81–100. ISBN 978-1-4615-0377-4.
19. Nigrin, A. *Neural Networks for Pattern Recognition*; MIT Press: Cambridge, MA, USA, 1993. ISBN 978-0-262-14054-6.
20. Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding Deep Learning (Still) Requires Rethinking Generalization. *Commun. ACM* **2021**, *64*, 107–115. [CrossRef]
21. Oymak, S.; Soltanolkotabi, M. Toward Moderate Overparameterization: Global Convergence Guarantees for Training Shallow Neural Networks. *IEEE J. Sel. Areas Inf. Theory* **2020**, *1*, 84–105. [CrossRef]
22. Multilayer Perceptron—An Overview. ScienceDirect Topics. Available online: https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron (accessed on 9 November 2021).
23. Impact of Deep Learning-Based Dropout on Shallow Neural Networks Applied to Stream Temperature Modelling. Elsevier Enhanced Reader. Available online: https://reader.elsevier.com/reader/sd/pii/S0012825219305549?token=CF6227805A1730A859BAEE2F77EBDCB50B400FCE09411CAF51B569500652A4EE3FA108283C5770F90F0470C4BF3EE5EE&originRegion=eu-west-1&originCreation=20211109021345 (accessed on 9 November 2021).
24. Jonsson, P. Company-Wide Integration of Strategic Maintenance: An Empirical Analysis. *Int. J. Prod. Econ.* **1999**, *60–61*, 155–164. [CrossRef]
25. Carnero, M. An Evaluation System of the Setting up of Predictive Maintenance Programmes. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 945–963. [CrossRef]
26. Yamashina, H. Japanese Manufacturing Strategy Competing with the Tigers. *Bus. Strategy Rev.* **1996**, *7*, 23–36. [CrossRef]
27. Dong, W.; Huang, Y.; Lehane, B.; Ma, G. XGBoost Algorithm-Based Prediction of Concrete Electrical Resistivity for Structural Health Monitoring. *Autom. Constr.* **2020**, *114*, 103155. [CrossRef]

28. Qi, Y. Random Forest for Bioinformatics. In *Ensemble Machine Learning: Methods and Applications*; Zhang, C., Ma, Y., Eds.; Springer: Boston, MA, USA, 2012; pp. 307–323. ISBN 978-1-4419-9326-7.
29. Rodrigues, J.; Cost, I.; Farinha, J.T.; Mendes, M.; Margalho, L. Predicting Motor Oil Condition Using Artificial Neural Networks and Principal Component Analysis. *EiN* **2020**, *22*, 440–448. [CrossRef]
30. Allah Bukhsh, Z.; Saeed, A.; Stipanovic, I.; Doree, A.G. Predictive Maintenance Using Tree-Based Classification Techniques: A Case of Railway Switches. *Transp. Res. Part C Emerg. Technol.* **2019**, *101*, 35–54. [CrossRef]
31. Elhag, T.M.S.; Wang, Y.-M. Risk Assessment for Bridge Maintenance Projects: Neural Networks versus Regression Techniques. *J. Comput. Civ. Eng.* **2007**, *21*, 402–409. [CrossRef]
32. Balluff, S.; Bendfeld, J.; Krauter, S. Short Term Wind and Energy Prediction for Offshore Wind Farms Using Neural Networks. In Proceedings of the 2015 International Conference on Renewable Energy Research and Applications (ICRERA), Palermo, Italy, 22–25 November 2015; pp. 379–382.
33. Deepika, T.; Prakash, P. Power Consumption Prediction in Cloud Data Center Using Machine Learning. *Int. J. Electr. Comput. Eng.* **2020**, *10*, 1524–1532. [CrossRef]
34. Hongxiang, T.; Yuntao, L.; Xiangjun, W. Application of Neural Network to Diesel Engine SOA. In Proceedings of the 2011 Third International Conference on Measuring Technology and Mechatronics Automation, Shanghai, China, 28 January 2011; Volume 1, pp. 555–558.
35. Kittisupakorn, P.; Thitiyasook, P.; Hussain, M.A.; Daosud, W. Neural Network Based Model Predictive Control for a Steel Pickling Process. *J. Process Control* **2009**, *19*, 579–590. [CrossRef]
36. Okoh, C.; Roy, R.; Mehnen, J. Predictive Maintenance Modelling for Through-Life Engineering Services. *Procedia CIRP* **2017**, *59*, 196–201. [CrossRef]
37. Makridis, G.; Kyriazis, D.; Plitsos, S. Predictive Maintenance Leveraging Machine Learning for Time-Series Forecasting in the Maritime Industry. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020; pp. 1–8.
38. Zhagparov; Buribayev, Z.; Joldasbayev, S.; Yerkosova, A.; Zhassuzak, M. Building a System for Predicting the Yield of Grain Crops Based on Machine Learning Using the XGBRegressor Algorithm. In Proceedings of the 2021 IEEE International Conference on Smart Information Systems and Technologies (SIST), Nur-Sultan, Kazakhstan, 28–30 April 2021; pp. 1–5.
39. Ayvaz, S.; Alpay, K. Predictive Maintenance System for Production Lines in Manufacturing: A Machine Learning Approach Using IoT Data in Real-Time. *Expert Syst. Appl.* **2021**, *173*, 114598. [CrossRef]
40. Fonseca, D.J.; Navaresse, D.O.; Moynihan, G.P. Simulation Metamodeling through Artificial Neural Networks. *Eng. Appl. Artif. Intell.* **2003**, *16*, 177–183. [CrossRef]
41. Suomela, J. *Condition Monitoring of Paper Machine with Thermal Imaging*; Maldague, X.P., Rozlosnik, A.E., Eds.; SPIE: Orlando, FL, USA, 2002; pp. 143–150.
42. Bissessur, Y.; Martin, E.B.; Morris, A.J. Machine Condition Monitoring for Consistent Paper Production. *Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng.* **1999**, *213*, 141–151. [CrossRef]
43. Mateus, B.C.; Mendes, M.; Farinha, J.T.; Cardoso, A.M. Anticipating Future Behavior of an Industrial Press Using LSTM Networks. *Appl. Sci.* **2021**, *11*, 6101. [CrossRef]
44. Mateus, B.; Mendes, M.; Farinha, J.; Assis, R.; Cardoso, A.J.M. Comparing LSTM and GRU Models to Predict the Condition of a Pulp Paper Press. *Energies* **2021**, *14*, 6958. [CrossRef]
45. Mateus, B.; Farinha, J.T.; Mendes, M.; Martins, A.B.; Cardoso, A.M. Data Analysis for Predictive Maintenance Using Time Series and Deep Learning Models—A Case Study in a Pulp Paper Industry. *TEPEN* **2021**, *IncoME-VI 2021*. in press.
46. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
47. Rodrigues, J.A.; Farinha, J.T.; Cardoso, A.J.M.; Mendes, M.; Mateus, R.J.G.M. Prediction of Sensor Values in Paper Pulp Industry Using Neural Networks. *TEPEN* **2021**, *IncoME-VI 2021*. in press.
48. Lara-Benítez, P.; Carranza-García, M.; Luna-Romera, J.M.; Riquelme, J.C. Temporal Convolutional Networks Applied to Energy-Related Time Series Forecasting. *Appl. Sci.* **2020**, *10*, 2322. [CrossRef]
49. Simulating Time-Series Data for Improved Deep Neural Network Performance. IEEE Jorunal Magazine. IEEE Xplore. Available online: https://ieeexplore.ieee.org/document/8835043 (accessed on 21 July 2022).
50. Frontiers. Continuous Timescale Long-Short Term Memory Neural Network for Human Intent Understanding. Available online: https://www.frontiersin.org/articles/10.3389/fnbot.2017.00042/full (accessed on 21 July 2022).
51. Santra, A.S.; Lin, J.-L. Integrating Long Short-Term Memory and Genetic Algorithm for Short-Term Load Forecasting. *Energies* **2019**, *12*, 2040. [CrossRef]
52. Li, D.; Sun, G.; Miao, S.; Gu, Y.; Zhang, Y.; He, S. A Short-Term Electric Load Forecast Method Based on Improved Sequence-to-Sequence GRU with Adaptive Temporal Dependence. *Int. J. Electr. Power Energy Syst.* **2022**, *137*, 107627. [CrossRef]
53. Liu, X.; Lin, Z.; Feng, Z. Short-Term Offshore Wind Speed Forecast by Seasonal ARIMA—A Comparison against GRU and LSTM. *Energy* **2021**, *227*, 120492. [CrossRef]

Appendix D

energies

IMPACT FACTOR 3.252

CITESCORE 5.0

Article

Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning

João Antunes Rodrigues, Alexandre Martins, Mateus Mendes, José Torres Farinha, Ricardo J. G. Mateus and António J. Marques Cardoso

Special Issue
Modeling and Optimization of Electrical Systems

Edited by
Prof. Dr. José Torres Farinha

MDPI

https://doi.org/10.3390/en15249387

energies

Article

Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning

João Antunes Rodrigues [1,2,*], Alexandre Martins [1,2], Mateus Mendes [3,4,*], José Torres Farinha [3,5], Ricardo J. G. Mateus [2] and António J. Marques Cardoso [1]

1 CISE, Electromechatronic Systems Research Centre, University of Beira Interior, Calçada Fonte do Lameiro, 6201-001 Covilhã, Portugal
2 EIGeS—Research Centre in Industrial Engineering, Management and Sustainability, Universidade Lusófona, Campo Grande 376, 1749-024 Lisboa, Portugal
3 Polytechnic of Coimbra— ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
4 Department of Electrical and Computer Engineering, Institute of Systems and Robotics, University of Coimbra, 3030-194 Coimbra, Portugal
5 Department of Mechanical Engineering, Centre for Mechanical Engineering, Materials and Processes, University of Coimbra, 3030-290 Coimbra, Portugal
* Correspondence: p5942@ulusofona.pt (J.A.R.); mmendes@isec.pt (M.M.)

Abstract: Monitoring the condition of industrial equipment is fundamental to avoid failures and maximize uptime. The present work used supervised and unsupervised learning methods to create models for predicting the condition of an industrial machine. The main objective was to determine when the asset was either in its nominal operation or working outside this zone, thus being at risk of failure or sub-optimal operation. The results showed that it is possible to classify the machine state using artificial neural networks. K-means clustering and PCA methods showed that three states, chosen through the Elbow Method, cover almost all the variance of the data under study. Knowing the importance that the quality of the lubricants has in the functioning and classification of the state of machines, a lubricant classification algorithm was developed using Neural Networks. The lubricant classifier results were 98% accurate compared to human expert classifications. The main gap identified in the research is that the found classification works only carried out classifications of present, short-term, or mid-term failures. To close this gap, the work presented in this paper conducts a long-term classification.

Keywords: maintenance; neural networks; k-means; MLPClassifer; unsupervised learning; supervised learning

check for updates

Citation: Rodrigues, J.A.; Martins, A.; Mendes, M.; Farinha, J.T.; Mateus, R.J.G.; Cardoso, A.J.M. Automatic Risk Assessment for an Industrial Asset Using Unsupervised and Supervised Learning. *Energies* **2022**, *15*, 9387. https://doi.org/10.3390/en15249387

Academic Editor: Marcin Kaminski

Received: 31 October 2022
Accepted: 29 November 2022
Published: 12 December 2022

1. Introduction

Management of the life cycle of assets is important for the economic success of any company. Equipment availability and its condition must reach levels of excellence, namely in industries that cannot have their assets stopped or operating at sub-optimal levels.

