UNIVERSIDADE
BEIRA INTERIOR

Context-aware algorithms for Diabetes or Prediabetes prediction and diagnosis support in Ambient Assisted Living.

Virginie dos Santos Felizardo

Tese para obtenção do Grau de Doutor em
Engenharia Informática
(3º ciclo de estudos)

Orientador: Prof. Doutor Nuno Manuel Garcia dos Santos
Co-orientador: Prof. Doutor Miguel Castelo-Branco Sousa

Júri:
Prof. Doutor Hugo Pedro Martins Carriço Proença
Prof. Doutora Ana Cristina Costa Aguiar
Prof. Doutor Hugo Humberto Plácido da Silva
Prof. Doutor Pedro Miguel Alves Brandão
Prof. Doutora Imen Megdiche Bousarasr
Prof. Doutor Nuno Gonçalo Coelho Costa Pombo
Prof. Doutor Nuno Manuel Garcia dos Santos

Março 2023

UNIVERSIDADE BEIRA INTERIOR

Declaração de integridade

Eu, *Virginie dos Santos Felizardo*, que abaixo assino, estudante com número de inscrição *d1607* do Doutoramento em *Engenharia Informática* da Faculdade de *Engenharia*, declaro ter desenvolvido o presente trabalho e elaborado o presente texto em total consonância com o **Código de integridade da Universidade da Beira Interior**.

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

Universidade da Beira Interior, Covilhã.

Data: 2/3/2023

Assinatura: Virginie Felizardo

Dedication

"You lose your curiosity when you stop learning." - Katherine Johnson

To my family and friends,

To my supervisors, colleagues and professors,

who allowed me to do what I like and enjoy what I do.

"Pour ce qui est de l'avenir, il ne s'agit pas de le prévoir, mais de le rendre possible." - Antoine de Saint-Exupéry

Acknowledgements

Thesis prepared at the Ambient Assisted Living Computing and Telecommunications Laboratory (ALLab), Instituto de Telecomunicações, Universidade da Beira Interior, and submitted to the Universidade da Beira Interior for defense in a public examination session.

I would like to thank my supervisors Dr. Nuno M. Garcia and Dr. Miguel Castelo-Branco Sousa for all their help and advice with this PhD, their knowledge and experience have encouraged me in all research activities. I would also like to thank Dr. Nuno Pombo, Dr. Imen Megdiche, Dr. František Babič and Dr. Pedro Brandão for their technical support on my study. I would like to thank my friends and colleagues from ALLab for a cherished time spent together in the lab. Finally, I would like to express my gratitude to my family, mainly to my husband and son, without their understanding and encouragement in the past few years, it would be impossible for me to complete my study.

The research has been partially funded by the FCT/MCTES through national funds, and when applicable, co-funded EU funds under the project UIDB/50008/2020 and Operação Centro 01-0145-FEDER-000019 – C4 – Centro de Competências em Cloud Computing, co-financed by the Programa Operacional Regional do Centro (CENTRO 2020), through the Sistema de Apoio à Investigação Científica e Tecnológica – Programas Integrados de IC&DT. I would also like to acknowledge the contribution of the COST Action IC1303: AAPELE—Archi- tectures, Algorithms and Protocols for Enhanced Living Environments and COST Action CA16226; SHELD-ON—Indoor living space improvement: Smart Habitat for the Elderly, supported by COST (European Cooperation in Science and Technology).

Foreword

This thesis presents the research work performed in the doctoral research programme in computer science. The doctoral research was conducted at the Ambient Assisted Living Computing and Telecommunications Laboratory (ALLab), Universidade da Beira Interior and Instituto de Telecomunicações, Covilhã, Portugal. The research work was supervised by Prof. Dr. Nuno M. Garcia (Universidade da Beira Interior), and co-supervised by Prof. Dr. Miguel Castelo-Branco Sousa (Universidade da Beira Interior). Several research activities are supported by the contributions of Prof. Dr. Nuno Pombo (Universidade da Beira Interior), Prof. Dr. Imen Megdiche (Institut National Universitaire Champollion, France), Prof. Dr. František Babič (Technical University of Kosice, Slovakia) and Prof. Dr. Pedro Brandão (Universidade do Porto). This work was financially supported in part by the FCT/MCTES through national funds, and when applicable, co-funded EU funds under the project identified by UIDB/50008/2020 and some activities were supported by the COST Action IC1303–AAPELE–Architectures, Algorithms and Protocols for Enhanced Living Environments, funded by The COST Association, through the European Science Foundation and the H2020 program.

In this thesis, we tackle different subjects that are connected to context-aware algorithms for diabetes prediction for Ambient Assisted Living solutions. First, the preparatory study resulted in the decision of the prediction target. Hypoglycaemia (or hypoglycemia in US English) is one of the short-term complications associated with bad control of blood glucose. The prediction of this unannounced event can improve the quality of life of these patients. Second, there are many methods for glycaemia prediction based on data collected with continuous glucose monitoring sensors, but these methods are not extensible to patients using finger sticks to monitor blood glucose. The research performed in this thesis includes a study of the state-of-the-art, the design and development of the predictive framework, and the experiment's evaluation using patients' real data. The study of the state-of-the-art comprised a systematic literature review on data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction. This study's findings supported the design of the prediction framework, showing that there are some open possibilities in data- and model-fusion. The research performed in this thesis resulted in the definition and design of context-aware framework for hypoglycaemia prediction using sparse data, information fusion and classifiers consensus decision in a 24h hour time frame. The proposed method includes the integration of contextual and time-based features with the conventional features and a classifier's consensus decision to catching appropriate patterns associated with hypoglycaemia prediction. The results obtained are promising, presenting several contributions to enhanced hypoglycaemia predictions using discrete data.

List of publications

Articles included in the thesis resulting from this doctoral research programme

[1] V. Felizardo, N. M. Garcia, N. Pombo & I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review," in Artificial Intelligence In Medicine, 118 (2021). https://doi.org/10.1016/j.artmed.2021.102120.

[2] V. Felizardo, D. Machado, N. M. Garcia, N. Pombo and P. Brandão, "Hypoglycaemia Prediction Models With Auto Explanation," in IEEE Access, vol. 10, pp. 57930-57941, 2022, doi: 10.1109/ACCESS.2021.3117340.

[3] V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, & F. Babič, "A novel context-aware framework for Hypoglycaemia Events Prediction integrating context and time-based features," (Under review).

[4] V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, M. Sousa & F. Babič, "Hypoglycaemia prediction using information fusion and classifiers consensus," (Under review).

[5] V. Felizardo, N. M. Garcia, N. Pombo, & I. Megdiche, "Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos," 2021 (patent application n.o 117301).

Other publications

[1] S. Merilampi, A. Koivisto, M. Leino, N. Pombo, V. Felizardo, J. Lu, A. Poberznik & J. Virkki, "Implementing Mobile Games into Care Services—Service Models for Finnish and Chinese Elderly Care," in Information 10(2), 55, 2019. https://doi.org/10.3390/info10020055.

[2] S. Pirbhulal, N. Pombo, V. Felizardo, N. Garcia, A. H. Sodhro and S. C. Mukhopadhyay, "Towards Machine Learning Enabled Security Framework for IoT-based Healthcare," 2019 13th International Conference on Sensing Technology (ICST), 2019, pp. 1-6, doi: 10.1109/ICST46873.2019.9047745.

[3] L. Souza-Pereira, N. Pombo, S. Ouhbi, V. Felizardo, N. Garcia, "Clinical decision support systems for chronic diseases: A Systematic literature review", Comput. Methods Programs-Biomed. 195 (2020) 105565.
https://doi.org/10.1016/j.cmpb.2020.105565.

[4] I. Matias, N. Garcia, S. Pirbhulal, V. Felizardo, N. Pombo, H. Zacarias M. Sousa, E. Zdravevskide, "Prediction of Atrial Fibrillation using artificial intelligence on Electrocardiograms: A systematic review", Computer Science Review, 39 (2021).
https://doi.org/10.1016/j.cosrev.2020.100334.

Resumo

A necessidade de controlar a diabetes oferece a oportunidade para o desenvolvimento de novas soluções tecnológicas para autogestão e previsões reais. Essas previsões podem ser úteis na prevenção de eventos indesejados, como a hipoglicemia. Os pacientes com diabetes tipo 1 e alguns pacientes com diabetes tipo 2 geralmente temem a hipoglicemia. O objetivo desta tese é desenvolver um algoritmo sensível ao contexto para a previsão de hipoglicemia recorrendo a dados discretos, fusão de informação e decisão consensual de classificadores em um período de 24 horas. Com esta abordagem, contribuímos propondo um algoritmo de previsão de hipoglicemia, num cenário de autogestão, permitindo que seja utilizado por pacientes que realizam a sua monitorização com um glicosímetro. A literatura propõe algoritmos de previsão de glicemia usando dados de sistemas contínuos de monitorização, mas essas abordagens não são extensíveis a pacientes sem esses sistemas. Algoritmos de previsão baseados em informações discretas são um desafio, por isso propusemos uma novo algoritmo sensível ao contexto para previsão de hipoglicemia com base na fusão de dados e decisão de consenso de classificadores. A fusão de informações de contexto com as variáveis convencionais pode contribuir para diminuir o efeito da inter- e intra-variabilidade nos padrões de previsão. Além disso, a decisão de previsão baseada no consenso dos classificadores pode contribuir para a criação de algoritmos preditivos adequados e generalizados. A integração de variáveis contextuais e baseadas no tempo melhora a precisão e a previsão de hipoglicemia. Usando a decisão de consenso dos classificadores, 66% dos pacientes têm mais de 90% de hipoglicemias previstas (37,7% dos pacientes com 100% de hipoglicemias previstas) sem o aumento de falsos positivos (alarmes falsos). Este trabalho mostra a importância da fusão de dados e decisão consensual para capturar os padrões associados ao risco de hipoglicemia e sua previsão, no entanto, é necessário aprofundar mais a questão da fusão de dados e modelos e explorar a interpretabilidade dos modelos preditivos.

Palavras-chave

Hipoglicemia, previsão, algoritmo sensivel ao contexto, fusão de dados, variáveis baseadas no tempo, decisão de consenso de classificadores.

Resumo Alargado

Introdução

Esta secção vem resumir, em língua Portuguesa, o trabalho de investigação da tese de doutoramento intitulada de "Context-aware algorithms for Diabetes or Prediabetes prediction and diagnosis support in Ambient Assisted Living". A Diabetes mellitus (DM) é caracterizada pela incapacidade ou insuficiência de produzir insulina resultando na dificuldade de controlar os níveis de glicose. Existem diferentes condições de diabetes, entre elas, diabetes tipo 1 (T1D) que requer terapia com insulina e diabetes tipo 2 (T2D) que requer insulina alguns anos após o diagnóstico. Em qualquer uma das situações, um excesso ou défice de insulina causa um mau controlo dos níveis de glicose no sangue, resultando em algumas complicações de curto e longo prazo, entre elas, a hipoglicemia. Esta tese aborda o trabalho de investigação sobre algoritmos sensíveis ao contexto para a previsão de hipoglicemia baseada em dados discretos, fusão de dados e decisão de consenso de classificadores numa janela de 24h. Neste capítulo são apresentados:

1) enquadramento da tese, definindo o problema abordado;

2) os objetivos da investigação;

3) a definição do argumento;

4) as principais contribuições deste estudo, e por fim, as últimas secções do capítulo trazem de forma resumida;

5) as principais conclusões e perspectivas para investigação futura.

Enquadramento da Tese

De acordo com o relatório da federação internacional de diabetes [1], em 2021, aproximadamente 537 milhões de adultos em todo o mundo são afetados pelo DM, cuja prevalência é estimada em atingir 783 milhões de adultos em 2045. A hipoglicemia é a condição mais temida que pode surgir comumente em pacientes que necessitam de terapia com insulina. A queda do nível de glicose no sangue abaixo de 70 mg/dL ou 3.9 mmol/L [2, 3] caracteriza esta condição. Os sintomas de hipoglicemia podem levar à perda de consciência e, nos piores casos, podem causar a morte. Na maioria dos casos, a hipoglicemia chega sem aviso prévio e é essencialmente assintomática. Essa situação pode diminuir seriamente a qualidade de vida dos pacientes com diabetes, o que leva a sentimentos como insegurança e medo. A hipoglicemia pode ter vários desencadeantes, como por exemplo, uma overdose de insulina, a baixa ingestão de hidratos de carbono, o excesso de atividade física ou uso de alguns medicamentos [4]. Atualmente, a hipoglicemia pode ser classificada em diferentes grupos de acordo com a hora da ocorrência ou a um limiar que pode variar de acordo com as características

do paciente: idade, desenvolvimento de doenças e outros. Com base na hora da ocorrência, a hipoglicemia pode ser hipoglicemia diurna [5, 6, 7, 8, 9, 10, 11, 12, 13], hipoglicemia pós-prandial [14, 15, 16], hipoglicemia noturna [17, 18, 19, 20, 21] e pós-exercício [22]. Os riscos de hipoglicemia são menores para pacientes equipados com sistemas de monitorização contínua de glicose (CGM), às vezes usados em conjunto com uma bomba de insulina, denominada bomba aumentada por sensores. O número de pacientes a usar esse tipo de sistemas tem vindo a aumentar, mas alguns motivos levam à descontinuidade desses sistemas [19, 23]. Alguns pacientes continuam a usar sistemas de auto monitorização recorrendo ao seu glicosímetro e os métodos de previsão dos níveis de glicose com base na CGM não são extensíveis nesses casos. A necessidade de controlar os níveis de glicose oferece uma oportunidade para desenvolver novas soluções tecnológicas para autogestão e previsões. Essas soluções podem propor abordagens incorporando o conhecimento de especialistas com o lado do conhecimento do paciente. Os dados de um indivíduo mudam com diferentes ações e condições médicas, portanto, é necessário recolher dados reais para fornecer informações contextuais cruciais para aplicações num cenário de ambiente assistido. Os dados de contexto fornecem conhecimento individualizado [24] e permitem o desenvolvimento de melhores algoritmos preditivos. Modelos preditivos baseados em aprendizagem de máquina (*machine learning*) têm sido cada vez mais adotados para previsões de níveis de glicose [4, 25]. As características do *machine learning* tornam essas técnicas uma ótima ferramenta para prever eventos futuros. O foco desta tese inclui o projeto e desenvolvimento de um novo algoritmo sensível ao contexto para a previsão de hipoglicemia. Há várias questões de investigação em aberto envolvidas no desenvolvimento do algoritmo proposto, como se enunciam a seguir:

Questão 1 - Como converter os dados recolhidos em conhecimento de contexto útil?

Esta questão de investigação aborda a forma de recolher os dados, a heterogeneidade das fontes de informação, a fusão de dados e como converter esta informação em conhecimento de contexto útil e adequado para a previsão de hipoglicemia. Para responder a esta questão, foram identificados algoritmos preditivos existentes para hipoglicemia, a fim de entender o tipo de dados mais adequado para nossa abordagem.

Questão 2 - Quais as técnicas de pré-processamento mais adequadas para o nosso cenário?

Da solução esperada para a primeira questão, a fusão de dados adequados e a conversão de informações de contexto foram selecionados. A segunda questão de investigação aborda a questão da necessidade ou não de técnicas de pré-processamento para melhor preparação da matriz de variáveis de entrada para o algoritmo preditivo. Para responder a esta questão de investigação, foram estudados mais detalhadamente os algoritmos preditivos existentes em termos de técnicas de pré-processamento para entender quais os métodos que contribuem para a eficiência do algoritmo preditivo.

Questão 3 - Como preparar o alvo (variável de saída) para o processo de aprendizagem?

Em linha com a solução anterior, a matriz de variáveis de entrada foi selecionada para o algoritmo preditivo. A terceira questão de investigação aborda a escolha do alvo adequado para o processo de aprendizagem. Para responder a esta questão de investigação, foi definido o alvo em função da matriz de variáveis de entrada e da decisão final do algoritmo preditivo.

Questão 4 - Quais os classificadores mais adequados para o cenário proposto?

De acordo com a solução das questões anteriores, a matriz de variáveis de entrada e o alvo serão utilizados para o processo de aprendizagem para a descoberta de conhecimento e, usando esse conhecimento, prever eventos de hipoglicemia. Para responder a esta questão, foram explorados diferentes classificadores a fim de verificar quais obtêm uma boa eficiência.

Descrição do problema e objetivos de investigação

A DM é um grande problema de saúde a nível global: 783 milhões de adultos são estimados em 2045 [1]. Os cuidados de saúde com a diabetes é dos mais avançados no comprometimento do paciente devido à sua prevalência e devido ao papel crucial dos cuidados de saúde efetuados pelo próprio paciente. Isso torna os cuidados de saúde para a diabetes especialmente interessantes também com recurso à tecnologia. Uma gestão correta da diabetes concede a um paciente com diabetes alguma qualidade de vida, sem complicações. Infelizmente, um mau controlo dos níveis de glicose pode causar algumas complicações a curto e longo prazo. O problema a curto prazo mais indesejado é a hipoglicemia, definida como o estado que corresponde a um valor de glicose no sangue que está abaixo de 70 mg/dL. Pacientes com diabetes requerem um controlo muito rigoroso da sua doença. A recolha de dados é fundamental para esses pacientes, como por exemplo, níveis de glicose, terapia, conteúdo das refeições, hábitos de vida ou histórico de saúde. Existem alguns trabalhos na literatura que propõem algoritmos de previsão dos níveis de glicose usando dados de sistemas de CGM que permitem alarmes diminuindo a incidência de hipoglicemia e o tempo gasto quando ocorre este evento. No entanto, esses resultados não são extensíveis a pacientes que usam sistemas de monitorização recorrendo ao seu glicosímetro. A escassez de dados deste tipo de medição de níveis de glicose leva a dados ainda não explorados na literatura. Essas limitações combinadas com a falta de variáveis contextuais de acesso público, descumprimento dos pacientes com os protocolos do estudo e o uso de informações discretas em trabalhos preditivos, são desafios para soluções inovadoras e disruptivas. O problema abordado nesta tese é propor um algoritmo preditivo para hipoglicemia usando dados discretos e informações de contexto. A abordagem proposta pode ser embutida num equipamento de monitorização que, combinados com tecnologia móvel (por exemplo, uma aplicação móvel de autogestão) pode ajudar pacientes e cuidadores no controlo da diabetes e pode prevenir eventos de hipoglicemia. A maioria dos especialistas de saúde concorda que a automonitorização dos níveis de glicose realizada com o glicosímetro é essencial para controlar a hipoglicemia [26]. Dessa forma,

esta abordagem também estaria acessível a pacientes sem sistemas de CGM. No geral, esta tese teve como objetivo o desenho e desenvolvimento de um algoritmo sensível ao contexto para a previsão de hipoglicemia recorrendo a dados discretos e informações contextuais. Este objetivo principal foi transposto em quatro objetivos secundários:

1. Identificação dos algoritmos preditivos existentes para hipoglicemia.
O objetivo foi identificar os algoritmos preditivos existentes para hipoglicemia e identificar possíveis desafios em aberto. Para isso, foram estudados os dados e métodos utilizados na literatura para a previsão de hipoglicemia. Além disso, as suas contribuições foram analisadas e comparadas, e algumas lacunas e oportunidades de pesquisa foram identificadas para trazer o conhecimento adequado para nossa abordagem.

2. Desenho do algoritmo sensível ao contexto para previsão de hipoglicemia usando dados discretos.
O objetivo foi o desenho do algoritmo preditivo sensível ao contexto para hipoglicemia. Para isso, diferentes bases de dados foram identificadas e analisadas para definir qual a mais adequada para nossa abordagem. Além disso, os algoritmos preditivos existentes foram estudados em mais detalhes para identificar diferentes métodos e técnicas para o pré-processamento e processo de aprendizagem. A matriz de variáveis de entrada e o alvo foram selecionados para o processo de aprendizagem.

3. Desenvolvimento de algoritmo preditivo de hipoglicemia sensível ao contexto.
O objetivo foi desenvolver o algoritmo preditivo de hipoglicemia sensível ao contexto que compreende a implementação métodos de pré-processamento e processo de aprendizagem. Para o processo de aprendizagem, diferentes classificadores foram implementados.

4. Avaliação da fusão de dados e decisão multi-classificador.
O objetivo foi comparar o desempenho entre os modelos usando variáveis convencionais e os modelos usando fusão de contexto com as características convencionais. Também teve como objetivo a comparação de desempenho entre modelos individuais e modelos de multi-classificadores para mostrar qual abordagem executa melhor as previsões.

Argumento da tese

Esta tese propõe um algoritmo sensível ao contexto para previsão de hipoglicemia usando fusão de dados e uma decisão de consenso de classificadores, tendo como tese o seguinte:

Os algoritmos de previsão de níveis de glicemia, desenvolvidos a partir de dados de sistemas de monitorização contínuo de glicose, são amplamente estudados na literatura, mas esse tipo de solução não é acessível para todos os pacientes com diabetes e a hipoglicemia raramente é abordada como objetivo de previsão. Algoritmos de previsão baseados em informações discretas são um desafio, mas permitem o desenvolvimento de soluções de autogestão. Há diferentes informações de contexto adicionais que podem ser fusionadas com as variáveis convencionais para diminuir o efeito da inter e intra-variabilidade dos sujeitos nos padrões de previsão. Definindo um novo algoritmo sensível ao contexto, que compreende fusão de dados e fusão de modelos, contribui significativamente para a criação de algoritmos preditivos generalizados. Esta solução poderá ser incorporada em equipamentos de monitorização que, combinados com tecnologia móvel (por exemplo, uma aplicação de autogestão) poderia ajudar os pacientes e cuidadores no controlo da diabetes e poderia prevenir eventos de hipoglicemia.

Para apoiar esta tese, foi feita uma revisão sistemática de literatura, a fim de identificar e analisar os algoritmos preditivos existentes para a previsão de hipoglicemia. A partir dos métodos e técnicas identificados, um novo algoritmo sensível ao contexto foi projetado e desenvolvido para prever hipoglicemia usando dados discretos. O algoritmo foi aplicado usando dados de pacientes reais para avaliar a fusão de dados e a decisão de multiclassificadores.

Principais contribuições

A consecução dos quatro objetivos desta tese resultou nas seguintes contribuições principais:

1. Foi realizada uma revisão sistemática de literatura abordando os algoritmos existentes para previsão de hipoglicemia de modo a identificar os dados e métodos utilizados na literatura. Das descobertas desta revisão, foram identificadas algumas lacunas e oportunidades de forma a trazer o conhecimento adequado para nossa abordagem. Esta contribuição foi publicada em Artifcial Intelligence In Medicine (Elsevier, Q1, *Computer Science*):

V. Felizardo, N. M. Garcia, N. Pombo & I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review," in Artificial Intelligence In Medicine, 118 (2021). https://doi.org/10.1016/j.artmed.2021.102120.

2. Foi projetado, desenvolvido e testado um novo algoritmo sensível ao contexto para a previsão de hipoglicemia que compreende a fusão de dados e a aprendizagem supervisionada de vários classificadores. Os principais resultados da aplicação do algoritmo mostram a contribuição de integrar variáveis baseadas em contexto e no tempo com as variáveis conven-

cionais no aperfeiçoamento da previsão de hipoglicemia. Também, os resultados mostram que foi possível melhorar a taxa de acerto da previsão sem o aumento de falsos positivos usando a decisão de consenso de diferentes classificadores em vez de uma decisão de um classificador individual. Esta contribuição foi proposta em dois artigos científicos e uma patente:

V. Felizardo, D. Machado, N. M. Garcia, N. Pombo and P. Brandão, "Hypoglycaemia Prediction Models With Auto Explanation," in IEEE Access, vol. 10, pp. 57930-57941, 2022, doi: 10.1109/ACCESS.2021.3117340. (IEEE, Q1, *Computer Science*)

V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, & F. Babič, "A novel context-aware framework for Hypoglycaemia Events Prediction integrating context and time-based features," (em revisão, Q1, *Computer Science*).

V. Felizardo, N. M. Garcia, N. Pombo, & I. Megdiche, "Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos," 2021 (Submissão para Patente Portuguesa n.o 117301).

3. Foi avaliado o novo algoritmo sensível ao contexto para a previsão de hipoglicemia realizando previsões usando dados de pacientes reais. Para isso, foi aplicado o algoritmo, em que, diferentes modelos e diferentes combinações de classificadores foram desenvolvidos para fornecer previsões personalizadas para cada paciente. Os resultados mostram que a fusão de dados e informações pode dar uma boa contribuição para representar pacientes com diabetes em diferentes cenários e dar novos conhecimentos. O algoritmo preditivo oferece uma boa previsão do risco de hipoglicemia em 24 horas recorrendo a dados discretos. A contribuição foi submetida no formato de artigo científico para uma revista de Ciências da Computação (Q1, em revisão):

V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, M. Sousa & F. Babič, "Hypoglycaemia prediction using information fusion and classifiers consensus," (em revisão, Q1, *Computer Science*).

Estado da arte

No segundo capítulo da tese são apresentadas as principais conclusões da revisão sistemática de literatura acerca dos algoritmos existentes para a previsão de hipoglicemia dando ênfase às metodologias que recorrem a dados reais de pacientes com diabetes. A previsão de eventos de hipoglicemia desempenha um papel importante na gestão da diabetes, sendo capaz de reduzir o número de situações perigosas. Assim, é relevante apresentar uma revisão sistemática sobre os algoritmos e modelos de previsão atualmente disponíveis para a previsão de hipoglicemia. Este estudo teve como objetivo identificar, avaliar e analisar os algoritmos e modelos existentes que recorrem a dados reais de pacientes com diabetes para previsão de hipoglicemia. Assim, foram selecionados cinco bases de dados para análise de estudos publicados entre janeiro de 2014 a junho de 2020: ScienceDirect, IEEE Xplore, ACM Digi-

tal Library, SCOPUS e PubMed. Sessenta e três estudos elegíveis preencheram os critérios de inclusão. De acordo com os estudos selecionados, a hipoglicemia raramente é tratada como o objetivo da previsão, sendo que o alvo são os níveis de glicose, que são estimados através de métodos de regressão, e após a previsão, os valores geralmente inferiores a 70 mg/dl são categorizados com hipoglicemia. Essas abordagens refletem a introdução de sistemas CGM que permitem várias previsões de glicose a curto prazo. A maioria dos estudos usa os valores anteriores das medições de CGM como entrada para prever os valores futuros com base no perfil de glicose num determinado horizonte de previsão (ou janela temporal), que pode variar entre 5 minutos e 180 minutos. O horizonte de previsão mais utilizado é o de 30 minutos por ser considerado o intervalo de tempo mínimo para que um paciente tenha uma intervenção eficiente para evitar um evento indesejado [9]. Os resultados dos estudos mostram que um aumento do horizonte de previsão leva a uma deterioração na taxa de acerto da previsão [5, 27]. No entanto, alguns estudos propõem resolver o problema em termos de classificação em vez de regressão utilizando classificadores para detetar padrões. Além disso, existem abordagens que combinam variáveis como níveis de glicose, unidades de insulina, informação acerca das refeições e exercício físico para aprender e mapear variáveis de entrada em classes ou intervalos pré-estabelecidos. Então o problema de classificação foi abordado como uma previsão baseada em intervalos de níveis de glicose ou no risco de hipoglicemia/hiperglicemia. Alguns autores limitaram o estudo para uma previsão específica de hipoglicemia, por exemplo, noturna ou pós-prandial o que levou a uma previsão de médio prazo devido à redução da amostra de dados. Outra forma de resolver o problema de previsão é em termos de métodos que recorrem a dados de uma população de pacientes com diabetes ou apenas a dados do próprio paciente. Normalmente, os modelos baseados em dados do próprio paciente apresentam melhor desempenho do que os modelos baseados na população, mas são sensíveis a variabilidade intra-paciente. Por outro lado, os modelos que usam dados populacionais são adequados para a generalização da previsão. A seleção de variáveis de entrada adequadas, de acordo com o problema de previsão, é crucial para criar um modelo que funcione bem. Nesta revisão limitamos as variáveis de entrada para pelo menos uma destas categorias de dados: níveis de glicose no sangue (de monitorização contínua ou utilizando glicosímetro), informações sobre insulina, refeição ou exercício físico. A razão é que a maioria dos pacientes com diabetes são pacientes em regime ambulatório e, portanto, algumas variáveis como electrocardiograma (ECG), electroencefalograma (EEG), ou parâmetros clínicos, quando usados isoladamente, não são adequados para uma solução de autogestão. Através da revisão de literatura foram identificados trabalhos que combinam variáveis de diferentes fontes de dados, no entanto, existem alguns estudos que utilizam apenas os níveis de glicose devido à variabilidade em variáveis como refeição, medicação e exercício. Por isso destacamos alguns estudos que utilizam métodos de seleção de variáveis para realizar seus modelos [6, 28, 7, 29, 30, 31, 19, 32, 15, 33, 18, 10]. A predominância das variáveis baseadas nos níveis de glicose dos últimos 30 minutos é clara para janelas de previsão de 30 e 60 minutos [30, 31]. Além disso, a contribuição de outras variáveis (refeição, insulina, exercício) é menor, mas não insignificante, a sua importância aumenta para janelas de previsão de 60 minutos [31]. Zecchin *et al.* [32] mostram que, para previsões pós-prandiais

com janela de previsão acima de 15 minutos, a combinação dos valores de CGM com refeição e insulina alcança melhores resultados. Nesta revisão, destacamos ainda a integração do exercício e informações contextuais. Em relação ao exercício, Zarkogianni *et al.* [7] alcançou melhores resultados adicionando exercício às variáveis associadas aos níveis de glicose. A informação acerca da atividade física combinada com informação acerca de insulina e refeições pode aumentar o desempenho para janelas de previsão acima de 45 minutos [29]. Midroni *et al.* [34] alcançou os melhores resultados usando uma combinação de CGM com SMBG, refeição, estresse e exercício. Reddy *et al.* [22] mostram que a frequência cardíaca mais alta no início do exercício, o aumento do gasto energético e a diminuição dos níveis de glicose no início do exercício tende a aumentar a probabilidade de hipoglicemia [22]. As variáveis do exercício baseadas no tempo explicam melhor os níveis de glicose nas predições de curto prazo [31]. Este último estudo mostra que variáveis baseados no tempo podem ser relevantes para as predições. Além disso, o uso da hora do dia inclui algumas variáveis ligadas à dinâmica do metabolismo da glicose [31]. No entanto, em He *et al.* [33] os resultados mostram que a hora, refeição e insulina proporcionam mais correlações causais dominantes do que o sono e o exercício na inferência de glicemia. Com base em métodos de seleção, Dave *et al.* [10] mostra a importância de variáveis contextuais como a hora a que é feita a recolha de dados. Comparando com a revisão apresentada por Oviedo *et al.* [35] (de 2010 a 2016), nos modelos baseados em dados notamos o aumento de modelos baseado em ANN, nomeadamente, RNN bem como modelos baseados em aprendizagem supervisionada. A tendência em abordagens baseadas em dados de pacientes reais mostra que a comunidade científica continua a testar várias técnicas de *machine learning*. Alguns estudos começaram a usar aprendizagem profunda (*deep learning*) e algoritmos evolutivos. Em relação aos estudos que usam os mesmos conjuntos de dados, DirectNet [36], composto por dados de crianças e adolescentes, os modelos baseados em *deep learning* e ANN mostram melhores resultados do que o modelos probabilísticos/estatísticos. Os estudos realizados com o conjunto de dados OhioT1DM [37] alcança resultados semelhantes usando modelos baseados em *deep learning*, ANN, algoritmos evolutivos e conjuntos (*ensembles*). Em estudos usando o conjunto de dados METABO [38] os melhores resultados são alcançados pelos modelos híbridos, em que as variáveis extraídas usando um modelo fisiológico provavelmente contribui para um melhor desempenho dos modelos. Relativamente à previsão noturna, destacam-se os estudos [21, 32, 17] e para as predições pós-prandiais, os estudos [14, 32, 15]. Com o crescente número de métodos complexos que usam várias fontes de dados, a disponibilidade de algoritmos e conjuntos de dados pode ser muito útil para a comunidade científica que trabalha em previsão de níveis de glicose ou de eventos futuros. De acordo com a revisão sistemática da literatura desenvolvida identificamos três conjuntos de dados que estão disponíveis e de acesso público: 1) conjunto de dados OhioT1DM, 2) a base de dados DirectNet e 3) conjunto de dados UCI Diabetes. O conjunto de dados OhioT1DM permite uma perfeita replicabilidade de métodos sendo que está dividido em conjuntos de dados de treino e teste. A base de dados DirectNet tem oito estudos com conjuntos de dados públicos e adequados para este campo. No entanto, dos artigos selecionados que usam estes conjuntos de dados não especificam qual foi usado, ou quais pacientes são considerados. O conjunto de dados UCI Diabetes precisa de alguma ma-

nipulação de dados, portanto, pode ser mais difícil replicar a proposta dos artigos científicos. Em relação aos algoritmos, a maioria dos estudos apresenta uma descrição detalhada de seus métodos, incluindo fontes, parâmetros, algoritmos e bibliotecas. No entanto, poucos trabalhos [39, 40] partilham sua implementação de código-fonte, portanto, é recomendável solicitar permissão de acesso. Para concluir, notamos que desde 2017 foram publicados mais artigos nesta área sendo que este fenómeno poderá estar ligado ao primeiro desafio de previsão do nível de glicose promovido por um seminário de descoberta de conhecimento em dados de saúde. O segundo desafio de previsão do nível de glicose foi realizado em 2020, portanto, esperamos que o número de trabalhos cresça novamente. Em comparação com as revisões de Oviedo *et al.* [35] e de Woldaregay *et al.* [25] relativas aos períodos entre 2010 e 2016 e entre 2000 e 2018, destacamos que a tendência é adicionar novas variáveis de entrada aos modelos em particular variáveis associadas ao exercício. Na presente revisão destacamos a integração de informações contextuais e características temporais (características definidas em relação ao tempo [31]). A disponibilidade de dispositivos de monitorização como *smartphone*, *smartwatch*, pulseira, e suas aplicações contribuem para uma maior disponibilidade de dados fisiológicos e contextuais. A revisão de Oviedo *et al.* [35] concluiu que a maioria dos estudos carecia de evidência clínica porque os modelos são validados com dados simulados. A nossa revisão incluiu 63 estudos, todos eles usando dados reais de pacientes para validação dos modelos. A limitação apontada pelas revisões de Oviedo *et al.* [35] e de Woldaregay *et al.* [25] é a baixa disponibilidade de dados. Em relação à replicabilidade dos estudos, o conjunto de dados OhioT1DM permite uma perfeita replicabilidade dos métodos razão pela qual este conjunto de dados foi sugerido para o desafio de previsão do nível de glicose no sangue.