Equipment stopped due to breakdowns increases production costs, decreases productivity, increases the stock of raw materials and increases the number of semi-finished products. All these negative consequences make predictive maintenance an essential investment.

1.1. Framework

Kumar defined maintenance as the combination of all technical and administrative activities necessary to maintain equipment, facilities, and other physical assets in the desired operational conditions, to fulfil their function with quality [1]. Many companies have selected predictive maintenance as their maintenance strategy to increase the safety, quality, and availability of their assets while also fostering environmental sustainability [2].

It is crucial to predict the future behaviour of equipment in order to make timely management decisions to reduce failure and downtime and achieve the highest level of equipment availability. A way to anticipate the behaviour of an asset is to forecast its sensor values, using time series and machine learning prediction algorithms. This can be achieved if the equipment has the necessary sensors and enough historical data are available to train the algorithms.

Sensing techniques are increasingly cheaper, more precise, and less invasive. The recent evolution of sensing technology opens new avenues for training predictive algorithms with a higher degree of reliability [3]. Such algorithms usually require large amounts of data and long periods of training and optimization, to generate forecasts with low error. Therefore, it is necessary to have reliable data history. For this to happen, it is essential to have sensors correctly calibrated, installed, and connected to a functional data system.

Data processing capacity is constantly evolving, enabling industries to improve many of their production, maintenance, and even logistics processes. The new industrial generation requires changes in its processes in virtually all areas, including monitoring and forecasting of production [4], quality control, or maintenance based on operating conditions [5].

Currently, the industrial maintenance sector is in a phase of reorganization, exploration, and research, where Artificial Intelligence (AI) is increasingly used [6]. The main advantages of using AI are the reduction in maintenance costs and the increase in the asset's availability, due to the use of intelligent machine learning algorithms to solve complex problems.

The main goal is to make the most of the assets' potential, lengthen their useful lives, and maintain their value and sustainability through the use of the proper maintenance [7]. This change will be a step toward improving the sustainability of the organizations, as the lifespan of their physical assets increases. On the other hand, the overall life and results of the organization will be improved, due to good maintenance of the equipment [8].

Clustering is applied in several areas, especially pattern recognition, data mining, and decision support [9]. One of the most well-known non-hierarchical data grouping techniques is K-means [10].

The present work describes a case study where data grouping methods are applied to determine distinct states of operation of a paper press. The method used for this task was K-means. The dataset collected for this purpose is composed of the values of the following equipment variables: Electric current; Rotation speed; Torque; Pressure; Temperature; Oil level. The unsupervised approach is to find the optimal number of operating states of the equipment, while the supervised approach is to perform classification and prediction.

The major limitation found in the research conducted is that the found classifications works only performed classifications of failures at the moment or in the short or medium term. The method presented in this paper performs a long-term classification in order to bridge this gap.

Due to the large correlation between the state of a piece of equipment and its lubrication quality, it was decided to develop a classification algorithm for the lubricants of the press. In this study, we analysed the following oil parameters: Viscosity at 100 °C, PQ Index, TAN (Total Acid Number), Al, Cr, Cu, Fe, Na, Ni, Pb, Si, and Sn.

1.2. Objectives

The present research aims to propose a model for determining probable future states of a piece of equipment, based on a two-and-a-half-year dataset of sensed values of an industrial pulp paper press. State classification is performed through a classification neural network, which has an input vector that uses threshold lines in the various sensors according to the manufacturer's recommendations to classify the asset's condition.

The equipment's states of operation must be determined using a clustering algorithm, such as k-means.

Another objective of this study is to develop a lubricant classifier algorithm based on neural networks with errors below 5% compared to the results of human experts.

1.3. Contributions

The contributions can have a positive impact on the availability of equipment and consequently reduce the company's production costs.

The importance of this classification work is reinforced by several authors who claim that the focus should be shifted from short-term to long-term maintenance policies, and this work manages to classify the state of a paper press up to 30 days in advance. It is mentioned that this same methodology can be applied to other sensing equipment.

Another important contribution is the lubricant classifier, developed using neural networks. The results of that classifier are compared to the classifications of human experts. They had a hit rate close to 98%.

1.4. Motivation and Innovation

Industries are increasingly focusing on the long term, and the maintenance department has to be fully aligned with this long-term vision.

In view of the above, it is important to study and investigate the long-term risk assessment of equipment in order to classify the long-term status of an asset. No classification of an equipment's status with an equal or superior time frame was found in the state of the art.

The classification works discovered only performed failure classifications in the present, in the short, or in the medium term, which is a limitation of the state of the art. To close this gap, the method described in this paper conducts a long-term classification.

Regarding the classification of the lubricants condition, the algorithm created has a 98% accuracy rate compared to the results of human experts, which is an increase of about 8% compared to the algorithm previously developed.

1.5. Paper Structure

The paper structure is as follows: Section 1 introduces the study; Section 2 presents a theoretical framework overview on the techniques used; Section 3 presents the state of the art; Section 4 presents the data used in this study; Section 5 explains the methodology of the work; Section 6 details how data clustering was performed; Section 7 describes neural network architectures; Section 8 presents the equipment status classification results; Section 9 shows the results of the press lubricant classification; Section 10 presents the limitations of this study; Section 11 concludes this paper.

2. Theoretical Framework

This section aims to briefly explain the main theoretical principles behind the techniques used in this study.

Machine learning methods can be classified as: 1. Supervised learning: requires a training set of inputs and outputs to learn a function that minimizes its prediction error; 2. Unsupervised Learning: tries to discover more compact representations from a dataset, without knowing their outputs; 3. Reinforcement learning: prescribes decisions based on their feedback on the objective it tries to maximize [11].

2.1. Artificial Neural Networks

Artificial neural networks (ANNs) are sophisticated adaptive systems that have the ability to modify their internal structure, which consists of a collection of interconnected nodes (neurons) between layers, in response to the input they receive.

The strength of the signal between any two nodes is represented by the weight assigned to each connection between them.

Supervised learning is achieved by adjusting the weights of these connections to minimize the error between the predicted output from the network and the target output.

ANN can have different applications, such as: forecast values [12], medical [13], business applications [14], pharmaceutical science [15], and speech recognition [16,17].

2.2. Data Grouping (Clustering)

Cluster analysis is an unsupervised learning technique used to group elements into groups (clusters), so that elements within the same group (cluster) are as similar as possible, while elements from distinct groups are as different as possible.

To define the similarity—or difference—between the elements, a distance function is used, which needs to be defined considering the context of the problem in question.

Methods of this type have applications in various fields, such as data visualization, pattern recognition, learning theory, computer graphics, identification, or classification.

2.3. K-Means

K-means is one of the most popular unsupervised learning algorithms used for data clustering [18]. This algorithm assigns each data point to one of the K groups (clusters) that minimize the square of the distance between that point and the centroid of each cluster.

The application of K-means suffers from some difficulties, such as the requirement that the number K of groups to be formed or their sensitivity to the initial conditions are provided a priori. Therefore, they must be determined experimentally during the data analysis process.

2.4. Principal Component Analysis (PCA)

Principal component analysis (PCA) is another unsupervised learning method that identifies a reduced set of the most significant components (transformed features) that explain the variance of the dataset [19].

The main objective of PCA is to condense the information contained in several original variables into a smaller set of statistical variables (principal components) with the least loss of information. The first components are those that explain most of the total variance of the original variables. By limiting the number of components, often outliers are also reduced or eliminated. The reduced set of selected variables is, therefore, easier to analyse, while still explaining much of the variance of the original data.

3. Related Work

This section presents several relevant works that use neural networks to classify equipment and lubricants, namely in predictive maintenance.

Transmission line faults are common in long-distance power transmission systems, so their classification is crucial. Mukherjee et al. proposed a method for classifying faults in the transmission lines using an approach based on PCA. This study extracted failure characteristics in terms of a Principal Component Index (PCI), followed by a threshold-based analysis of PCI values. The development of two threshold values helps to segregate the three distinct levels of fault disturbance in terms of PCI values, thus developing fault signatures for classification. According to the authors, this classification method presented a 99.78% accuracy.

The ability to group similar data is becoming more and more important, as the amount of data generated and used for analysis grows. Seal et al. proposed a non-Euclidean similarity measure, which is based on the non-linear Jeffreys divergence (JS). They then developed c-means using the proposed JS (Jc-means). The various properties of JS and Jc-means were discussed. All the analyses were carried out. The results demonstrated that Jc-means outperforms some state-of-the-art c-means algorithms [20].

Clustering is a crucial unsupervised machine learning technique used to find some underlying structure in a collection of patterns or objects. Karlekar et al. proposed the distance S, which is derived from the newly introduced divergence S, replacing the Euclidean distance of the conventional Fuzzy k-means (FKM) algorithm. With the aid of various datasets, the performance of the proposed FKM was compared to that of the traditional FKM with Euclidean distance and its variations. The comparative study demonstrated that the outcomes are solid. Additionally, the results showed that the modified FKM outperforms some cutting-edge FKM algorithms [21].

The harsh operating conditions are a common cause of some failures in industrial equipment. Analysing a large amount of data is how faults in mechanical gearbox systems are found and diagnosed. Sharma et al. created a new simple and effective peak density clustering algorithm based on an adaptive mixing distance for handling mixed data, as real-world datasets that encompass numerical and categorical attributes, in order to acquire more distinguishable fault characteristics under various conditions [22].

Due to the search for a more sustainable world, wind energy emerges as one of the most important sources of energy production. Zang et al., proposed a fault detection method for main bearing wind turbines based on SCADA data using an RNA artificial neural network. This algorithm makes it possible to identify the initial stage of main bearing failures, allowing for early intervention [23].

Ertun et al., created an algorithm using a neuro-fuzzy ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults [24].

Rodrigues et al., used a neural network to predict and classify the degradation state of diesel engine oils from laboratory analysis data on 21 oils' parameters, achieving an accuracy over 90% [25].

Lubricating oil plays an important role in vehicle maintenance, and good lubrication can extend engine life as well as reduce maintenance costs. Le et al., through machine learning models, classified the condition of military vehicle engine lubricating oils. Oil condition was classified into three categories: normal, degraded, and unsuitable [26].

Kittisupakorn et al., proposed to control a steel pickling process, using an algorithm based on a multilayer feed-forward neural network model [27].

Gajewski et al., presented a study focusing on transport system engines. The types of oils were obtained from heavy track engines. They used the data with neural networks to identify the patterns that model the deterioration of the system [28].

Using the data, manufacturers can produce goods of higher quality while spending less money by using predictive models for quality control. Zhang et al., proposed a two-stage method for doing this, first clustering the data into clusters based on the manufacturing process and then using supervised learning to predict the failed product in each cluster. Their goal was to predict manufacturing failures using the anonymous features. The final model was decided, based on the Random Forest algorithm's performance [29].

Mazumder et al. used machine learning to develop a viable alternative to computationally intensive analytical approaches to assessing the failure risk of oil and gas pipelines. The conclusion is that XGBoost is the optimal algorithm to predict failure and is recommended for future analysis [30].

The works cited above show that neural networks can improve support in decision making in the maintenance and condition monitoring field.

4. Data Processing

4.1. Sensor Data Collection

An industrial press's three-year history of six variables, including electric current, rotation speed, temperature, pressure, oil level, and torque, was provided for this study by a pulp company. The sampling frequency used by all sensors was one minute [31].

The recorded data must be of the highest calibre. Poor data can lead to models that are incorrect and have biased results. To increase confidence in the findings, the data were previously processed and analysed.

The different variables' measurement units are as follows: Oil Level is measured in percentage of full tank (% Tank); Electric Current is measured in Amperes (A); Pressure is measured in Pascal (Pa); Temperature is measured in Degrees Celsius (°C); Torque is measured in Newton-meter (N × m); Rotation Velocity is measured in Rotations Per Minute multiplied by 1000 (RPM × 1000). Figure 1 presents a boxplot of the original data according to the six variables.

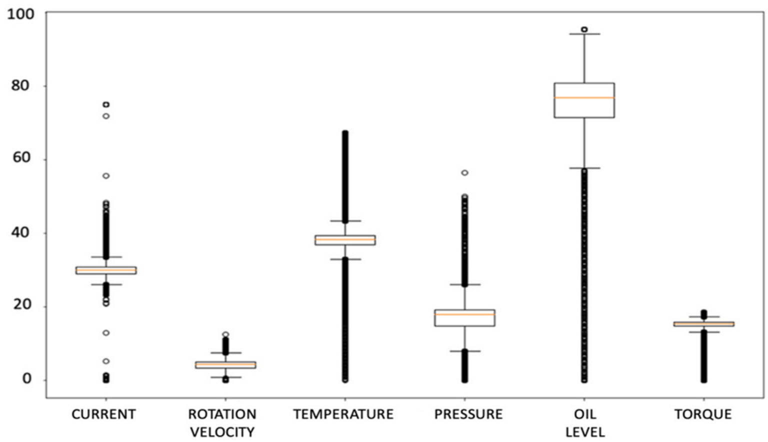

Figure 1. Boxplot of the original data.

Figure 1 allows us to conclude that the collected data contained some discrepant points, including outliers such as null and repeated values, due to stoppages of the equipment under analysis. Outliers with higher values were probably the result of errors in the sensor reading or recording. Outliers with lower values probably resulted, in addition to the causes mentioned above, from scheduled and unscheduled stops of the equipment.

Outliers were replaced by the average value of the variable in the sliding window before the outlier. This method has been described in more detail by Mateus et al. [32].

Figure 2 shows the boxplot of the dataset after being filtered and treated as described above. As the figure shows, most of the discrepant samples were removed and there was more compactness of the data.

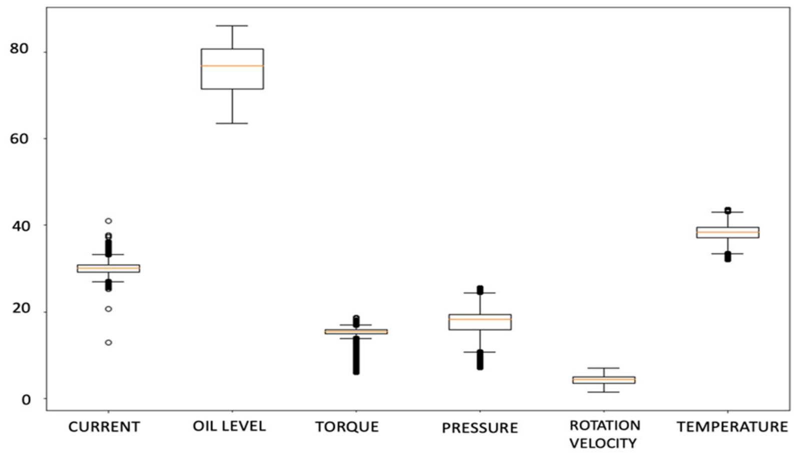

Figure 2. Boxplot of processed data.

4.2. Data Enrichment

4.2.1. Equipment Nominal Operation Zones

According to the manufacturer's user manual, the equipment is recommended to work between a predefined range for all six variables. For instance, Figure 3 depicts yellow lines representing the lower and upper temperature thresholds of the equipment when working in its normal functioning zone. Likewise, when the equipment works beyond the red lines, it generates a red alert indicating that is working in its failure zone and hence needs urgent attention. When the equipment is working between the yellow and red lines, it generates a yellow alert informing that the equipment needs to be checked for possible overload or anomaly. The dataset was enriched with this classification, namely indicating whether each variable value was within the range of normal, alert, or high risk of failure operation values provided by the equipment's manufacturer.

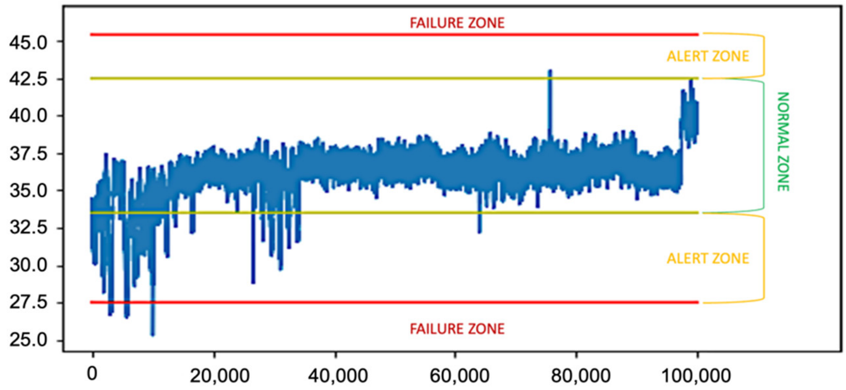

Figure 3. Normal, alert, and failures zones for the variable Temperature.

4.2.2. Sensor Values Predicted at 30 Days

The aim of this study is to classify the state of the equipment 30 days in advance. The dataset was enriched with 30-day forecasts of each variable from a previously developed and validated neural network that can predict sensor variable values at 30 days with a MAPE error of less than 10% [33,34]. Figure 4 illustrates the respective predicted time series for the six variables (sensors).

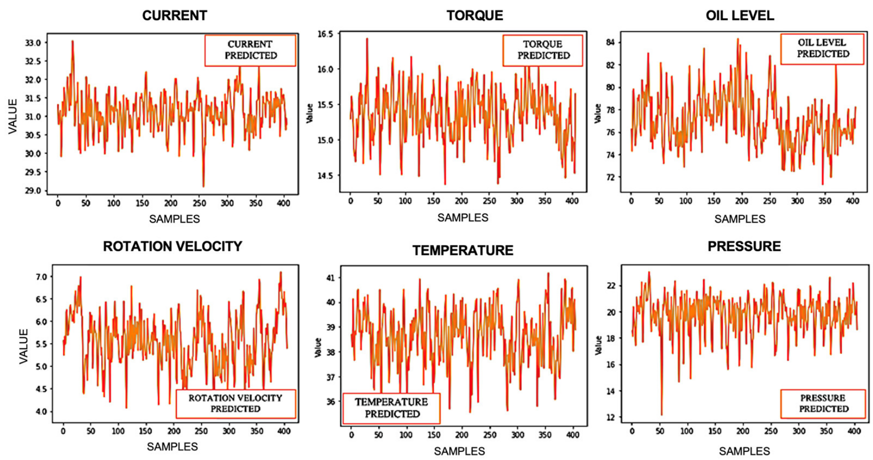

Figure 4. Predicted time series of sensor values.

4.3. Lubricating Oil Database

Oil analysis is an extremely important tool in predictive maintenance. Through this, it is possible to evaluate the conditions of the fluids and of the equipment.

The dataset of the lubricating oils in question was supplied by a company and contains all the lab results of oil analyses carried out on the press, as well as the classification of all the analyses by a human expert.

To assess the reliability of assets, increase their availability and clarify the condition of equipment, it is extremely important to know the condition of its lubricants.

It should be noted that through oil analysis, it is also possible to identify problems early, before they turn into serious failures. Due to the aforementioned reasons, the authors decided to develop an algorithm for the classification of the condition of press lubricating oil.

The present database contains 179 oil samples, and each oil sample contains the analysis of 12 lubricant parameters. The parameters analysed are Viscosity at 100 °C PQ Index, TAN (Total Acid Number), Al, Cr, Cu, Fe, Na, Ni, Pb, Si, and Sn. Table 1 presents a summary of statistical parameters of the variables used in this study.

Table 1. Metrics of Oil Analysis.

	Units	Mean	Min	Max	Var	Std
TAN (Total Acid Number)	mgKOH/g	1.26	0.18	2.85	0.26	0.52
PQIndex	ppm	131.78	0.00	6732.00	396,718.55	631.63
Al Content	ppm	1.30	0.00	15.00	8.00	2.84
Cr Content	ppm	5.59	0.00	2.00	34.02	5.85
Cu Content	ppm	9.16	0.00	243.00	815.87	28.65
Fe Content	ppm	260.17	2.00	1231.00	91,004.30	302.55
Na Content	ppm	5.21	0.00	38.00	25.82	5.10
Ni Content	ppm	4.20	0.00	26.00	17.16	4.16
Pb Content	ppm	0.51	0.00	30.00	6.25	2.51
Si Content	ppm	2.39	0.00	22.00	8.10	2.85
Sn Content	ppm	1.07	0.00	8.00	2.62	1.62
Viscosity at 100 °C	m^2/s	3035.76	954.40	4146.20	168,647.64	436.90

Analysing Table 1, key variables such as Index PQ, TAN, Fe, and Viscosity at 100 °C have a large volatility. This volatility is a consequence of the differences in the state of the oil in the various analyses. Figure 5 graphically shows the variability in those parameters. The x-axis represents the oil analysis samples, and the y-axis represents the results of each parameter normalized between zero and one.

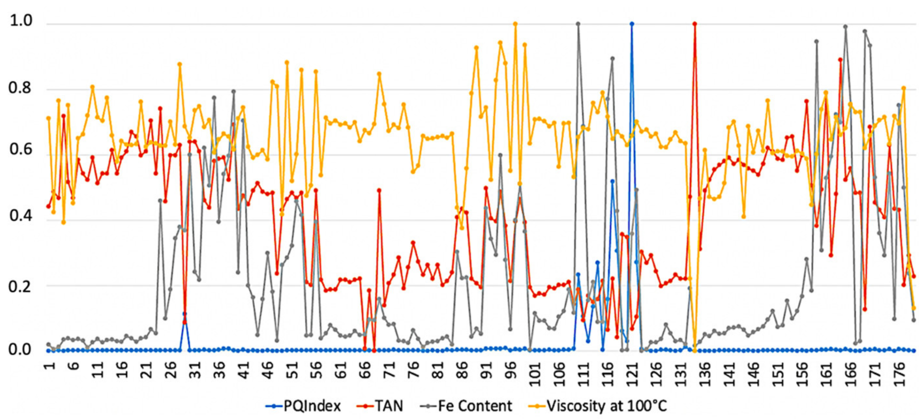

Figure 5. Graph of all oil analyses (Index PQ, TAN, Fe, and Viscosity at 100 °C).

5. Data Processing

This section describes the methodology used to carry out this study.

To elucidate the readers, it was decided to make two flowcharts, one containing the methodology used to classify the state of the press (Figure 6) and the other one presenting the methodology used to classify the lubricant in the machine (Figure 7).