Algoritmo preditivo de eventos de hipoglicemia

No terceiro capítulo da tese são apresentadas as principais conclusões da proposta do algoritmo para a previsão de eventos de hipoglicemia dando ênfase à estrutura e técnicas utilizadas. Ainda, neste capítulo, são apresentados alguns resultados que visam mostrar a contribuição de integrar variáveis baseadas em contexto e no tempo com as variáveis convencionais e de usar a decisão de consenso de diferentes classificadores em vez de uma decisão de um classificador individual. A abordagem proposta apoia o desenvolvimento de soluções preditivas abordando a heterogeneidade desencadeante de eventos de hipoglicemia num cenário de autogestão. O desenho experimental apresentado mostra o passo a passo do algoritmo de previsão de hipoglicemia. As variáveis e dimensões de contexto podem descrever adequadamente pacientes com diabetes com diferentes características e comportamentos, o que é importante, pois permite a generalização dos modelos num cenário de autogestão. As informações de contexto auxiliam no reconhecimento de padrões distintos associados ao risco de hipoglicemia e fornecem informações para a tomada de decisão numa janela de 24h. Os resultados sobre o treino de modelos individuais mostram uma precisão acima de 80% suportando que foi possível realizar previsões usando dados discretos. A literatura mostra que é comum o uso de informações complementares como unidades de insulina, refeição e atividade física em combinação com níveis de glicose para previsão dos níveis de glicose. No entanto, algumas dessas

características apresentam dinâmicas diferentes para cada pessoa. Além disso, na mesma pessoa, há mudanças ao longo do tempo. Assim pretendemos a integração de informações contextuais para diminuir o efeito da variabilidade inter e intra-sujeitos nos padrões de previsão. Adicionamos a hora do dia, os comportamentos das pessoas relacionados à refeição e à atividade física e a variabilidade da glicose baseada no tempo. A dimensão do tempo é muito importante na criação de soluções para pacientes com diabetes, pois é a que mais afeta as variáveis na inter e intra-variabilidade. Determinar a ação do paciente num momento específico afeta diretamente o desempenho para um cenário específico. A dimensão do tempo conecta operativamente as informações das outras dimensões do contexto. Por exemplo, a dimensão comportamento qualifica a refeição e a atividade física, mas o maior impacto na previsão é quando conectado com outra dimensão de contexto ou variáveis como unidades de insulina ou nível de glicose no sangue. Neste trabalho, a maior contribuição da realização de modelos preditivos são as variáveis baseadas no tempo, ou seja, a variabilidade da glicose baseada no tempo. Essas novas variáveis mostram uma boa previsibilidade para eventos de hipoglicemia, fornecendo novos conhecimentos sobre as alterações do perfil glicémico em função do contexto. A integração de variáveis contextuais e temporais permite o desenvolvimento de um maior número de soluções adequadas a uma intervenção (taxa de acerto superior a 60% e falsos positivos inferior a 30%) melhorando a taxa de acerto e a previsão de hipoglicemia. Estas experiências mostram que 76% destas soluções atingiram a previsão de mais de 70% dos eventos de hipoglicemia. Além disso, essas soluções mostram melhores previsões para 58.5% dos pacientes quando comparadas com as previsões usando variáveis convencionais. Estes resultados sugerem que as informações contextuais e variáveis baseados no tempo melhoraram as previsões e permitiram o desenvolvimento de melhores soluções para a maioria dos pacientes usados neste estudo. O contexto de pacientes com diabetes é desafiante de monitorizar e embora a tecnologia móvel e vestível esteja mais acessível, tal como soluções comerciais já no mercado e nas lojas de aplicações móveis, ainda não há um uso generalizado delas, e a disponibilidade de seus dados geralmente não é público. Assim, há a necessidade de criar redes de dispositivos de monitorização que possam fornecer informações contextuais em combinação com dados de glicosímetros pessoais. Estamos cientes de que também a usabilidade deste tipo de solução pode estar comprometida, pois os utilizadores deixariam de usá-la se estivessem sobrecarregados com previsões irrelevantes devido à alta probabilidade de alarmes falsos. Os resultados da validação para cada paciente mostram que a previsão melhorou com a integração de variáveis contextuais e baseados no tempo, mas nenhum modelo individual permitiu uma boa solução para todos os pacientes e, em alguns casos, nenhuma solução única foi apropriada para um paciente. Assim, para mitigar o problema, adicionamos ao nosso algoritmo uma decisão consensual de diferentes modelos individuais para melhorar a taxa de acerto sem o aumento de alarmes falsos. Esperávamos que com uma decisão consensual de classificadores fosse possível diminuir mais o efeito da inter e intra-variabilidade e apresentar aos pacientes uma melhor solução personalizada captando padrões apropriados associados ao risco de hipoglicemia. Para este trabalho, usamos apenas a combinação de três ou cinco modelos para a decisão de consenso. Os resultados dos testes mostram que esta abordagem alcançou uma melhor taxa de acerto, diminuição de falsos positivos e aumento na

previsão de eventos de hipoglicemia. Além disso, a decisão de consenso melhorou a previsão para 57.4% dos pacientes oferecendo a melhor solução. A proposta de soluções personalizadas é uma parte importante da nossa abordagem. A criação de bases de dados com dados de uma população de pacientes com diabetes representando os padrões anteriores de alguém associados ao risco de hipoglicemia traz outro desafio potencial de uma generalização ou de personalização dos modelos. Portanto, o comprometimento do utilizador é muito importante, pois ele deve estar ciente de que quanto mais dados fornecer ao algoritmo, mais precisa será esta abordagem.

Avaliação do algoritmo preditivo

No quarto capítulo da tese são apresentadas as principais conclusões da avaliação do algoritmo preditivo dando ênfase aos resultados da aplicação do método e de diferentes combinações de classificadores para a previsão personalizada de hipoglicemia. O principal objetivo deste trabalho foi mostrar a eficiência da previsão de hipoglicemia em um período de 24 horas, usando dados discretos, fusão de informações e decisão de consenso de classificadores. A nossa abordagem apoia o desenvolvimento de soluções preditivas abordando a fusão de dados e informações para viabilizar seu uso em um cenário de autogestão em que os diabéticos são pacientes em regime ambulatório realizando a monitorização da sua doença. A hipoglicemia pode ter vários desencadeantes, por exemplo, uma overdose de insulina, a baixa ingestão de hidratos de carbono, o excesso de atividade física ou uso de alguns medicamentos. Esta é uma das razões para o uso da fusão de dados que permitirá descrever os vários cenários de um paciente com diabetes. No nosso trabalho anterior [41], mostramos que variáveis e dimensões de contexto, também usadas neste capítulo, podem descrever diferentes características e comportamentos de pacientes com diabetes e melhorar a taxa de acerto da previsão. Isso é importante para a generalização dos modelos e para a captura de vários padrões associados ao risco de hipoglicemia. Além disso, o algoritmo permite a previsão de hipoglicemia com fusão sequencial de informações. Primeiramente, os dados recolhidos foram utilizados para extração de variáveis onde foi realizada uma fusão de variáveis e de dimensões de contexto. A fusão da dimensão tempo com a glicemia fornece um novo conhecimento sobre as alterações do perfil glicémico em função do tempo. Em segundo lugar, os métodos de imputação e reamostragem também contribuem para a adição de informações. Terceiro, usamos multi-classificadores que é um meio promissor para fundir informações de dados influenciados por diferentes perfis, inter e intra-variabilidade e capturando padrões apropriados associados à previsão de hipoglicemia. A fusão de dados de pacientes com diabetes pode ser um desafio, mas hoje em dia, algumas tecnologias móveis e vestíveis permitem monitorizar diferentes fontes de informação. No entanto, há a necessidade de criar redes de dispositivos de monitorização que combinem dados fisiológicos ou clínicos com diferentes dimensões de contexto. A nossa abordagem usa a fusão de informações de multi-modelos para dar uma decisão de consenso dos classificadores para prever o risco de hipoglicemia para um paciente específico. Considerando que os eventos de hipoglicemia são comumente temidos em pacientes com insulinoterapia e que podem causar perda de consciência do paciente, privilegiamos soluções

que tenham um maior valor de eventos previstos. Para realizar a avaliação do nosso algoritmo foram usados dados de 54 pacientes, sendo que, para todos eles apresentamos uma previsão personalizada. Para cada paciente são treinados seis modelos: árvore de decisão (CART), *support vector machine* (SVM), KNN, *Adaboost Tree*, *Bagged Tree*, e *Subspace* KNN (SKNN). Assim, as combinações de três ou cinco classificadores mostram resultados com mais de 80% das predições de hipoglicemia. No entanto, sabemos que este tipo de solução necessita de aceitação por parte dos utilizadores, pelo que é crucial considerar os falsos alarmes. O resultado médio de falsos alarmes mostra que as combinações de três ou cinco classificadores atingiram falsos alarmes entre 25-30%, uma combinação de quatro e seis classificadores valores percentuais entre 20-25% e combinações de dois classificadores entre 15-20%. Embora os resultados para a combinação de dois classificadores mostrem menos falsos alarmes, sua capacidade preditiva também é menor. Para esses testes, a decisão de consenso alcançou uma solução melhor para 81.5% dos pacientes. Apenas para dez pacientes (P7, P13, P14, P17, P24, P25, P27, P30, P35 e P39) a melhor solução foi obtida com um classificador individual. No entanto, não observamos nenhum padrão significativo entre as combinações que fornecem melhores soluções. Isso é explicado pela inter e intra-variabilidade dos pacientes. A maior percentagem de melhor solução dos pacientes foi alcançada com combinações de dois e três classificadores. Como mencionamos anteriormente, a solução que oferece melhores resultados varia de paciente para paciente, mas destacamos três combinações que proporcionam melhores resultados para quatro pacientes cada. A combinação C2.2 (CART+KNN) fornece melhores resultados para P6, P10, P40 e P43; a combinação C2.5 (CART+SKNN) para P4, P6, P37 e P70; e combinação C2.12 (KNN+SKNN) para P19, P34, P36 e P46. A partir desses resultados, podemos concluir que os *ensembles* baseados em KNN apresentam algumas potencialidades para alguns *clusters* de pacientes. Além disso, considerando as soluções que usam a decisão de consenso e atingiram uma previsão de 100% dos eventos de hipoglicemia, a maioria das soluções combinou classificadores baseados em árvores de decisão e classificadores baseados em KNN. A previsão de hipoglicemia é feita pela decisão dos multimodelos em vez da decisão do modelo individual. Em relação aos melhores resultados com um modelo individual, a solução atingiu mais de 90% da hipoglicemia prevista para 54.7% dos pacientes (30.2% com 100% da hipoglicemia prevista). Usando a decisão de consenso dos classificadores, 66% dos pacientes têm mais de 90% da hipoglicemia prevista (37.7% com 100% da hipoglicemia prevista). Considerando ambas as abordagens, a hipoglicemia prevista foi alcançada com mais de 90% para 66% das melhores soluções dos pacientes (41.5% das melhores soluções dos pacientes com 100% da hipoglicemia prevista). Destas soluções, 69.8% são baseadas em KNN. Considerando os resultados com melhor classificador individual, os resultados mostram que os falsos alarmes estão entre 20-30% para 56.6% dos pacientes. Utilizando o consenso dos classificadores, o intervalo de falsos alarmes 20-30% é o de maior representatividade, 62.3% dos pacientes. Considerando ambas as abordagens, para 63% dos pacientes, os falsos alarmes estão entre 20-30%. Além disso, os achados mostram que nosso método apresenta melhor desempenho em pacientes com recolhas de dados em intervalos de tempo regulares e apresentam menos alarmes falsos. Para concluir, este trabalho mostra a importância da fusão de dados e decisão consensual para capturar os padrões associados ao

risco de hipoglicemia e sua previsão. A fusão de dados de diferentes fontes e contextos pode representar pacientes com diabetes em diferentes cenários. Uma fusão sequencial de informações pode ajudar a entender a complexidade desse problema e a decisão de consenso dos classificadores pode permitir a proposta de uma melhor solução para um paciente específico, diminuindo os efeitos da inter e intra-variabilidade.

Principais conclusões

No quinto capítulo da tese são apresentadas as principais conclusões. Esta tese propõe um novo algoritmo contextual para a previsão de hipoglicemia, que compreende a fusão de dados e a decisão de consenso de classificadores em um período de 24 horas. A literatura mostra uma abordagem extensivamente estudada para a previsão de glicemia usando dados de sistemas de CGM, mas essas abordagens são inacessíveis para pacientes que fazem a sua auto-gestão da diabetes. O desenvolvimento de soluções de autogestão com propósitos de previsão baseados em informações discretas é um desafio. No entanto, a integração de informações adicionais de contexto com as variáveis convencionais pode contribuir de forma complementar, reduzindo a variabilidade inter e intra-paciente. A definição de um algoritmo de previsão que compreende a fusão de dados e decisão de multiclassificadores contribui significativamente para a criação de soluções adequadas e generalizadas para a prevenção da hipoglicemia de forma a auxiliar os pacientes e cuidadores nos cuidados de saúde com a diabetes. Para apoiar esta tese, quatro objetivos principais foram alcançados: **1)** Identificação dos algoritmos preditivos existentes para hipoglicemia; **2)** Desenho do algoritmo sensível ao contexto para previsão de hipoglicemia usando dados discretos; **3)** Desenvolvimento de algoritmo preditivo de hipoglicemia sensível ao contexto,e, **4)** Avaliação da fusão de dados e decisão multi-classificador. A seguir, estão resumidas as contribuições do trabalho de investigação por objetivo.

O primeiro objetivo apresentado no capítulo 2 consiste numa revisão sistemática da literatura abordando os modelos baseados em dados existentes utilizando dados reais para previsão de hipoglicemia. Esta revisão analisa sessenta e três estudos em que os modelos foram abordados massivamente, como a previsão de glicose no sangue a curto prazo apresentada em um horizonte de previsão entre 5 minutos e 180 minutos de acordo com a hora do dia, por exemplo, previsões diurnas, previsões noturnas ou pós-prandiais. Os modelos preditivos recorrem a diferentes variáveis de entrada como dados relativos à refeição, insulina e atividade física combinados com dados de glicose, mas alguns estudos usam apenas dados de glicose devido à inter-variabilidade nas refeições do dia, medicação e exercício. Os estudos utilizaram diferentes métodos e técnicas apresentando bons resultados preditivos. Existem algumas tendências relacionadas aos conjuntos de dados usados, métodos e introdução de novas variáveis tal como exercício e dados de contexto.

O segundo e terceiro objetivos, apresentados no capítulo 3, consistem no desenho e desenvolvimento do algoritmo sensivel ao contexto para a previsão de hipoglicemia. O algoritmo integra fusão de dados e decisão de multiclassificadores, fornecendo uma boa previsão do risco de hipoglicemia nas próximas 24 horas suportado por dados discretos. Com o cumprimento

desses objetivos, mostramos que a integração de variáveis contextuais e temporais com as variáveis convencionais melhora a precisão e o número de soluções adequadas (precisão superior a 60% e alarmes falsos inferiores a 30%). Além disso, a decisão de consenso melhora o desempenho das previsões oferecendo boas soluções para 57.4% dos pacientes.

O quarto objetivo foi apresentado no capítulo 4, onde se pretende avaliar o algoritmo utilizando dados de pacientes reais. O algoritmo preditivo oferece uma boa previsão do risco de hipoglicemia em um período de 24 horas usando dados discretos. A decisão de consenso dos classificadores utilizando fusão de dados mostrou potencialidades para as predições oferecendo soluções adequadas para 81.5% dos pacientes atingindo mais de 90% das hipoglicemias previstas para 66% dos pacientes.

Conforme apresentado na Introdução, existem várias questões de investigação em aberto envolvidas no desenvolvimento do algoritmo proposto. A seguir, discutimos como abordamos e respondemos às perguntas apresentadas na Introdução: **1)** Questão 1 - Como converter os dados recolhidos em conhecimento de contexto útil?; **2)** Questão 2 - Quais as técnicas de pré-processamento mais adequadas para o nosso cenário?; **3)** Questão 3 - Como preparar o alvo (variável de saída) para o processo de aprendizagem?, e, **4)** Questão 4 - Quais classificadores são mais adequados para o cenário proposto?.

Respondemos à Questão 1 no Capítulo 2, onde analisamos as variáveis de entrada de cada estudo identificado na revisão sistemática da literatura. Os algoritmos preditivos são baseados principalmente em dados de sistemas de CGM usados de forma isolada ou outros estudos usam esses dados em combinação com informações sobre refeições, terapia e exercício físico. Para nossa abordagem, com foco numa solução de autogestão da diabetes, optamos pela recolha de dados discretos por meio de uma aplicação móvel. Portanto, a fonte de dados é semelhante à encontrada na literatura. No Capítulo 3, o desenho do algoritmo é apresentado detalhadamente, sendo a principal contribuição a proposta de variáveis de entrada contextuais. Como os eventos de hipoglicemia podem ter diferentes desencadeantes, a informação contextual pode ser útil para definir e apoiar o processo de decisão e é dada por três dimensões de contexto, como 1) tempo, 2) comportamento e 3) variáveis baseadas no tempo. A dimensão do tempo é dada pela variável hora do dia ou hora do registo, permitindo fornecer informações sobre quando algumas ações foram realizadas. A dimensão comportamento é dada por informações sobre a refeição ou atividade física e permite responder como essas ações foram realizadas. As variáveis baseados no tempo são a ligação entre a variável tempo e o nível de glicose, dando uma noção de como o nível de glicose varia ao longo do tempo. Assim, o contexto funciona como uma informação complementar para fazer uma ponte entre os dados discretos. Os resultados apresentados no capítulo 3 mostram a contribuição da fusão de informações contextuais com as variáveis convencionais como nível de glicose e terapia com insulina para as previsões de hipoglicemia.

Para responder à Questão 2, apresentamos no capítulo 3 o algoritmo de previsão de hipoglicemia que compreende duas etapas de pré-processamento: 1) Imputação e 2) Reamostragem.

A imputação de dados é necessária devido à impossibilidade de substituir os dados em falta de glicemia por zero. A outra opção pode ser excluir a observação com o valor ausente, mas isso pode resultar na perda de informações relevantes. Como os eventos de hipoglicemia são sempre uma classe minoritária em comparação com a totalidade dos dados, os métodos de reamostragem são cruciais para dados não balanceados. Assim, para nossa abordagem, optamos pela técnica SMOTE. Essas opções não são novas, mas mostram que quando propomos um novo algoritmo, é fundamental focar no problema e tentar sintetizar informações e conhecimentos para chegar a uma solução.

Para responder à questão 3, no capítulo 3, o *target* (*output*) para o processo de aprendizagem possui três classes: sem risco (classe 0), risco de hipoglicemia (classe 1) e evento de hipoglicemia (classe 2). A classe de risco é o risco de ter um evento de hipoglicemia nas próximas 24 horas ou um valor de glicose inferior a 75 mg/dl. Assim, para as observações com valor inferior a 70 mg/dl consideramos todas as observações das últimas 24h como classe 1. Além disso, para as observações com valor menor ou igual a 75 mg/dl consideramos todas as observações das últimas 24h como classe 1 para mitigar o risco de um possível evento de hipoglicemia.

Para responder à questão 4, no Capítulo 3, propusemos uma decisão consensual de diferentes modelos individuais para mitigar o fato de que nenhum modelo individual permite uma boa solução para todos os pacientes. Os resultados apresentados no capítulo 3 mostram que esta abordagem alcançou melhor precisão, falsos positivos e hipoglicemias previstas. No capítulo 4, apresentamos a avaliação do algoritmo usando dados reais de pacientes em que a decisão de consenso oferece uma boa previsão de risco de hipoglicemia em um período de 24 horas usando dados discretos.

Direções para trabalho futuro

Além das ideias gerais para trabalhos futuros na previsão de hipoglicemia sensível ao contexto, descrevemos várias linhas de investigação que surgem do desenvolvimento desta tese e que correspondem a perspetivas de trabalho futuro. Organizamos essas ideias em torno da fusão de dados, fusão de modelos e interpretabilidade do modelo.

Nesta tese, quando identificamos vários algoritmos preditivos existentes para hipoglicemia, destacamos a falta de conjuntos de dados acessiveis e de acesso público. Primeiro, o desenvolvimento e validação de algoritmos beneficiariam de conjuntos de dados públicos. Em segundo lugar, os registros de informações contextuais são de grande interesse, pois fornecem uma melhor perspectiva sobre o perfil e o comportamento do paciente. Terceiro, o conjunto de dados da UCI Diabetes continha uma população com eventos de hipoglicemia frequentes. Seria interessante testar o algoritmo numa população com menos eventos. Quarto, seria interessante explorar diferentes fontes de dados para definir diferentes dimensões de contexto e variáveis baseadas no tempo.

No capítulo 4, o algoritmo de previsão foi aplicado fornecendo previsões personalizadas para cada paciente. Embora tenham sido obtidos resultados promissores, devemos reduzir ainda mais os falsos alarmes e, assim, melhorar a precisão. Primeiro, existem diferentes classificadores que não foram testados. Em segundo lugar, os *ensembles* baseados em KNN apresentaram algumas potencialidades para alguns grupos de pacientes, a aplicação de técnicas de *clustering* pode ser interessante. Terceiro, integrar dados anteriores do paciente em conjunto com os dados populacionais no conjunto de dados de treino.

O objetivo desta tese não foi dar interpretabilidade ao algoritmo desenvolvido, no entanto, é importante enfatizar a importância desta caracteristica para esta abordagem. Em primeiro lugar, seria interessante explorar modelos personalizados considerando redes de similaridade de pacientes. Em segundo lugar, explorar e implementar um sistema de recomendação para fornecer informações adequadas ao paciente e aos cuidadores.

Referências

Bibliography

[1] I. D. Federation, "IDF Diabetes Atlas 10th edition 2021," Tech. Rep., 2021. [Online]. Available: https://diabetesatlas.org/atlas/tenth-edition/ xv, xvii

[2] NICE UK, "Diabetes (type 1 and type 2) in children and young people: diagnosis and management," National Institute for Health and Care Excellence (NICE) guideline, Tech. Rep., 2020. [Online]. Available: https://www.nice.org.uk/guidance/ng18 xv

[3] ——, "Type 1 diabetes in adults: diagnosis and management," National Institute for Health and Care Excellence (NICE) guideline, Tech. Rep., 2020. [Online]. Available: https://www.nice.org.uk/guidance/ng17 xv

[4] O. Mujahid, I. Contreras, and J. Vehi, "Machine learning techniques for hypoglycemia prediction: Trends and challenges," 2021. xv, xvi

[5] R. H. Botwey, E. Daskalaki, P. Diem, and S. G. Mougiakakou, "Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014, pp. 4843–4846, 2014. xvi, xxi

[6] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, "A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management," Modern Artifical Intelligence for Health Analytics, pp. 35–39, 2014. [Online]. Available: https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/viewFile/8737/8308 xvi, xxi

[7] Zarkogianni, K., Mitsis, K., Litsa, E., Arredondo, M. T., Fico, G., Fioravanti, A., and Nikita, K. S., "Comparative assessment of glucose prediction models for patients with

type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring," Medical and Biological Engineering and Computing, vol. 53, no. 12, pp. 1333–1343, 2015. xvi, xxi, xxii

[8] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A Deep Learning Approach to Diabetic Blood Glucose Prediction," Frontiers in Applied Mathematics and Statistics, vol. 3, no. July, pp. 1–11, 2017. xvi

[9] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting learning to accurately predict blood glucose trajectories," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1387–1395, 2018. xvi, xxi

[10] D. Dave, D. J. DeSalvo, B. Haridas, S. McKay, A. Shenoy, C. J. Koh, M. Lawley, and M. Erraguntla, "Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction," Journal of Diabetes Science and Technology, 2020. xvi, xxi, xxii

[11] X. Yu, K. Turksoy, M. Rashid, J. Feng, N. Hobbs, I. Hajizadeh, S. Samadi, M. Sevil, C. Lazaro, Z. Maloney, E. Littlejohn, L. Quinn, and A. Cinar, "Model-fusion-based online glucose concentration predictions in people with type 1 diabetes," Control Engineering Practice, vol. 71, no. April 2017, pp. 129–141, 2018. [Online]. Available: https://doi.org/10.1016/j.conengprac.2017.10.013 xvi

[12] J. Yang, L. Li, Y. Shi, and X. Xie, "An ARIMA Model with Adaptive Orders for Predicting Blood Glucose Concentrations and Hypoglycemia," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1251–1260, 2019. xvi

[13] M. Gadaleta, A. Facchinetti, E. Grisan, and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 650–659, 2019. xvi

[14] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, "Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning," Health Informatics Journal, 2019. xvi, xxii

[15] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," International Journal of Medical Informatics, vol. 126, no. March, pp. 1–8, 2019. [Online]. Available: https://doi.org/10.1016/j.ijmedinf.2019.03.008 xvi, xxi, xxii

[16] F. Dubosson, N. Mordvanyuk, B. López Ibáñez, and M. Schumacher, "Negative results for the prediction of postprandial hypoglycemias from insulin intakes and carbohydrates: analysis and comparison with simulated data," pp. 25–29, 2017. [Online]. Available: https://dugi-doc.udg.edu/handle/10256/14859 xvi

[17] A. Bertachi, L. Biagi, N. Luo, and J. Veh, "Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks," in KHD@ IJCAI, 2018, pp. 85–90. xvi, xxii

[18] M. H. Jensen, C. Dethlefsen, P. Vestergaard, and O. Hejlesen, "Prediction of Nocturnal Hypoglycemia From Continuous Glucose Monitoring Data in People With Type 1 Diabetes: A Proof-of-Concept Study," Journal of Diabetes Science and Technology, vol. 14, no. 2, pp. 250–256, 2019. xvi, xxi

[19] P. Tkachenko, G. Kriukova, M. Aleksandrova, O. Chertov, E. Renard, and S. V. Pereverzyev, "Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application," Computer Methods and Programs in Biomedicine, vol. 134, pp. 179–186, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2016.07.003 xvi, xxi

[20] G. Kriukova, N. Shvai, and S. V. Pereverzyev, "Application of regularized ranking and collaborative filtering in predictive alarm algorithm for nocturnal hypoglycemia prevention," in 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), vol. 2, 2017, pp. 634–638. xvi

[21] F. Stahl, R. Johansson, and M. L. Olsson, "Predicting Nocturnal Hypoglycemia Using a Non-parametric Insulin Action Model," Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015, pp. 1583–1588, 2016. xvi, xxii

[22] R. Reddy, N. Resalat, L. M. Wilson, J. R. Castle, J. El Youssef, and P. G. Jacobs, "Prediction of Hypoglycemia During Aerobic Exercise in Adults With Type 1 Diabetes," Journal of Diabetes Science and Technology, vol. 13, no. 5, pp. 919–927, 2019. xvi, xxii

[23] D. Rodbard, "Continuous Glucose Monitoring: A Review of Recent Studies Demonstrating Improved Glycemic Outcomes," Diabetes Technology & Therapeutics, 2017. xvi

[24] A. R. M. Forkan and W. Hu, "A context-aware, predictive and protective approach for wellness monitoring of cardiac patients," 2016 Computing in Cardiology Conference (CinC), vol. 43, pp. 369–372, 2016. xvi

[25] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Botsis, and G. Hartvigsen, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," 2019. xvi, xxiii

[26] E. M. Benjamin, "Self-Monitoring of Blood Glucose: The Basics," Clinical Diabetes, 2002. xvii

[27] A. Aliberti, I. Pupillo, S. Terna, E. Macii, S. Di Cataldo, E. Patti, and A. Acquaviva, "A Multi-Patient Data-Driven Approach to Blood Glucose Prediction," IEEE Access, vol. 7, pp. 69 311–69 325, 2019. xxi

[28] J. I. Hidalgo, J. M. Colmenar, G. Kronberger, S. M. Winkler, O. Garnica, and J. Lanchares, "Data Based Prediction of Blood Glucose Concentrations Using Evolutionary Methods," Journal of Medical Systems, vol. 41, no. 9, 2017. xxi

[29] S. Faccioli, B. Ozaslan, J. F. Garcia-Tirado, M. Breton, and S. Del Favero, "Black-box Model Identification of Physical Activity in Type-l Diabetes Patients," in 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 3910–3913. xxi, xxii

[30] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, "Online prediction of glucose concentration in type 1 diabetes using extreme learning machines," in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015, pp. 3262–3265. xxi

[31] ——, "Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models," Medical and Biological Engineering and Computing, vol. 53, no. 12, pp. 1305–1318, 2015. [Online]. Available: http://dx.doi.org/10.1007/s11517-015-1263-1 xxi, xxii, xxiii

[32] C. Zecchin, A. Facchinetti, G. Sparacino, and C. Cobelli, "How much is short-term glucose prediction in type 1 diabetes improved by adding insulin delivery and meal content information to CGM data? A proof-of-concept study," Journal of Diabetes Science and Technology, vol. 10, no. 5, pp. 1149–1160, 2016. xxi, xxii

[33] M. He, W. Gu, Y. Kong, L. Zhang, C. J. Spanos, and K. M. Mosalam, "CausalBG: Causal Recurrent Neural Network for the Blood Glucose Inference with IoT Platform," IEEE Internet of Things Journal, p. 1, 2019. xxi, xxii

[34] C. Midroni, P. J. Leimbigler, G. Baruah, M. Kolla, A. J. Whitehead, and Y. Fossat, "Predicting glycemia in type 1 diabetes patients: Experiments with XGBoost," CEUR Workshop Proceedings, vol. 2148, pp. 79–84, 2018. xxii

[35] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for t1dm patients," International Journal for Numerical Methods in Biomedical Engineering, vol. 33, no. 6, pp. 1–21, 2017. xxii, xxiii

[36] JAEB, "Diabetes Research in Children Network (DirecNet)," accessed: April 27, 2021. [Online]. Available: https://public.jaeb.org/direcnet xxii

[37] C. Marling and R. C. Bunescu, "The ohiot1dm dataset for blood glucose level prediction," in KHD@ IJCAI, 2018. [Online]. Available: http://smarthealth.cs.ohio.edu/OhioT1DM-dataset.html xxii

[38] E. Georga, V. Protopappas, A. Guillen, G. Fico, D. Andigo, M. T. Arredondo, T. P. Exarchos, D. Polyzos, and D. I. Fotiadis, "Data mining for blood glucose prediction and knowledge discovery in diabetic patients: The METABO diabetes modeling and management system," in Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, EMBC 2009, 2009. xxii

[39] S. Mirshekarian, H. Shen, R. Bunescu, and C. Marling, "LSTMs and Neural Attention Models for Blood Glucose Prediction: Comparative Experiments on Real and Synthetic Data," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 706–712. xxiii

[40] J. Martinsson, A. Schliep, B. Eliasson, and O. Mogren, "Blood Glucose Prediction with Variance Estimation Using Recurrent Neural Networks," Journal of Healthcare Informatics Research, vol. 4, no. 1, pp. 1–18, 2020. xxiii

[41] V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, and F. Babič, "A novel context-aware framework for Hypoglycaemia Events Prediction," Under review for publication and patent being submitted, 2022. xxv

Abstract

The need to control diabetes provides an opportunity to develop new technological solutions for self-management and real predictions. These predictions can be useful in preventing unwanted events, such as hypoglycaemia. The patients with diabetes type 1 and some patients with diabetes type 2 commonly fear hypoglycaemia. The aim of this thesis is to develop a context-aware framework for hypoglycaemia prediction using sparse data, information fusion and classifiers' consensus decision in a 24h hour time frame. With this approach, we contribute by proposing a hypoglycaemia prediction algorithm, in a self-management scenario, allowing it to be used by patients who perform their monitorization using a glucometer. The literature proposes glycaemia prediction algorithms using data from Continuous Glucose Monitoring (CGM) systems, but these approaches are not extensible to patients without these systems. Prediction algorithms based on discrete information are a challenge, so we proposed a novel context-aware framework for hypoglycaemia prediction based on data fusion and classifiers' consensus decision. The fusion of additional context information with the conventional features can contribute to decrease the effect of inter- and intra-subject variability on prediction patterns. Also, the prediction decision based on classifiers' consensus can contribute to the creation of suitable and generalised predictive algorithms. Integrating contextual and time-based features improves the accuracy on predicted hypoglycaemia. Using the classifiers' consensus decision, 66% of the researched patients have over 90% of hypoglycaemia predicted (with 37.7% with 100% of hypoglycaemia predicted), without the increase of false positives (false alarms). This work shows the importance of data fusion and consensus decision to handle the patterns associated with hypoglycaemia risk and its prediction, however, further research is necessary to provide the necessary interpretability to the predictive models.