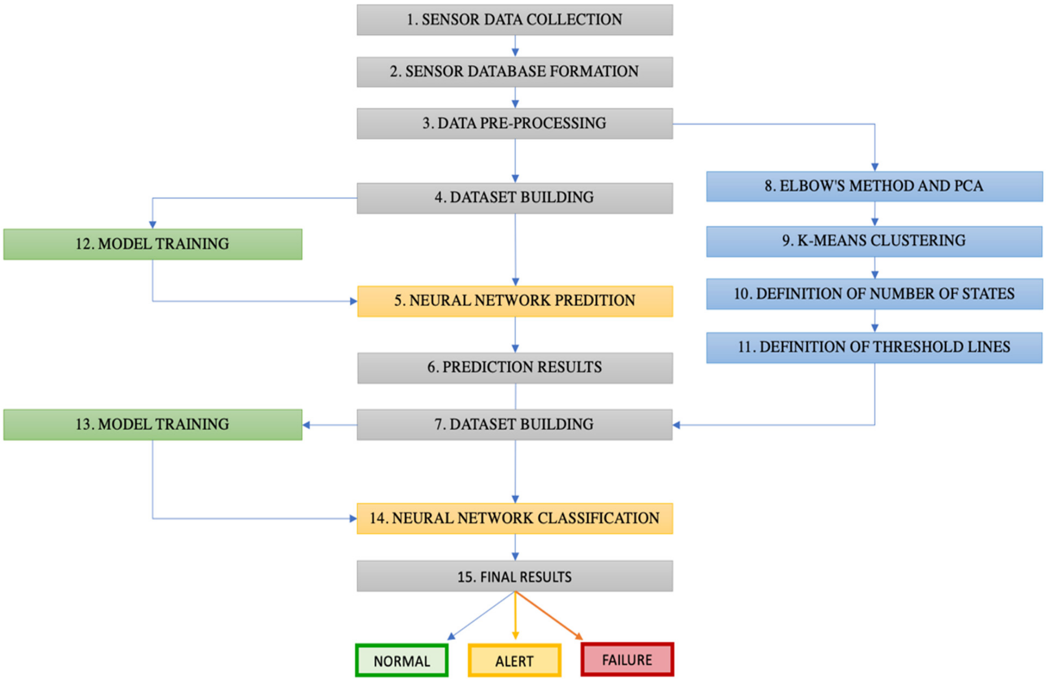

Figure 6. Methodology used to classify press condition.

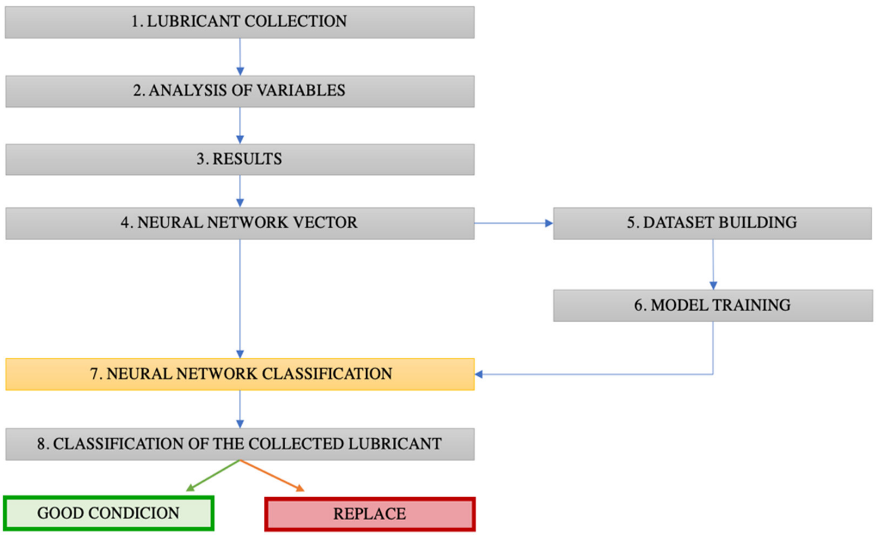

Figure 7. Methodology used to classify the lubricating oil of the press.

6. Clustering (Operating States)

The methodology applied for grouping the dataset into clusters, representing the operating states of the equipment, is as follows:

a. Application of the K-means method described for different K values between 1 and 10;
b. Specifying the number of clusters K;
c. Initialization of the centroids of each cluster, randomly selecting, without repetition, a data point for each of the centroids of the K clusters;
d. Calculation of the square of the distance between each of the remaining data points and each of the K centroids;

e. Assignment of each of these data points to the cluster whose centroid is closest;
f. Calculation of the new position of the centroids of each cluster according to the average position of all data points belonging to each cluster;
g. Repetition of the last three steps, until the position of the centroids no longer changes;
h. Determining the optimal number of clusters based on the elbow method [35];
i. Use of PCA to reduce the number of variables to two and, thus, be able to view the data points classified according to the cluster to which they belong.

After applying the k-means method for K values between 1 and 10, the ideal number of clusters was determined using the Elbow method. For this purpose, the relationship between the number K (1,10) of clusters and the sum of the squares of the distances between each point and the centroid of the cluster where it was grouped was graphically represented, and then the ideal number of clusters is where the value of the dependent variable begins to stabilize, which visually resembles the shape of an elbow, hence the name of the method. From this value of K, the function starts to move almost parallel to the abscissa axis. The K value corresponding to this point where the error starts to stabilize is the optimal K value, that is, it represents the ideal number of clusters (Figure 8).

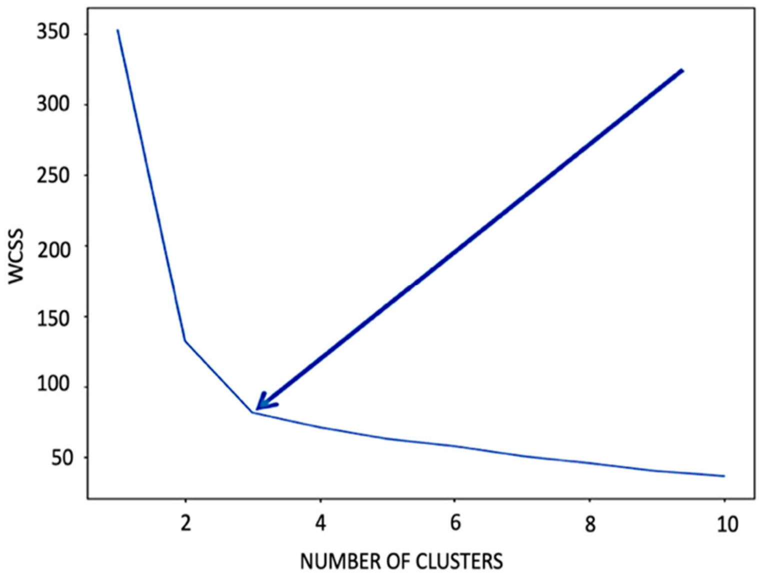

Figure 8. Application of the elbow method to determine the optimal number of clusters.

From the analysis of the graph presented in Figure 8, it was concluded that the ideal number of clusters was K = 3. The next step was to convert the multidimensional dataset (6 variables) into 2 dimensions (variables), just to be able to visualize the distribution of the dataset more easily by the three defined clusters. To this end, principal components analysis was applied to the initial dataset to represent the data in the two most representative principal components (PC1 and PC2) in terms of explained variance (Figure 9).

Figure 10 presents the dataset classified according to the three defined clusters, as well as their centroids. The *x*-axis represents Principal Component 1 and the *y*-axis represents Principal Component 2. The respective centroids are marked with a cross (x). The first two Principal Components aggregate about 90% of the total variance of the data. Due to this, it can be said that there is a minimal loss of information when using the two main components to form the 3 clusters as indicated by the Elbow Method.

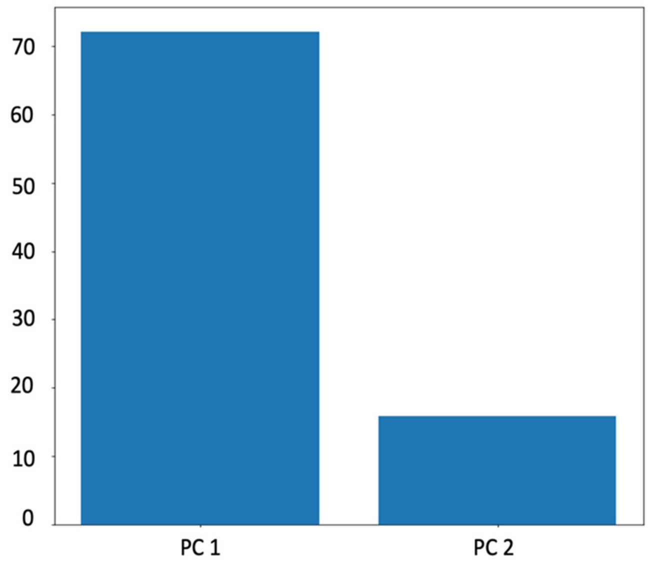

Figure 9. Variance in CP1 and CP2.

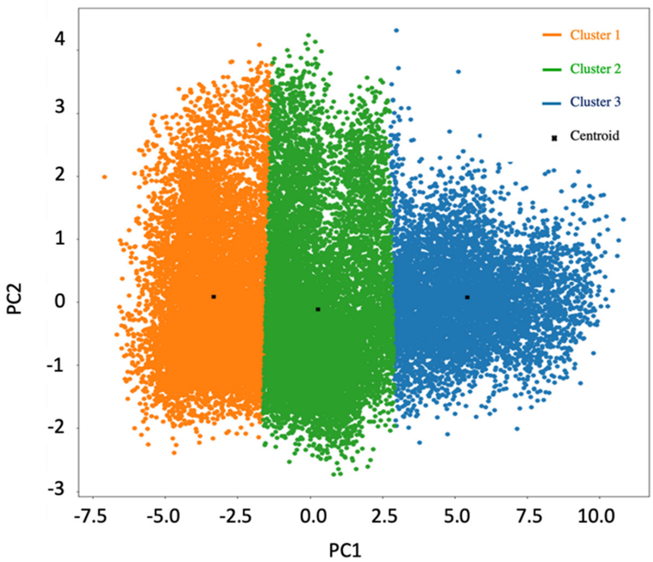

Figure 10. Representation of the data in the three clusters according to CP1 (*x*-axis) and CP2 (*y*-axis).

Finally, the data were represented in Figure 11 according to several scatter plots that allow a visual analysis of the degree of association between the variables under analysis.

The last row and column in Figure 11 show the cluster rating of each data point. Analysing the various dispersion plots, 3 distinct states of asset functioning are unequivocally identified according to the various variables under analysis. This reinforces the idea that the equipment has three distinct operating states.

Analysing Figure 11, it is easy to identify three distinct groups of data in all ratios between variables. Temperature has higher values when current or torque is higher. The temperature is lower when the electric current is lower, the pressure is higher, and when the speed assumes its nominal value. It can be said that temperature and electric current are directly proportional. Electric current is higher when torque is higher. The circle of different colours represents the different clusters.