Keywords

Hypoglycaemia, Prediction, Context-aware framework, data fusion, time-based features, classifiers consensus decision.

Contents

List of Figures

List of Tables

Acronyms

AAL	Ambient Assisted Living
ACM	Association for Computing Machinery
ANN	Artificial Neural Network
ALLab	Assisted Living Computing and Telecommunications Laboratory
BG	Blood glucose
CART	Classification and Regression Trees
CGM	Continuous glucose monitoring
DM	Diabetes Mellitus
ECG	Electrocardiogram
EEG	Electroencephalogram
IEEE	Institute of Electrical and Electronic Engineers
KNN	k-nearest neighbors
ML	Machine learning
RNN	Recurrent neural network
SKNN	Subspace k-nearest neighbors
SMBG	Self monitoring blood glucose
SMOTE	Synthetic Minority Over-sampling Technique
SVM	Support vector machine
T1D	Type 1 diabetes
T2D	Type 2 diabetes
UBI	Universidade da Beira Interior
UCI	University of California Irvine
US	United States

Chapter 1

Introduction

This chapter presents an introduction to the thesis entitled "Context-aware algorithms for Diabetes or Prediabetes prediction and diagnosis support in Ambient Assisted Living.". Diabetes mellitus (DM) is characterized by the inability to produce enough insulin or insulin at all, resulting in difficulty to control blood glucose. There are different diabetes conditions, among them, diabetes type-1 (T1D) that requires insulin therapy and diabetes type-2 (T2D) that requires insulin some years after the diagnosis. In either situation, an excessive or insufficient insulin intake causes a bad control of blood glucose, resulting in some short-term and long-term complications such as hypoglycaemia (or hypoglycemia in US English). This thesis addresses research about context-aware algorithms for hypoglycaemia prediction based on discrete data, data and information fusion and classifiers' consensus decision in a 24h hour time frame. The remaining chapter presents the following topics: **1)** the thesis focus and scope; **2)** the problem definition and objectives of the research; **3)** the definition of the statement; **4)** the main contributions of this study, and, **5)** the thesis organisation.

1.1 Thesis Focus and Scope

According to International Diabetes Federation report [1], in 2021, approximately 537 million adults around the world are affected by DM, which prevalence is estimated to reach 783 million adults in 2045. Hypoglycaemia is the most feared condition that may arise commonly in patients that need insulin therapy. The drop in blood glucose (BG) level below 70 mg/dL or 3.9 mmol/L [2, 3] characterises this condition. The symptoms of hypoglycaemia may lead to loss of consciousness and, in the worst cases, can cause death. In most cases, hypoglycaemia arrives unannounced and is essentially asymptomatic. This situation can seriously decrease the quality of life of patients with diabetes, which in turn leads to feelings like insecurity and fear. Hypoglycaemia may be several triggering, *e.g.*, an overdose of insulin, a low intake of carbohydrates, an excess of physical activity or use of some medicines [4]. Nowadays, hypoglycaemia may be classified into different groups according to time of day of occurrence or to threshold that may vary according to the patient's characteristics: age, disease development and others. Based on time of day of occurrence, hypoglycaemia can be daytime hypoglycaemia [5, 6, 7, 8, 9, 10, 11, 12, 13], postprandial hypoglycaemia [14, 15, 16], nocturnal hypoglycaemia [17, 18, 19, 20, 21], and post-exercise [22]. Hypoglycaemia risks are smaller for patients equipped with Continuous Glucose Monitoring (CGM) systems, sometimes used with an insulin pump, named sensor-augmented pump. The number of patients using these systems is growing, but some reasons lead to discontinuing these systems [19, 23]. Some patients continue to use BG monitoring performed using finger sticks and the methods for glycaemia prediction based on CGM are not extensible in these cases. The need to

control BG levels provides an opportunity for developing new technological solutions for self-management and predictions. These solutions may propose approaches embedding expert knowledge with the patient's knowledge side. Data from an individual changes with different actions and medical conditions, so collecting real data is necessary to give crucial contextual information for Ambient Assisted Living (AAL) applications. The context of the collected data provides individualised knowledge solutions [24] and allows for the development of better predictive algorithms. Predictive models based on machine learning (ML) have been increasingly adopted for BG predictions [4, 25]. The features of ML make these techniques an adequate tool for predicting future events. The focus of this thesis includes the design and development of a novel context-aware framework for hypoglycaemia prediction. There are several open research questions involved in the development of the proposed framework:

Question 1 - How to convert collected data into useful context-knowledge?

This research question addresses the issue regarding the data collected, the heterogeneous sources of information, data fusion and how to convert this information into useful context-knowledge suitable for hypoglycaemia prediction. To answer to this question, it was identified existing predictive algorithms for hypoglycaemia in order to understand the type of data most suitable for our approach.

Question 2 - Which pre-processing techniques are most suitable for our scenario?

From the expected solution to the first question, the suitable data fusion and the conversion of context information were selected. The second research question addresses the issue regarding the need or not of pre-processing techniques to best prepare the inputs matrix to the predictive algorithm. To address this research question, it was studied in more detail the existing predictive algorithms in terms of pre-processing techniques to understand what are the methods that contribute to the efficiency of the predictive algorithm.

Question 3 - How to prepare the target (output) for the learning process?

In line with the previous solution, the inputs matrix was selected for the predictive algorithm. The third research question addresses the issue regarding the target suitable for the learning process. To address this research question, it was defined the target in function of inputs matrix and the final decision from the predictive algorithm.

Question 4 - Which classifiers are most suitable for the proposed scenario?

According to the previous questions' solution, the inputs matrix and the target will be used for the learning process for the knowledge discovery and, using that knowledge, predict hypoglycaemia events. To address this question, it was explored different classifiers in order to verify which ones obtain a good efficiency.

1.2 Problem statement and research objectives

Diabetes is a major global health problem: 783 million adults are estimated by 2045 [1]. Diabetes care is one of the most advanced in patient engagement due to its prevalence and because of self-care's crucial role in it. This makes diabetes care especially interesting also with the use of technology in the self-care process. A correct management of diabetes gives to a patient with diabetes some life quality, without complications. Unfortunately, a bad glycaemic control can cause some short-term and long-term complications. The most unwanted short-term issue is hypoglycaemia, i.e., a value of blood glucose that is below the acceptable health threshold (70 mg/dL). Patients with diabetes require a very strict control of their disease. The collection of data is crucial for these patients, such as glucose levels, therapy, meal content, life-style habits, or health history. There have been some works in literature that propose glycaemia prediction algorithms using data from CGM systems that allow for alarms decreasing the incidence of hypoglycaemia and time spent when hypoglycaemia occurs. However, these results are not extensible to patients who use finger sticks to monitor BG. The sparsity of data from this type of BG measurements leads to discrete data, which are not yet explored in the literature. These limitations combined with the lack of contextual open resources, noncompliance of patients with the study protocols and the use of discrete information in predictive works, challenges for innovative and disruptive solutions. The problem addressed in this thesis is to propose a predictive algorithm for hypoglycaemia using discrete data and context information. The proposed approach could be embedded in monitoring equipment that, combined with mobile technology (for example, a self-management application) could assist patients and caregivers in managing diabetes, and could prevent hypoglycemia events. Most medical experts agree that self-monitored blood glucose (SMBG) performed using finger sticks is essential to manage hypoglycaemia [26]. In this way, this approach would also be accessible to patients without CGM systems. Overall, this thesis aimed to advance design and develop a context-aware algorithm for hypoglycaemia prediction using discrete data and contextual information. This goal was translated into four objectives:

1. Identification of the existing predictive algorithms for hypoglycaemia.
The aim was to identify the existing predictive algorithms for hypoglycaemia and identify possible open challenges. To achieve this, the current findings about the data and methods described in the literature for hypoglycaemia prediction were reviewed. Also, their contributions were analysed and compared, and some research gaps and opportunities were identified to bring the knowledge suitable for our approach.

2. Design of context-aware framework for hypoglycaemia prediction using discrete data.
The aim was to design the framework for the context-aware predictive algorithm for hypoglycaemia. To achieve this, different databases were identified and analyzed to define which one is most suitable for our approach. Also, the existing predictive algorithms were studied

in more detail to identify different methods and techniques for pre-processing and learning process. The inputs matrix and the target were selected for the learning process.

3. Development of context-aware predictive algorithm for hypoglycaemia.
The aim was to develop the context-aware predictive algorithm for hypoglycaemia, that comprises implementing pre-processing methods and learning process. In the learning process, different classifiers were implemented.

4. Evaluation of data fusion and multi-classifier decision.
The aim was the performance comparison between the models using conventional features and model using context fusion to conventional features. Also as aim, was the performance comparison between individual models and multi-classifiers models in order to show which approach performs better the predictions.

1.3 Thesis statement

This thesis proposes a context-aware algorithm for hypoglycemia prediction using discrete data, data fusion and multi-classifiers decision, having as thesis statement following:

The algorithms for glycaemia prediction based on data from continuous glucose monitoring systems, are extensively studied in literature, but this kind of solution is not usable by all patients with diabetes and the prediction of hypoglycaemia is rarely addressed as the aim on these systems. Prediction algorithms based on discrete information are a challenge but allow for the development of self-management solutions. There is different additional context information that can be fused with the conventional features to decrease the effect of inter- and intra-subject variability on prediction patterns. Defining a new context-aware framework, that comprises data fusion and multi-classifiers learning, significantly contributes to the creation of suitable and generalised predictive algorithms. This solution could be embedded in monitoring equipment that, combined with mobile technology (for example, a self-management application) could assist patients and caregivers in managing diabetes, and could prevent hypoglycemia events.

To support this thesis statement, a systematic literature review was made in order to identify and analyse the existing predictive algorithm for hypoglycaemia prediction. From the methods and techniques identified, a novel context-aware framework was designed and developed to predict hypoglycaemia using discrete data. The framework was applied using real patient data to evaluate the data fusion and the multi-classifiers decision.

1.4 Main contributions

The achievement of the four objectives of this thesis resulted in the following main contributions:

1. A systematic literature review addressing the existing algorithm for hypoglycaemia prediction was performed to identify data and methods used in literature. From the findings of this review, some research gaps and opportunities were identified to bring the knowledge suitable for our approach. This contribution has been published in Artificial Intelligence In Medicine (Elsevier, Q1, Computer Science):

V. Felizardo, N. M. Garcia, N. Pombo & I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review," in Artificial Intelligence In Medicine, 118 (2021). https://doi.org/10.1016/j.artmed.2021.102120.

2. A novel context-aware framework for hypoglycaemia prediction that comprises data fusion and multi-classifiers learning was designed, developed and tested. The main results of the tested framework show the contribution of integrating context and time-based features with the conventional features for the improvement of hypoglycaemia prediction. Also, the results show that was possible to improve the accuracy of the prediction without the increase of false positives using a classifier's consensus decision rather than an individual classifier decision. This contribution has been published in two research papers and one portuguese patent submission:

V. Felizardo, D. Machado, N. M. Garcia, N. Pombo and P. Brandão, "Hypoglycaemia Prediction Models With Auto Explanation," in IEEE Access, vol. 10, pp. 57930-57941, 2022, doi: 10.1109/ACCESS.2021.3117340. (IEEE, Q1, Computer Science).

V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, & F. Babič, "A novel context-aware framework for Hypoglycaemia Events Prediction integrating context and time-based features ," (under review, Q1, Computer Science).

V. Felizardo, N. M. Garcia, N. Pombo, & I. Megdiche, "*Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos*," 2021 (patent application n.o PT117301).

3. The novel context-aware framework for hypoglycaemia prediction was evaluated by performing predictions for real patients' data. For this, the framework was applied, different models and different classifiers combinations were developed to provide personalised predictions for each patient. The results show that data and information fusion can give a good contribution to represent patients with diabetes in different scenarios and giving new knowledge. The predictive algorithm offers a good prediction of risk of hypoglycaemia in a 24-hour

time frame using sparse data. The contribution has been submitted to a Computer Science journal (under review, Q1, Computer Science):

V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, M. Sousa & F. Babič, "Hypoglycaemia prediction using information fusion and classifiers consensus," (under review, Q1, Computer Science).

1.5 Thesis organisation

This Thesis is organized as follows:

Chapter 1: Introduction A brief introduction to the thesis is presented, including the focus and scope, problem definition and thesis objectives, thesis statement, and main contributions of the research work.

Chapter 2: State-of-the-art This chapter comprises the article titled "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction—A systematic literature review" published in Artificial Intelligence In Medicine journal. This research paper aim review and identify the existing algorithm and methods used for hypoglycaemia prediction and allowed to identify some research opportunities.

Chapter 3: Hypoglycaemia prediction framework The third chapter of this thesis comprises two research papers. First article titled "Hypoglycaemia prediction models with auto explanation", published in IEEE Access journal, presents an update of the state-of-the-art and proposes a general framework for the hypoglycaemia prediction. The second research paper titled "A novel context-aware framework for Hypoglycaemia Events Prediction", submitted to a Computer Science journal, presents each method and technique used for the design of this framework. Also, presents some results obtained from the framework application.

Chapter 4: Framework evaluation This chapter comprises the research paper titled "Hypoglycaemia prediction using information fusion and classifiers consensus", submitted in a Computer Science journal, presents the evaluation of the framework for hypoglycaemia prediction, performing personalised predictions for real patients.

Chapter 5: Conclusion and Future work To conclude this thesis, the main achievements and contributions are summarised and the future research directions are presented.

Bibliography

[1] I. D. Federation, "IDF Diabetes Atlas 10th edition 2021," Tech. Rep., 2021. [Online]. Available: https://diabetesatlas.org/atlas/tenth-edition/ 1, 3

[2] NICE UK, "Diabetes (type 1 and type 2) in children and young people: diagnosis and management," National Institute for Health and Care Excellence (NICE) guideline, Tech. Rep., 2020. [Online]. Available: https://www.nice.org.uk/guidance/ng18 1

[3] ——, "Type 1 diabetes in adults: diagnosis and management," National Institute for Health and Care Excellence (NICE) guideline, Tech. Rep., 2020. [Online]. Available: https://www.nice.org.uk/guidance/ng17 1

[4] O. Mujahid, I. Contreras, and J. Vehi, "Machine learning techniques for hypoglycemia prediction: Trends and challenges," 2021. 1, 2

[5] R. H. Botwey, E. Daskalaki, P. Diem, and S. G. Mougiakakou, "Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014, pp. 4843–4846, 2014. 1

[6] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, "A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management," Modern Artifical Intelligence for Health Analytics, pp. 35–39, 2014. [Online]. Available: https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/viewFile/8737/8308 1

[7] Zarkogianni, K., Mitsis, K., Litsa, E., Arredondo, M. T., Fico, G., Fioravanti, A., and Nikita, K. S., "Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring," Medical and Biological Engineering and Computing, vol. 53, no. 12, pp. 1333–1343, 2015. 1

[8] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A Deep Learning Approach to Diabetic Blood Glucose Prediction," Frontiers in Applied Mathematics and Statistics, vol. 3, no. July, pp. 1–11, 2017. 1

[9] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting learning to accurately predict blood glucose trajectories," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1387–1395, 2018. 1

[10] D. Dave, D. J. DeSalvo, B. Haridas, S. McKay, A. Shenoy, C. J. Koh, M. Lawley, and M. Erraguntla, "Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction," Journal of Diabetes Science and Technology, 2020. 1

[11] X. Yu, K. Turksoy, M. Rashid, J. Feng, N. Hobbs, I. Hajizadeh, S. Samadi, M. Sevil, C. Lazaro, Z. Maloney, E. Littlejohn, L. Quinn, and A. Cinar, "Model-fusion-based online glucose concentration predictions in people with type 1 diabetes," Control

Engineering Practice, vol. 71, no. April 2017, pp. 129–141, 2018. [Online]. Available: https://doi.org/10.1016/j.conengprac.2017.10.013 1

[12] J. Yang, L. Li, Y. Shi, and X. Xie, "An ARIMA Model with Adaptive Orders for Predicting Blood Glucose Concentrations and Hypoglycemia," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1251–1260, 2019. 1

[13] M. Gadaleta, A. Facchinetti, E. Grisan, and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 2, pp. 650–659, 2019. 1

[14] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, "Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning," Health Informatics Journal, 2019. 1

[15] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," International Journal of Medical Informatics, vol. 126, no. March, pp. 1–8, 2019. [Online]. Available: https://doi.org/10.1016/j.ijmedinf.2019.03.008 1

[16] F. Dubosson, N. Mordvanyuk, B. López Ibáñez, and M. Schumacher, "Negative results for the prediction of postprandial hypoglycemias from insulin intakes and carbohydrates: analysis and comparison with simulated data," pp. 25–29, 2017. [Online]. Available: https://dugi-doc.udg.edu/handle/10256/14859 1

[17] A. Bertachi, L. Biagi, N. Luo, and J. Veh, "Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks," in KHD@ IJCAI, 2018, pp. 85–90. 1

[18] M. H. Jensen, C. Dethlefsen, P. Vestergaard, and O. Hejlesen, "Prediction of Nocturnal Hypoglycemia From Continuous Glucose Monitoring Data in People With Type 1 Diabetes: A Proof-of-Concept Study," Journal of Diabetes Science and Technology, vol. 14, no. 2, pp. 250–256, 2019. 1

[19] P. Tkachenko, G. Kriukova, M. Aleksandrova, O. Chertov, E. Renard, and S. V. Pereverzyev, "Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application," Computer Methods and Programs in Biomedicine, vol. 134, pp. 179–186, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cmpb.2016.07.003 1

[20] G. Kriukova, N. Shvai, and S. V. Pereverzyev, "Application of regularized ranking and collaborative filtering in predictive alarm algorithm for nocturnal hypoglycemia prevention," in 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), vol. 2, 2017, pp. 634–638. 1

[21] F. Stahl, R. Johansson, and M. L. Olsson, "Predicting Nocturnal Hypoglycemia Using a Non-parametric Insulin Action Model," Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015, pp. 1583–1588, 2016. 1

[22] R. Reddy, N. Resalat, L. M. Wilson, J. R. Castle, J. El Youssef, and P. G. Jacobs, "Prediction of Hypoglycemia During Aerobic Exercise in Adults With Type 1 Diabetes," Journal of Diabetes Science and Technology, vol. 13, no. 5, pp. 919–927, 2019. 1

[23] D. Rodbard, "Continuous Glucose Monitoring: A Review of Recent Studies Demonstrating Improved Glycemic Outcomes," Diabetes Technology & Therapeutics, 2017. 1

[24] A. R. M. Forkan and W. Hu, "A context-aware, predictive and protective approach for wellness monitoring of cardiac patients," 2016 Computing in Cardiology Conference (CinC), vol. 43, pp. 369–372, 2016. 2

[25] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Botsis, and G. Hartvigsen, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," 2019. 2

[26] E. M. Benjamin, "Self-Monitoring of Blood Glucose: The Basics," Clinical Diabetes, 2002. 3

Chapter 2

State-of-the-art

The chapter presents and discusses concepts of state-of-the-art of data-based algorithms and models for blood glucose and hypoglycaemia prediction. The chapter is presented as an article titled "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review."

2.1 Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review.

Virginie Felizardo, Nuno M. Garcia, Nuno Pombo and Imen Megdiche.

Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review.

Published in Artificial Intelligence In Medicine, 2021, in the area of Computer Science.
H−Index 87, Q1, by Scimago Journal & Country Rank.

Artificial Intelligence In Medicine 118 (2021) 102120

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review

Virginie Felizardo [a,b,*], Nuno M. Garcia [a,b], Nuno Pombo [a,b], Imen Megdiche [c]

[a] Instituto de Telecomunicações, Covilhã, Portugal
[b] Universidade da Beira Interior, Covilhã, Portugal
[c] IRIT, Institut de Recherche en Informatique de Toulouse, Toulouse University, France

ARTICLE INFO

Keywords:
Hypoglycaemia or hypoglycemia
Blood glucose level
Prediction
Data-based algorithms or models
Diabetics real data

ABSTRACT

Background and aim: Hypoglycaemia prediction play an important role in diabetes management being able to reduce the number of dangerous situations. Thus, it is relevant to present a systematic review on the currently available prediction algorithms and models for hypoglycaemia (or hypoglycemia in US English) prediction.
Methods: This study aims to systematically review the literature on data-based algorithms and models using diabetics real data for hypoglycaemia prediction. Five electronic databases were screened for studies published from January 2014 to June 2020: ScienceDirect, IEEE Xplore, ACM Digital Library, SCOPUS, and PubMed.
Results: Sixty-three eligible studies were retrieved that met the inclusion criteria. The review identifies the current trend in this topic: most of the studies perform short-term predictions (82.5%). Also, the review pinpoints the inputs and shows that information fusion is relevant for hypoglycaemia prediction. Regarding data-based models (80.9%) and hybrid models (19.1%) different predictive techniques are used: Artificial neural network (22.2%), ensemble learning (27.0%), supervised learning (20.6%), statistic/probabilistic (7.9%), autoregressive (7.9%), evolutionary (6.4%), deep learning (4.8%) and adaptative filter (3.2%). Artificial Neural networks and hybrid models show better results.
Conclusions: The data-based models for blood glucose and hypoglycaemia prediction should be able to provide a good balance between the applicability and performance, integrating complementary data from different sources or from different models. This review identifies trends and possible opportunities for research in this topic.

1. Introduction

Diabetes mellitus (DM) is a chronic metabolic disorder, characterized by the inability to produce enough insulin or insulin at all, that results in difficulty to control blood glucose. Managing and living with DM becomes a radical challenge for patient with diabetes because it starts by changing the patients' lifestyle, leading the person to develop their autonomy skills to manage the control of blood glucose levels.

According to the World Health Organization (WHO) report, DM is one of four major non-communicable diseases causing more deaths than all other causes combined [1]. Worldwide, in 2017 approximately 425 million people had DM. In 2045, the number of DM patients is expected to be closer to 629 million people [2]. There are different diabetes conditions, among them, diabetes type-1 (T1D) that requires insulin treatment and diabetes type-2 (T2D) that requires insulin some years after the diagnosis. In either situation, an excessive or insufficient insulin intake causes a bad glycaemic control resulting in some short-term and long-term complications such as hypoglycaemia (or hypoglycemia in US English). Hypoglycaemia is when blood glucose (BG) level falls low than 70 mg/dl. This complication is commonly feared in patients with insulin therapy, as a Hypoglycaemia crisis can result in death either directly or indirectly, for example, because the patient loses consciousness.

Traditionally, the monitoring of patients with diabetes includes the measurement and collection of their glycaemic levels' values using a finger-stick blood glucose meter, often multiple times in a day. This type of monitoring has a major disadvantage, in that it cannot detect fluctuations of the glycaemia levels that may be caused by intense physical activity or food/insulin intake between measurements.

Since the introduction of Continuous glucose monitoring (CGM) systems, literature has proposed several approaches for the prediction of BG levels or associated future events. These solutions enable the patient

* Corresponding author.
E-mail addresses: virginie@it.ubi.pt (V. Felizardo), ngarcia@di.ubi.pt (N.M. Garcia), ngpombo@ubi.pt (N. Pombo), imen.megdiche@irit.fr (I. Megdiche).

https://doi.org/10.1016/j.artmed.2021.102120
Received 11 February 2020; Received in revised form 24 May 2021; Accepted 25 May 2021
Available online 28 May 2021

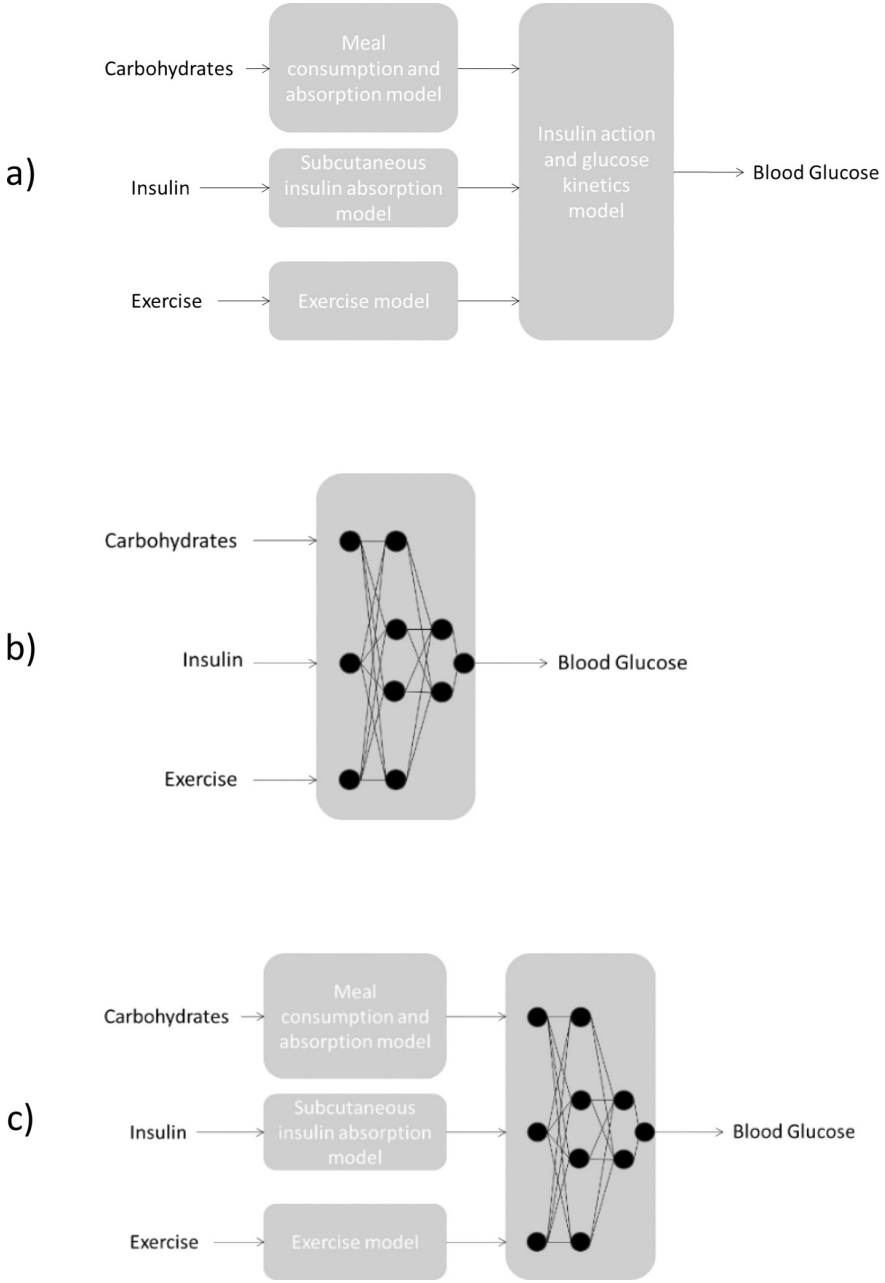

Fig. 1. Diagram of Blood Glucose prediction models: a) Physiological models, b) Data-based models and c) Hybrid models (adapted from [16]).

to take appropriate actions to prevent the oncoming hypoglycaemia event, *e.g.* carbohydrate intake or the suspension of insulin intake [3,4]. Also, these systems allow the deployment of warning alerts that would enable the implementation of applications such as "what if" analysis to forecast the effects of the patient's lifestyle or therapy choices [4].

Some years ago, physiological models were the most common approach, however, in the last years, some data-based approaches have emerged as well as a combination of both called the hybrid models [5]. Fig. 1 shows a diagram of each prediction model.

A physiological model is a continuous dynamic model in which a state transition function computes the next state of the system given the current state and input variables. The overall blood glucose dynamics is usually characterized by four compartments: meal dynamics [6,7], insulin dynamics [8,9], exercise model [10] and glucose dynamics [9,11,12]. These models are generally not very accurate and require previous knowledge to set the physiological constants [5,13]. In order to overcome these limitations, the data-based models have been proposed using pattern recognition techniques [14]. These models are created using experimental data and, for this reason, are in general more accurate [13]. The major challenge for the predictive models is that the effects of insulin intake, meals, physical activity and other events in glucose dynamics are different for every person, this effect being termed inter-subject variability, and even in the same patient there are changes over time, termed intra-subject variability [15]. According to this, the models for blood glucose or hypoglycaemia prediction can be categorised into two groups: 1) personalized-, individual- or subject-based models trained on an individual's specific data, and 2) general- or population-based models trained on pooled data from many people expecting that they will generalize well enough to allow them to be used in previously unseen individuals.

Table 1
Keywords used in each database.

Database	Keywords
ScienceDirect IEEE Xplore PubMed SCOPUS ACM Digital Library	"hypoglycemia prediction" OR "hypoglycemia detection" OR "blood glucose prediction" OR "blood glucose estimation" OR "hypoglycemia estimation" OR "blood glucose detection" (hypoglycemia OR "blood glucose") AND (prediction OR detection OR estimation)

Both type of models present some drawbacks. In individual-based models, the accuracy can be negatively affected by intra-subject variability if the model is trained on limited data sets. The population-based models trained on pooled data may not generalize well to specific groups of individuals if they are not appropriately represented in the training data set.

Regarding the prediction outcome, the models can be categorised into three groups according to the duration of the prediction horizon (PH). The PH is the future time window in which the predicted glucose concentration is determined by a model. The groups are as follows: 1) short-term, when the prediction horizon (PH) is less than 180 min; 2) mid-term, when the PH is more than 180 min and less than 24 h; and 3) long-term, when the PH is days, months, or years. The thresholds of each group were designed after merging information in several studies as no standard definition of thresholds exists.

To conclude, no specific method provides 100% reliable event predictions. This is the reason why an increasing number of approaches have been proposed by combining multiple predictors and techniques with the hope to represent spatial and temporal input–output dependencies.

This paper presents a systematic review on data-based algorithms and models using patients' real data for hypoglycaemia prediction covering the past six years, including studies published until June 2020. We found two reviews that partially cover this topic, one between 2000 and 2018 [17] and the other between 2010 and 2016 [16]. Several papers deal with the prediction of BG values in general, labelling hypoglicemic events after prediction is performed. Taking this into account, to present the most recent advances in the field, this paper has considered both specific research on hypoglycemic events prediction and more general works on BG levels prediction. We make the choice to present hybrid models as these are a combination of data-based models with physiological models.

The main contributions of this paper are: 1) a discussion on how the hypoglycaemia prediction problem has been presented previously and how it is currently addressed; 2) a clear identification of the features used in the predictive models based on real data; 3) an assessment of the most used data-based or hybrid algorithms; and 4) a comparative analysis to conclude which models performs better.

The remainder of this paper is organized as follows: Section 2 presents a description of the method that was designed for the eligibility selection and extraction of information. Section 3 is dedicated to the results where all studies and their features were presented in summary tables. In Section 4 we discuss and answer the research questions. We conclude this review by presenting the highlights and limitations from this study in Section 5.

2. Methods

We conducted a systematic review informed by recommendations from the Cochrane Handbook for Systematic Reviews of Interventions, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [18–20]. The aim of this systematic review is to identify, assess and analyse the recent state-of-the-art of data-based or hybrid models to predict BG levels or hypoglycaemia events using real data from patients with diabetes. The review covers different approaches to predict BG levels or hypoglycaemia events, such as regression or classification, relevant to population- or individual-based solutions.

2.1. Search strategy

For this study, we searched the databases ScienceDirect, IEEE Xplore, ACM Digital Library, SCOPUS and PubMed for relevant peer-reviewed publications published between January 2014 and June 2020. We searched titles and abstracts using the keywords presented in Table 1. The list of references from the selected articles were manually screened for inclusion of additional relevant articles.

2.2. Study selection

We screened the titles and abstracts of all identified publications for eligibility, using the web application Rayyan QCRI [21]. The inclusion criteria were broadly defined as to increase the sensitivity of the search. We aimed to identify the articles that applied data-based or hybrid models to BG or hypoglycaemia prediction using real data from diabetic people. Additional inclusion criteria are summarized as follows:

- Participants: patients diagnosed with diabetes, without age or diabetes type restrictions;
- Intervention: real data from participants, collected on clinical or daily living contexts, with at least one of this data categories, comprising blood glucose levels (from CGM or Self-monitoring Blood glucose (SMBG)), insulin-, meal-, or exercise information;
- Comparison: different approaches of data-based models and hybrid models (physiological and data-based methods);
- Outcome: different outcomes for BG levels or hypoglycaemia prediction (short-, mid-, or long-term predictions).