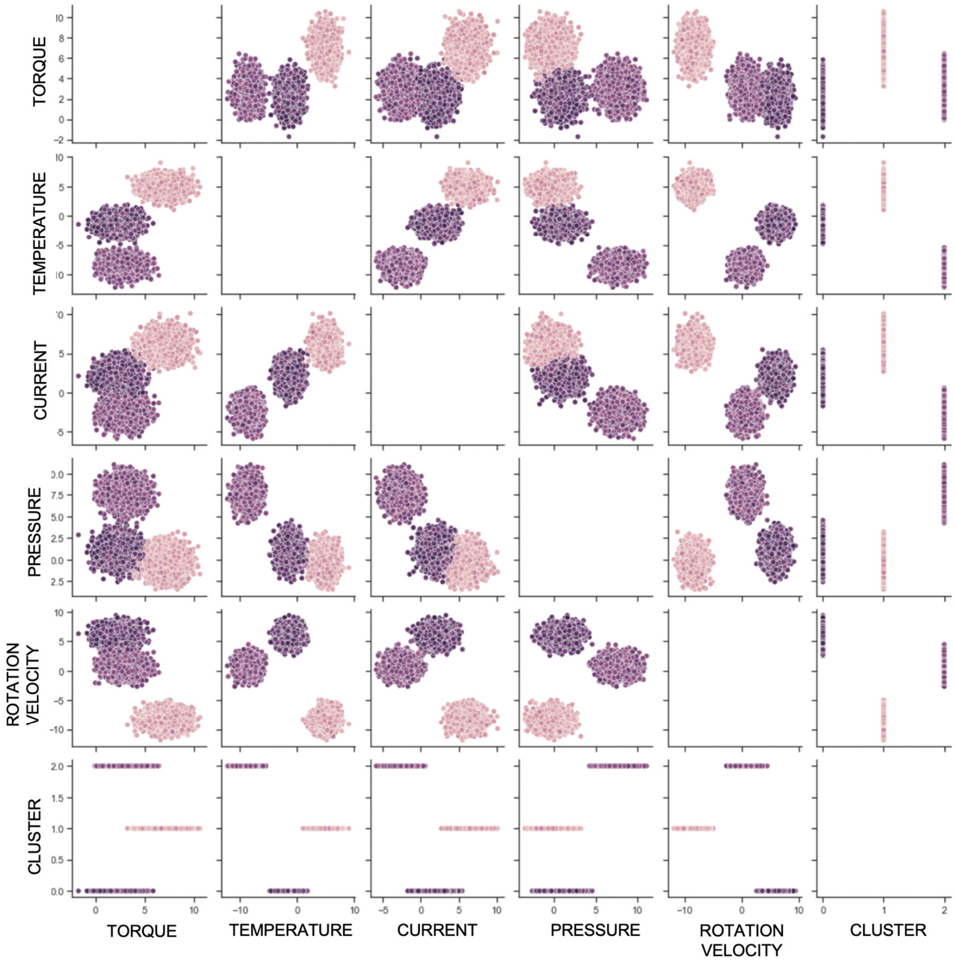

Figure 11. Matrix of the scattering plots of five variables and the state of operation.

7. Neural Networks Architecture

7.1. Network Classification for the State of the Paper Press

A neural model was developed to automatically classify each data sample into one of the three operational states. To achieve this classification, a 30-day data prediction performed by a neural network using MLPRegressor was used [33,34]. This prediction database was separated into two parts: the first 80% was used for training the model, and the remaining 20% was for carrying out the tests.

In this classification, we chose to work with feedforward architectures (MultiLayer Perceptron), using the Sklearn Python library model called MLPClassifier. Knowing that the dataset is very large, we chose to use a graph-based optimization algorithm named "adam", using a logistic sigmoid as activation function [36].

Several architecture combinations were tested to find the best possible network configuration in terms of accuracy, resulting in a final architecture with an accuracy above 96%.

Knowing that three clusters cover almost all the variance of the data, as indicated in the previous section, the authors defined that the neural network would classify the machine in three states.

The network is composed of a first layer with 6 neurons that receive information from the press sensors, then the information is processed by 2 layers of hidden neurons (100, 10). The output of this neural network is the following: Good, Alert, Failure. Figure 12 depicts the chosen ANN architecture for classification for the state of the paper press.

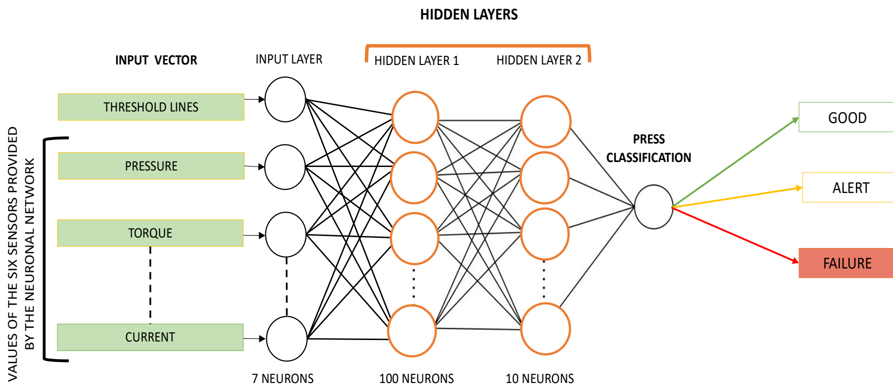

Figure 12. Architecture of the ANN for classification of the state of the paper press.

Training this classification network took approximately 4 min using Apple's M1 processor. The neural network required 150 iterations and had a final loss of 0.061

7.2. Neural Network for Press Lubricant Classification

The network is composed of a first layer with 12 neurons that receive information from the press sensors, then the information is processed by 3 layers of hidden neurons (500, 100, 10). The output of this neuronal network is the following: Oil in good Condition or Replace the oil.

The "lbfgs" solvers were chosen for this architecture, using "relu" as the activation function. The network needed 455 iterations for training. Training this classification network takes approximately 2 min using Apple's M1 processor. Figure 13 represents the ANN architecture chosen for oil classification.

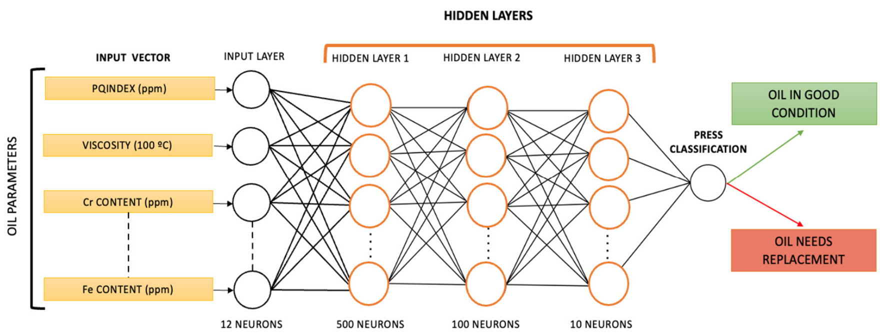

Figure 13. Architecture ANN for classification for oil.

7.3. Evaluation Models

To fully evaluate the effectiveness of a model, you must examine both precision and recall. Precision quantifies the number of positive class predictions that belong to the positive class. Recall quantifies the number of positive class predictions made from all positive examples in the dataset. Recall can also be called True Positive Rate (TPR), or Sensitivity.

Accuracy is a metric that summarizes the performance of a classification model as the number of correct predictions divided by the total number of predictions.

The next equations show the formulas for these metrics, where TP is True Positives, FN is False Negatives, FP is False Positives, and TN is True Negatives.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{\text{TP+FN}} \quad (2)$$

$$Accuracy = \frac{CorrectPredictions}{TotalPrediction} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$F1Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (4)$$

Macro AVG is the arithmetic mean of the individual classes' score in relation to precision, recall, and F1-score.

Weighted average considers how many of each class there were in its calculation, so fewer of one class means that its precision/recall/F1 score has less of an impact on the weighted average for each of those things.

8. Press State Classification Results

The classifier algorithm showed very good results, as the error was below the defined *p*-value of 5%. The results of the press state of the classification network were compared with the results of the classification network, the convergence was 96%.

Table 2 presents the different metrics and accuracy in the different states of the press. It was in the Failure classification that the network had its highest success rate.

Table 2. Press classification results.

Classification	Precision	Recall	F1-Score
Normal	0.96	1.00	0.98
Alert	0.98	0.85	0.91
Failure	1.00	0.90	0.94
Accuracy			0.96
Macro AVG	0.98	0.91	0.94
Weighted AVG	0.96	0.96	0.96

Figure 14 shows the results of the classification network. It is possible to observe the prediction of the press states in the future.

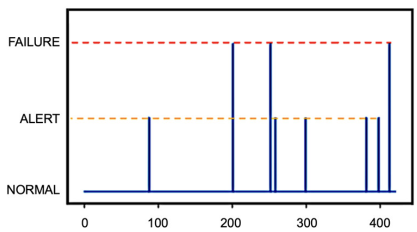

Figure 14. Time series of Press classification results.

Analysing Figure 14, it is observed that the press in the future will work mostly in the normal state, with five alerts and three expected malfunctions.

9. Lubricating Oil Classification Results

The condition of the press and any equipment depends a lot on the condition of the lubricant, so it is important to know its state of degradation. The algorithm created had an error well below the defined *p*-value of 5%. The accuracy of this classifier was 98%. Thus, the algorithm proved to be quite reliable to classify the oil of any equipment if they have a robust oil analysis database. Table 3 presents the classification of the lubricating oil results.

Table 3. Classification of the lubricating oil results.

Classification	Precision	Recall	F1-Score	Support
Oil in good Condition	0.96	1.00	0.98	27
Replace the oil	1.00	0.94	0.97	18
Accuracy			0.98	45
Macro AVG	0.98	0.97	0.98	45
Weighted AVG	0.98	0.98	0.98	45

Table 4 presents a confusion matrix where it can be observed that the classification network only missed a classification comparing its results with the results of human experts.

Table 4. Lubricating oil results (Confusion Matrix).

		Predictive Value	
		Oil in Good Condition	**Replace the Oil**
Real	Oil in good Condition	27	0
	Replace the oil	1	17

10. Limitations

One major limitation of the present approach is that it is based on machine learning, which uses inductive reasoning. There are no guarantees regarding the certainty of fault detection. However, the quality of the predictive data that are fed into the press condition classification network is quite good. The classification network learned according to the limit lines that were stipulated based on the history of the equipment and the recommendations of the technicians.

Finding the hyperparameters to use in the lubricant classification network and machine status classification network, so that the results meet the predefined objectives and errors, was the biggest challenge to overcome in this work. Obtaining a reliable and good-quality long-term forecast was the other difficulty encountered in this study [33,34]. To achieve small error margins, it is necessary to use deep knowledge of the machine being modeled, as well as machine learning methods. The parameters and methods are only valid for the machine being studied, even though similar procedures may be followed to pursue similar or better results for other machines.

11. Conclusions

Through clustering using Kmeans, it was shown that it is possible to identify equipment operating zones and, thus, define how many states the classifier network should have at its output.

The present developed model can classify future failures in a paper press considering a long-term forecast database of values.

The contribution of a 30-day classification is innovative and provides a great advantage in industrial planning, as it allows you to schedule stops a month in advance.

This methodology is very important in this area, as it can be applied to monitor this and other equipment automatically, if these assets have a robust database of their sensors.

The lubricant classifier was developed using neural networks. The results of this classifier, compared to the human expert classifications, had a hit rate close to 98%.

These contributions can have a significant impact on the quality of operation and availability of assets, thus reducing maintenance costs.

Author Contributions: Conceptualization, J.A.R. and M.M.; Methodology, M.M.; Software, J.A.R., A.M. and M.M.; Formal analysis, J.A.R.; Investigation, J.A.R.; Data curation, J.A.R.; Writing—original draft, J.A.R.; Writing—review & editing, M.M., R.J.G.M. and A.J.M.C.; Supervision, J.T.F.; Project administration, J.T.F. All authors have read and agreed to the published version of the manuscript.

Funding: The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement 871284 project SSHARE and the European Regional Development Fund (ERDF) through the Operational Programme for Competitiveness and Internationalization (COMPETE 2020), under Project POCI-01-0145-FEDER-029494, and by National Funds through the FCT—Portuguese Foundation for Science and Technology, under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020, and UIDP/04131/2020. This research is sponsored by FEDER funds through the program COMPETE—Programa Operacional Factores de Competitividade—and by national funds through FCT—Fundação para a Ciência e a Tecnologia—under the project UIDB/00285/2020. This work was produced with the support of INCD funded by FCT and FEDER under the project 01/SAICT/2016 n° 022153.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
AVG	Average
FF	Feed Forward
ITER	Iterations
MAPE	Mean Absolute Percentage Error
MLP	Multi-Layer Perceptron
MSE	Mean Square Error
PC	Principal Component
PCA	Principal Component Analysis
PCI	Principal Component Index
RF	Random Forest
RNN	Recurrent Neural Network
TAN	Total Acid Number
TPR	True Positive Rate

References

1. Kumar, U.; Galar, D.; Parida, A.; Stenström, C.; Berges, L. Maintenance Performance Metrics: A State-of-the-art Review. *J. Qual. Maint. Eng.* **2013**, *19*, 233–277. [CrossRef]
2. Selcuk, S. Manutenção Preditiva, Sua Implementação e Últimas Tendências. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2017**, *231*, 1670–1679. [CrossRef]
3. Martins, A.; Fonseca, I.; Farinha, J.; Reis, J.; Cardoso, A.J.M. Maintenance Prediction through Sensing Using Hidden Markov Models-A Case Study. *Appl. Sci.* **2021**, *11*, 7685. [CrossRef]
4. Uygun, Y. *Industry 4.0: Principles, Effects and Challenges—Nova Science Publishers*; Nova Science Publishers: Hauppauge, NY, USA, 2020; ISBN 1-5361-8423-3.
5. Cachada, A.; Barbosa, J.; Leitño, P.; Gcraldcs, C.A.S.; Deusdado, L.; Costa, J.; Teixeira, C.; Teixeira, J.; Moreira, A.H.J.; Moreira, P.M.; et al. Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. In Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, 4–7 September 2018; Volume 1, pp. 139–146.
6. Rodrigues, J.; Torres Farinha, J.; Marques Cardoso, A. Predictive Maintenance Tools—A Global Survey. *WSEAS Trans. Syst. Control* **2021**, *16*, 96–109. [CrossRef]
7. de-Almeida-e-Pais, J.E.; Raposo, H.; Farinha, J.; Cardoso, A.J.M.; Marques, P. Optimizing the Life Cycle of Physical Assets through an Integrated Life Cycle Assessment Method. *Energies* **2021**, *2021*, 6128. [CrossRef]
8. de-Almeida-e-Pais, J.E.; Farinha, J.; Cardoso, A.J.M.; Raposo, H. Optimizing the Life Cycle of Physical Assets -a Review. *WSEAS Trans. Syst. Control* **2020**, *15*, 417–430. [CrossRef]
9. Jain, A.K.; Murty, M.N.; Flynn, P.J. Data Clustering: A Review. *ACM Comput. Surv.* **1999**, *31*, 264–323. [CrossRef]
10. Xie, J.; Jiang, S.; Xie, W.; Gao, X. An Efficient Global K-Means Clustering Algorithm. *JCP* **2011**, *6*, 271–279. [CrossRef]
11. Bala, R.; Kumar, D.D. Classification Using ANN: A Review. *Int. J. Comput. Intell. Res.* **2017**, *13*, 10.

12. Rodrigues, J.; Farinha, J.; Mendes, M.; Mateus, R.; Cardoso, A.J.M. Short and Long Forecast to Implement Predictive Maintenance in a Pulp Industry. *Eksploat. Niezawodn.-Maint. Reliab.* **2021**, *24*, 33–41. [CrossRef]
13. Shenbagarajan, A.; Ramalingam, V.; Balasubramanian, C.; Palanivel, S. Tumor Diagnosis in MRI Brain Image Using ACM Segmentation and ANN-LM Classification Techniques. *Indian J. Sci. Technol.* **2016**, *9*. [CrossRef]
14. Rajab, S.; Sharma, V. Performance Evaluation of ANN and Neuro-Fuzzy System in Business Forecasting. In Proceedings of the 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 11–13 March 2015; pp. 749–754.
15. Agatonovic-Kustrin, S.; Beresford, R. Basic Concepts of Artificial Neural Network (ANN) Modeling and Its Application in Pharmaceutical Research. *J. Pharm. Biomed. Anal.* **2000**, *22*, 717–727. [CrossRef] [PubMed]
16. Paul, A.K.; Das, D.; Kamal, M.M. Bangla Speech Recognition System Using LPC and ANN. In Proceedings of the 2009 Seventh International Conference on Advances in Pattern Recognition, Kolkata, India, 4–6 February 2009; pp. 171–174.
17. Wahyuni, E.S. Arabic Speech Recognition Using MFCC Feature Extraction and ANN Classification. In Proceedings of the 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 1–3 November 2017; pp. 22–25.
18. Sinaga, K.P.; Yang, M.-S. Unsupervised K-Means Clustering Algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [CrossRef]
19. Kurita, T. Principal Component Analysis (PCA). In *Computer Vision: A Reference Guide*; Springer International Publishing: Cham, Switzerland, 2019; pp. 1–4, ISBN 978-3-030-03243-2.
20. Seal, A.; Karlekar, A.; Krejcar, O.; Herrera-Viedma, E. Performance and Convergence Analysis of Modified C-Means Using Jeffreys-Divergence for Clustering. *IJIMAI* **2021**, *7*, 141. [CrossRef]
21. Karlekar, A.; Seal, A.; Krejcar, O.; Gonzalo-Martin, C. Fuzzy K-Means Using Non-Linear S-Distance. *IEEE Access* **2019**, *7*, 55121–55131. [CrossRef]
22. Sharma, K.K.; Seal, A.; Yazidi, A.; Krejcar, O. A New Adaptive Mixture Distance-Based Improved Density Peaks Clustering for Gearbox Fault Diagnosis. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–16. [CrossRef]
23. Zhang, Z.-Y.; Wang, K.-S. Wind Turbine Fault Detection Based on SCADA Data Analysis Using ANN. *Adv. Manuf.* **2014**, *2*, 70–78. [CrossRef]
24. Ertunc, H.M.; Ocak, H.; Aliustaoglu, C. ANN- and ANFIS-Based Multi-Staged Decision Algorithm for the Detection and Diagnosis of Bearing Faults. *Neural Comput. Appl.* **2013**, *22*, 435–446. [CrossRef]
25. Rodrigues, J.; Cost, I.; Farinha, J.T.; Mendes, M.; Margalho, L. Predicting Motor Oil Condition Using Artificial Neural Networks and Principal Component Analysis. *Eksploat. Niezawodn.* **2020**, *22*, 440–448. [CrossRef]
26. Le, V.T.; Lim, C.P.; Mohamed, S.; Nahavandi, S.; Yen, L.; Gallasch, G.E.; Baker, S.; Ludovici, D.; Draper, N.; Wickramanayake, V. Condition Monitoring of Engine Lubrication Oil of Military Vehicles: A Machine Learning Approach. *AIAC* **2017**, 8.
27. Kittisupakorn, P.; Thitiyasook, P.; Hussain, M.A.; Daosud, W. Neural Network Based Model Predictive Control for a Steel Pickling Process. *J. Process Control* **2009**, *19*, 579–590. [CrossRef]
28. Gajewski, J.; Vališ, D. The Determination of Combustion Engine Condition and Reliability Using Oil Analysis by MLP and RBF Neural Networks. *Tribol. Int.* **2017**, *115*, 557–572. [CrossRef]
29. Zhang, D.; Xu, B.; Wood, J. Predict Failures in Production Lines: A Two-Stage Approach with Clustering and Supervised Learning. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data), Washington, DC, USA, 5–8 December 2016; pp. 2070–2074.
30. Mazumder, R.K.; Salman, A.M.; Li, Y. Failure Risk Analysis of Pipelines Using Data-Driven Machine Learning Algorithms. *Struct. Saf.* **2021**, *89*, 102047. [CrossRef]
31. Mateus, B.C.; Mendes, M.; Farinha, J.T.; Cardoso, A.M. Anticipating Future Behavior of an Industrial Press Using LSTM Networks. *Appl. Sci.* **2021**, *11*, 6101. [CrossRef]
32. Mateus, B.; Farinha, J.T.; Mendes, M.; Martins, A.B.; Cardoso, A.M. Data Analysis for Predictive Maintenance Using Time Series and Deep Learning Models—A Case Study in a Pulp Paper Industry. 2021. Available online: https://www.researchgate.net/publication/363646966_Data_Analysis_for_Predictive_Maintenance_Using_Time_Series_and_Deep_Learning_Models-A_Case_Study_in_a_Pulp_Paper_Industry (accessed on 1 December 2022).
33. Rodrigues, J.A.; Farinha, J.T.; Mendes, M.; Mateus, R.J.G.; Cardoso, A.J.M. Comparison of Different Features and Neural Networks for Predicting Industrial Paper Press Condition. *Energies* **2022**, *15*, 6308. [CrossRef]
34. Rodrigues, J.A.; Farinha, J.T.; Cardoso, A.M.; Mendes, M.; Mateus, R. Prediction of Sensor Values in Paper Pulp Industry Using Neural Networks. In *Proceedings of IncoME-VI and TEPEN 2021*; Zhang, H., Feng, G., Wang, H., Gu, F., Sinha, J.K., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 281–291.
35. Thorndike, R.L. Who Belongs in the Family? *Psychometrika* **1953**, *18*, 267–276. [CrossRef]
36. Sklearn.Neural_network.MLPClassifier. Available online: https://scikit-learn/stable/modules/generated/sklearn.neural_network.MLPClassifier.html (accessed on 1 December 2022).

Appendix E

Prediction of Sensor Values in Paper Pulp Industry Using Neural Networks

João Antunes Rodrigues [1,2], José Torres Farinha[3,4], António Marques Cardoso[1], Mateus Mendes [3,5], Ricardo Mateus [2]

[1] CISE, Univ. Beira Interior, Covilhã, 6201-001, Portugal
[2] Research Centre in Industrial Engineering, Management and Sustainability, Universidade Lusófona, Campo Grande 376, 1749-024, Lisboa, Portugal
[3] Polytechnic of Coimbra – ISEC, Quinta da Nora, 3030-199 Coimbra, Portugal
[4] CEMMPRE, Coimbra University, DEM, Polo 2, 3030-290 Coimbra, Portugal.
[5] University of Coimbra – ISR, Coimbra, Portugal

p5942@ulusofona.pt; torres.farinha@dem.uc.pt; ajmcardoso@ieee.org; mmendes@isec.pt, p5768@ulusofona.pt

Abstract. The economic sustainability of any industry is directly linked to the management and efficiency of its physical assets. The maintenance of these assets is one of the key elements for the success of a company since it represents a relevant part of its Capital and Operational Expenses (CAPEX and OPEX). Due to the importance of maintenance, a lot of research has been done to improve the methodologies aiming to maximize physical assets' availability at the most rational costs. The introduction of Artificial Intelligence in the world of maintenance increased the quality of prediction on equipment failures, namely when associated to continuous equipment monitoring. This paper presents a case study where a neural network is proposed to predict the future values of various sensors installed on a paper pulp press. Data from the following variables is processed: electric current; pressure; temperature; torque; and speed.

Keywords: Predictive Maintenance; Condition Monitoring; Neural Networks; Forecasting; Artificial Intelligence.

1 Introduction

Cost reductions along with regulations and concerns for safety and environmental impacts play a major role in the success of industries. [1]

Industries have been looking for and investigating new techniques and equipment management tools that provide a competitive advantage in the quality or cost of their products, processes or services. Industrial equipment maintenance is therefore a key issue [1].

Kumar and the British Standards Institute (2015) define maintenance as the combination of all technical and administrative activities necessary to keep equipment, facilities and other physical assets in the desired operational condition or to restore them to comply with its function with quality [2] .

The four main objectives of maintenance are: safety, quality, cost and availability.