The publications that met any of the following exclusion criteria upon title and abstract review from subsequent evaluations were excluded from the study:

- Participants: patients not diagnosed with diabetes or in-silico patients;
- Intervention: data from simulations, data from electrocardiogram (ECG), electroencephalogram (EEG) or other biosignals when not combined with the data classes mentioned in inclusion criteria;
- Comparison: approaches based only on physiological models or control-oriented models;
- Outcome: Artificial pancreas control system;
- Manuscripts not written in English, not the final version, reviews, chapters, abstracts and short papers;
- Oldest work from an author with the same methodology purpose or when the same work with different years of publication.

Full texts of all remaining identified articles were reviewed for eligibility.

2.3. Extraction of study characteristics

Information from the selected publications was extracted based on the pre-defined categories. The authors of this study defined the following categories to collect the relevant data and to assess, analyse the model characteristics and its experimental setup:

- Study Information: This category defines the study citation and year of publication;
- Data: This category defines the type of diabetes, sample (number of participants, age range or average age and gender distribution) and source (collected, clinical trial, EHR or dataset);

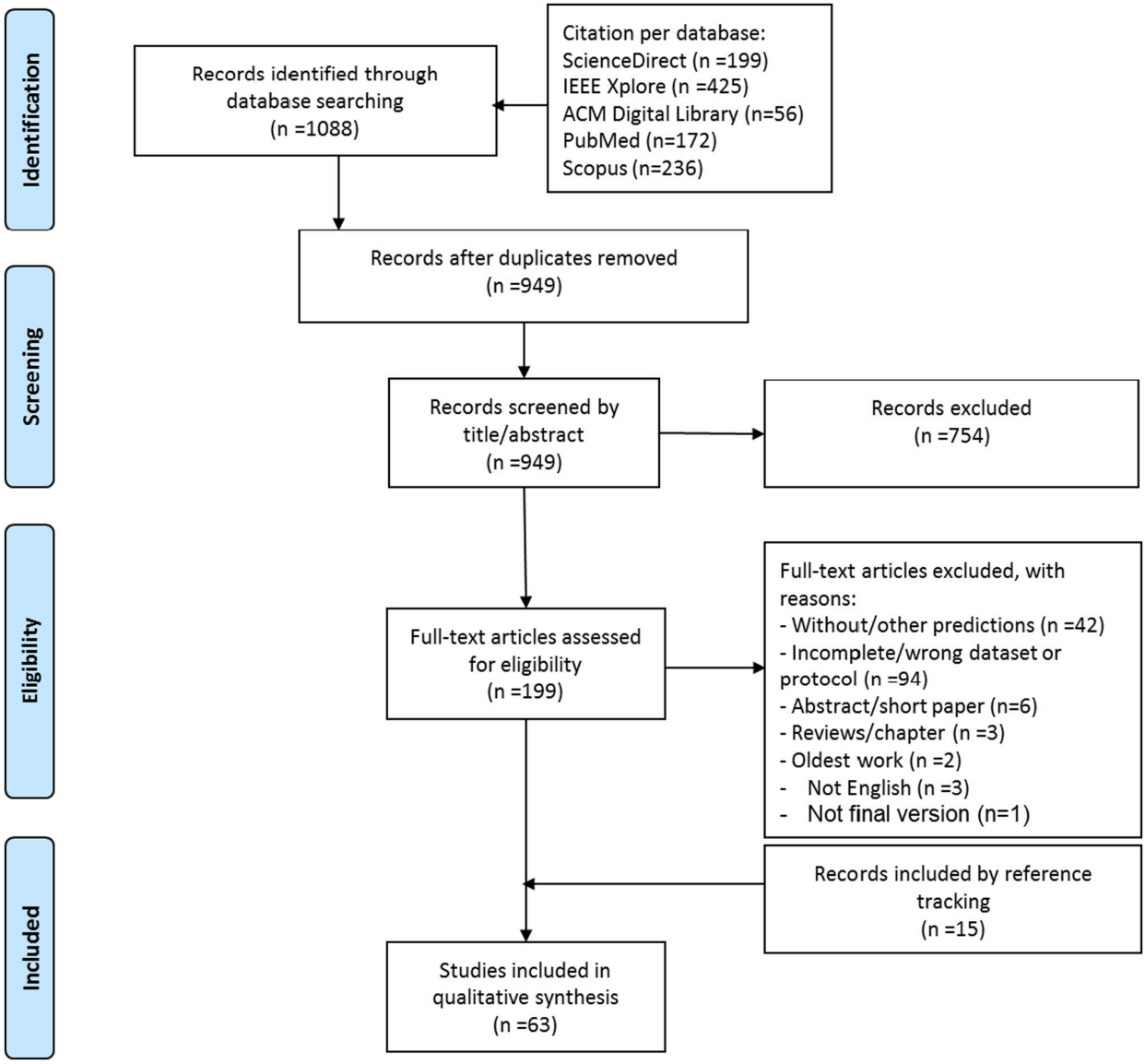

Fig. 2. Flow diagram of identification and inclusion papers.

- Prediction: This category defines the application where the algorithm is being exploited. It can be short-, mid- or long-term predictions, nocturnal, postprandial or other. The prediction can also be BG in general prediction, hypoglycaemia prediction or a combination of BG prediction and, after the prediction, hypoglycaemia labelled in BG values.
- Algorithms: This category defines the different approaches used to train the model;
- Inputs: This category was defined to assess the inputs used to develop the algorithm. This includes BG, meal, insulin, exercise or others;
- Performance: This category defines the evaluation metrics used to assess the predictions.

2.4. Research questions

The research questions of this review are:

- (RQ1) How is the prediction problem addressed?
- (RQ2) What are the features used based on real data?
- (RQ3) What are the most used data-based algorithms?
- (RQ4) Which are the models that perform better?
- (RQ5) What about replicability and availability of methods of the studies to scientific community?

The motivations behind the different research questions are as follows:

- The motivation for RQ1 is to identify the trends and possible opportunities for the research topic focus.
- The motivation for RQ2 is to identify the recent advances in new feature combination for hypoglycaemia prediction.
- The motivation for RQ3 is to identify the trend for the use of research methods.
- The motivation for RQ4 is to be able to assess the models' performance and identify trends and opportunities for the use of research methods.
- The motivation of RQ5 is to identify studies with a reproducible methodology and share with the readers the access to algorithms and data.

Table 2
An analysis of the selected studies.

Study	Year	Dataset			Prediction					
		Type diabetes	Sample[a]	Source	Short-term	Mid-term	Long-term	Nocturnal	Postprandial	Approach[b]
Botwey et al. [3]	2014	T1D	23 participants, 17–70 years	Collected, daily living [22]	X					BG/H
Zhao et al. [23]	2014	T1D	17 participants, 49 years	Collected, ambulatory clinical data	X					BG
Plis et al. [4]	2014	T1D	5 participants	Collected	X					BG/H
Efendic et al. [24]	2014	T1D	14 participants, 48 years (6 M/6F)	Collected, clinical context	X					BG
Dasanayake et al. [25]	2015	T1D	15 participants	Collected	X					BG
Georga et al. [26]	2015	T1D	15 participants	METABO dataset [27], daily living	X					BG
Zarkogianni et al. [14]	2015	T1D	10 participants, 42 years (7M/3F)	METABO dataset [27], daily living	X					BG/H
Georga et al. [28]	2015	T1D	15 participants	METABO dataset [27], daily living	X					BG
Cescon et al. [29]	2015	T1D	14 participants, 45 years (9M/5F)	DIAdvisor dataset [30], clinical context	X					BG
Stahl et al. [31]	2015	T1D	29 participants, 19–68 y (15M/14F)	Collected, daily living	X			x		BG/H
Novara et al. [13]	2016	T1D	5 participants	Collected, daily living	X					BG
Li and Fernando [32]	2016	T1D	70 participants	Diabetes UCI dataset [33], daily living		x				BG
Tkachenko et al. [34]	2016	T1D	34 participants, 18–65 years and 179 participants, children s	DIAdvisor [30] and ChildrenDat [35] datasets, clinical context		x		x		H
Georga et al. [36]	2016	T1D	15 participants	METABO dataset [27], daily living	X					BG
Zecchin et al. [37]	2016	T1D	15 participants	DIAdvisor dataset [30], clinical context	X			x	x	BG
Hidalgo et al. [5]	2017	T1D	10 participants, 42 years (2M/8F)	Collected	X					BG
Hamdi et al. [38]	2017	T1D	12 participants	Collected, clinical context	X					BG
Kriukova et al. [39]	2017	T1D	34 participants and 70 participants	DIAdvisor [30], clinical context and Diabetes UCI datasets [33], daily living		x		x		H
Schroeder et al. [40]	2017	T1D/T2D	82,474 participants, 59 years	EHR records			x			H
Khan et al. [41]	2017	T1D	70 participants	Diabetes UCI dataset [33], daily living		x				BG
Mirshekarian et al. [42]	2017	T1D	10 participants	Collected	X					BG
Dubosson et al. [43]	2017	T1D	10 participants, 24–74 years	ABC4D dataset [44], both context	X				x	H
Mhaskar et al. [45]	2017	T1D	25 participants, <18 years	DirecNet dataset [46], clinical context	X					BG/H
Ali et al. [47]	2018	T1D	12 participants	Collected, daily living	X					BG
Hamdi et al. [48]	2018	T1D	12 participants	Collected, daily living	X					BG
Zhu et al. [49]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	X					BG
Chen et al. [51]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	X					BG
Li et al. [52]	2018	T1D	8 participants	Collected, clinical context	X					BG
Fox et al. [53]	2018	T1D	40 participants	Collected, daily living	X					BG/H
Griva & Basualdo [54]	2018	T1D	16 participants	Collected, daily living	X					BG
Faccioli et al. [15]	2018	T1D	6 participants, 21–65 years	Clinical trial [55]	X					BG
Vahedi et al. [56]	2018	T1D	93 participants, 8–84 years (50M/43F)	Collected, daily living	X					BG
Yu et al. [57]	2018	T1D	9 participants, 18–35 years	Collected, clinical context	x					BG/H
Contreras et al. [58]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	x					BG
Midroni et al. [59]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	x	x				BG
Bertachi et al. [60]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	x			x		H
Xie & Wang [61]	2018	T1D	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset [50], daily living	x					BG

(continued on next page)

Table 2 (*continued*)

Study	Year	Dataset			Prediction					
		Type diabetes	Sample[a]	Source	Short-term	Mid-term	Long-term	Nocturnal	Postprandial	Approach[b]
Oviedo et al. [62]	2019	T1D	10 participants, 41 years (8M/2F)	Collected, daily living		x			x	H
Li et al. [63]	2019	T1D	10 participants, 47 years and 6 participants, 40–60 years (2M/4F)	ABC4D [44], both context and OhioT1DM datasets [50], daily living	x					BG
Yang et al. [64]	2019	T1D/T2D	100 participants, 45 years	Collected, daily living	x					BG/H
Gadaleta et al. [65]	2019	T1D	89 participants	Collected, both context	x					H
Aliberti et al. [66]	2019	T1D	451 participants, 8–25 years and > 25 years (55%F)	Clinical trial [67]	x					BG
Dong et al. [68]	2019	T1D/T2D	80 participants	Collected	x					BG
He et al. [69]	2019	T1D/T2D	112 participants, >35 years (49.1%F)	Collected [70]	x					BG
Idriss et al. [71]	2019	T1D	10 participants	DirecNet dataset [46]	x					BG
Koutny et al. [72]	2019	TID	14 participants	Collected	x (blind window)					BG
Mirshekarian et al. [73]	2019	T1D	6 participants, 40–60 years (4F/2M)	OhioT1DM dataset [50], daily living	x					BG
He et al. [74]	2019	T1D	10 participants	DirecNet dataset [46]	x					BG
Chen et al. [75]	2019	T1D	5 participants	DirecNet dataset [46]	x					BG
Reddy et al. [76]	2019	T1D	55 participants, 33 ± 6 years (33F/22M)	Collected, controlled exercise environment	x					H
Jensen et al. [77]	2019	T1D	463 participants	Clinical trial, daily living (NCT 02825251)		x		x		H
Vu et al. [78]	2019	T1D	9800 participants, avr 45 years, 51%F/49%M	Collected, daily living		x		x		H
Alfian et al. [79]	2020	T1D	5 participants, 3–17 years	DirecNet dataset [46]	x					BG
Anand et al. [80]	2020	T1D/T2D	91 participants, 51.0 ± 30.0 years (41F/50M)	Clinical trial [81]	x					BG/H
Contador et al. [82]	2020	T1D	10 participants	Collected, daily living	x	x				BG
Dave et al. [83]	2020	T1D	112 participants, 1–21 years (48F/64M)	Collected, daily living	x					H
Hidalgo et al. [84]	2020	T1D	5 participants	Collected, clinical context	x					BG
Li et al. [85]	2020	T1D	10 participants	Collected, clinical context [86]	x					BG
Li et al. [87]	2020	T1D	10 participants	Collected, clinical context [88]	x					BG
Martinsson et al. [89]	2020	T1D	6 participants, 40–60 years (4F/2M)	OhioT1DM dataset [50], daily living	x					BG
Misra-Hebert et al. [90]	2020	T2D	1876 participants, avr 65 years, 41.6%M	EMR			x			H
Vehi et al. [10]	2020	T1D	10 participants, 41 years and 6 participants, 40–60 years	Collected, daily living		x		x	x	H
Xie & Wang [91]	2020	T1D	6 participants, 40–60 years (4F/2M)	OhioT1DM dataset [50], daily living	x					BG

[a] Gender distribution: F for female and M for male.

[b] Approach: BG for blood glucose prediction, BG/H for hypoglycaemia labelling after blood glucose prediction, H for hypoglycaemia prediction.

Finally, to conclude our discussion, a comparison between the past reviews in this subject and the present review is disclosed.

3. Results

As depicted in Fig. 2, the initial search yielded 949 unique records After title and abstract review, 754 records were excluded; 199 full-text publications were assessed for eligibility and after full-text review, 149 records were excluded.

The excluded records are described as follows: (i) forty-two studies reported research on diabetes, but their main goal was not to apply models for BG levels or hypoglycaemia predictions; (ii) Ninety-four studies applied prediction models for BG levels or hypoglycaemia prediction but did not present enough information on the experimental protocol or on the dataset; (iii) Six studies presented only abstracts or were short papers; (iv) Three studies were a review paper or chapter, two other paper were oldest than the publication of an author already considered, three papers were not in English and one were not the final version. Among the remaining 48 records, reference tracking was performed, and fifteen studies were added, making a total of 63 studies to be included for the data extraction and the qualitative synthesis stage (Table 2).

3.1. Studies eligibility

All selected studies (Table 2) meet the inclusion and exclusion criteria, but we need to clarify some studies selection. Khan et al. [41] present a BG prediction for type-2 diabetics, but the data used from Diabetes UCI Dataset [33] belongs to diabetics type-1. Despite this error we consider the study as eligible. In Yu et al. [57], Li et al. [63] and Li

Table 3
Data-based models: Algorithms and inputs used to perform the predictions, and the best results performed.

Data-based model												
Study	Algorithm	Inputs					Performance					
							Prediction horizon (min)/RMSE (mg/dl), *Other metric					
		BG	I	CHO/ M	E	Other	15	30	45	60	120	Other PH
Botwey et al. [3]	Autoregressive with exogenous inputs, recurrent neural network, and Genetic Algorithm	x	x				8.5	13.8	19.2			
Zhao et al. [23]	Latent-variable (LV) and Latent variable with exogenous inputs (LVx)	x	x	x				18.5 (LV)		29.2 (LVX)		
Efendic et al. [24]	Gaussian Mixture Model	x	x	x				Cor/Inc.* 76.4/ 23.6%				
Dasanayake et al. [25]	Linear model based on AIC and subspace identification.				x		9.22	19.70				
Zarkogianni et al. [14]	Self-organizing map (SOM)	x			x			11.42		19.58	31.00	
Novara et al. [13]	Blind identification with an Autoregressive integrated moving average with exogenous inputs		x					FIT indexes* 90%		FIT indexes* 54%		
Li and Fernando [32]	Pooled-panel-data regression model with Random forest	x										39.73
Tkachenko et al. [34]	Stochastic model using aggregation and interferential theory	x										Se/Sp* 73.4/ 87.8%
Georga et al. [36]	Quantized Kernel Least Mean Square (QKLMS-FB)	x					11.1	18.7	24.7	30.0		
Zecchin et al. [37]	Jump Neural Network: 1) CGM; 2) CGM + I;3) CGM + CHO; 4) CGM + I + CHO (Nocturnal prediction)	x	x				2)6.3	1)10.7	2)14.5	2)17.8		
	Jump Neural Network: 1) CGM; 2) CGM + I;3) CGM + CHO; 4) CGM + I + CHO (postprandial prediction)	x	x	x			3) 13.5	4)27.6	4)40.2	4)51.3		
Hidalgo et al. [5]	Grammatical evolution	x	x	x		x		EGA* 80.7%%		EGA* 58.7%	EGA* 41.7%	
Hamdi et al. [38]	Artificial Neural Network	x					6.44					
Kriukova et al. [39]	Collaborating filtering	x										F1* 32–53%
Schroeder et al. [40]	Cox regression model (16-variable model)		x			x						c-statistic* 0.84
Khan et al. [41]	Support vector regression (with best feature vector size)	x						acc* 76%				
Dubosson et al. [43]	Support vector machine	x	x	x								F1* 13.70%
Mhaskar et al. [45]	Deep network (with 50% training data)	x						EGA* 96.43%				
Ali et al. [47]	Artificial Neural Network	x					6.43	7.45	8.13	9.03		
Hamdi et al. [48]	Support Vector Regression based on Differential evolution	x					9.44	10.78	11.82	12.95		
Zhu et al. [49]	Convolutional Neural Network	x	x	x		x		21.73				
Chen et al. [51]	Dilated Recurrent Neural Network.	x	x	x		x		19.04				
Li et al. [52]	Echo state network optimized with leakage integral and ridge regression	x					12		36.35	52.05		
Fox et al. [53]	Sequential polynomial multi-output recurrent neural network	x						4.59				
Griva & Basualdo [54]	Autoregressive with exogenous inputs	x	x	x		x				0.5–0.7	6.4–7.3	
Faccioli et al. [15]	Black-box linear model		x	x	x							PA increase COD* in 18.5% for PH 180 min
Vahedi et al. [56]	Random Forest	x			x			MAPE* 27.9%				
Yu et al. [57]	Model-fusion-based on recursive least squares	x			x			EGA* 90.57%				
Midroni et al. [59]	Regression-based gradient-boosted decision tree (XGBoost)	x		x	x	x						19.32
Xie & Wang [61]	Autoregressive with exogenous inputs	x	x	x		x		19.59				
Oviedo et al. [62]	Support vector regression (model C1)	x	x	x								Se/Sp* 73.0/ 72.8%
Li et al. [63]	dilated convolution neural network (using two datasets)	x	x	x				19.19/ 19.28		31.78/ 31.83		

(*continued on next page*)

Table 3 (*continued*)

Data-based model												
Study	Algorithm	Inputs					Performance					
							Prediction horizon (min)/RMSE (mg/dl), *Other metric					
		BG	I	CHO/M	E	Other	15	30	45	60	120	Other PH
Yang et al. [64]	Autoregressive integrated moving average	x						SE//FPR* 100/10.7 (I) 100/8.0 (II)				
Gadaleta et al. [65]	Support vector Regression (static model)	x	x	x				Se/PPV* 75.0/ 51.0%				
Aliberti et al. [66]	Long short-term memory network (with filtered training set)	x						EGA* (A + B) 98.1%	EGA* (A + B) 94.4%	EGA* (A + B) 93.3%		
Dong et al. [68]	Clustering pre-process based recurrent neural network	x						MSE* 0,104(I) 0.060(II)	MSE* 0,318 (I) 0.143 (II)	MSE* 0,556(I) 0.306(II)		
Idriss et al. [71]	Long short-term memory neural network (Inertial scenario)	x										12.38
Mirshekarian et al. [73]	Long short-term memory neural network	x	x	x		x		17.99		28.19		
He et al. [74]	Canonical Correlation Analysis	x					17.38					
Chen et al. [75]	Grey Wolf Optimization evolving Kernel Extreme Learning Machine	x					4.82	13.37				
Reddy et al. [76]	Random Forest	x	x		x	x						acc* 86.67%, Se* 86.21%, Sp* 86.89%
Jensen et al. [77]	Linear discriminant analysis	x	x	x		x						Se* 75% Sp* 70%
Vu et al. [78]	Random Forest	x										PH 6 h: AUC* 84% PH 3 h: AUC* 90%
Alfian et al. [79]	Extreme Gradient Boosting	x						23.22		35.80		
Anand et al. [80]	AdaBoost+k-means clustering	x						MARD* 25.1				
Contador et al. [82]	Genetic Programming+clustering (CHAID-GP)	x	x	x				EGA* (A + B) 91.41%		EGA* (A + B) 90.95%	EGA* (A + B) 89.52%	
Dave et al. [83]	Random Forest (with VIF selected features, and insulin and carb on board)	x	x	x		x		Se/Sp* 97.04/ 95.23		Se/Sp* 96.21/ 95.73		
Hidalgo et al. [84]	Random-GE Bagging (synthetic data generated by the SAX-SIR Method)	x	x							EGA* (A + B) 97.45%%	EGA* (A + B) 95.63%	
Li et al. [85]	Convolution neural network	x	x	x				21.07		33.27		
Li et al. [87]	Echo state networks embedded with incremental learning	x						13.57	13.67			
Martinsson et al. [89]	Recurrent neural network	x						18.87		31.40		
Misra-Hebert et al. [90]	Logistic regression		x			x						Se/Sp* 82%/ 79%
Xie & Wang [91]	Autoregressive with exogenous inputs and Ridge regression (direct method)		x	x		x		19.48				

et al. 2020 [85] the models were performed on real patients and in-silico patients independently. In Vehi et al. [10] the short-term prediction model was excluded because it uses data from in-silico patients to train the system. For this study, we include only the models for postprandial prediction and nocturnal prediction, both trained and tested with real data.

In Tables 3 and 4 we presented the studies categorised in data-based models and hybrid models, respectively. These tables show the algorithm and inputs (blood glucose (BG), insulin (I), meal information (CHO/M), Exercise information (E) and Other) used to perform the predictions. The last columns show the best results obtained by the selected studies (the acronyms of the metrics presented in the tables are covered in the Section 3.7).

3.2. Sources of evidence

From our selected studies, forty-one studies (65.1%) were published in the past three years, more than the number of studies published between 2014 and 2017 (34.9%). Thirty-two studies (50.8%) are research articles published in a peer-reviewed journal or magazine, seventeen studies (27%) were published in computer science journals, six of them in currently 2nd quarter rank magazines (between 2014 and 2017) and eleven of them in 1st quarter rank magazines (between 2018 and 2020); thirty-one studies (49.2%) are a research article in a conference

Table 4
Hybrid models: Algorithms and inputs used to perform the predictions, and the best results performed.

Hybrid model												
Study	Algorithm	Inputs					Performance					
							Prediction horizon (min)/RMSE (mg/dl), *Other metric					
		BG	I	CHO/M	E	Other	15	30	45	60	120	Other PH
Plis et al. [4]	Support vector machine (using Physiological features) + Autoregressive integrated moving average (additional features)	x	x	x				19.5		35.7		
Georga et al. [26]	Kernels extreme learning machines	x	x	x	x			6.1				
Georga et al. [28]	Gaussian Process+Random Forest (feature selection)	x	x	x	x	x		5.6		6.3		
Cescon et al. [29]	Subspace-based multi-step linear predictors		x	x				15.58		39.44	59.56	
Stahl et al. [31]	Autoregressive moving average	x	x					Se/ FPR* 96/12%		Se/ FPR* 96/16%	Se/ FPR* 74/ 21%	
Hidalgo et al. [5]	Genetic Programing		x	x				EGA* (A + B) 83.0%		EGA* (A + B) 63.3%	EGA* (A + B) 50.3%	
Mirshekarian et al. [42]	Recurrent neural network using Long short-term memory	x	x	x				21.4		38.0		
Contreras et al. [58]	Grammatical Evolution		x	x	x			21.19		31.34		
Bertachi et al. [60]	Artificial neural network	x	x	X	x			19.33		31.72		NH: acc* 90.1%, Se* 72.2% and Sp* 91.9%
He et al. [69]	Sparse-group lasso based recurrent neural network		x	x	x	x	0.29 mmol/L	0.47 mmol/L		0.91 mmol/L		
Koutny et al. [72]	Meta-Differential Evolution	x										93% of glucose levels with a Relative error* less than or equal to 15%
Vehi et al. [10]	Postprandial hypoglycemia prediction support vector machine	x	x	x								Se/Sp* 69.0/80.0%
	Nocturnal hypoglycemia prediction artificial neural network	x	x	x	x							Se/Sp* 44.0//85.9%

proceeding, seventeen studies (27%) were published in proceedings of conferences in the ranks of the magazines provided by Scimago Journal and Country Rank [92]. The majority of the studies (52.4%) were conducted in Europe.

Twenty-eight studies (44.4%) were cited by at least one or more papers selected in this study. Fig. 3 shows the number of intra-study citations per reference.

3.3. Study participants and design

Twenty-seven studies (42.9%) present small samples (less than 10 participants), nineteen studies (30.2%) with small-medium samples (between 11 and 50 participants) and seventeen studies (27%) with samples including over 50 participants. In thirty-five studies (55.6%) the data were collected or partially collected in free-living conditions. Twenty-nine studies (46.0%) deal with adult subjects, two studies (3.2%) deal with adult, adolescent, and children subjects and five studies (7.9%) deal with children and adolescent subjects. Only five studies treat the diabetes type-1 and type-2 and one treats the diabetes type-2. Some studies used datasets to apply their models: OhioT1DM (15.9%), DirectNet (7.9%), METABO (6.4%), DIAdvisor (6.4%), Diabetes UCI (4.8%), ABC4D (3.2%) and ChildrenDat (1.2%). Table 5 shows detailed information concerning these datasets and their availability.

3.4. Prediction outcome

The most popular topic is a short-time blood glucose prediction, which is addressed in fifty-two studies (82.5%). Regarding short-time predictions, three studies (4.8%) present nocturnal predictions and two studies (3.2%) a postprandial prediction. Nine studies (14.3%) present mid-term approaches and two studies present a long-term prediction. For mid-term prediction, one study was limited to fasting blood glucose predictions, five studies (7.9%) to nocturnal predictions and two studies (3.2%) to postprandial predictions.

3.5. Trends in the application of predictive methods

As shown in Tables 3 and 4, we clustered the included articles in terms of their models: 1) Data-based models and 2) Hybrid models. Fifty-one studies (80.9%) present data-based models and twelve studies (19.1%) present hybrid models. Table 6 shows the distribution for physiological models used in hybrid models.

Twenty-five (39.7%) were population-based models, thirty-six studies (57.1%) were individual-based, and two studies present both models. Table 7 shows the distribution of studies that are population-based or individual-based. This table also shows some advantages and disadvantages of each approach. The last column shows how the limitations may be overcome.

There are several approaches for the predictive task, including Artificial neural network (ANN) (22.2%), supervised learning (20.6%), statistics or probabilistic models (7.9%), autoregressive models (7.9%), evolutionary models (6.4%), adaptative filter models (3.2%), deep learning (4.8%), hybrid and ensemble models (27.0%).

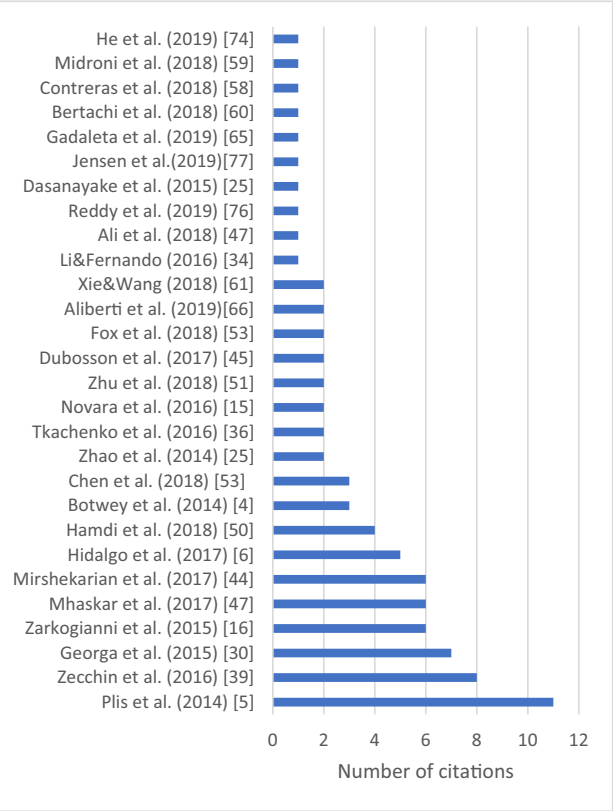

Fig. 3. Intra-study citations.

3.6. Inputs

The proper input selection is a crucial strategy design for successful predictions. Twenty-six studies (41.3%) used only glucose levels to support the predictions, thirty studies (47.6%) used meal and/or insulin information combined with glucose levels, twelve studies (19.1%) used exercise information combined with the inputs previously mentioned and seven studies (11.1%) used also time-based features. Eight studies (12.7%) used other inputs combined with the previous referred such as clinical data, circadian information, heart rate, skin temperature, skin conductance and sleep, stress, and work information. Table 8 shows the features extracted from original datasets.

For short-term predictions, it is desirable to find a reasonable compromise between accuracy and an adequate prediction horizon. In general, the results for a PH of 5 min are better than for a PH of 30 min but this does not mean that it presents a better intervention. So, most of short-term studies (79.4%) included in this study used PH between 5 min and 180 min. The mostly used PH are 15-, 30-, 45-, 60- and 120-min (reason why we present only these results in Tables 2 and 3). The mid-term prediction uses glucose profile in 4-,6-,9- and a 24-hour window. Only one study uses blind window.

3.7. Performance metrics

Most of work use error rate metrics to evaluate the performance of their models. Thirty-four studies (54.4%) used Root-mean-square error (RMSE) to present the error between observed and predicted BG. Thirteen studies (20.6%) used EGA (Clarke error grid analysis) to quantify the accuracy of patient estimates of their current BG as compared to the BG value obtained in their meter. Table 9 shows the different metrics used to evaluate the prediction performance. When a study compares different models, we report the metrics of the most accurate model (Tables 3 and 4).

Table 5
Detailed information concerning studies datasets and their availability.

Databases/ datasets	Content	Access
OhioT1DM	Eight weeks' worth of data for each of 12 people with type 1 diabetes; CGM blood glucose level every 5 min; blood glucose levels from periodic self-monitoring of blood glucose (finger sticks); insulin doses, both bolus and basal; self-reported meal times with carbohydrate estimates; self-reported times of exercise, sleep, work, stress, and illness; and physiological data from fitness bands.	Requires Permission Access [93]
DirectNet	Different protocols and different data are available. Data collected from children and adolescents with T1D. The Accuracy of Continuous Glucose Monitors in Children with Type 1 Diabetes (89 participants, CGM and CWB) A Randomized Trial to Assess the Effectiveness of the GlucoWatch Biographer in the Management of Type 1 Diabetes in Children (200 participants, CW2B and SMBG) The Effect of Exercise on the Development of Hypoglycemia in Children with Type 1 Diabetes (15 participants, nocturnal BG and during exercise) The Effect of Basal Insulin During Exercise on the Development of Hypoglycemia in Children with Type 1 Diabetes (50 participants, CGM during exercise) A Pilot Study to Evaluate the Navigator Continuous Glucose Sensor in the Management of Type 1 Diabetes in Children (30 participants, CGM) Evaluation of Counter-regulatory Hormone Responses during Hypoglycemia and the Accuracy of Continuous Glucose Monitors in Children with T1DM (25 participants, CGM) A Randomized Clinical Trial to Assess the Efficacy and Safety of Real-Time Continuous Glucose Monitoring in the Management of Type 1 Diabetes in Young Children (4 to <10 Year Olds) (140 participants, CGM) A Pilot Study to Assess the Feasibility of Real-Time Continuous Glucose Monitoring in the Management of Infants and Toddlers with Type 1 Diabetes (35 participants, CGM)	Requires Permission Access [46]
METABO	15 participants TD1 were monitored from 5 to 22 days in free living conditions (CGM, energy expenditure, food intake, insulin regime).	Not accessible [27]
DIAdvisor	34 participants, 3 days in-hospital trial (CGM, food intake and insulin).	Not accessible [30]
Diabetes UCI	70 participants T1D, SMBG, insulin, meal and exercise information.	Open Access [33]
ABC4D	75 participants T1D, more info in [44], [94] (CGM, food intake, insulin, alcohol intake, physical activity).	Not accessible
ChildrenDat	179 participants, 3.5 months in-hospital context (SMBG).	Not accessible [35]

4. Discussion

The principal goal of this systematic review was to identify, assess and analyse the state-of-the-art of data-based algorithms or models using diabetics real data for hypoglycemia and BG prediction. The following paragraphs discuss the previously defined research questions.

Table 6
Distribution for physiological models used in hybrid models.

Physiological model	Studies
Lehmann & Deutsch (1992) [11] - Glucose dynamics	[26], [28]
Hovorka et al. (2004) [7] - Meal dynamics	[10], [58], [60]
Wilinska et al. (2005) [8] - Insulin dynamics	[5], [10], [58], [60]
Tarin et al. (2005) [12] - Glucose dynamics	[26], [28]
Dalla Man et al. (2007) [6]/modified - Meal dynamics	[29]
Kovatchev et al. [95] (2009) - Glucose-Insulin dynamics	[29]
Duke 2009 [9]/modified - Glucose dynamics	[4], [42]
Ozaslan (2017) - Exercise dynamics	[10], [58], [60]
Other Glucose dynamics models	[31], [69], [72]

4.1. RQ1 – How is the prediction problem addressed?

According to the studies selected, the hypoglycaemia is rarely addressed as the objective of prediction. The prediction is performed on BG levels where low values, usually less than 70 mg/dl, were predicted using regression approaches. These approaches reflect the introduction of CGM devices which allow several short-term glucose predictions. Most studies use the past values of CGM measurements as input to predict the future values based on glucose profile in a determined PH. The PH used in the selected studies varies between 5 min and 180 min, however, the most used PH is 30 min because it is considered as the minimum time interval for a patient to have an efficient intervention in order to prevent an unwanted event [53]. The results of the studies show that an increase of PH leads to a deterioration in the accuracy of the predictions [3,66]. Nevertheless, some studies propose to resolve the problem in term of classification instead of regression. They use classifiers to detect some patterns. In addition, there are approaches that combine inputs such as glucose, meal, insulin, exercise, and others to learn and mapping inputs to pre-established classes or ranges. So, the classification problem was addressed as a prediction based on BG ranges or on hypoglycaemia/hyperglycaemia risk. Some authors limited the studies to a specific hypoglycaemia prediction, *e.g.*, nocturnal, or postprandial which led to a mid-term prediction due to the reduced sampling. The model proposed in Khan et al. [41] uses current and previous BG levels to predict the next day fasting BG. Schroeder et al. [40] present a long-term approach to predict the 6-months risk of severe hypoglycaemia. Misra-Hebert et al. [90] propose a long-term risk prediction for severe hypoglycaemia in a type-2 diabetes population. In Reddy et al. [76], two models were proposed to predict hypoglycaemia at the start of exercise. Another way to address the prediction problem is in terms of population-based or individual-based. Usually, the subject-based models perform better than population-based models, but they are sensitive to intra-patient variability. On the other hand, population-based models are suitable for the generalization of the prediction.

4.2. RQ2 – What are the features used based on real data?

The selection of proper inputs, according to the prediction problem, is crucial to create a model that performs well. In this review we limited the inputs to at least one of this data categories: blood glucose (from CGM or SMBG), insulin-, meal-, or exercise information. The reason is that most diabetics are outpatients and therefore some inputs like ECG, EEG, or clinical parameters, when used alone, are not suitable for a self-management solution. The distribution of inputs presented in Table 8 shows that several papers combine features from a different type of input. But there are some studies that use BG alone due to inter-subject variability in variables such as daytime meal, medication, and exercise. Therefore we highlight some studies that use feature selection methods to perform their models [4,5,62,69,73,77,83,14,15,26,28,34,37,56,59]. The predominance of the features based on BG in last 30 min is clear to PH of 30 and 60 min [26,28]. In addition, the contribution of other features (meal, insulin, exercise) is lower but not insignificant; their importance increases for predictions with PH of 60 min [28]. Zecchin et al. [37] show that, for postprandial predictions with a PH above 15 min, the combination of CGM values with meal and insulin intake achieves better results.

In this review, we highlight the integration of exercise and contextual information. Regarding exercise, Zarkogianni et al. [14] achieved better results adding exercise to glucose features. Physical activity combined with insulin and meal intake can increase the performance for predictions with a PH above 45 min [15]. Midroni et al. [59] achieved the better results using a combination of CGM with SMBG, meal, stress and exercise. Reddy et al. [76] show the most important features within the RF model for hypoglycaemia prediction. Higher heart rate at start of exercise, the increase of energy expenditure and the lower glucose values at the start of exercise tends to increase likelihood of hypoglycaemia [76]. Exercise features based on time explain better the glucose concentrations in the short-term predictions [28]. This latest study shows that time-based features can be relevant for predictions. Also, the use of the time of the day includes some novel features highly connected to glucose dynamics [28]. However, in He et al. [69] the results show that meal and insulin intake and time providing more dominant causal correlations than sleep and exercise on blood glucose inference. In Jensen et al. [77], the feature subset with the highest ROC-AUC is reached using Slope of linear regression and minimum value the evening (9–12 pm), minimum value of the first, second, or third night and body mass index at baseline. Based on selection methods, Dave et al. [83] show the importance of contextual features such as hour and day when observation was made.

4.3. RQ3 – What are the most used data-based algorithms?

Comparing with the review presented by Oviedo et al. [16] (from 2010 to 2016), on data-based models we notice the increase of models based on ANN [10,14,75,87,26,37,38,47,60,66,71,73], namely RNN [51,53,68,69,89] as well as models based on supervised learning [4,10,76–80,83,84,90,32,41,43,48,56,59,62,65]. The trend in data-based approaches shows that scientific community is still doing experiments with several classifiers of machine learning techniques. Some

Table 7
Distribution of studies per model approach.

Model approach	Studies	Advantages	Disadvantages	How to overcome limitations
Population-based: models trained on pooled data from many people.	[3], [10], [43], [45], [53], [56], [65], [66], [68], [72], [76], [77], [23], [78], [80], [82–85], [90], [31], [32], [34], [37], [39], [40], [42]	Generalize well. Catch different patterns for the same event. Increase the robustness of the model [3], [25], [66].	May not generalize well to specific groups of individuals if they are not appropriately represented in the training data set.	Clustering [32], [68].
Individual-based: models trained on an individual's specific data.	[4], [5], [29], [36], [38], [41], [47–49], [51], [52], [54], [10], [57–64], [69], [71], [13], [73–75], [79], [82], [87], [89], [91], [14], [15], [24–26], [28]	High performance. Personalized prediction model with continuous adaptation to changes [28]; Suitable in Artificial pancreas context.	The accuracy can be negatively affected by intra-subject variability if the model is trained on limited data sets. Solution limits the usability to one individual [66]. Increase the risk of overfitting [66].	Adaptative models updated with more data from the individual.

Table 8
Feature extraction.

Input	Features	Studies
BG	CBG and CGM-based features	[3], [4], [48], [52], [53], [59], [60], [71]– [73], [76], [77], [5], [78]– [80], [82]– [85], [87], [89], [10], [31], [37], [38], [43], [45], [47]
	SMBG	[32], [34], [39]– [41], [56], [59]
	Frequency bands from CGM	[23]
	Previous BG	[14], [24], [42], [49], [54]
	Current BG	[54]
	12 h before BG	[24]
	24 h before BG	[24]
	BG within the last 20, 30 or 60 min	[26], [28], [36], [74], [75]
	Total glucose inserted into plasma	[26]
	BG change	[14]
Insulin	Insulin pump data	[3], [76]
	Bolus insulin delivery or	[15], [23], [42], [49]
	Rapid acting insulin (units)	[4], [5], [54], [58], [59], [61], [73], [77], [82]– [85], [10], [91], [13], [24], [29], [31], [37], [40], [43]
	Insulin estimations	[26], [28], [60], [69]
	Insulin from basal rate	[42]
Meal	CHO estimate	[23], [43], [60], [69]
	CHO intake (g)	[4], [5], [58], [59], [61], [73], [82]– [85], [91], [10], [15], [24], [29], [37], [42], [49], [54]
	Number of meals	[77]
	Rate of appearance of meal-derived glucose in plasma	[26], [28]
Exercise	Euclidean norm of triaxial accelerations	[25]
	Energy expenditure	[14], [26], [28], [56], [76]
	Exercise estimations	[60], [69]
	Step counts	[15], [56], [58]
Other	Hour/Time index/other time-based features	[5], [28], [49], [59], [73], [77], [83]
	Circadian rhythm effect	[54]
	Sleep quality score and sleep estimations	[69]
	Heart rate	[56], [61], [73], [76], [91]
	Skin conductance	[73]
	Skin temperature	[73]
	Clinical parameters	[40], [90]
	Stress and work information	[59]
	Anthropometric/ demographic/historical information	[76], [77], [83], [90]
	Glucagon (dual hormone or not)	[76]

studies started to use deep learning [45,49,63,85] and evolutionary algorithms [5,48,72,82].

4.4. RQ4 – Which are the models that perform better?

To address a comparative evaluation, we undertake the discussion in terms of 1) studies using the same datasets, 2) studies presenting nocturnal prediction, 3) studies presenting postprandial prediction and 4) studies with same prediction horizon results.

1) Regarding the studies performed on the DirectNet dataset [46], composed of data from children and adolescents, the models based on deep learning [45] and ANN [71,75] show better results than the probabilistic/statistic models [74]. The studies performed on the OhioT1DM dataset [50] achieve similar results using models based on deep learning [49,63], ANN [51,73], evolutionary algorithms [58] and ensemble model [59]. In studies using METABO dataset [27], the better results are achieved by the hybrid models [26,28], the features extracted using the physiological model probably contribute to better perform models.

Table 9
Prediction metrics used in selected studies.

Metric	Definition	Studies
Root mean square error (RMSE) and Mean squared error (MSE)	Measure of the differences between values predicted by a model and the values observed.	[4], [14], [41], [47]– [49], [51], [52], [54], [57]– [59], [23], [60], [61], [63], [68–69], [71], [73]– [75], [79], [80], [25], [85], [87], [89], [91], [26], [28], [32], [36], [38], [40]
Time lag or Time gain	Interval of time between one event and another related event.	[3], [26], [37], [52], [63], [64,85], [91]
Correlation coefficient	The strength of the relationship between the relative movements of two variables.	[3], [14], [31], [64], [79], [85]
Relative error	Measure of the uncertainty of measurement compared to the size of the measurement.	[29], [38], [72]
Accuracy (acc)	Refers to how close a measurement is to the true value	[10], [41], [60], [76]
Sensibility (se)	Measures the proportion of actual positives that are correctly identified.	[3], [4], [10], [31], [34], [60], [62], [64], [65], [76], [77], [83]
Specificity (sp)	Measures the proportion of actual negatives that are correctly identified.	[10], [34], [60], [62], [76], [77], [83]
Positive predictive value (PPV)	Proportions of positive results in statistics and diagnostic tests that are true positive results.	[34], [39], [76], [77]
Negative predictive value (NPV)	Proportions of negative results in statistics and diagnostic tests that are true negative results.	[34], [39], [76], [77]
False positive rate (FPR)	Ratio between the number of negative events wrongly categorised as positive (false positives) and the total number of actual negative events.	[3], [4], [31], [64]
F1-score	The harmonic means of the precision and recall.	[4], [34], [39], [43], [65]
F2-score	The weighted average of precision and recall.	[34]
Correct/incorrect	Correct (true positive + true negative) and incorrect (False positive + false negative)	[24]
FIT indexes	Measures the percentage of data that are correctly explained by the model.	[13], [66]
c-statistic	Measure of goodness of fit for binary outcomes in a logistic regression model.	[40]
Odd ratio	Quantifies the strength of the association between two events	[90]
Clarke Error Grid analysis (EGA)	Quantify clinical accuracy of patient estimates of their current blood glucose as compared to the blood glucose value obtained in their meter.	[5], [29], [82], [84], [87], [36], [45], [52], [54], [57], [58], [66], [68]
Mean absolute percentage error (MAPE)	Measure of prediction accuracy of a forecasting method in statistic	[47], [56], [74], [79], [80], [85], [87]
Mean absolute relative difference (MARD)	Mean absolute difference divided by the arithmetic mean.	[16], [34], [59], [63], [67], [76]
Coefficient of Determination (COD)	Proportion of the variance in the dependent variable	[15], [32], [47], [66], [74]

(continued on next page)

Table 9 (*continued*)

Metric	Definition	Studies
	that is predictable from the independent variable(s)	
Sum squares of the prediction error	Summary measure of the fit of a model to a sample of observations that were not themselves used to estimate the model.	[38]
Area under the ROC Curve (AUC)	Measures the entire two-dimensional area underneath the entire ROC curve	[62], [76], [78]
Surveillance Error Grid	Define an evaluation criterion that accounts for the clinical risk of making an incorrect prediction	[89]

2) Relatively to nocturnal prediction, we highlight the studies [31,37,60]: the authors of [31] propose an hybrid model based on Autoregressive moving average (ARMA) [31], the authors of based on BG and Insulin as inputs. And the authors of [60] propose an ANN model taking as inputs the BG, Insulin, meal and physical activity.
3) For the postprandial predictions, the studies [10,37,43,62] use BG, insulin, and meal as inputs. Zecchin et al. [37] achieve good results for a PH of 15 min using ANN. Oviedo et al. [62] and Vehi et al. [10] obtain similar results using Support vector Regressor (SVR) and support vector machine (SVM) respectively for a 4-hour PH window. Dubosson et al. [43] also used SVM but the model does not perform well for a 2-hour PH window.
4) In terms of PH, all studies presenting results for PH of 15 min achieved good results, excepted for He et al. [74] using canonical-correlation analysis. This kind of model find linear combinations between input and output, we can explain their bad results with a non-linear profile of BG. For a 30 min PH, the models based on ANN [14,37,47,66,75], namely recurrent neural network (RNN) [55], deep learning [45], ensemble models [3,48] and hybrid models [5,26,28], [69] achieve good results. The models based on ANN [37,47,66], ensemble models [48] and hybrid models [28,69] present also good results for PH ranging between 45 and 60 min. Some work use time lag to evaluate their method, we highlight the results of Aliberti et al. [66]. This work achieved a time lag of zero minutes for PH of 30 and 45 min, 5 min for PH of 60 and 30 min for PH of 90. In terms of sensibility Yang et al. [64] shows excellent performance results above 99%.

4.5. RQ5 – What about replicability and availability of methods of the studies to scientific community?

With the increasing number of complex methods using several sources of data, the availability of algorithms and datasets can be very useful for the scientific community working n future events or glucose concentration prediction. According to this systematic literature review we identify three databases/datasets that are available on free access: 1) OhioT1DM Dataset, 2) DirectNet datasets and 3) UCI Diabetes dataset. OhioT1DM Dataset allow a perfect replicability of methods as already divided in train and test datasets. DirectNet database has eight studies with public datasets and adequate for this field. However, the selected papers using these database does not precise which dataset was used, or which patients are used. UCI Diabetes Dataset need some data manipulation so, it can be more difficult to replicate the proposal of the scientific papers.

Regarding algorithms, most studies present a detailed description of their methods, including sources, parameters, frameworks, and libraries. However, only few authors [73,89] share their source code implementation, so, its recommendable to request permission access.

To conclude the discussion, we notice that since 2017 much more papers were published. We link that to the first Blood Glucose Level Prediction Challenge promoted by a Workshop on Knowledge Discovery in Healthcare Data. The second Blood Glucose Level Prediction Challenge was realized in 2020, so, we expect that the number of works will grow again.

Compared to the reviews of [16,17] published respectively between 2010 and 2016 and between 2000 and 2018, we highlight that the trend is to add new inputs to models in particular exercise input. In the present review we underline the integration of contextual information and time-based features (features defined with respect to the time [28]). The availability of monitoring devices such as smartphone, smartwatch, wristband, and their applications contribute to a greater availability of physiological and contextual data.

The review of [16] concluded that most studies lacked clinical evidence because the models are validated with simulated data. Our review included 63 studies, all of them using real patient data to model validation.

The limitation pointed out by the reviews [16,17] is the low free access data availability. Regarding the replicability of studies, we suggest that OhioT1DM Dataset allow a perfect replicability of methods reason why this dataset was suggested for the Blood Glucose Level Prediction Challenge.

5. Conclusion

This systematic review summarized the existing data-based models using real data for hypoglycemia prediction. Sixty-three studies were analyzed, and the main findings are summarized as follows:

- (RQ1) The hypoglycaemia prediction has been addressed, massively, as a blood glucose prediction. Most studies used short-term glucose predictions to show the model performance based on a prediction horizon between 5 min and 180 min. Some studies limited the hypoglycaemia or blood glucose prediction according to the time of day, *e.g.* nocturnal or postprandial.
- (RQ2) The use of inputs like meal, insulin and physical activity data combined with glucose data is common but some studies use only glucose data due to inter-variability in day-time meals, medication, and exercise. Some studies showed the importance of some features relatively to others using feature selection techniques or testing different combinations of features.
- (RQ3) The trend in the application of predictive methods based on ANN, namely RNN and supervised learning methods is increasing. We also noticed that some studies started to use deep learning and evolutionary algorithms.
- (RQ4) Generally, the models based on ANN, namely RNN and hybrid models showed good predictive results. Deep learning and evolutionary models showed potentialities.
- (RQ5) From the data and datasets used in the studies only OhioT1DM Dataset allow a perfect replicability and although the methods are well detailed, few studies share their algorithm.

Comparing our results with the results of other reviews, we see a trend of increasing studies in this area promoted by workshops and conferences, an increase in studies that use exercise, time-based and contextual information as inputs as well as a greater motivation to use real data for model's validation. However, the problem of available and public access data remains.

The growth in data-based models for blood glucose and hypoglycaemia prediction is encouraging and studies show promising results. The trend in the application of predictive methods show that there are some open possibilities on data- and model-fusion.

Finally, some limitations of this review should be mentioned. First, we limited this review in terms of real participants. There are interesting approaches using in-silico patients is not yet considered Second, we exclude studies using inputs when not combined with the data as

mentioned in inclusion criteria. Currently, some biosignals can be collected in a home-care context. Third, only English-language publications were included.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by FCT/MEC through national funds and co-funded by the FEDER—PT2020 partnership agreement under the project UIDB/50008/2020 (Este trabalho é financiado pela FCT/MEC através de fundos nacionais e cofinanciado pelo FEDER, no âmbito do Acordo de Parceria PT2020 no âmbito do projeto UIDB/50008/2020). This article is based upon work from COST Action IC1303–AAPELE–Architectures, Algorithms and Protocols for Enhanced Living Environments and COST Action CA16226–SHELD-ON–Indoor living space improvement: Smart Habitat for the Elderly, supported by the European Cooperation in Science and Technology (COST). More information can be found at www.cost.eu (accessed on 31 March 2021).

References

[1] W. H. Organization. Global status report on noncommunicable diseases 2014. 2014.
[2] I. D. Federation. IDF Diabetes Atlas 8th edition [Online]. Available, https://diabetesatlas.org/IDF_Diabetes_Atlas_8e_interactive_EN/. [Accessed 27 June 2019].
[3] Botwey RH, Daskalaki E, Diem P, Mougiakakou SG. Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events. In: 2014 36th annu. int. conf. IEEE eng. med. biol. soc. EMBC 2014; 2014. p. 4843–6.
[4] Plis K, Bunescu R, Marling C, Shubrook J, Schwartz F. A machine learning approach to predicting blood glucose levels for diabetes management. Mod Artifical Intell Heal Anal 2014:35–9.
[5] Hidalgo JI, Colmenar JM, Kronberger G, Winkler SM, Garnica O, Lanchares J. Data based prediction of blood glucose concentrations using evolutionary methods. J Med Syst 2017;41(9).
[6] Man CD, Rizza RA, Cobelli C. Meal simulation of glucose-insulin system. IEEE Trans Biomed Eng 2007;54(10):1–33.
[7] Hovorka R, et al. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. In: Physiological measurement; 2004.
[8] Wilinska ME, Chassin LJ, Schaller HC, Schaupp L, Pieber TR, Hovorka R. Insulin kinetics in type-1 diabetes: continuous and bolus delivery of rapid acting insulin. IEEE Trans Biomed Eng 2005;52(1):3–12.
[9] Duke DL. Intelligent diabetes assistant: a telemedicine system for modeling and managing blood glucose. 2010.
[10] Vehí J, Contreras I, Oviedo S, Biagi L, Bertachi A. Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning. Health Informatics J 2019:703–18. https://journals.sagepub.com/doi/10.1177/1460458219850682.
[11] Lehmann ED, Deutsch T. A physiological model of glucose-insulin interaction in type 1 diabetes mellitus. J Biomed Eng 1992;14:235–42.
[12] Tarín C, Teufel E, Picó J, Bondia J, Pfleiderer H. Comprehensive pharmacokinetic model of insulin glargine and other insulin formulations 2005;52(12):1994–2005.
[13] Novara C, Pour NM, Vincent T, Grassi G. A nonlinear blind identification approach to modeling of diabetic patients. IEEE Trans Control Syst Technol 2016;24(3):1092–100.
[14] Zarkogianni K, et al. Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring. Med Biol Eng Comput 2015;53(12):1333–43.
[15] Faccioli S, Ozaslan B, Garcia-Tirado JF, Breton M, Favero S Del. Black-box model identification of physical activity in type-l diabetes patients. In: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2018. p. 3910–3.
[16] Oviedo S, Vehí J, Calm R, Armengol J. A review of personalized blood glucose prediction strategies for T1DM patients. Int J Numer Methods Biomed Eng 2017;33(6):1–21.
[17] Woldaregay AZ, et al. Data-driven modeling and prediction of blood glucose dynamics: machine learning applications in type 1 diabetes. Artif Intell Med 2019;98:109–34.
[18] Moher D, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7).
[19] Liberati A, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7).
[20] Higgins U, et al. "Cochrane handbook for systematic reviews of interventions version 6.0," Cochrane. In: [Online]. Available; 2019. www.training.cochrane.org/handbook.
[21] Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5(1):1–10.
[22] Daskalaki E, Nrøgaard K, Zuğer T, Prountzou A, Diem P, Mougiakakou S. An early warning system for hypoglycemic/hyperglycemic events based on fusion of adaptive prediction models. J Diabetes Sci Technol 2013;7(3):689–98.
[23] Zhao C, Dassau E, Zisser HC, Jovanovič L, Doyle FJ, Seborg DE. Online prediction of subcutaneous glucose concentration for type 1 diabetes using empirical models and frequency-band separation. AIChE J 2014;60(2):574–84.
[24] Efendic H, Kirchsteiger H, Freckmann G, Re L del. Short-term prediction of blood glucose concentration using interval probabilistic models. In: 22nd Mediterranean conference on control and automation; 2014. p. 1494–9.
[25] Dasanayake IS, Seborg DE, Pinsker JE, Doyle FJ, Dassau E. Empirical dynamic model identification for blood-glucose dynamics in response to physical activity. In: Proc. IEEE conf. decis. control, vol. 54rd IEEE, no. Cdc; 2015. p. 3834–9.
[26] Georga EI, Protopappas VC, Polyzos D, Fotiadis DI. Online prediction of glucose concentration in type 1 diabetes using extreme learning machines. In: Proc. annu. int. conf. IEEE eng. med. biol. soc. EMBS, vol. 2015-Novem; 2015. p. 3262–5.
[27] Georga E, et al. Data mining for blood glucose prediction and knowledge discovery in diabetic patients: the METABO diabetes modeling and management system. In: Proceedings of the 31st annual international conference of the IEEE engineering in medicine and biology society: engineering the future of biomedicine, EMBC 2009; 2009.
[28] Georga EI, Protopappas VC, Polyzos D, Fotiadis DI. Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models. Med Biol Eng Comput 2015;53(12):1305–18.
[29] Cescon M, Johansson R, Renard E. Subspace-based linear multi-step predictors in type 1 diabetes mellitus. Biomed Signal Process Control 2015;22:99–110.
[30] F. E. project. DIAdvisor, "personal glucose predictive diabetes advisor" [Online]. Available, https://cordis.europa.eu/project/rcn/85459/factsheet/en. [Accessed 3 September 2019].
[31] Stahl F, Johansson R, Olsson ML. Predicting nocturnal hypoglycemia using a non-parametric insulin action model. In: Proc. - 2015 IEEE int. conf. syst. man, cybern. SMC 2015; 2016. p. 1583–8.
[32] Li J, Fernando C. Smartphone-based personalized blood glucose prediction. ICT Express 2016;2(4):150–4.
[33] U. Machine Learning Repository. Diabetes data set [Online]. Available, https://archive.ics.uci.edu/ml/datasets/diabetes. [Accessed 20 July 2005].
[34] Tkachenko P, Kriukova G, Aleksandrova M, Chertov O, Renard E, Pereverzyev SV. Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application. Comput Methods Programs Biomed 2016;134:179–86.
[35] E. H. Funded M. AMMODIT, "ChildrenData dataset.". 2020.
[36] Georga EI, Principe JC, Polyzos D, Fotiadis DI. Non-linear dynamic modeling of glucose in type 1 diabetes with kernel adaptive filters. In: 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2016. p. 5897–900.
[37] Zecchin C, Facchinetti A, Sparacino G, Cobelli C. How much is short-term glucose prediction in type 1 diabetes improved by adding insulin delivery and meal content information to CGM data? A proof-of-concept study. J Diabetes Sci Technol 2016;10(5):1149–60.
[38] Hamdi T, et al. Artificial neural network for blood glucose level prediction. In: 2017 int. conf. smart, monit. control. cities, SM2C 2017; 2017. p. 91–5.
[39] Kriukova G, Shvai N, Pereverzyev SV. Application of regularized ranking and collaborative filtering in predictive alarm algorithm for nocturnal hypoglycemia prevention. In: Proc. 2017 IEEE 9th int. conf. intell. data acquis. adv. comput. syst. technol. appl. IDAACS 2017, vol. 2, no. September; 2017. p. 634–8.
[40] Schroeder EB, Xu S, Goodrich GK, Nichols GA, O'Connor PJ, Steiner JF. Predicting the 6-month risk of severe hypoglycemia among adults with diabetes: development and external validation of a prediction model. J Diabetes Complications 2017;31(7):1158–63.
[41] Khan T, Masud MA, Mamun KA. Methods to predict blood glucose level for type 2 diabetes patients. In: 2017 IEEE region 10 humanitarian technology conference (R10-HTC); 2017. p. 392–5.
[42] Mirshekarian S, Bunescu R, Marling C, Schwartz F. Using LSTMs to learn physiological models of blood glucose behavior. In: 2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2017. p. 2887–91.
[43] Dubosson F, Mordvanyuk N, Ibáñez B López, Schumacher M. Negative results for the prediction of postprandial hypoglycemias from insulin intakes and carbohydrates: analysis and comparison with simulated data. 2017. p. 25–9.
[44] Reddy M, et al. Clinical safety and feasibility of the advanced bolus calculator for type 1 diabetes based on case-based reasoning: a 6-week nonrandomized single-arm pilot study. Diabetes Technol Ther 2016;18(8):487–93.
[45] Mhaskar HN, Pereverzyev SV, van der Walt MD. A deep learning approach to diabetic blood glucose prediction. Front Appl Math Stat 2017;3, no. July:1–11.
[46] J. center for health Research, "Diabetes research in children network (DirecNet)." [Online]. Available: https://public.jaeb.org/direcnet.
[47] Ali J Ben, Hamdi T, Fnaiech N, Costanzo V Di, Fnaiech F, Ginoux JM. Continuous blood glucose level prediction of type 1 diabetes based on artificial neural network. Biocybern Biomed Eng 2018;38(4):828–40.
[48] Hamdi T, Ali J Ben, Costanzo V Di, Fnaiech F, Moreau E, Ginoux JM. Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm. Biocybern Biomed Eng 2018;38(2):362–72.

[49] Zhu T, Li K, Herrero P, Chen J, Georgiou P. A deep learning algorithm for personalized blood glucose prediction. In: CEUR workshop proc. vol. 2148; 2018. p. 64–78.
[50] Marling C, Bunescu R. The OhioT1DM dataset for blood glucose level prediction. In: CEUR Workshop Proceedings; 2018.
[51] Chen J, Li K, Herrero P, Zhu T, Georgiou P. Dilated recurrent neural network for short-time prediction of glucose concentration. In: CEUR workshop proc. vol. 2148; 2018. p. 69–73.
[52] Li N, Tuo J, Wang Y. Chaotic time series analysis approach for prediction blood glucose concentration based on echo state networks. In: 2018 Chinese control and decision conference (CCDC); 2018. p. 2017–22.
[53] Fox I, Ang L, Jaiswal M, Pop-Busui R, Wiens J. Deep multi-output forecasting learning to accurately predict blood glucose trajectories. In: Proc. ACM SIGKDD int. conf. knowl. discov. data min; 2018. p. 1387–95.
[54] Griva LO, Basualdo MS. Evaluating clinical accuracy of models for predicting glycemic behavior for diabetes care. In: 2018 Argentine conference on automatic control (AADECA); 2018. p. 1–6.
[55] U. of Virginia. Feasibility of a decision support system to reduce glucose variability in subject with T1DM [Online]. Available, https://clinicaltrials.gov/ct2/show/NCT02558491?term=NCT02558491&draw=2&rank=1.
[56] Vahedi MR, et al. Predicting glucose levels in patients with type1 diabetes based on physiological and activity data. In: Proc. 8th ACM MobiHoc 2018 work. pervasive wirel. healthc. work. mobilehealth 2018; 2018. p. 1–5.
[57] Yu X, et al. Model-fusion-based online glucose concentration predictions in people with type 1 diabetes. Control Eng Pract 2018;71(April 2017):129–41.
[58] Contreras I, Bertachi A, Biagi L, Oviedo S, Vehí J. Using grammatical evolution to generate short-term blood glucose prediction models. In: CEUR workshop proc. vol. 2148; 2018. p. 91–6.
[59] Midroni C, Leimbigler PJ, Baruah G, Kolla M, Whitehead AJ, Fossat Y. Predicting glycemia in type 1 diabetes patients: experiments with XGBoost. In: CEUR workshop proc. vol. 2148; 2018. p. 79–84.
[60] Bertachi A, Biagi L, Luo N, Veh J. Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks. 2018.
[61] Xie J, Wang Q. Benchmark machine learning approaches with classical time series approaches on the blood glucose level prediction challenge. In: CEUR workshop proc. vol. 2148; 2018. p. 97–102.
[62] Oviedo S, Contreras I, Quirós C, Giménez M, Conget I, Vehi J. Risk-based postprandial hypoglycemia forecasting using supervised learning. Int J Med Inform 2019;126, no. March:1–8.
[63] Li K, Liu C, Zhu T, Herrero P, Georgiou P. GluNet: a deep learning framework for accurate glucose forecasting. IEEE J Biomed Health Inform 2019:1.
[64] Yang J, Li L, Shi Y, Xie X. An ARIMA model with adaptive orders for predicting blood glucose concentrations and hypoglycemia. IEEE J Biomed Health Inform 2019;23(3):1251–60.
[65] Gadaleta M, Facchinetti A, Grisan E, Rossi M. Prediction of adverse glycemic events from continuous glucose monitoring signal. IEEE J Biomed Health Inform 2019;23(2):650–9.
[66] Aliberti A, et al. A multi-patient data-driven approach to blood glucose prediction. IEEE Access 2019;7:69311–25.
[67] JDRF continuous glucose monitoring clinical trial [Online]. Available, http://diabetes.jaeb.org/RT_CGMRCTProtocol.aspx.
[68] Dong Y, Wen R, Li Z, Zhang K, Zhang L. Clu-RNN: a new RNN based approach to diabetic blood glucose prediction. In: 2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB); 2019. p. 50–5.
[69] He M, Gu W, Kong Y, Zhang L, Spanos CJ, Mosalam KM. CausalBG: causal recurrent neural network for the blood glucose inference with IoT platform. IEEE Internet Things J 2019:1.
[70] Gu W, et al. SugarMate: non-intrusive blood glucose monitoring with smartphones. 2017. p. 1–27.
[71] Idrissi T El, Idri A, Abnane I, Bakkoury Z. Predicting blood glucose using an LSTM neural network. In: Proc. 2019 fed. conf. comput. sci. inf. syst. vol. 18; 2019. p. 35–41.
[72] Koutny T, Cioppa A Della, Falco I De, Tarantino E, Scafuri U, Krcma M. De-randomized meta-differential evolution for calculating and predicting glucose levels. In: 2019 IEEE 32nd international symposium on computer-based medical systems (CBMS); 2019. p. 269–74.
[73] Mirshekarian S, Shen H, Bunescu R, Marling C. LSTMs and neural attention models for blood glucose prediction: comparative experiments on real and synthetic data. In: 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC); 2019. p. 706–12.
[74] He J, He T, Wang Y. Blood glucose concentration prediction based on canonical correlation analysis. In: Chinese control conf. CCC, vol. 2019-July; 2019. p. 2942–7.
[75] Chen X, Tuo J, Wang Y. A prediction method for blood glucose based on grey wolf optimization evolving kernel extreme learning machine. In: Chinese control conf. CCC, vol. 2019-July; 2019. p. 3000–5.
[76] Reddy R, Resalat N, Wilson LM, Castle JR, Youssef J El, Jacobs PG. Prediction of hypoglycemia during aerobic exercise in adults with type 1 diabetes. J Diabetes Sci Technol 2019;13(5):919–27.
[77] Jensen MH, Dethlefsen C, Vestergaard P, Hejlesen O. Prediction of nocturnal hypoglycemia from continuous glucose monitoring data in people with type 1 diabetes: a proof-of-concept study. J Diabetes Sci Technol 2019;14(2):250–6.
[78] Vu L, et al. Predicting nocturnal hypoglycemia from continuous glucose monitoring data with extended prediction horizon. AMIA Ann Symp Proc 2019;2019:874–82.
[79] Alfian G, Syafrudin M, Rhee J, Anshari M, Mustakim M, Fahrurrozi I. Blood glucose prediction model for type 1 diabetes based on extreme gradient boosting. IOP Conf Ser Mater Sci Eng 2020;803(1).
[80] Anand PK, Shin DR, Memon ML. Adaptive boosting based personalized glucose monitoring system (PGMS) for non-invasive blood glucose prediction with improved accuracy. Diagnostics 2020;10(5):1–22.
[81] Harman-Boehm I, Gal A, Raykhman AM, Naidis E, Mayzel Y. Noninvasive glucose monitoring: increasing accuracy by combination of multi-technology and multi-sensors. J Diabetes Sci Technol 2010;4(3):583–95.
[82] Contador S, Velasco JM, Garnica O, Hidalgo JI. Profiled glucose forecasting using genetic programming and clustering. Proc ACM Symp Appl Comput 2020:529–36.
[83] Dave D, et al. Feature-based machine learning model for real-time hypoglycemia prediction. J Diabetes Sci Technol 2020:1–13.
[84] Hidalgo JI, et al. Glucose forecasting combining Markov chain based enrichment of data, random grammatical evolution and bagging. Appl Soft Comput J 2020;88:105923.
[85] Li K, Daniels J, Liu C, Herrero P, Georgiou P. Convolutional recurrent neural networks for glucose prediction. IEEE J Biomed Health Inform 2020;24(2):603–13.
[86] Reddy M, et al. Clinical safety and feasibility of the advanced bolus calculator for type 1 diabetes based on case-based reasoning: a 6-week nonrandomized single-arm pilot study. Diabetes Technol Ther 2012;66(0):37–9.
[87] Li N, Tuo J, Wang Y, Wang M. Prediction of blood glucose concentration for type 1 diabetes based on echo state networks embedded with incremental learning. Neurocomputing 2020;378:248–59.
[88] Tamborlane WV, et al. JDRF randomized clinical trial to assess the efficacy of real-time continuous glucose monitoring in the management of type 1 diabetes: research design and methods. Diabetes Technol Ther 2008;10(4):310–21.
[89] Martinsson J, Schliep A, Eliasson B, Mogren O. Blood glucose prediction with variance estimation using recurrent neural networks. J Healthc Informatics Res 2020;4(1):1–18.
[90] Misra-Hebert AD, et al. Risk prediction for severe hypoglycemia in a type 2 diabetes population with previous non-severe hypoglycemia. J Diabetes Complications 2020;34(1).
[91] Xie J, Wang Q. Benchmarking machine learning algorithms on blood glucose prediction for type i diabetes in comparison with classical time-series models. IEEE Trans Biomed Eng 2020;67(11):3101–24.
[92] (n.d.) SCImago, "SJR — SCImago Journal & Country Rank [portal]." [Online]. Available: http://www.scimagojr.com.
[93] The OhioT1DM dataset [Online]. Available, http://smarthealth.cs.ohio.edu/OhioT1DM-dataset.html.
[94] I. C. London. Advanced bolus calculator for type 1 diabetes (ABC4D) (ABC4D) [Online]. Available, https://clinicaltrials.gov/ct2/show/record/NCT02053051.
[95] Kovatchev BP, Breton M, Man C Dalla, Cobelli C. In silico preclinical trials: a proof of concept in closed-loop control of type 1 diabetes. J Diabetes Sci Technol 2009;3(1):44–55.