There are many different approaches to maintenance. The predictive approach is one of the most important and effective, aiming to maximize the equipment's availability at a minimum cost.

This type of maintenance mainly involves predicting system failures, the main task of which is to detect the first signs of failure. Taking these signs into account, predictive maintenance aims to warn when faults are likely to occur and thus to suggest making scheduled stops on the asset. The advancement of technology and new techniques in the field of predictive maintenance have been making predictive maintenance more efficient, applicable and accessible to industries.

Predictive maintenance aims to predict the occurrence of failures before they happen, using data from consistent and constant monitoring of the conditions and operation of the target equipment.

Unwanted conditions, such as wear and tear of equipment components, are observed and/or predicted using forecasting algorithms to optimize when preventive interventions in the asset should be scheduled, thus avoiding breakdowns and reducing repair costs and production losses, and consequently increasing the asset availability.

Predictive maintenance has been adopted by various sectors in the manufacturing and service industries, in order to improve reliability, safety, availability, quality, as well as promoting environmental sustainability, since predictive maintenance reduces production surpluses as well as non-compliant products [3].

Quality measures how well an asset performs its function properly, while reliability measures how that asset maintains its original level of quality steady over time, under the various operating conditions to which it is exposed.

Predictive maintenance techniques are increasingly associated with sensor technologies because to make good predictive maintenance decisions it is necessary to have good quality information regarding the past and current operations of the asset. Therefore, it is paramount to calibrate sensors properly and to process the resulting data in a reliable way [4].

The present paper focus on a predictive maintenance approach aiming to identify the current equipment state and predict its future operating conditions through collected past data processed using Artificial Intelligence (AI) algorithms.

Prediction is made for the next 90 days, a timeframe which allows the industry to adequately prepare and schedule maintenance interventions, thereby avoiding loss of production and optimizing stopping time. The company's competitiveness advantage is achieved by reducing maintenance downtime and increasing production time.

A multi-layer perceptron Artificial Neural Networks (ANN) model is presented for predicting future data from various variables. The algorithm presented in this paper is implemented in Python, using the supervised learning model MLP Regressor from Scikit-learn (Sklearn).

Sklearn is an open-source machine learning library for the Python programming language. If offers numerous functions for processing data, performing operations such as sorting, regression and grouping algorithms, including Support Vector Machines (SVM), Random Forest, gradient augmentation and k-means. It is designed to work with the numerical and scientific Python libraries NumPy and SciPy. Sklearn is written in Python and uses Numpy extensively for high-performance linear algebra and array operations.

2 Literature Review

Rodrigues *et al*. (2019) use Feed Forward Neural Networks to classify the level of degradation of lubricants of Diesel engines. The results show that neural network models can classify oil conditions, achieving more than 90% precision compared to the performance of human experts, and thus allowing the process to be automated in the future [5].

It is important to have continuous and efficient maintenance in order to keep assets as available as possible and with no accidents. Bukhsh (2019) developed predictive models that used existing data from a railway and produced interpretable results. Such predictive models were supported by classification techniques based on Decision Trees, Random Forest and Fault Trees with gradient increase; these tools allow to predict the need for maintenance on the railway [6].

Hongxiang *et al*. develop an algorithm using Artificial Neural Networks to analyze spectroscopy data from an oil. Results show that the mining of oil spectroscopy data by ANN methods can be used to classify types of lubricant and distinguish routine conditions of a Diesel engine from operating conditions [7].

Prediction in maintenance area support making better decisions. Okoh *et al*. (2017) present an approach to determine when a system needs to undergo maintenance, repair, and overhaul, before a failure occurs. The novelty in this study is the development of the through-life performance approach [8].

One of the main maintenance challenges is to increase equipment availability. Makridis *et al*. (2020) describe that predictive maintenance extends vessel lifetimes in the marite sector, while reducing overall maintenance costs as well. The authors present a machine learning approach for detecting anomalies in the data collected through sensors installed on the vessels, hence predicting the condition of specific parts of the vessel's main engine [9].

Upgrading equipment and making it smarter is a common goal for many managers around the world. However, leveraging Artificial Intelligence methods from older equipment is sometimes extremely difficult, as these assets must be equipped with technical diagnostic tools and sensors for data collection. Vlasocv et *al*. (2018) discuss methods for maintaining industrial equipment, with a focus on predictive maintenance and the principles for building wireless sensor networks and data transmission protocols to collect information. The purpose of this study is to demonstrate the feasibility and reliability of using wireless sensors as technical diagnostic tools and as decision support tools for prediction. Main advantages include cost reductions and real-time information and analysis of equipment's state [10].

The use of Internet of Things (IoT) technologies to allow the exchange of information among sensors, machines, servers, and processing units is currently revolutionizing the industrial world, also providing innovation to the maintenance sector. Fernandes *et al*. (2020) describe a related system for the prediction of failures in a metallurgical industry. There was no history of failures, so, learning takes place in an unsupervised way. Failures are predicted through moving average models, integrated in autoregressive models, using data from the sensors installed in the equipment, thus allowing the monitoring of different machine components and parameters [11].

3 Methods

3.1 Dataset and data preprocessing

The present case study focuses on the prediction of target variables on an industrial paper pulp press. The company in question provided a three-year dataset containing the history of six variables: Electric Current Intensity (Sensor 1); Pressure (Sensor 2); Rotation Speed (Sensor 3); Temperature (Sensor 4); Torque (Sensor 5); and Velocity (Sensor 6). All sensors collected data with a sample frequency of one minute.

The dataset contains several repeated values as well as discrepant samples which may be due to reading errors: the upper outliers may be the result of errors in the sensors reading; the lower outliers are a consequence of the same causes previously mentioned, as well as possibly downtimes programmed or not programmed.

Repeated values were removed using a Python algorithm developed by Balduino *et al*. (2021)[12]. The lower and upper outliers were also removed and subsequently replaced by the average of each variable in question.

Data was then transformed into histogram bins and predefined statistics computed from sliding windows. These statistics are the input of the ANN model Network. It should be noted that all data were normalized using the StandardScaler library from Sklearn before feeding into the ANN model.

To evaluate the performance of the Neural Network prediction algorithm, two different evaluation metrics were used: Mean Absolute Percentage Error (MAPE); and Mean Squared Error (MSE).

3.2 ANN Architecture

For predicting future values, the chosen architecture model was the Multi-Layer Perceptron (MLP), which is one of the most popular feed forward architectures, using Sklearn's MLPRegressor.

MLPRegressor uses various hyper parameters to optimize the generalization capacity of the network model for prediction. Various combinations of hyper-parameters were tested during training to find the best possible network configuration, so that the best forecast could be achieved.

To avoid overfitting, MLPRegressor might include a regularization term added to the loss function to reduce the number of model parameters.

The algorithm chosen for weight optimization was Adam solver. The Adam solver is an optimizer algorithm based on a stochastic gradient proposed by King-ma, Diederik and Jimmy Ba. This solver is recommended for large data sets [13].

In order to choose the ideal number of hidden layers, tests were carried out with one, two, and three hidden layers.

The initial approach for the input vector of the ANN considered only data from histogram bins of a sliding window, one for each sensor, along with their average values. However, the prediction results were very irregular. Therefore, it was decided to introduce additional metrics regarding the variance and median in the input vector, which made the prediction results more stable.

Results from tests carried out with two and three hidden layers were very similar and better than with one layer only. Hence, a network with two hidden layers (150, 75) was chosen. Figure 1 depicts the chosen architecture.

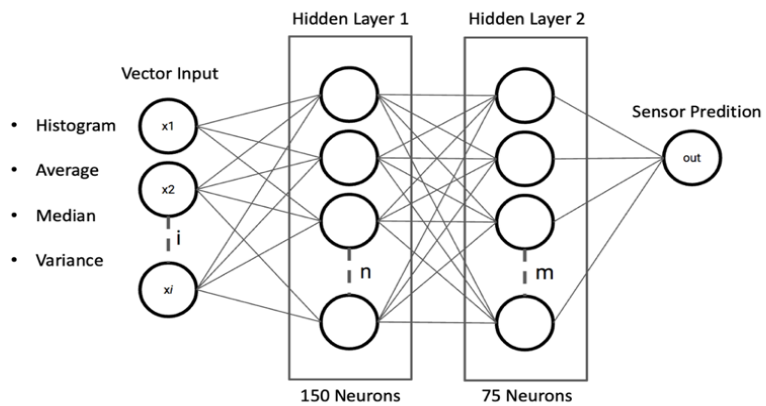

Fig. 1. Network architecture. The network receives histograms, average, median and variance from each sensor for a given sliding window, and outputs predictions for that sensor.

4 Experiments and results

This section describes the experiments carried out during the development of this research study along with some of the best results obtained through several tests.

Tests were carried out with various window sizes applying. Histograms computed from those sliding windows, along with the respective mean, median, and variance, compose the input vector $(S_{11}, S_{12}, \ldots, S_{1w}, \ldots, S_{nw}, AVG_1, \ldots AVG_n, M_1, \ldots M_n, V_1, \ldots V_n)$ of the ANN model. S_{ij} represents the value of sensor i at time j of the window of size w. The window started with the first w samples of the time series and slided to the end of the series, in steps of 1, for an overlapping window, or steps of w samples for a non-overlapping window. AVG_k is the mean value of sensor k in the window. M_k is the median value of sensor k for the window and V_k is its variance.

Data were resampled to speed up processing. Results shown in Table 1 and Table 2 were obtained with a period of 10 samples. The Neural Network performed up to 200 learning epochs in all tests. Table 1 shows the results of the tests developed using the overlapping window.

Table 1. Results of the prediction tests with overlapping sliding window

Window Size (Samples)	Oil		Current		Temperature		Torque		Pressure		Velocity	
	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss
2440	10.51	0.70	1.73	0.13	3.26	0.49	0.62	0.06	9.53	1.69	2.14	0.25
1440	7.23	0.48	1.67	0.11	3.67	0.29	0.80	0.06	7.57	1.18	2.84	0.19
720	5.94	0.19	1.56	0.09	3.55	0.19	0.54	0.06	4.74	0.78	1.84	0.14
360	5.22	0.07	1.45	0.07	3.28	0.13	0.56	0.05	5.39	0.51	1.68	0.04
180	5.44	0.13	1.15	0.06	2.59	0.10	0.55	0.04	4.94	0.36	1.43	0.08
144	5.69	0.12	1.15	0.06	2.46	0.09	0.54	0.04	4.62	0.34	1.51	0.08
90	5.02	0.11	1.08	0.05	2.46	0.09	0.56	0.04	4.53	0.31	1.43	0.07
45	5.24	0.20	1.08	0.06	2.23	0.11	0.56	0.03	4.97	0.37	1.46	0.07
24	5.60	0.41	1.14	0.07	2.29	0.19	0.52	0.03	4.96	0.69	1.49	0.08
12	5.76	0.86	1.10	0.11	2.16	0.34	0.53	0.03	4.70	1.25	1.61	0.13

It can be seen from Table 1 that the window size of 2440 samples is the one that obtains the worst loss and MSE results for all variables. In general, the window can be reduced to six hours (360 samples) or in some cases up to 12 minutes (12 samples), since they all present good prediction results, which indicate that these will be the best window sizes to use for 10 samples resample rate in this data set.