Chapter 3

Hypoglycaemia prediction framework

The chapter presents and discusses concepts related to hypoglycaemia predictive framework and is presented as two articles, one titled "Hypoglycaemia prediction models with auto explanation.", and the other one is titled "A novel context-aware framework for hypoglycaemia events prediction.".

3.1 Hypoglycaemia prediction models with auto explanation.

Virginie Felizardo, Diogo Machado, Nuno Garcia, Nuno Pombo and Pedro Brandão.

Hypoglycaemia prediction models with auto explanation.

Accepted for publication in IEEE Access, 2021, in the area of Computer Science.
H−Index 127, Q1, by Scimago Journal & Country Rank.

This article presents a updated state-of-the-art on reasoning and prediction models related with either blood glucose level or hypoglycaemia events in order to discuss an architecture that combines a predictive model and a reasoning system. For the purpose of this thesis, only what is related to predictive models fits the proposed objectives.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 13, 2021, accepted September 25, 2021, date of publication October 4, 2021, date of current version June 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2021.3117340

Hypoglycaemia Prediction Models With Auto Explanation

VIRGINIE FELIZARDO[1,2], DIOGO MACHADO[1,3], NUNO M. GARCIA[1,2], NUNO POMBO[1,2], AND PEDRO BRANDÃO[1,3], (Member, IEEE)

[1]Instituto de Telecomunicações, 1049-001 Lisboa, Portugal
[2]Universidade da Beira Interior, 6201-001 Covilhã, Portugal
[3]Universidade do Porto, 4099-002 Porto, Portugal

Corresponding author: Virginie Felizardo (virginie@it.ubi.pt)

This work was supported in part by FCT/MCTES through National Funds, and in part by the FEDER-PT2020 partnership agreement under Project UIDB/50008/2020.

ABSTRACT World-wide statistics show a considerable growth of the occurrence of different types of Diabetes Mellitus, posing diverse challenges at many levels for public health policies. Some of these challenges may be addressed by means of computerised systems which may pave the way to provide practitioners with insight on their patient's conditions anywhere and at anytime, but also to empower Diabetes patients as managers of their health. These systems for disease management come in many shapes and sizes, being the most promising trends the ones that involve expert systems that comprise specialised knowledge, use predictive models, feature engineering and reasoning. This study presents the state-of-the-art on reasoning and prediction models related with either blood glucose level or hypoglycaemia events. The main findings revealed are that there is room for improvement on predictive models, namely to enhance its accuracy and ability to forecast future events into a wider time frame. On the other hand, reasoning models are understudied and its usage in Diabetes management is reduced. We discuss an architecture that combines a predictive model and a reasoning system, with the objective of alerting of impending occurrences and interpret the current situation to accurately advise the diabetic user.

INDEX TERMS AI, eHealth, expert-systems, data-mining, prediction models, diabetes mellitus.

I. INTRODUCTION

Diabetes is a chronic disease that in 2019 affected approximately 463 million adults around the world [36]. Despite all the efforts to stale the growth of this disease, the number of people with diabetes is continuously increasing and it is estimated to reach 700 million in 2045 [36]. Alongside these numbers there is the economic impact. Diabetes is reported to have caused expenditures of 727 billion USD dollars in 2017 [36]. These facts challenge for innovative solutions capable not only mitigating the origin of new cases, but also providing sustainable medical practices for those already affected by diabetes. Diabetes is characterised by high levels of glucose in the blood, caused by the person's pancreas only producing little or no insulin. If glycaemic values are not controlled, diabetes can seriously decrease the quality of life. In the worst cases, this disease leads to blindness, amputation or heart problems. There are different types of diabetes where type 2 and type 1 are the most common. Each type of diabetes has a specific treatment associated. Given the diabetic's inability to produce insulin, a strict regime and periodic or even continuous verification of glycaemic values is recommended. This regime evolves and adapts each time the diabetic consults a doctor. In these appointments, the medical expert evaluates the evolution of the glycaemic values, among other relevant annotated information, and personalises the current regime, to the diabetic's needs. This process is not optimal since it requires the availability of a medical expert to receive the person, evaluate the data and refine the treatment. Real-time events such as hyperglycaemia (when concentration of Blood Glucose (BG) is high) or hypoglycaemia (when BG is low),[1] require immediate action which poses a possible problem in the absence of expert knowledge.

Research has tried to provide support directly to the patient, embedding this knowledge at the patient side.

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

[1]These thresholds may vary according to the patient's characteristics: before/meal, age, disease development and others. Usually hypoglycaemia is a BG below 70 mg/dL (3.9 mmol/L) and hyperglycaemia a BG higher than 170 mg/dL (9.4 mmol/L) after a meal [56], [57].

Predictive models have been increasingly adopted to cope with hypoglycaemia events. The objective is to empower patients before an event occurs and thus to support them on the decision-making required to mitigate or to avoid the subsequent symptoms. However, in current research, the reasons for the hypoglycaemia to occur are not addressed. Reasoning about why a hypoglycaemia will occur, may help avoid it and educate the patient on a better control. This last point concerns the area of expert systems for health.

The objective of predictive algorithms can be divided into three purposes: improving the health care of the population, increasing patient experience, and adding value to health care. A predictive algorithm can be applied to some challenging healthcare scenarios such as: promoting self-management by status interpretation and health education, individual change detection, future event prediction or assisting medical decision-making. Predictive algorithms are based on various techniques, including data mining, statistical, modelling and artificial intelligence. For developing these models different pre-processing techniques, various feature extraction techniques and classifiers are applied.

This approach attempts to glance the future. Considering the user's past data, it trains specialised models to foresee the consequences of the user's current actions. Doing so, these systems create a window of opportunity for the user to avoid possible unwanted occurrences. Therefore, the discovery of patterns associated with future events in combination with feature engineering to understand which features allow reasoning about this field is crucial development for diabetes management.

Reasoning systems, are a part of computer science that given a goal attempts to make logical inferences automatically [32], [73]. In the case of diabetes management, these systems can be viewed as a medical support tool. Diabetes management is a complex task that requires not only expert knowledge, but also experience. There are multiple factors that influence glycaemic values, and often these factors vary from person to person. With this in mind, many projects try to specialise in a particular aspect of diabetes management, insulin calculation being the most popular. In general, reasoning systems, applied to diabetes, seek to advise and assist their users as an expert would.

In this article we will first describe what is the current state of the art for hypoglycaemia prediction models and reasoning approaches for diabetes management. We will then discuss the shortcomings of the current methodologies. After which we will argue that both techniques need to be improved and coupled to provide the diabetic patient with a good support for diabetes management.

II. METHODOLOGY

In this work, the authors searched and reviewed the state of the art of both prediction and reasoning models applied to diabetes. The models focus on different subjects, as a consequence, search parameters differ, although the search engines used were the same. The authors used the ScienceDirect, IEEE Xplore, ACM Digital Library, SCOPUS and PubMed databases to retrieve relevant peer-reviewed publications.

This study aims to provide an up-to-date state-of-the-art on data-based or hybrid models applied to predict either blood glucose levels or hypoglycaemia events by means of data collected from patients with diabetes.

We searched titles and abstracts using: "*hypoglycemia prediction*" OR "*hypoglycemia detection*" OR "*blood glucose prediction*" OR "*blood glucose estimation*" OR "*hypoglycemia estimation*" OR "*blood glucose detection*", for the databases ScienceDirect, IEEE Xplore, SCOPUS and PubMed; and (*hypoglycemia* OR "*blood glucose*") AND (*prediction* OR *detection* OR *estimation*), for ACM Digital Library. We used the US version of hypoglycemia for searching as it provides more results. In this paper we use the UK version to conform to UK English.[2] The inclusion criteria were: real data from participants; collected on clinical or daily living contexts; with at least one of these data categories, comprising BG levels (from Continuous Glucose Monitoring (CGM) or Self-Monitoring Blood Glucose (SMBG)), insulin, meal, or exercise information. Information from the selected publications was extracted taken into account the following characteristics:

- Study Information: This defines the study's citation and year of publication;
- Data: This defines the type of diabetes, sample (number of participants, age range or average age and gender distribution) and source (collected, clinical trial, EHR or dataset);
- Prediction: This defines the application where the algorithm is being exploited. It can be short or mid-term predictions, nocturnal, postprandial or other;
- Algorithms: This defines the different approaches used to train the model;
- Input: This was defined to assess the inputs used to develop the algorithm. This includes BG, meal, insulin, exercise or others.

After a search and an eligibility review, 33 studies were included for data extraction and qualitative synthesis.

The reasoning models that fit our criteria are expert-systems with the ability to support patients on the decision-making. In this case, the authors did not filter works considering implementation platform or published date. For the reasoning models, we searched titles and abstracts considering the keywords: "*expert-system diabetes*" OR "*expert system diabetes*" OR "*reasoning diabetes*". A careful review of all the titles and abstracts associated with this search was performed. To filter out non-pertinent studies the following inclusion criteria were applied: original studies, studies directed for management of diabetes type I containing user advice. We have obtained 50 results related to expert systems and diabetes management in the initial search. From these we have considered 19 to fit our inclusion criteria.

[2]We added hypoglycemia to this article's keywords.

Information from the selected publications was extracted taken into account the following characteristics:

- Algorithm: This defines the different methods used by the given approach;
- Input: The input required by the given approach in order to function correctly;
- Output: Defines the application's feedback to the users. It can be insulin recommendations, glycaemia read recommendations, generic recommendations or treatment recommendations.

III. MODELS AND SYSTEMS FOR GLYCAEMIA PREDICTION

The majority of models described in this section aim to predict glycaemic values. The accuracy and time frame varies depending on the predictive model. Some of these works [10], [25], [51], [61], [74] predict, classify and label glycaemic values. Others aim to predict hypoglycaemia events [9], [16], [37], [63], [68].

BG prediction is the estimation of future BG values based on present and past inputs. For this, several approaches have been proposed by the scientific community: physiological models, data-based approaches, and a combination of both, the hybrid models [34].

In the physiological model, the overall BG dynamic is usually characterised into: meal dynamics [15], [35], insulin dynamics [19], [71], exercise model [68] and glucose dynamics [19], [39], [66]. These models require previous knowledge to set the physiological constants [34], [58].

The data-based models are in general more accurate [58]. These models have been proposed using pattern recognition techniques and experimental data. However, there may be exceptions as shown in Mirshekarian *et al.* [52].

The major challenge for the predictive models is related to the variable impact that the collected variables have on the glycaemia value. The effects of insulin intake, meals, physical activity and other events in glucose dynamics are different for every person (inter-subject variability) and even in the same person there are changes over time (intra-subject variability) [22]. According to this, the models for BG can be categorised into two groups: individual-based models trained on an individual's specific data, and population-based models trained on pooled data from many people expecting they will generalise well enough to allow them to be used in previously unseen individuals.

Despite the efforts in this area, no specific method provides total reliable event predictions, thus an increasing number of approaches have been proposed for inputs and techniques fusion in the hope to represent spatial and temporal input–output dependencies. The availability of personal devices and wearable devices allowed the data collection of several individual inputs provided by the user. Among them, we find CGM systems, smartphones, smartwatches, wristbands, among others. It allows the collection of real data on clinical or free-living context. Some studies have used simulated data (from virtual patients) using BG software simulators AIDA[3] and UVA/Padova.[4] CGM systems usually measure BG levels at a fixed instant time, e.g., every 3 or 5 minutes, as such, the majority of studies work with time-series approaches. Past reviews [23], [59], [72] in this subject show that uni-variate approaches are commonly used predicting the future BG only based on past BG data. But these reviews also show that studies started to combine different data such as therapies, meal, exercise, and contextual information. Different datasets were used but few are accessible, e.g., OhioT1DM [46] and DirectNet datasets.[5] There are several approaches for the predictive task, including Artificial Neural Networks (ANN), supervised learning, statistics or probabilistic models, auto-regressive models, evolutionary models, adaptive filter models, deep learning, hybrid, and ensemble. Table 1 shows some information about each study: year of publication, data information and prediction model. Most studies between 2014 and 2020 used ANN based models, deep learning, and ensemble models. Table 2 shows detailed information about the prediction model, inputs category and prediction outcome of ANN based models. ANN approaches are used due to its great capacity to model the various non-linear and non-stationary glucose dynamics. Several approaches of ANN are addressed, including extreme learning machine (ELM), jump neural networks (JNN), multi-layer perception (MLP), recurrent neural network (RNN), and self-organising map (SOM). As an emerging theory, ELM presents a better generalisation performance, a faster learning speed, a good performance in regression applications and in large datasets or multi-class application. According to Mirshekarian *et al.* [52], ANN approaches, with gradient-based learning methods and back propagation, are difficult to train due to vanishing gradients, a problem that is often compounded by small datasets. So, some approaches start to use Long Short Term Memory (LSTM) units, which are not affected by the vanishing gradient problem, embedded into Deep Learning (DL) frameworks they can capture the complex dynamics system, particularly, when it is difficult to derive the mathematical expressions of the system. Table 2 show detailed information about the prediction model, inputs category and prediction outcome of DL based models.

Despite the good results and potentialities of ANN based and DL based approaches, some authors continue to use supervised learning (SL) approaches. Random Forest (RF) and Support vector machine (SVM) are commonly used for glucose concentration prediction [16], [63], [68]. RF presents some advantages useful for this kind of predictions: reduce overfitting, flexible to both classification and regression problems and works well with both categorical and numeric attributes. SVMs also present interesting advantages for this task, e.g., the reliability of Support Vector Recursion (SVR) for predictions [30]. Table 2 shows detailed information about the prediction model, inputs category and prediction

[3] AIDA http://www.2aida.org/
[4] UVA/Padova https://tegvirginia.com/software/t1dms/
[5] DirectNet https://public.jaeb.org/datasets/diabetes

TABLE 1. Studies information, year, data and prediction model.

Study	Data			Prediction model	
	DM Type	Sample	Source	Category	Approach
Botwey et al. (2014)	TDI	23 participants, 17–70 years	Collected (daily living)	Ensemble	(cARX+RNN)+data fusion
Plis et al. (2014)	TDI	5 participants	Collected	Ensemble	SVR+ARIMA
E I Georga et al. (2015)	TDI	15 participants	METABO dataset (daily living)	ANN	KELM
Zarkogianni, K. et al. (2015)	TDI	10 participants, 42 years (7M/3F)	METABO dataset (daily living)	ANN	SOM
Eleni I. Georga et al. (2015)	TDI	15 participants	METABO dataset (daily living)	Ensemble	Genetic Programming+RF
Zecchin et al. (2016)	TDI	15 participants	DIAdvisor dataset (clinical context)	ANN	JNN
Mirshekarian et al. (2017)	TDI	10 participants	Collected	Deep Learning (DL)	RNN using LSTM
Mhaskar, Pereverzyev and Walt (2017)	TDI	25 participants, <18 years	DirecNet dataset (clinical context)	DL	Deep Network
Ben Ali et al. (2018)	TDI	12 participants	Collected (daily living)	ANN	ANN
Fox et al. (2018)	TDI	40 participants	Collected (daily living)	ANN	Sequential polynomial multi-output RNN
Bertachi et al. (2018)	TDI	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset (daily living)	ANN	ANN
Hamdi et al. (2018)	TDI	12 participants	Collected (daily living)	Ensemble	SVR+Differential evolution
He et al. (2019)	TDI/TDII	112 participants, >35 years (49.1%F)	Collected	Ensemble	Sparse-group lasso based RNN
Li et al. (2019)	TDI	10 participants, 47 years and 6 participants, 40–60 years (2M/4F)	ABC4D (both context) and OhioT1DM datasets (daily living)	DL	dilated convolutional ANN
Aliberti et al. (2019)	TDI	451 participants, 8–25 years and >25 years (55%F)	Clinical trial	DL	LSTM network
Dong et al. (2019)	TDI/TDII	80 participants	Collected	Ensemble	Clustering pre-process based RNN
Chen, Tuo and Y. Wang (2019)	TDI	5 participants	DirecNet dataset (clinical context)	ANN	kernel ELM
Reddy et al. (2019)	TDI	55 participants, 33 ± 6 years (33F/22M)	Collected (controlled exercise environment)	Supervised Learning (SL)	Random Forest (RF)
Jensen et al. (2019)	TDI	463 participants	Clinical trial (daily living)	SL	linear discriminant function
De Bois, Yacoubi and Ammi (2019)	TDI/TDII	11 participants	Collected and OhioT1DM dataset (daily living)	DL	pcLSTM
W. Wang, Tong and Yu (2020)	TDI	56 participants	RTCGM dataset (daily living)	Ensemble	VMD+improved PSO+LSTM
Alessandro Aliberti et al. (2020)	TDI	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset (daily living)	DL	LSTM ANN
Alfian et al. (2020)	TDI	12 participants, <18 years	DirecNet dataset (clinical context)	ANN	MLP
Mohebbi et al. (2020)	TDI	50 participants	Collected (daily living)	DL	LSTM RNN
Pavan et al. (2020)	TDI	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset (daily living)	Ensemble	Shallow ANN +Imputation
Gu, Dang and Prioleau (2020)	TDI	34 participants and 6 participants, 40–60 years (2M/4F)	Collected and OhioT1DM dataset (daily living)	DL	LSTM Neural physiological encoder
Nemat et al. (2020)	TDI	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset (daily living)	Ensemble	MLP+LSTM+PLSR
Montaser et al. (2020)	TDI	18 participants	Collected (clinical context)	Ensemble	Global Seasonal Model (fuzzy approach)
Cappon et al. (2020)	TDI	6 participants, 40–60 years (2M/4F)	OhioT1DM dataset (daily living)	DL	bidirectional LSTM
Contador et al. (2020)	TDI	10 participants	Collected (daily living)	Ensemble	Genetic programming+clustering
Vehí et al. (2019)	TDI	10 participants, 41 years/ 6 participants, 40–60 years (2M/4F)	Collected and OhioT1DM dataset (daily living)	SL	ANN and SVR
D. Dave et al. (2020)	TDI	112 participants, 1–21 years (48F/64M)	Collected (daily living)	SL	RF
Kriventsov, Lindsey and Hayeri (2020)	TDI	1200 participants and 6 participants, 40–60 years (2M/4F)	Collected (daily living) and OhioT1DM dataset (daily living)	Ensemble	boosted Tree+SVR

outcome of SL-based models. Nevertheless, every prediction algorithm has its own advantages. So, sometimes it is necessary to combine different prediction methods to cover the disadvantages. Table 2 shows detailed information about the prediction model, inputs category and prediction outcome of ensemble approaches. In this category we find different approaches of ensembles. In order to address the complexity of factors affecting glucose response some studies [13], [18], [54] used clustering associated with other kind of models to better characterise scenarios with similar responses. Other studies used regression models and auto-regressive models (Auto-regressive model with output correction (cARX) or Auto-Regressive Integrated Moving Average (ARIMA)) in combination with other models [10], [25], [33], [54], [61]. He *et al.* [33] used a sparse group lasso to mine the underlying causal features. Georga *et al.* [27], [38], [60] used ensembles

TABLE 2. Algorithms, inputs and prediction outcome.

Approach	Study	Algorithm	Inputs					Predict. Outcome
			BG	I	CHO/M	PA	P	
ANN-Based	E I Georga et al. (2015)	ELM, kernel ELM, online sequential ELM and online sequential ELM kernels.	x	x	x	x	x	Short term
	Zarkogianni, K. et al. (2015)	Neuro-fuzzy network (seven layers) using a gradient-based algorithm with an adaptative learning rate; activation function: wavelets; membership function: Gaussian function.	x			x		Short term
	Zecchin et al. (2016)	JNN that is a feedforward neural network with inputs directly connected to both the first hidden layer and the output layer.	x	x			x	Nocturnal
	Zecchin et al. (2016)	JNN that is a feedforward neural network with inputs directly connected to both the first hidden layer and the output layer.	x	x	x		x	Postprandial
	Ben Ali et al. (2018)	ANN using algorithm Levenberg-Marquardt; input activation function: tansig function and output activation function: linear function.	x					Short term
	Bertachi et al. (2018)	Eight-layer network using a hyperbolic tangent as hidden-layer activation model and linear function as output-layer activation, the training algorithm used was Levenberg-Marquardt.	x	x	x	x	x	Short term
	Bertachi et al. (2018)	Sixteen-layer network using a hyperbolic tangent as hidden-layer activation model and logarithmic sigmoid function as output-layer activation, the training algorithm used was Scaled conjugate gradient.	x	x	x	x	x	Nocturnal
	Chen, Tuo and Y. Wang (2019)	Kernel ELM combined with gray wolf optimization	x					Short term
	Alfian et al. (2020)	MLP using a grid search algorithm to select the best parameters: two hidden layer network with 100 neurons each, activation function: rectified linear unit (ReLu); weight optimization: Adam.	x					Short term
	Vehí et al. (2019)	ANN-based model with adaptative synthetic sampling algorithm; 6-hour window	x	x	x	x	x	Nocturnal
DL-Based	Mirshekarian et al. (2017)	RNN architecture that uses a single LSTM layer in the hidden layer, with 5 nodes.	x	x	x		x	Short term
	Mhaskar, Pereverzyev and Walt (2017)	Semi-supervised deep learning neural network with a judge predictor based on the function approximation on data-defined manifolds, using diffusion polynomials.	x					Short term
	Li et al. (2019)	GluNet architecture, based on five-layer dilated convolutional neural network. For the first three layers, the hidden units are set to 32, while the top two layers have 64 neurons. The sliding window with a size of 16 time-steps.	x	x	x			Short term
	Aliberti et al. (2019)	LSTM network consisted of a layer of 30 LSTM units and a single output layer, with a number of units equal to the future glucose samples that need to be predicted.	x					Short term
	De Bois, Yacoubi and Ammi (2019)	Prediction-Coherent LSTM RNN consisted of a single hidden layer of 128 LSTM units, two-output architecture and its associated cMSE loss function. The coherence factor of 2 has been optimized through grid search to ensure a good trade-off between the accuracy of the predictions and the accuracy of the predicted variations.	x	x	x			Short term
	Alessandro Aliberti et al. (2020)	The structure consists of 30 inputs, a layer composed of 50 cells and an output layer. The Adaptive Moment Estimation (Adam) was used as optimization algorithm obtained the optimal number of 2000 iterations, and the learning rate 0.001.	x					Short term
	Mohebbi et al. (2020)	population-based LSTM using a sliding window (using 20-minute steps) of BG values was used as input in batches of 64 windows. A Bayesian optimization implemented in SigOpt was used to efficiently explore a predefined bounded hyperparameter space for LSTMs, dropout level of the MLP layer, and historical window size. Networks are optimized by the Rectified Adam (RAdam) gradient descent method	x					Short term
	Gu, Dang and Prioleau (2020)	Neural Physiological encoder (NPE), which can combine with any downstream neural network model and trained end-to-end. NPE was combined with LSTM, gated recurrent unit and RNN.	x	x	x			Short term
	Cappon et al. (2020)	bidirectional LSTM architecture, that consists of a four-layer neural network: a bidirectional LSTM input layer composed of 128 cells having a look back period of 15 minutes (i.e. 3 samples), two LSTM layers respectively composed of 64 and 32 cells, and a fully connected layer consisting of a single neuron computing the BG level prediction.	x	x	x	x		Short term
SL-Based	Reddy et al. (2019)	RF	x	x		x		Short-term during exercise protocol
	Jensen et al. (2019)	LDA	x	x	x		x	Mid-term nocturnal
	Vehí et al. (2019)	SVM	x	x	x			Postprandial
	D. Dave et al. (2020)	RF	x	x	x		x	Short term
Ensemble-Based	Botwey et al. (2014)	Autoregressive model with output correction and an RNN in combination with data-fusion techniques based on Dempster-Shafer evidential theory (DST), genetic algorithm (GA) and genetic programming (GP).	x	x				Short term
	Plis et al. (2014)	SVR (using Physiological features) + ARIMA (additional features)	x	x	x		x	Short term
	Eleni I. Georga et al. (2015)	Gaussian Process and Random Forest	x	x	x	x	x	Short term
	Fox et al. (2018)	Sequential polynomial multi-output RNN	x					Short term
	Hamdi et al. (2018)	Support Vector Regression based on Differential evolution	x					Short term
	Dong et al. (2019)	Clustering pre-process based RNN	x					Short term
	He et al. (2019)	Sparse-group lasso based RNN		x	x	x	x	Short term
	W. Wang, Tong and Yu (2020)	Variational modal decomposition+ improved Particle swarm optimization + LSTM	x					Short term
	Nemat et al. (2020)	Multilayer perceptron + LSTM+ Partial least squares regression.	x			x		Short term
	Pavan et al. (2020)	shallow neural network and an error imputation module (NN-EIM); a Bayesian optimization procedure returns the best method to train the ensemble of trees (i.e., Bagging or Boosting).	x	x				Short term
	Contador et al. (2020)	Genetic Programming+clustering (CHAID-GP)	x	x	x			Short and Mid-term
	Kriventsov, Lindsey and Hayeri (2020)	gradient boosted decision trees and Support Vector Machine (SVM) regression	x	x	x	x		Short term
	Montaser et al. (2020)	Seasonal AutoRegressive Integrated Moving Average including eXogenous variables (SARIMAX) local model is built for each cluster (Fuzzy C-Means). Box-Jenkins methodology is used to identify a seasonal model for each set.	x	x	x	x		Postprandial

Notes: This table shows the algorithm and inputs (blood glucose (BG), insulin (I), meal information (CHO/M), Exercise information (E) and features from physiological models (P)) used to perform the predictions.

based on decision trees because this kind of ensemble handles higher dimensional data very well (Bagging Tree, Boosting Tree or RF). Wang *et al.* [70] used variational mode decomposition (VMD) to reduce non-stationarity of BG time series and the Particle Swarm Optimization (PSO) algorithm is used to optimise these LSTM parameters. Nemat *et al.* [55]

applied a stacked regression to enhance the performance of BG prediction. This technique uses predictions from several models (MLP, LSTM and Partial Least Squares Regression (PLSR)) as features to train a new model based on PLSR.

The feature extraction and feature selection are crucial techniques to obtain good and optimised prediction algorithms and to understand the effects of some inputs in the predictions in order to later try to reason about these effects. The meals content of individual diets and how carbohydrates affect BG is one of the basis of diabetes treatment. Proper exercise plays an important role in BG control and reduces the risk of cardiovascular events. Reasoning is targeted inference where inputs are carefully selected, organised, and the inference is used to generate the desired result. Table 3 shows different approaches of how BG prediction studies select features and the improvements achieved.

From the studies with feature selection, we can conclude that some information like insulin, meals, exercise, and time of the day are relevant when combined with BG to improve the performance of the models. Regarding the prediction outcome, it is common to use the Prediction Horizon (PH). This can define the models according to two categories: *short term*, when the PH is less than 180 minutes; and *mid-term*, when the PH is more than 180 minutes and less than 24 h. PH is the future time window in which the predicted glucose concentration is determined by a model. The thresholds of each category do not have a standard definition so, in this literature study we used the above thresholds. Some recent studies focused on specific periods of the day, providing prediction of glucose concentration or events: nocturnal [9], [37], [68], [75], postprandial [54], [68], [75] and during exercise protocol [63].

IV. REASONING APPROACHES

As important to know that something will happen, is to know why it will happen. By analysing the user's data thoroughly, it is possible to retrieve knowledge that can help users to improve their glycaemic control. If available to medical experts, the conclusions obtained by these methods, can even be used to better understand the user's needs, and adapt the user's current diabetes' monitoring plan.

Expert knowledge is an ever-increasing necessity. In various areas such as engineering, finance, science and medicine, experts are a fundamental piece. As our base knowledge increases, the necessary expertise to handle this knowledge also increases. In some cases, teams of experts are needed to handle and work with a particular system. Expert knowledge is therefore an uttermost valuable asset. With today's demand of experts, it is unreasonable to expect an expert to be perpetually available. Nonetheless, certain occurrences require the continuous attention of an expert.

Expert systems are automatic consulting systems [1] that attempt to emulate an expert's decision-making [67]. Expert systems have been implemented in numerous areas from agriculture [21], [62] to computer security [43] and health [49], [69]. In the scope of diabetes, expert-systems have been, for a long time, a useful tool to interpret data and retrieve information. For instance, the SESAM-DIABETE project, an interactive educational expert-system capable of providing personalised advice and therapeutic recommendations for insulin diabetic patients was available in 1989 [41]. Table 4 shows details about different studies: the used algorithm, the required input, and the output.

Generally, expert-systems related to diabetes focus on bolus adjustments. These systems, evaluate the user's data, trace the user's profile, calculate and recommend insulin dosage adjustments [5], [6], [14], [20], [24], [40], [65]. Projects like the DIABETES [6] system can even explain, to users, the therapeutic recommendations given. The treatment recommendations made by the system were evaluated against the recommendations of an expert. The results showed that in 22% of the cases there was a full disagreement; 31% of the cases had one parameter in disagreement and finally, in 47% of the cases the medical expert fully confirmed the expert system's conclusions. The authors explain that in some cases, a conclusion that has resulted in a total disagreement may not be wrong, in some cases the system suggests a correct, but more complex route that diverges from the more direct expert's approach. Other expert-system approaches, related to diabetes, can also advise basal rate adjustments [64] and advise changes to meal and insulin administration scheduling [40].

Less frequent, there are approaches that aid users to better their glycaemic management, giving general advice, which is usually abstract or related to meal, exercise and glucose value test guidelines [7]. Advising in the diabetes context is a complex task, since it will influence the user's health. Possibly, because of this, most approaches avoid specifying what the user should do, instead, the application will show the user general best practices as advice [28], [31], [48], [50]. The project 4 Diabetes Support System [47] tries a different approach. This system is a case-based approach to diabetes self-management. In this project several problems related to diabetes are identified and defined as cases. Cases are constituted by an identified problem, a determined solution to the problem and a verification of the problem's resolution. This case is then translated to logical rules. The created rules identify the problem, apply the determined solution and verify its success. In order to update and evolve this system, the impact of the solution is verified. If the known solution does not solve the identified problem or if a better solution is uncovered, then it is possible to reach the case and update its parameters to translates the new, better, solution. In this manner, it is possible to advise diabetics with a good level of confidence. This project was developed during seven years and had three clinical research studies [47]. The project MyDiabetes [44][6] is a mobile application that contains an expert-system. This system did not follow a case-based approach, instead, the authors translated existing medical rules and guidelines to logical rules, that compose the

[6]MyDiabetes https://mydiabetes.dcc.fc.up.pt/ (in Portuguese).

TABLE 3. Feature selection techniques.

Study	Technique	Feature selected	Improvements
Plis et al. (2014)	SVR using physiological features and other using also ARIMA features.	Physiological features and ARIMA features	Performance
E I Georga et al. (2015)	Different combination of inputs: 1) glucose concentration within the last 30 minutes; 2) with meal-derived glucose rate, total glucose, plasma insulin concentration; and 3) with energy expenditure	Case 3: all features	Performance
Eleni I. Georga et al. (2015)	Random forest and RReliefF are used to rank the 46 candidate features.	20 features: CGM data, glucose rate after meal, plasma insulin concentration, energy expenditure and time of the day.	Feature importance/ Performance
Zarkogianni, K. et al. (2015)	Different combination of inputs: 1) glucose concentration and glucose change; and 2) with energy expenditure.	Case 2: all features	Performance
Zecchin et al. (2016)	Different combination of inputs: 1) CGM, 2) CGM and insulin, 3) CGM and CHO and 4) CGM, insulin and CHO	Postprandial prediction: case 4 and Nocturnal prediction: case 2	Performance/ Target importance
He et al. (2019)	10 physiological features and CHO, insulin, energy expenditure and sleep quality score. Causal feature learning using sparse group lasso.	Food, insulin and time features provide more dominant causal correlations than those of sleep and exercise.	Feature importance/ Performance
Jensen et al. (2019)	forward selection to select features from 32 candidate features	Slope of linear regression and minimum value the evening (9–12 pm) before the event, minimum value the first, second, or third night before the event and body mass index at baseline.	Feature importance/ Performance
Alfian et al. (2020)	MLP using last 30 minutes BG or with addiction of time domain features	last 30 minutes BG and time domain features	Performance
Pavan et al. (2020)	CGM readings, CGM slope and IOB and additional features selected with ReliefF: first order differences, at several time lags, of: CGM, IOB, COB, sleep/work period, skin temperature and acceleration data.	CGM readings and additional features selected with ReliefF.	Feature importance/ performance
Cappon et al. (2020)	generated the power set, dynamic risk; insulin; correction boluses; CHO; physical activity	Personalised feature set	Personalised features/ performance

Notes: The features in the table are: carbohydrate ingestion (CHO), insulin on board (IOB), carbohydrates on board (COB)

expert-system. This approach takes advantage of the records introduced by the user in the mobile application, to evaluate the user's state and advise accordingly. Although there are important contributions made by the different presented projects, there is a lack of results on the impact these systems have on glycaemic control.

A limitation regular expert systems have is their static nature. The rules and knowledge integrated in the expert system are not dynamic and can only be changed after a software update. Considering the learning abilities of data-mining and machine-learning, these systems, if integrated within an expert system could give it the opportunity to adapt and better the knowledge contained in it. The "4 diabetes project" [47] has two branches dedicated to the use of data-mining for prediction. During the project's second study, a common diabetes problem was addressed, glycaemic variability. This blood control problem is linked to hypoglycaemia unawareness and to oxidative stress which leads to long-term diabetic complications. The system can identify twelve types of glycaemic variability, using multiple different rules. It was not possible to create a generic rule capable of identifying this problem. The authors refer that there are numerous metrics for glycaemic variability characterisation, but there is no consensus on a technique to be used in clinical practice. Still, for physicians, the identification of this problem becomes trivial when in the presence of a BG plot. In order to reach a global method for identification of glycaemic variability, the authors decided to apply machine learning. To create a system capable of recognising glycaemic variability, the authors considered the quantifiable aspects of this problem. The metric chosen to quantify the glycaemic variability was the Mean Amplitude of Glycaemic Excursion (MAGE). This metric captures the distance between the local maximum and minimum of a BG plot. Considering this metric insufficient, the authors devised two other metrics: the distance travelled and the excursion frequency. Distance travelled captures the overall daily fluctuations, and excursion frequency counts the number of significant glucose deviations in a day. Two physicians, also authors in this work, classified BG plots as excessively variable or not, based on their expert knowledge. The same plots were also scored by MAGE and the two other metrics. The classifications obtained for 218 BG plots were then used to train multiple machine learning algorithms. Among the tested algorithms, the naive Bayes classifier used was able to match 85% of the physicians' ratings.

The "4 Diabetes Support System" also addressed glycaemic value prediction. To achieve this goal, SVR was used. The choice of SVR is bound to its ability of incorporating contextual features, without the assumption of feature independence. The results obtained showed that SVR is capable of achieving better results than a baseline model that uses the present BG as the prediction for future values.

These prediction systems' results, if integrated in an expert-system, could be used to prevent future occurrences, but also as input data. For the expert system, knowing that certain occurrences are reasonably believed to happen, can be even more meaningful than simple data. In the case of the referred systems, the expert-system would be able to obtain data about glycaemic variability, and a projection of future glycaemic values. With this information, the system could be pro-active and better advise the user. Additionally, the system could use this information to refine variables such as insulin sensitivity or carbohydrate ratios.

V. DISCUSSION

The field of prediction models and reasoning applied to diabetes is rich and varied in approaches. Considering the described research, in general, most options tend to focus on

TABLE 4. Reasoning approaches.

Study	Algorithm	Input	Output
Ambrosiadou, Alevizos and Ziakas (1993) and Ambrosiadou, Goulis and Pappas (1996)	Rule-based	BG, administrated insulin (type/quantity), hypo/hyper episodes, sleep	insulin recommendations, glycaemia read recommendations
Cosenza (2012)	Fuzzy-based	carbohydrates, proteins, lipids, prepandrial glycaemia, insulin	insulin recommendations
Fortwaengler et al. (2013)	undisclosed	carbohydrates	insulin recommendations
Lehmann, T. Deutsch et al. (1994)	Rule-based reasoning and mathematical modelling	clinical data, BG data, insulin data, special event data, nutritional data	insulin recommendations
El Fathi et al. (2020)	Maximum-a-posteriori method	insulin basal dose Units (U), Insulin bolus dose Units (U), Amount of carbohydrates in consumed meal (g), Sensor glucose mmol/L, Optimal basal insulin dose Units (U), Optimal carbohydrate ratio g/U, Basal insulin sensitivity mmol/L/U, Time-to-peak of insulin action min, Carbohydrate sensitivity mmol/L/g, Time-to-peak of carbohydrate absorption min, Duration effect of insulin basal dose min, Transfer-rate constant between plasma and interstitial glucose	Insulin doses recommendations
Rudi and Celler (2006)	Rule-based	glycaemic values, Insulin, Exercise	Insulin recommendations, Generic recommendations
Al-Ghamdi et al. (2011)	Undisclosed, cloud-computing	age, gender, weight, diabetes type, glycaemic values, HbA1c	Generic recommendations
McCausland et al. (1999)	Rule-based, Fuzzy-based	current BG level, anticipated food intake, exercise level, prescribed insulin dosage	Generic recommendations
Hashemi (2012)	Rule-based	hypoglycaemic symptoms, hyperglycaemic symptoms, diabetes type, fasting glycaemia, postprandial (2h), random glycaemia, HbA1c	Generic recommendations
Mbogho, J. Dave and Makhubele (2013)	Rule-based	diabetes related questions	Generic recommendations
Marling, Wiley et al. (2011)	Case-based	daily BG, life-event data, insulin	Treatment recommendations
Machado et al. (2017)	Rule-based	Glycaemic values, insulin, carbohydrates, exercise, HbA1c, weight, cholesterol, blood pressure, disease events	Treatment recommendations, general recommendations

aiding, or solving a particular diabetes related issue. Only a few, such as the 4diabetes project attempt to include different modules that encompass different diabetes related issues. The prediction models approaches contribute to identify the features that affect the patterns associated with risk of future events.

A. PREDICTION MODELS

Prediction models have made significant progress in transforming available data and clinical information into valuable new knowledge allowing new patterns discovery and what-if scenarios definition. These new findings can give some reasoning about the effects of concurrent actions. The distribution of inputs presented in the tables of the different approaches shows that several papers combined features from different types of input. However, there are some studies that use BG alone due to inter-subject variability in variables. Therefore, we highlight some studies that use different sources for feature and contextual information integration. Features based on BG in the last 30 minutes is clear to have a PH of 30 and 60 minutes [2], [26], [27]. The contribution of other features, like meal, insulin or exercise, are lower but not insignificant [27]. Their importance increases for predictions with a PH of 60 minutes [27]. Zecchin *et al.* [75] show that, for postprandial predictions, the combination of CGM values with meal and insulin intake achieves better results. Regarding exercise, Zarkogianni *et al.* [74] achieved better results adding exercise to glucose features. Reddy *et al.* [63] show that the most important features for hypoglycaemia prediction are higher heart rate at start of exercise, the increase of energy expenditure and the lower glucose values at the start of exercise. Exercise time-based features explain better the glucose concentrations in short-term predictions [27]. The study by Pavan *et al.* [60] used some additional features like insulin-on-board and carbohydrates-on-board to carry information about the dynamics of slow insulin absorption and carbohydrates slow impact on BG. Similarly, this study used physical exercise-on-board to describe physical activity intensity. The integration of the Error Imputation Module using these additional features improved the predictions. Cappon *et al.* [11] presented personalised features to give personalised predictions. The most important features are CGM and insulin or the correction boluses (all patients have these features). For some patients other features are added depending on what PH (30 or 60 minutes) is aimed for. For one patient the combination of features is the same for the two PH (CGM, insulin and reported meals). In Alfian *et al.* [2] the contribution of time domain features shows potentialities for predictions. Also, the use of *time of day* includes some novel features highly connected to glucose dynamics [27]. However, in He *et al.* [33], the results show that meal and insulin intake and time provide more dominant causal correlations than sleep and exercise on BG inference. In Jensen *et al.* [37] the feature subset with the highest results is reached using four features: slope of linear regression and minimum value of the evening (9–12 pm) before the event, minimum values of the three nights before the event and body mass index at baseline. Based on selection methods, Dave *et al.* [16] show the importance of contextual features such as hour and day when observation was made. Few studies [47], [63] address reasoning or explanations about predictions results. This is crucial to understand how other features like insulin, meal and exercise influence the BG dynamics and their effects on prediction models. From selected studies, Reddy *et al.* [63] present a rule for recommendation during aerobic exercise, which they identified after the prediction during physical exercise. This recommendation consists of: if the glucose is below 180 at the beginning of the physical exercise and the heart rate above 120 during the exercise, there is risk of hypoglycaemia. We will discuss reasoning and its importance in the next section.

B. REASONING APPROACHES

Traditional standalone expert-systems for diabetes management and counselling that are comprised of static systems either rule-based, case-based, or fuzzy, are currently in decline.

Recent approaches to this topic utilise data-mining and machine learning algorithms, more autonomous and dynamic to create modules that can then be used to advise diabetic patients. These approaches are usually specialised, attempting to tackle specific diabetes related issues. Regarding the approaches presented previously, they focus on two topics: bolus adjustments [5], [6], [14], [20], [24], [40], [65] and glycaemic advice [7], [28], [31], [48], [50] for better glycaemic control. Most do not have concrete evaluations of the impact of their approach. Nonetheless, the ones that disclose their results, prove that these systems have a positive impact on the user's diabetes management.

The ''4 diabetes project'' [45], [47], in contrast, possesses reasoning models and prediction models. The system is composed of a case-based reasoning system for glycaemic control and therapeutic adjustments, a glycaemic variability classifier, and a blood-glucose prediction component. The different modules were not tested as a unit. Marling *et al.* [45] report accuracy levels of 77.5% in a first test, and 97.9% in a second test, for the case-based system. The glycaemic variability classifier, the best classifier, also evaluated in two tests, obtained a 93.8% accuracy. Finally, the prediction model, being a work-in-progress, was only described, not mentioning practical results. Nonetheless, as the authors convey, this module, once available, could be valuable to take preventative actions.

Reasoning modules can be accurate and beneficial to detect and act on particular occurrences. Despite this, diabetes management consists of more than singular occurrences. The authors believe that, in order to help diabetic patients to better manage their diabetes, it is necessary to combine both reasoning and prediction in a more complete approach.

C. IMPROVE AND COMBINE APPROACHES

In our view, what is needed is the combination of glycaemia prediction and explaining the potential reason for the prediction of a dangerous event.

While managing diabetes, diabetic patients should avoid episodes of hypoglycaemia or hyperglycaemia. Hyperglycaemia episodes on the long term can damage the nerves, blood vessels, tissues, and even organs. Hypoglycaemia is significantly more dangerous on the short-term. Reaching this glycaemic state, in severe cases, can lead to loss of consciousness, seizures and ultimately to death. Given its severity, it is important to not only predict hypoglycaemic episodes, but also to understand the actions that led to this consequence. Uncovering the causes of hypoglycaemia for a given user, can be the first step to educate and ultimately change the user's routine for the better. To achieve this, the authors propose combining a predictive module with a reasoning module.

Our current reasoning component is a rule-based system, in Prolog, composed of logical rules, obtained from medical protocols and guidelines. The current system has access to the MyDiabetes app's database and, after each new record, evaluates the user's current situation. If it concludes that the user requires guidance, it sends an advice through the mobile app as a notification, indicating the recommended medical approach for the current occurrence.

Our current prediction component, for hypoglycemia events, uses discrete information fusion and a predictive model consensus. This component uses as data the glucose levels, the insulin therapy, meals, exercises, and other information related to time-dependent information, for example, the record's date and time, the type of meal, and the glucose level variability, considering the previous records. Predictive models are trained and tested on pooled population data using machine learning. The machine learning techniques are a decision tree, support vector machine, k-nearest neighbours, adaboost tree, random forest and subspace k-nearest neighbours. Each model decision is used, by majority vote, for the predictive decision. The consensus decision of the predictive models is given in a personalised way to the patient, indicating the risk of a hypoglycemic event that may occur within the next 24 hours (window).

Our goal is to connect both components. The predictive model receives data from users and detects possible future hypoglycaemia occurrences. Then it supplies this information to the reasoning system that, knowing that a hypoglycaemia will occur, searches the user's data for possible motives for this occurrence. The motives can be simple daily actions such as exercise or incorrect insulin intake, or connected to deeper patterns related to the day of the week and the user's routine, that culminates in a hypoglycaemic occurrence.

We aim to develop this approach in the MyDiabetes smartphone application. Using a mobile application not only facilitates the access to the user's data, but also provides a convenient mean to alert and advice the user.

The first steps will involve adapting the prediction work done[7] to run on a mobile platform. The proposed architecture is shown in Fig. 1. The offline trained model would run online on the smartphone, in the *machine learning system*. Based on the input data it will predict potential hypoglycaemias and inform the *reasoning system*. The *reasoning system*, based on the available data, will warn the user providing advice and a possible reason for the hypoglycaemia. The current reasoning engine [44] is to be tuned to use the more relevant variables for explaining the future hypoglycaemia result.

VI. CONCLUSION

In this article we reviewed the work being done for glycaemia prediction and for reasoning in diabetes management for patients. As discussed, there is still a need to improve the predictions, to make them more useful within pragmatic horizon times and accuracy. Regarding reasoning systems they are still lacking in its usage for diabetes management.

[7]Under review for publication and patent being submitted.

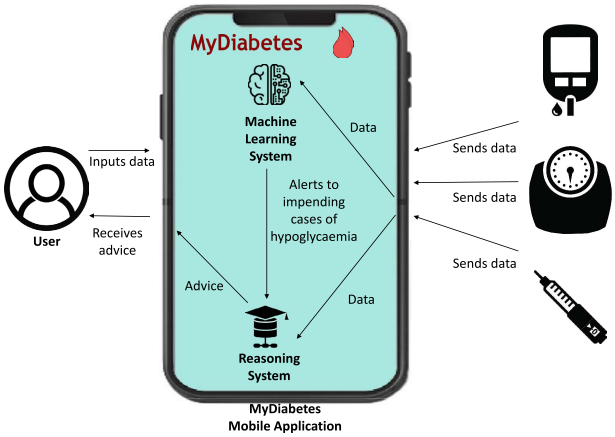

FIGURE 1. Proposal's architecture.

Our focus is on type I diabetics, as there is more information being collected by these patients than type II, and they are in need of a more permanent advising. Type I disease management has several decisions to be taken along the day, whereas for type II this decision taking is less stringent. In some cases diabetics type II are treated with type I methodologies, namely when the disease control deteriorates. In those cases, the proposed approach can also be applied to them.

We briefly propose an approach to incorporate prediction and reasoning, to provide a more whole counselling to patients. They would be able to, using their regular smartphone, receive warnings regarding future episodes and explanations with possible reasons that lead to those events. This future work aims to not only interconnect both approaches but also to improve the predictions and the accuracy of the reasoning in the provided advice.

ACKNOWLEDGMENT

This article is based on work from COST Action IC1303-AAPELE-Architectures, Algorithms and Protocols for Enhanced Living Environments and COST Action CA16226-SHELD-ON-Indoor living space improvement: Smart Habitat for the Elderly, supported by COST (European Cooperation in Science and Technology).

REFERENCES

[1] R. Akerkar and P. Sajja, *Knowledge-Based Systems*. Burlington, MA, USA: Jones & Bartlett Learning, 2010.
[2] G. Alfian, M. Syafrudin, M. Anshari, F. Benes, F. T. D. Atmaji, I. Fahrurrozi, A. F. Hidayatullah, and J. Rhee, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," *Biocyber. Biomed. Eng.*, vol. 40, no. 4, pp. 1586–1599, Oct. 2020, doi: 10.1016/j.bbe.2020.10.004.
[3] A. Aliberti, I. Pupillo, S. Terna, E. Macii, S. Di Cataldo, E. Patti, and A. Acquaviva, "A multi-patient data-driven approach to blood glucose prediction," *IEEE Access*, vol. 7, pp. 69311–69325, 2019, doi: 10.1109/ACCESS.2019.2919184.
[4] A. Aliberti, A. Bagatin, A. Acquaviva, E. Macii, and E. Patti, "Data driven patient-specialized neural networks for blood glucose prediction," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2020, pp. 1–6, doi: 10.1109/ICMEW46912.2020.9105950.
[5] B. V. Ambrosiadou, M. Alevizos, G. Ziakas, and B. V. Ambrosiadou, "Decision support in diabetes management for optimal glycaemic control by insulin administration," in *Proc. Int. Conf. Syst., Man Cybern. Syst. Eng. Service Hum.*, vol. 5, Oct. 1993, pp. 391–396, doi: 10.1109/ICSMC.1993.390883.
[6] B. V. Ambrosiadou, D. G. Goulis, and C. Pappas, "Clinical evaluation of the DIABETES expert system for decision support by multiple regimen insulin dose adjustment," *Comput. Methods Programs Biomed.*, vol. 49, no. 1, pp. 105–115, 1996, doi: 10.1016/0169-2607(95)01711-9.
[7] K. J. Angelides, "Method and device for personalized interactive monitoring for diabetes," U.S. Patent 13 793 208, Jul. 25, 2013,
[8] J. B. Ali, T. Hamdi, N. Fnaiech, V. Di Costanzo, F. Fnaiech, and J.-M. Ginoux, "Continuous blood glucose level prediction of type 1 diabetes based on artificial neural network," *Biocybern. Biomed. Eng.*, vol. 38, no. 4, pp. 828–840, 2018, doi: 10.1016/j.bbe.2018.06.005.
[9] A. Bertachi, L. Biagi, N. Luo, and J. Veh, "Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks," in *Proc. IJCAI*, 2018, pp. 85–90.
[10] R. H. Botwey, E. Daskalaki, P. Diem, and S. G. Mougiakakou, "Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events," in *Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2014, pp. 4843–4846, doi: 10.1109/EMBC.2014.6944708.
[11] G. Cappon, L. Meneghetti, F. Prendin, J. Pavan, G. Sparacino, S. D. Favero, and A. Facchinetti, "A personalized and interpretable deep learning based approach to predict blood glucose concentration in type 1 diabetes," in *Proc. CEUR Workshop*, vol. 2675, 2020, pp. 75–79.
[12] X. Chen, J. Tuo, and Y. Wang, "A prediction method for blood glucose based on grey wolf optimization evolving kernel extreme learning machine," in *Proc. Chin. Control Conf.*, Jul. 2019, pp. 3000–3005, doi: 10.23919/ChiCC.2019.8866210.
[13] S. Contador, J. M. Velasco, O. Garnica, and J. I. Hidalgo, "Profiled glucose forecasting using genetic programming and clustering," *Proc. ACM Symp. Appl. Comput.*, 2020, pp. 529–536, doi: 10.1145/3341105.3374003.
[14] B. Cosenza, "Off-line control of the postprandial glycemia in type 1 diabetes patients by a fuzzy logic decision support," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 10693–10699, Sep. 2012, doi: 10.1016/j.eswa.2012.02.198.
[15] C. Dalla Man, R. A. Rizza, and C. Cobelli, "Meal simulation model of the glucose-insulin system," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 10, pp. 1740–1749, Sep. 2007.
[16] D. Dave, D. J. DeSalvo, B. Haridas, S. McKay, A. Shenoy, C. J. Koh, M. Lawley, and M. Erraguntla, "Feature-based machine learning model for real-time hypoglycemia prediction," *J. Diabetes Sci. Technol.*, vol. 15, no. 4, pp. 842–855, Jun. 2021, doi: 10.1177/1932296820922622.
[17] M. De Bois, M. A. E. Yacoubi, and M. Ammi, "Prediction-coherent LSTM-based recurrent neural network for safer glucose predictions in diabetic people," in *Proc. Int. Conf. Neural Inf. Process.*, 2019, pp. 510–521, doi: 10.1007/978-3-030-36718-3_43.
[18] Y. Dong, R. Wen, Z. Li, K. Zhang, and L. Zhang, "Clu-RNN: A new RNN based approach to diabetic blood glucose prediction," in *Proc. IEEE 7th Int. Conf. Bioinf. Comput. Biol. (ICBCB)*, Mar. 2019, pp. 50–55, doi: 10.1109/ICBCB.2019.8854670.
[19] D. L. Duke, "Intelligent diabetes assistant: A telemedicine system for modeling and managing blood glucose," Ph.D. dissertation, Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, USA, 2010.
[20] A. E. Fathi, R. E. Kearney, E. Palisaitis, B. Boulet, and A. Haidar, "A model-based insulin dose optimization algorithm for people with type 1 diabetes on multiple daily injections therapy," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 4, pp. 1208–1219, Apr. 2021, doi: 10.1109/TBME.2020.3023555.
[21] A. A. Elsharif and S. S. Abu-Naser, "An expert system for diagnosing sugarcane diseases," *Int. J. Academic Eng. Res.*, vol. 3, no. 3, pp. 19–27, 2019.
[22] S. Faccioli, B. Ozaslan, J. F. Garcia-Tirado, M. Breton, and S. Del Favero, "Black-box model identification of physical activity in type-l diabetes patients," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2018, pp. 3910–3913, doi: 10.1109/EMBC.2018.8513378.
[23] V. Felizardo, N. M. Garcia, N. Pombo, and I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – a systematic literature review," *Artif. Intell. Med.*, vol. 118, Aug. 2021, Art. no. 102120, doi: 10.1016/j.artmed.2021.102120.
[24] K. Fortwaengler, C. G. Parkin, W. Koehler, D. A. Cavan, R. Ziegler, I. Cranston, K. Barnard, O. Mast, M. A. Schweitzer, and R. S. Wagner, "Lower short-term health care cost with the Accu-Chek Aviva expert system in multiple daily insulin injection (MDI) treated diabetes patients—Learnings from the automated bolus advisor control and usability study (ABACUS)," *Value Health*, vol. 16, no. 7, p. A437, Nov. 2013, doi: 10.1016/j.jval.2013.08.658.

[25] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting learning to accurately predict blood glucose trajectories," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2018, pp. 1387–1395, doi: 10.1145/3219819.3220102.

[26] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, "Online prediction of glucose concentration in type 1 diabetes using extreme learning machines," in *Proc. 37th Annu. Int. Conf. Eng. Med. Biol. Soc. (EMBC)*, 2015, pp. 3262–3265, doi: 10.1109/EMBC.2015.7319088.

[27] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, "Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models," *Med. Biol. Eng. Comput.*, vol. 53, no. 12, pp. 1305–1318, Dec. 2015, doi: 10.1007/s11517-015-1263-1.

[28] D. A. A.-M. Al-Ghamdi, M. A. Wazzan, F. M. Mujallid, and N. K. Bakhsh, "An expert system of determining diabetes treatment based on cloud computing platforms," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 5, pp. 1982–1987, 2011.

[29] K. Gu, R. Dang, and T. Prioleau, "Neural physiological model: A simple module for blood glucose prediction," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 5476–5481, doi: 10.1109/EMBC44109.2020.9176004.

[30] T. Hamdi, J. B. Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, and J.-M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 362–372, 2018, doi: 10.1016/j.bbe.2018.02.005.

[31] B. Hashemi, "An approach for recommendations in self management of diabetes based on expert system," *Int. J. Comput. Appl.*, vol. 53, no. 14, pp. 887–975, 2012, doi: 10.5120/8487-2431.

[32] F. Hayes-Roth, D. Waterman, and D. Lenat, *Building Expert Systems*. Reading, MA, USA: Addison-Wesley, 1984.

[33] M. He, W. Gu, Y. Kong, L. Zhang, C. J. Spanos, and K. M. Mosalam, "CausalBG: Causal recurrent neural network for the blood glucose inference with IoT platform," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 598–610, Jan. 2020, doi: 10.1109/JIOT.2019.2946693.

[34] J. I. Hidalgo, J. M. Colmenar, G. Kronberger, S. M. Winkler, O. Garnica, and J. Lanchares, "Data based prediction of blood glucose concentrations using evolutionary methods," *J. Med. Syst.*, vol. 41, no. 9, pp. 1–20, Sep. 2017, doi: 10.1007/s10916-017-0788-2.

[35] R. Hovorka, V. Canonico, L. J. Chassin, U. Haueter, M. Massi-Benedetti, M. O. Federici, T. R. Pieber, H. C. Schaller, L. Schaupp, T. Vering, and M. E. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiol. Meas.*, vol. 25, no. 4, pp. 905–920, Aug. 2004, doi: 10.1088/0967-3334/25/4/010.

[36] *IDF Diabetes Atlas*, International Diabetes Federation, Belgium, Brussels, 2019. Accessed Jun. 2021. [Online]. Available: https://www.diabetesatlas.org/upload/resources/material/20200302_133351_IDFATLAS9e-final-web.pdf

[37] M. H. Jensen, C. Dethlefsen, P. Vestergaard, and O. Hejlesen, "Prediction of nocturnal hypoglycemia from continuous glucose monitoring data in people with type 1 diabetes: A proof-of-concept study," *J. Diabetes Sci. Technol.*, vol. 14, no. 2, pp. 250–256, Mar. 2020, doi: 10.1177/1932296819868727.

[38] S. Kriventsov, A. Lindsey, and A. Hayeri, "The diabits app for smartphone assisted predictive monitoring of glycemia in patients with diabetes: Retrospective observational study," *JMIR Diabetes*, vol. 5, no. 3, pp. 1–14, 2020, doi: 10.2196/18660.

[39] E. D. Lehman and T. Deutsch, "A physiological model of glucose-insulin interaction in type 1 diabetes mellitus," *J. Biomed. Eng.*, vol. 14, pp. 235–242, May 1992.

[40] E. Lehmann, T. Deutsch, E. Carson, and P. Sãnksen, "Combining rule-based reasoning and mathematical modelling in diabetes care," *Artif. Intell. Med.*, vol. 6, no. 2, pp. 137–160, 1994, doi: 10.1016/0933-3657(94)90042-6.

[41] M. Levy, P. Ferrand, and V. Chirat, "Sesam-diabete, an expert system for insulin-requiring diabetic patient education," *Comput. Biomed. Res., Int. J.*, vol. 22, no. 5, pp. 442–453, Oct. 1989, doi: 10.1016/0010-4809(89)90037-2.

[42] K. Li, C. Liu, T. Zhu, P. Herrero, and P. Georgiou, "GluNet: A deep learning framework for accurate glucose forecasting," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 2, pp. 414–423, Feb. 2020, doi: 10.1109/JBHI.2019.2931842.

[43] T. F. Lunt and R. Jagannathan, "A prototype real-time intrusion-detection expert system," in *Proc. IEEE Symp. Secur. Privacy*, Oakland, CA, USA, Apr. 1988, pp. 1–8.

[44] D. Machado, T. Paiva, I. Dutra, V. S. Costa, and P. Brandao, "Managing diabetes: Pattern discovery and counselling supported by user data in a mobile platform," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jul. 2017, pp. 296–299.

[45] C. Marling, R. Bunescu, J. Shubrook, and F. Schwartz, "System overview: The 4 diabetes support system," in *Proc. 20th Int. Conf. Case-Based Reasoning*, 2012, pp. 81–86.

[46] C. Marling and R. C. Bunescu, "The OhioT1DM dataset for blood glucose level prediction," in *Proc. IJCAI*, 2018, pp. 1–4.

[47] C. Marling, M. Wiley, T. Cooper, R. Bunescu, J. Shubrook, and F. Schwartz, "The 4 diabetes support system: A case study in CBR research and development," *Case-Based Reasoning Research and Development* (Lecture Notes in Computer Science), vol. 6880. Berlin, Germany: Springer, 2011, pp. 137–150, 2011, doi: 10.1007/978-3-642-23291-6_12.

[48] A. Mbogho, J. Dave, and K. Makhubele, "Diabetes advisor—A medical expert system for diabetes management," in *Proc. Int. Conf. E-Infrastruct. E-Services Developing Countries*. Blantyre, Malawi: Springer, 2013, pp. 140–144, doi: 10.1007/978-3-319-08368-1_16.

[49] P. D. McAndrew, D. L. Potash, B. Higgins, J. Wayand, and J. Held, "Expert system for providing interactive assistance in solving problems such as health care management," U.S. Patent 5 517 405, May 4, 1996.

[50] L. McCausland, I. M. Y. Mareels, R. W. Barnett, and D. Arad, "An adaptive expert system for blood glucose control in type 1 diabetes mellitus," in *Proc. 1st Joint BMES/EMBS Conf.*, Oct. 1999, p. 1209, doi: 10.1109/IEMBS.1999.804375.

[51] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A deep learning approach to diabetic blood glucose prediction," *Frontiers Appl. Math. Statist.*, vol. 3, pp. 1–11, Jul. 2017, doi: 10.3389/fams.2017.00014.

[52] S. Mirshekarian, R. Bunescu, C. Marling, and F. Schwartz, "Using LSTMs to learn physiological models of blood glucose behavior," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2887–2891, doi: 10.1109/EMBC.2017.8037460.

[53] A. Mohebbi, A. R. Johansen, N. Hansen, P. E. Christensen, J. M. Tarp, M. L. Jensen, H. Bengtsson, and M. Mørup, "Short term blood glucose prediction based on continuous glucose monitoring data," 2020, *arXiv:2002.02805*. [Online]. Available: http://arxiv.org/abs/2002.02805

[54] E. Montaser, J.-L. Diez, P. Rossetti, M. Rashid, A. Cinar, and J. Bondia, "Seasonal local models for glucose prediction in type 1 diabetes," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 2064–2072, Jul. 2020, doi: 10.1109/JBHI.2019.2956704.

[55] H. Nemat, H. Khadem, J. Elliott, and M. Benaissa, "Data fusion of activity and CGM for predicting blood glucose levels," in *Proc. CEUR Workshop*, vol. 2675, 2020, pp. 120–124.

[56] H. Nemat, H. Khadem, J. Elliott, and M. Benaissa, "Diabetes (type 1 and type 2) in children and young people: Diagnosis and management," Nat. Inst. Health Care Excellence (NICE) Guideline, London, U.K., Tech. Rep. NG18, 2020.

[57] *Type 1 Diabetes in Adults: Diagnosis and Management*, National Institute for Health and Care Excellence (NICE) Guideline, London, U.K., 2020.

[58] C. Novara, N. M. Pour, T. Vincent, and G. Grassi, "A nonlinear blind identification approach to modeling of diabetic patients," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 3, pp. 1092–1100, May 2016, doi: 10.1109/TCST.2015.2462734.