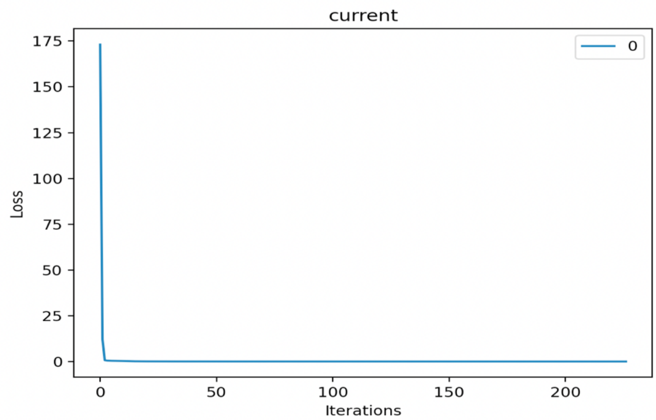

Fig. 2. Learning history of current's intensity, using an overlap window size of 12 samples. The curve shows that the model learns in the first epochs.

Figure 2 showd that the learning of the Neural Network is quite fast: in just three epochs the network has already learned up to a small error.

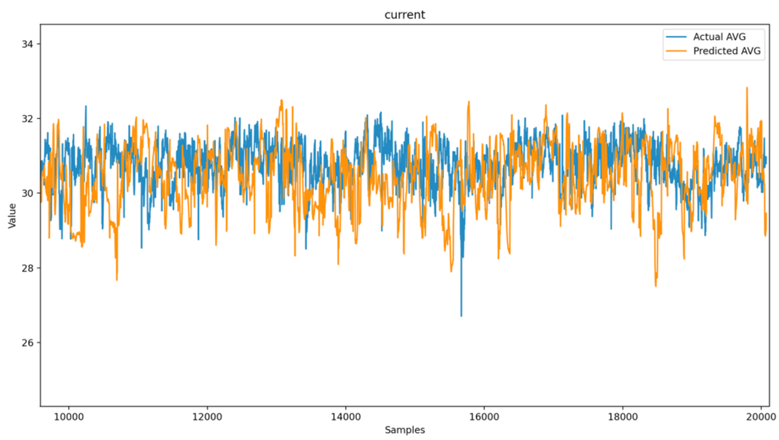

Fig. 3. Current's intensity prediction using an overlapping window size of 12 samples.

Fast learning and the low data loss make it possible to achieve a good prediction fit as shown in Figure 3. This prediction presents a Mean Squared Error of 1.10 (see Table

1). Table 2 shows corresponding results for the tests developed with non-overlapping sliding windows.

Table 2. Results of the prediction tests using non-overlapping sliding windows

Window Size (Samples)	Oil		Current		Temperature		Torque		Pressure		Velocity	
	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss	MSE	Loss
2440	66.93	2016.18	20.61	156.63	26.90	348.65	6.52	9.13	8.12	17.95	1.00	0.19
1440	66.19	1985.44	19.06	153.46	25.40	332.34	5.34	9.03	8.31	18.00	0.65	0.36
720	65.94	1945.29	18.79	153.35	25.32	329.65	4.99	9.08	8.21	20.09	1.19	0.59
360	54.69	1327.35	10.13	44.62	15.58	135.72	1.03	0.34	4.50	10.41	1.13	0.75
180	43.93	836.25	4.64	7.71	8.25	42.00	0.53	0.14	4.25	8.76	1.07	0.73
144	43.95	837.15	4.69	7.81	8.28	41.90	0.52	0.14	4.16	8.04	1.14	0.83
90	25.99	267.54	1.30	0.65	1.61	3.39	0.46	0.12	4.04	7.98	1.18	0.76
45	8.60	26.05	1.02	0.53	1.59	1.92	0.47	0.13	3.80	7.21	1.15	0.57
24	4.64	11.07	0.74	0.44	1.58	1.68	0.49	0.11	4.15	3.58	1.41	0.17
12	5.45	3.61	0.78	0.26	1.77	0.78	0.48	0.10	4.72	1.13	1.56	0.09

Table 2 shows the learning problems resulting from Neural Network when using large size windows without overlapping. However, window sizes with 24 and 12 samples generate results that are already satisfactory.

In the absence of the overlapping technique, the Neural Network receives much less input data per epoch, which delays the learning process. This problem is in part overcome by reducing the window size, for the Network receives more input data samples per epoch. However, when using smaller windows, it can be more difficult to catch larger patterns.

Figure 4 illustrates the network learning history for predicting velocity using a 12-sample overlapping window size. A slow slope is evident, demonstrating a slower learning rate when compared to other variables.

Note though that albeit learning is slow, the Network can learn, presenting a loss value of 0.09 in the window of 12 samples for the variable in question.

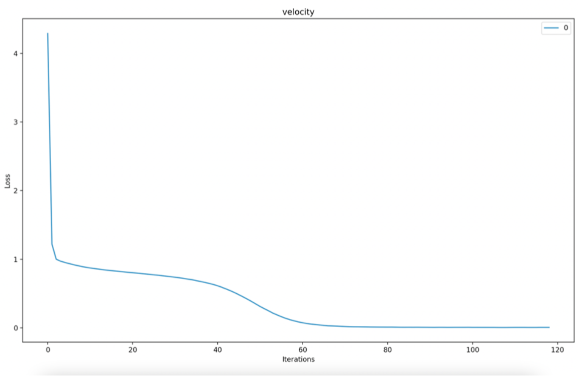

Fig. 4. Learning curve of the velocity using a non-overlapping window size of 12 samples

Figure 5 presents the corresponding results of the velocity prediction using a window size of 12 samples and 200 learning epochs. It should be noted that for the test shown in this figure the outliers from a programmed stop were not removed.

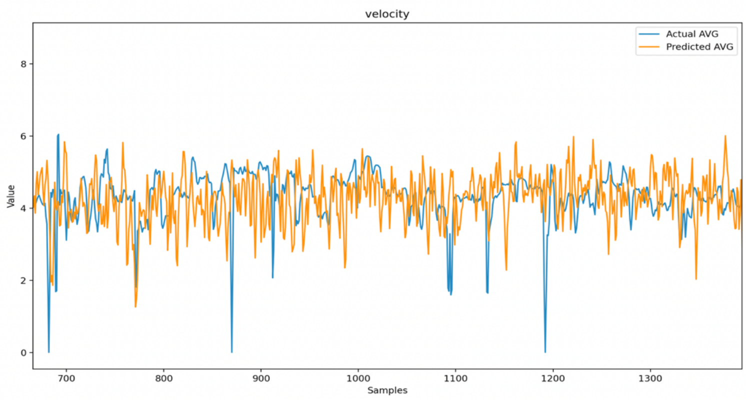

Fig. 5. Velocity prediction using a non-overlapping window size 12 samples

5 Conclusion

Prediction is very important for better decisions in maintenance and other areas. Data resampling can make the prediction process much faster since it reduces the dataset considerably.

The use of sliding windows over time series is necessary for training. Overlapping windows offer learning in less epochs. Larger windows make it easier to catch long trends, but the optimal window size needs to be determined experimentally. Overlapping windows offer more input data to the Neural Network in each epoch and thus generating faster learning rates and better prediction results.

Another drawback of non-overlapping windows is the limitation imposed by small datasets. Another major disadvantage of this type of sliding windows is the slower learning rate, especially for small window sizes. Its great advantage though is the speed of processing when the input vector is created.

Future work includes additional experiments to improve the input vectors and optimize other neural network hyperparameters.

Funding

The research leading to these results has received funding from the European Union's Hori- zon 2020 research and innovation programme under the Marie Sklodowvska-Curie grant agreement 871284 project SSHARE and the European Regional Development Fund (ERDF) through the Oper- ational Programme for Competitiveness and Internationalization (COMPETE 2020), under Project POCI-01-0145-FEDER-029494, and by National Funds through the FCT—Portuguese Foundation for Science and Technology, under Projects PTDC/EEI-EEE/29494/2017, UIDB/04131/2020, and UIDP/04131/2020.

Acknowledgments

This research is sponsored by FEDER funds through the program COMPETE— Programa Operacional Factores de Competitividade—and by national funds through FCT—Fundação para a Ciência e a Tecnologia—under the project UIDB/00285/2020. This work was produced with the support of INCD funded by FCT and FEDER under the project 01/SAICT/2016 no 022153.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] A. Cachada *et al.*, 'Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture', in *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2018, vol. 1, pp. 139–146. doi: 10.1109/ETFA.2018.8502489.

[2] U. Kumar, D. Galar, A. Parida, C. Stenström, and L. Berges, 'Maintenance performance metrics: a state-of-the-art review', *J. Qual. Maint. Eng.*, vol. 19, no. 3, pp. 233–277, Jan. 2013, doi: 10.1108/JQME-05-2013-0029.

[3] S. Selcuk, 'Predictive maintenance, its implementation and latest trends', *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.*, vol. 231, no. 9, pp. 1670–1679, Jul. 2017, doi: 10.1177/0954405415601640.

[4] A. B. Martins, J. Torres Farinha, and A. Marques Cardoso, 'Calibration and Certification of Industrial Sensors – a Global Review', *WSEAS Trans. Syst. CONTROL*, vol. 15, pp. 394–416, Sep. 2020, doi: 10.37394/23203.2020.15.41.

[5] J. Rodrigues, I. Costa, J. T. Farinha, M. Mendes, and L. Margalho, 'Predicting motor oil condition using artificial neural networks and principal component analysis', *Sci. Technol.*, vol. 22, no. 3, p. 9, 2020.

[6] Z. Allah Bukhsh, A. Saeed, I. Stipanovic, and A. G. Doree, 'Predictive maintenance using tree-based classification techniques: A case of railway switches', *Transp. Res. Part C Emerg. Technol.*, vol. 101, pp. 35–54, Apr. 2019, doi: 10.1016/j.trc.2019.02.001.

[7] T. Hongxiang, L. Yuntao, and W. Xiangjun, 'Application of Neural Network to Diesel Engine SOA', in *2011 Third International Conference on Measuring Technology and Mechatronics Automation*, Jan. 2011, vol. 1, pp. 555–558. doi: 10.1109/ICMTMA.2011.141.

[8] C. Okoh, R. Roy, and J. Mehnen, 'Predictive Maintenance Modelling for Through-Life Engineering Services', *Procedia CIRP*, vol. 59, pp. 196–201, Jan. 2017, doi: 10.1016/j.procir.2016.09.033.

[9] G. Makridis, D. Kyriazis, and S. Plitsos, 'Predictive maintenance leveraging machine learning for time-series forecasting in the maritime industry', in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, Sep. 2020, pp. 1–8. doi: 10.1109/ITSC45102.2020.9294450.

[10] A. I. Vlasov, P. V. Grigoriev, A. I. Krivoshein, V. A. Shakhnov, S. S. Filin, and V. S. Migalin, 'Smart management of technologies: predictive maintenance of industrial equipment using wireless sensor networks', *Entrep. Sustain. Issues*, vol. 6, no. 2, pp. 489–502, 2018, doi: 10.9770/jesi.2018.6.2(2).

[11] M. Fernandes, A. Canito, J. M. Corchado, and G. Marreiros, 'Fault Detection Mechanism of a Predictive Maintenance System Based on Autoregressive Integrated Moving Average Models', in *Distributed Computing and Artificial Intelligence, 16th International Conference*, Cham, 2020, pp. 171–180. doi: 10.1007/978-3-030-23887-2_20.

[12] B. C. Mateus, M. Mendes, J. T. Farinha, and A. M. Cardoso, 'Anticipating Future Behavior of an Industrial Press Using LSTM Networks', *Appl. Sci.*, vol. 11, no. 13, Art. no. 13, Jan. 2021, doi: 10.3390/app11136101.

[13] 'sklearn.neural_network.MLPRegressor — scikit-learn 0.24.2 documentation'. https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html (accessed May 16, 2021).