[59] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for t1dm patients," *Int. J. Numer. Methods Biomed. Eng.*, vol. 33, no. 6, pp. 1–21, 2017, doi: 10.1002/cnm.2833.

[60] J. Pavan, F. Prendin, L. Meneghetti, G. Cappon, G. Sparacino, A. Facchinetti, and S. Del Favero, "Personalized machine learning algorithm based on shallow network and error imputation module for an improved blood glucose prediction," in *Proc. CEUR Workshop*, vol. 2675, 2020, pp. 95–99.

[61] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, "A machine learning approach to predicting blood glucose levels for diabetes management," in *Proc. Mod. Artif. Intell. Health Anal.*, 2014, pp. 35–39.

[62] A. Al-Qumboz, N. Mohammed, and S. S. Abu-Naser, "Spinach expert system: Diseases and symptoms," *Int. J. Academic Inf. Syst. Res.*, vol. 3, no. 3, pp. 16–22, 2019.

[63] R. Reddy, N. Resalat, L. M. Wilson, J. R. Castle, J. El Youssef, and P. G. Jacobs, "Prediction of hypoglycemia during aerobic exercise in adults with type 1 diabetes," *J. Diabetes Sci. Technol.*, vol. 13, no. 5, pp. 919–927, Sep. 2019, doi: 10.1177/1932296818823792.

[64] R. E. Reinke, J. F. Price, and P. J. Galley, "Diabetes health management systems and methods," U.S. Patent 12 622 520, May 26, 2011.
[65] R. Rudi and B. G. Celler, "Design and implementation of expert telemedicine system for diabetes management at home," in *Proc. Int. Conf. Biomed. Pharmaceutical Eng.*, 2006, pp. 595–599. [Online]. Available: https://ieeexplore.ieee.org/document/4155973
[66] C. Tarin, E. Teufel, J. Pico, J. Bondia, and H.-J. Pfleiderer, "Comprehensive pharmacokinetic model of insulin Glargine and other insulin formulations," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 12, pp. 1994–2005, Dec. 2005.
[67] B. Todd, *An Introduction to Expert Systems* (An Introduction to Expert Systems). Oxford, U.K.: Oxford Univ. Computing Laboratory, 1992.
[68] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, "Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning," *Health Informat. J.*, vol. 26, no. 1, pp. 703–718, Mar. 2020, doi: 10.1177/1460458219850682.
[69] W. F. Velicer and J. O. Prochaska, "An expert system intervention for smoking cessation," *Patient Educ. Counseling*, vol. 36, pp. 119–129, Feb. 1999.
[70] W. Wang, M. Tong, and M. Yu, "Blood glucose prediction with VMD and LSTM optimized by improved particle swarm optimization," *IEEE Access*, vol. 8, pp. 217908–217916, 2020, doi: 10.1109/ACCESS.2020.3041355.
[71] M. E. Wilinska, L. J. Chassin, H. C. Schaller, L. Schaupp, T. R. Pieber, and R. Hovorka, "Insulin kinetics in type-1 diabetes: Continuous and bolus delivery of rapid acting insulin," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 3–12, Jan. 2005, doi: 10.1109/TBME.2004.839639.
[72] A. Z. Woldaregay, E. Årsand, S. Walderhaug, D. Albers, L. Mamykina, T. Botsis, and G. Hartvigsen, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," *Artif. Intell. Med.*, vol. 98, pp. 109–134, Jul. 2019, doi: 10.1016/j.artmed.2019.07.007.
[73] L. Wos, R. Overbeck, E. Lusk, and J. Boyle, *Automated Reasoning: Introduction Application*. Upper Saddle River, NJ, USA: Prentice-Hall, 1984.
[74] K. Zarkogianni, K. Mitsis, E. Litsa, M.-T. Arredondo, G. Fic, A. Fioravanti, and K. S. Nikita, "Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring," *Med. Biol. Eng. Comput.*, vol. 53, no. 12, pp. 1333–1343, Dec. 2015, doi: 10.1007/s11517-015-1320-9.
[75] C. Zecchin, A. Facchinetti, G. Sparacino, and C. Cobelli, "How much is short-term glucose prediction in type 1 diabetes improved by adding insulin delivery and meal content information to CGM data? A proof-of-concept study," *J. Diabetes Sci. Technol.*, vol. 10, no. 5, pp. 1149–1160, Sep. 2016, doi: 10.1177/1932296816654161.

VIRGINIE FELIZARDO graduated in biomedical sciences from the University of Beira Interior (UBI), Covilhã, Portugal. She received the M.Sc. degree in electrotechnical engineering—bionic systems from UBI, in 2010, where she is currently pursuing the Ph.D. degree in computer science with the Assisted Living Computing and Telecommunications Laboratory (ALLab), a research laboratory, Instituto de Telecomunicações. Since 2010, she has been a Researcher with the ALLab, Instituto de Telecomunicações, UBI. Her research interests include sensors for biomedical applications, biomedical instrumentation, health monitoring, medical signal acquisition, analysis, and processing, ambient assisted living, predictive algorithms, and machine learning.

DIOGO MACHADO received the M.Sc. degree from the Faculty of Sciences, University of Porto, where he is currently pursuing the Ph.D. degree and supported by a scholarship from the Fundação para a Ciência e Tecnologia of Portugal. He worked as a Researcher with the Instituto de Telecomunicações.

NUNO M. GARCIA received the B.Sc. degree (Hons.) in mathematics/informatics and the Ph.D. degree in computer science engineering from the Universidade da Beira Interior (UBI), Covilhã, Portugal, in 2004 and 2008, respectively. He was an Entrepreneur, from 1988 to 2004, a member of the Research Team at Siemens SA, from 2004 to 2007, and Nokia Siemens SA, from 2007 to 2008, and the Head of Research at PLUX SA, from 2008 to 2010. He was an Associate Professor with Habilitation at the Computer Science Department, UBI, in 2010. He was the Founder and a Researcher of the Assisted Living Computing and Telecommunications Laboratory (ALLab), a research group within the Instituto de Telecomunicações, UBI, in 2010. Since 2010, he has been an Invited Associate Professor with the Universidade Lusófona de Humanidades e Tecnologias, Lisbon, Portugal. He was also the Co-Founder and the Chair of the Executive Council of the BSAFE Laboratory—Law Enforcement, Justice and Public Safety Research and Technology Transfer Laboratory, a multidisciplinary research laboratory at UBI, in 2015. Since 2018, he has been working as the Vice Dean of the Faculty of Engineering, UBI. His main research interests include next-generation networks, predictive algorithms for healthcare and well-being, distributed and cooperative algorithms, and the battle for a free and open internet.

NUNO POMBO is currently an Assistant Professor with the University of Beira Interior (UBI), Covilhã, Portugal. He is also the Co-ordinator of the Assisted Living Computing and Telecommunication Laboratory (ALLab), UBI. He is also a member of the BSAFE Laboratory, and the Instituto de Telecomunicações—IT, UBI. His current research interests include information systems (with special focus on clinical decision support systems), data fusion, artificial intelligence, and software.

PEDRO BRANDÃO (Member, IEEE) received the Diploma and M.Sc. (ECE) degrees from the Engineering School, University of Porto, and the Ph.D. degree from the Computer Laboratory, University of Cambridge, under the topic of body sensor networks. The Ph.D. was supported by a scholarship from the Fundação para a Ciência e Tecnologia of Portugal. He worked in research with IT-Aveiro, PT Inovação (now Altice Labs), INESC-Porto, and LIACC. Since 2007, he has been with IT-Porto. During this time, he worked in identity management, communication networks, middleware development and currently on ad-hoc wireless networks for sensors/actuators and protocols for these networks; medical informatics and mHealth; and network security. He is also an Assistant Professor with the Computer Science Department, School of Sciences, University of Porto. He has participated in several national and international projects (DAIDALOS, Future Cities, Future Health, VR2Market, and NanoSTIMA); where in some there was development of prototypes (app for diabetes management, health kiosk, and ad-hoc network for smartphones).

3.2 A novel context-aware framework for hypoglycaemia events prediction.

Virginie Felizardo, Nuno Garcia, Imen Megdiche, Nuno Pombo and František Babič.

A novel context-aware framework for hypoglycaemia events prediction.

Submitted to a Computer Science journal, Q1, by Scimago Journal & Country Rank, currently in the second round of peer review.

This research paper presents and defines the context-aware framework for hypoglycaemia prediction, giving a detailed description of experimental design in a comprehensible step-by step method. Also, this research paper shows the contribution of integration of context, time-based features and multi-classifiers learning for hypoglycaemia prediction.

A novel context-aware framework for Hypoglycaemia Events Prediction integrating context and time-based features

Virginie Felizardo[a,b]*, Nuno Garcia[a,b], Imen Megdiche[c], Nuno Pombo[a,b], František Babič[d]

[a]*Universidade da Beira Interior, Covilhã, Portugal*
[b]*Instituto de Telecomunicações, Covilhã, Portugal*
[c] *Institut National Universitaire Champollion - Ecole d'ingénieur ISIS Campus Universitaire, Castres, France*
[d] *Department of Cybernetics and Artificial Intelligence of Technical University of Kosice, Košice, Slovakia*
virginie@it.ubi.pt, ngarcia@di.ubi.pt, imen.megdiche@univ-jfc.fr, ngpombo@di.ubi.pt, frantisek.babic@tuke.sk

ARTICLE INFO

Article history:
Received 00 December 00
Received in revised form 00 January 00
Accepted 00 February 00

Keywords:
Hypoglycaemia
Prediction
Context-aware framework
Time based features
Consensus decision

ABSTRACT

The prediction of relevant events in diabetes plays an important role in the management of the disease. This paper proposes a new framework for the prediction of hypoglycaemia using a 24-hour prediction window, sparse data, contextual and time-based features. Also, we propose an approach using a consensus decision to improve hypoglycaemia prediction without loss of accuracy and the increase in probability of false alarms. After performing the pre-processing steps (data handling, feature extraction, data imputation and re-sampling), we build a pooled population model with data from fifty-two patients, aiming to produce personalised predictions. For this, we proposed the integration of context and time-based features extracted from the University of California Irvine diabetes dataset. We proposed an ensemble approach based on consensus of different individual models whose experiments, using data collected from fifty-three patients, evidenced the effectiveness of the proposed approach. We observed that context data and time-based features improve the results for 58,5% of the patients. We also observed that using the consensus decision, the results improved on 57.4% of patients. The proposed solution achieved 100% of hypoglycaemia prediction on 39.6% of the patients along with a false alarm probability lower than 30%. In general, the solution achieved 90-100% of hypoglycaemia predicted for 58,5% of the patients. Our approach shows that integrating contextual and time-based features was relevant for the predictions, and the consensus decision improved the results for hypoglycaemia prediction without loss of accuracy and without the increase of probability of false alarms.

1. Introduction

Diabetes mellitus prevalence is growing, affecting an estimated 592 million adults by 2030 and 629 million by 2045 [1]. This chronic metabolic disorder is characterised by the difficulty to control blood glucose (BG) due to the inability to produce insulin. Unfortunately, the bad control of BG, results in some short-term and long-term health complications. These complications can seriously decrease the quality of life of diabetic's patients. Hypoglycaemia (or hypoglycemia in US English) is one of the most problematic short-term issues of a bad glycaemia control. This event occurs when BG levels fall below the acceptable health threshold (70 mg/dL). This crisis is commonly feared because it can result in blurred vision, confusion, erratic behaviour, garbled speech, clumsiness, seizures and even loss of consciousness, which in some cases can result in death.

Hypoglycaemia risks are smaller for patients equipped with Continuous Glucose Monitoring (CGM) systems that allow for alarms by using a predictive threshold. Yet some reasons lead to discontinuing these systems, these being: 1) inaccuracy [2], [3], 2) reimbursement [3], 3) cost [2], [4] and 4) frequent annoying false alarms [2]. Previous data from the T1D Exchange (T1DX), related to the period of 2013-2014, showed that only 9% of participants were using the CGMs [5]. New data from

2017 show that around 24% of patients in the T1D Exchange Clinic Registry were using CGM [6]. The number is growing, but some diabetic patients continue to use blood BG monitoring performed by the use of finger sticks. This may be because, despite the discomfort, this kind of measurement has a lower price and more accurate results of BG concentrations [2], [3].

Regarding the reimbursement status of CGM systems, many countries in Europe have no or partial coverage [7]. Countries such as Germany, Belgium, and France have national coverage for type 1 and type 2 diabetes patients on insulin treatment. Other European countries have National coverage for type 1 diabetes patients, is available in *e.g.* Switzerland, Netherlands, Czech Republic, or Regional funding is available in, e.g. Austria, Denmark, Italy and Spain [7]. Portugal has full reimbursement for Flash glucose monitoring (FGM) systems, but only for type 1 diabetes. Dunn et al. [8] analysed 50,831 glucose level readers using FGM systems and showed that 93% of these readers are from Germany, Spain, France, United Kingdom, Italy, Sweden and Netherlands.

Several self-management applications (apps), such as Glucose Buddy [9], Dottli: A [10], Sugar Sense [11], Social Diabetes [12], and Diabetes M [13], are available in online web stores. These applications allow the recording of contextual data (*e.g.* daytime of meal, exercise and stress) but sometimes are not supported by medical evidence. The lack of participation of medical experts and the lack of evidence of sound clinical trials can be a risk for the users of these apps. There are similar products in the market but these only work with specific glucose meters, *e.g.* Contour®Diabetes app [14] and OneTouch Reveal® [15].

Extensive work has been done on diabetes classification or prediction [16]–[19]. Most of the literature on the prediction of events in diabetes is focused on predictions using data from continuous glucose monitoring [20]–[36]. Using real-time CGM and the hypoglycaemic alarms reduces the incidence and time spent in the state of hypoglycaemia, but these results still scarce and are not extensible to patients using BG monitoring performed using finger sticks.

To the best of our knowledge, hypoglycaemia prediction using discrete data is not an approach usually considered by the research community, despite discrete measurement being the most common scenario for patients who try to control their BG levels. This may be because of sparsity of data from BG measurements [37], [38].

These limitations, combined with the lack of contextual open resources and the use of discrete information in predictive works, leaves an open challenge. The existence of predictive algorithms for the discovery of personalised knowledge can be used to predict possible changes in patients' health status. Machine learning algorithms can improve the accuracy of prediction over the use of conventional regression models by capturing complex, nonlinear relationships in the data.

In the present research, we propose a new hypoglycaemia predictive framework based on contextual and time-based features. This framework aims to predict hypoglycaemia events in a 24-hour window for different patients. Figure 1 shows a comparison between common approaches and the proposed approach. The main differences between the approaches are the integration of context and time-based features (in grey colour) and the use of different classifier to obtain a consensus decision using majority vote.

We summarise the main contributions of this paper:

- A new framework for the prediction of hypoglycaemia events using sparse data, in a 24h hour window;
- The integration of contextual and time-based features with the conventional features in order to improve the prediction;
- An ensemble approach for a consensus decision to improve the accuracy without increasing the number of false alarms.

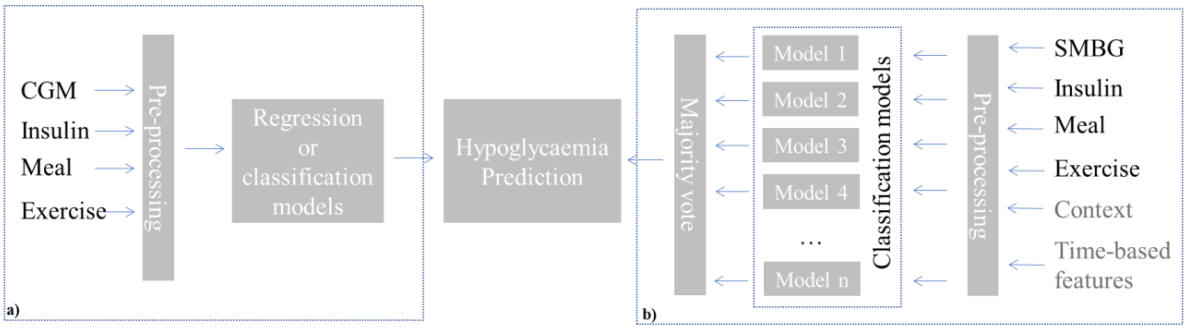

Fig. 1 - A comparison between common and proposed approach: a) common approach and b) proposed approach.

We organise the rest of the manuscript as follows. Section 2 presents the related work that inspired our approach. In section 3, we introduce step by step the framework proposal. Section 4 presents the dataset and the settings used in experiments. We present the results of the experiments in Section 5, giving a case study to our framework proposal. Section 6 presents the conclusion and a proposal for future research work.

2. Related work

In the last years, some data-based approaches have emerged using pattern recognition techniques for the prediction of diabetes related events [39]. The literature shows that BG prediction or future events prediction uses BG measurements, insulin units, meal information, and physical activity information [20], [22], [24], [26], [32], [35], [36], [39]–[48]. However, some of dynamics of these features are different for each person, this effect being termed inter-subject variability [49]. Also, even in the same patient there are changes over time, termed intra-subject variability [49]. The inter- and intra-subject variabilities, not being considered, contribute to inefficiencies in the prediction results.

The models for hypoglycaemia prediction can be categorised into two groups: 1) personalised-, individual- or subject-based models trained on an individual's specific data, and 2) general- or population-based models trained on pooled data from many people. In the first one, the accuracy can be negatively affected by intra-subject variability if the model is trained on limited data sets. The second one aims to generalise well enough to allow them to be used in previously unseen individuals [50]. However, these models may not generalise well to specific groups of individuals if they are not appropriately represented in the training data set [50].

2.1 Context-aware systems

Nowadays, the use of mobile devices is generalised [51], namely smartphones, as they can be used easily and allow for information exchange, some of which are represented by context dimensions. The dimension term is a mensurable extent or quantity, such as time, location, environment and personal information [52], [53]. The context is defined by several manners, but the most used definition is given by Abowd *et. al* [54]: "*Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered as relevant to the interaction between a user and an application, including the user and applications themselves*". The works focused on glucose or diabetes events prediction does not provide context-aware systems. Some other works propose context-aware systems for diabetes management or other disease [55]–[57]. Quinde *et al.* [58] proposed a context-aware solution for asthma management based on six context dimensions: triggers and symptoms, environment, internal, location and time, communication and medication. The context dimensions facilitate the development of preventive and reactive features supporting patients with personalised solutions [58].

2.2 Features and predictors

The literature shows that no specific method provides 100% reliable event predictions. In this sense, a growing number of approaches have been proposed by the scientific community to combine multiple predictors and techniques with the hope of representing spatial and temporal input and output dependencies.

The selection of proper features is crucial to create models that perform good results for future event predictions [59]. Some studies that use BG alone due to inter-subject variability [25], [27], [29]–[31], [33], [34], [60]–[65]. This variability can be represented

by daytime meal, medication, and exercise. So, the integration of contextual information can be relevant to minimise this interference.

Georga *et al.* [44], show that the contribution of features like meal intake, insulin and exercise are small but not irrelevant. Also [44] shows that exercise features based on time are more significant in short-term predictions than in long-term predictions. Zarkogianni *et al.* [39] show the importance of the integration of exercise features with glucose features and the time of the day includes some features connected to glucose dynamics. Also, He *et al.* [64] show that the addition of information on meals, insulin and time provide dominant causal correlations. Rodbard's work [3] presents a hypoglycaemia risk estimation based on features that show the variability and nature of BG. Kovatchev *et al.* [66] established that BG variability was correlated to hypoglycaemia and Monnier [67] found that this variability can increase the risk of asymptomatic hypoglycaemia. Klimontov *et al.* [68] showed that BG fluctuations over the previous day is a useful nocturnal hypoglycaemia prediction.

In short, literature shows that data fusion can be appropriate for hypoglycaemia prediction. The common features like BG levels, insulin, meal and exercise information together with other features like time of the day show correlations with glucose dynamics.

2.3 Methods of prediction

The advantage of using artificial intelligence to address the prediction of hypoglycaemia events is that the physicians could estimate the risk of hypoglycaemia, but they cannot explain the specific rules, and a prediction model that coded this was published in [37]. So, artificial intelligence, namely machine learning, besides the extraction of new knowledge can identify patterns and correlations between different features that can help to capture hypoglycaemia complexity.

In this sense, Sudharsan *et al.* [37] proposed a classification approach because it is easier and more efficient than the regression approach. This study proposed a mixed measure dataset (with numerical and categorical inputs) and a target based on binary classes (yes/no). The authors refer to data cleaning as very important to identify anomalies, errors and missing values. For train/test design, the authors proposed that a 24-hour prediction window would be the most useful time frame for delivering an intervention [37]. Li & Fernando [38] proposed an approach based on smartphone-collected patient data and population data to drive personalised prediction. This model uses pooled population data that produce a personalized BG prediction for an individual. The experimental results show that this approach can remedy the problem of sparse data. Both approaches achieved best results using random forest.

Due to individual inter- and intra-variability, no specific classifiers provide valuable solutions for all patients. So, a way to get a successful prediction model is using ensemble methods combining the output of individual classifiers [69].

Some studies propose approaches based on sparse data [37], [38]. According to these studies, the major limitation of discrete BG measurements is the sparse data. Other limitations include the non-linearity of BG profile, the complexity of hypoglycaemia events and the non-linear relationships between the different features [3], [39], [44], [64], [66], [68].

Table 1 summarises the literature characteristics that provides the theoretical support to design our approach. It is inspired by this set of ideas: 1) a pooled population model to produce a personalised prediction; 2) using a 24-hour prediction window; 3) using a mixed measure dataset (with numerical and categorical inputs); and 4) resort to an ensemble to improve the prediction performance.

Fewer studies in literature explored the idea of creating a glucose level prediction model using discrete data from a population of diabetics to provide a personalised prediction. In [37], [38] the authors propose approaches using discrete data but using only BG and Insulin features. In this work, we proposed a novel framework using a discrete data fusion and a multi-classifier decision, that we named consensus.

3. Framework

The proposed framework involves data acquisition, pre-processing, training/testing of individual models and an ensemble model for hypoglycaemia prediction (Figure 2). Figure 2 shows the flowchart of the proposed ensemble framework composed by a Training Process which incorporates data collection, feature extraction, data imputation, resampling, building and selection of individual classifiers models, and the combination of individual models. In Application process, new data are collected and used by ensemble models for hypoglycaemia prediction. This module includes a response defined by majority vote and a final model decision to perform a better solution for hypoglycaemia prediction. For the framework design we used the University of California Irvine (UCI) diabetes dataset [70], that to the best of our knowledge, is the unique free access dataset that satisfies our requirements, by providing information on 1) Time and time slots, 2) BG, Insulin, Meal and Exercise, 3) patients with hypoglycaemia events (more information about this dataset in section 3).

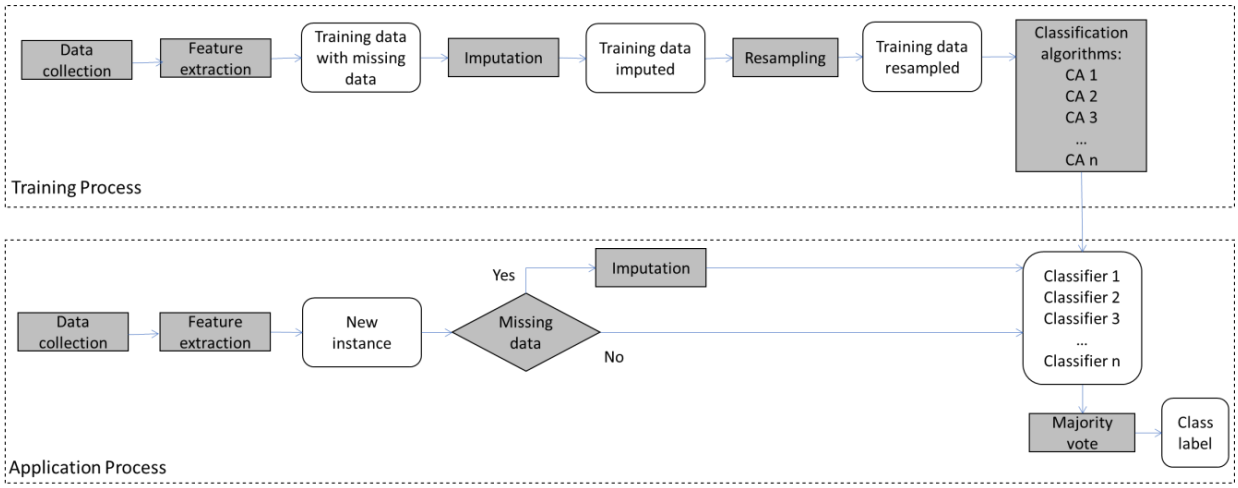

Fig. 2 – Flow chart of the proposed framework.

3.1 Data collection

Data acquisition may be performed via a mobile application for self-management and organised to allow that user can register their BG level giving a context to this measure, *e.g.* kind of meal (pre- or post-breakfast, pre- or post-lunch, pre- or post-dinner and break time) and time of register, information about meal (more, less or typical), information about physical activity (more, less or typical) and type and quantity of insulin.

The context information and the time of register give to this approach a new contribution using this information for the calculation of time-based features (subsection 3.2).

6

Table 1 – Summary table of literature characteristics for blood glucose or hypoglycaemia prediction.

Study (Year)	Data							Prediction models							
	Continuous	Discrete	Inputs					Prediction horizons			Target	Models Data-based		Classifier	
			BG	M	PA	I	Other	Short-term	Mid-term	Long-term		Personalised-based	Population-based	Individual	Ensemble or hybrid
[20]	X		x			x		x			H	x			x
[38]		X	x						X		BG	x	x		x
[37]		X	x			x			X		H		x		x
[2]	X		x						X		NH		x		x
[65]	X	X	x						X		NH		x	x	
[71]		X				x	x			X	H		x	x	
[26]	X		x	x		x			X		PH		x	x	
[72]		X	x					x			FBG		x	x	
[32]	X		x	x	x		x		X		BG	x			x
[48]			x	x		x			X		PH	x		x	
[22]	X		x	x	x	x			X		NH		x	x	
	X		x	x		x			X		PH		x	x	
[41]	X		x	x		x		x			H		x	x	
[50]	X		x			x		x			H		x	x	
[60]	X		x					x			H		x	x	
Our approach		X	x	x	x	x	x		X		H		x		x

H – Hypoglycaemia; BG – Blood glucose; NH – Nocturnal hypoglycaemia; PH – Postprandial hypoglycaemia; FBG – Fasting blood glucose.

3.2 Feature extraction

For the UCI diabetes dataset [70], first, we transformed each text file into a spreadsheet by converting the data formats. We organise our dataset as a mixed (numerical and categorical) matrix where each row is an instance. We divide the data into thirteen columns: the 1st to 12nd columns are input vectors and the 13rd is the target vector (Table 2). We extracted three additional features based on time, column 11th (difference between actual BG and previous register) and 12th (difference between actual BG and previous 24-hour register) give some information about variability of BG.

Table 2 – Dataset organisation.

Original Features	**Type**	**Description**
Hour	Numerical	1st column, register time, 24h per day presented with a difference of 30 minutes
Slot time	Categorical	2nd column, eight different slots: • No info – 0; • Pre-Breakfast – 1 • Post-Breakfast – 2 • Pre-Lunch – 3 • Post-Lunch – 4 • Break -5 • Pre-Dinner – 6 • Post-Dinner – 7
Blood Glucose	Numerical	3rd column, blood glucose level between 22-487 mg/dl
Regular Insulin	Numerical	4th column, regular insulin dose between 0-22 units
NHP Insulin	Numerical	5th column, NHP insulin dose between 0-40 units
Ultralente Insulin	Numerical	6th column, ultraslow insulin dose between 0-30 units
Meal	Categorical	7th column, meal information: • No information – 0 • More than usual – 1 • Typical meal – 2 • Less than usual – 3
Exercise	Categorical	8th column, exercise information: • No information – 0 • More than usual – 1 • Typical meal – 2 • Less than usual – 3
Extracted Features	**Type**	**Description**
Duration	Numerical	9th column, duration between consecutive registers
BG-1	Numerical	10th column, Blood Glucose of previous register
Diff_previous	Numerical	11nd column, Difference between actual Blood Glucose and previous register
Diff_24h	Numerical	12nd column, Difference between actual Blood Glucose and previous 24h register
Target	**Type**	**Description**
HypoPredict	Multiclass	13rd column, hypoglycaemia prediction classes: • No risk – 0 • Risk for the next 24h – 1 • Hypoglycaemia – 2.

As shown in Figure 3, the context information is given by three context dimensions, such as 1) time, 2) behaviour, and 3) time-based features. Each context interaction is given by different colours: blue for time, green for behaviour and orange for time-based context. The dimension of time is given by the feature hour of the day or time of register, allowing deliver information about when some actions were performed. The behaviour dimension is given by information about meal or physical activity and allow answer to how these actions were performed. The time-based features are the link between the time and glucose level feature, giving a notion of how the glucose level varies over time. In this approach, the context works like as supplementary information to bridge the discrete data and allow the recognition of distinct patterns associated to hypoglycaemia risk. Because it is an event that has heterogeneous triggering, the context information can be useful in defining and support the decision process.

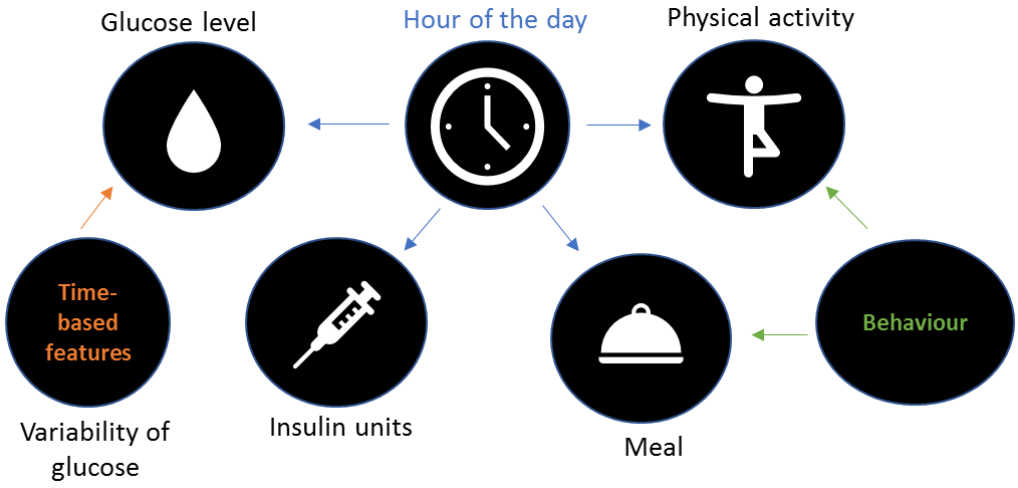

Figure 3 – Context dimensions: Time (blue), behaviour (green) and time-based (orange).

3.3 Data imputation and re-sampling

With the impossibility of replacing the BG missing data by zero, we can either replace the missing data using imputation techniques. The other option could be to delete the instance with the missing value, but this could result in losing relevant information. To justify our option, we performed a selection feature method comparison to show the feature ranking for hypoglycaemia prediction (see Table 3). We used six ranking feature selection methods [73]–[76]: in- finite feature selection (Inf-FS), ReliefF algorithm, Laplacian Score [77], Unsupervised Discriminative Feature Selection (UDFS) [78], Feature Selection and Kernel Learning for Local Learning-Based Clustering (Llcfs) [79] and Sort features according to pairwise correlations (Cfs). Table 3 shows the ranking and top 5 of the features. In this table, column 1 describes the name of the feature selection technique, column two describes the ranking of the features and column 3 describes the top 5 of the potential best features. Because of these findings, we conclude that BG and variables build with this one, present more predictive potential for hypoglycaemia prediction. Therefore, we opted to impute BG values using a ridge regression-based method [80]. We experimented using other methods, such as an iterative algorithm [81] and data augmentation [82], but only the ridge regression-based method imputed values of BG that were acceptable. The data imputation was performed for each patient data to perform more realistic imputations. Figure 4 illustrates the imputation procedure.

After imputation and feature extraction, our method proposal uses the Synthetic Minority Over-sampling Technique (SMOTE) [83] for re-sampling the dataset. Because hypoglycaemia events are always a minority class in comparison with the totality of the data, re-sampling methods are crucial for unbalanced data. We opted to use SMOTE method.

Table 3 – Feature selection for hypoglycaemia prediction.

Feature selection techniques	Ranking of the feature number	Top 5 # number of features
Inf-FS [77]	11 12 3 2 1 6 7 8 10 5 4 9	BG (#3) and context (#2)
ReliefF [73]–[76]	1 9 11 12 4 5 2 3 10 6 7 8	Context (#2), BG (#2) and Insulin (#1)
Laplacian [77]	11 12 3 10 5 4 2 1 9 6 7 8	BG (#4) and insulin (#1)
UDFS [78]	10 3 11 12 5 4 2 1 9 6 7 8	BG (#4) and insulin (#1)
Llcfs [79]	10 3 11 12 5 4 2 1 9 6 7 8	BG (#4) and insulin (#1)
Cfs	2 3 10 1 12 4 5 11 9 6 7 8	BG (#3) and context (#2)

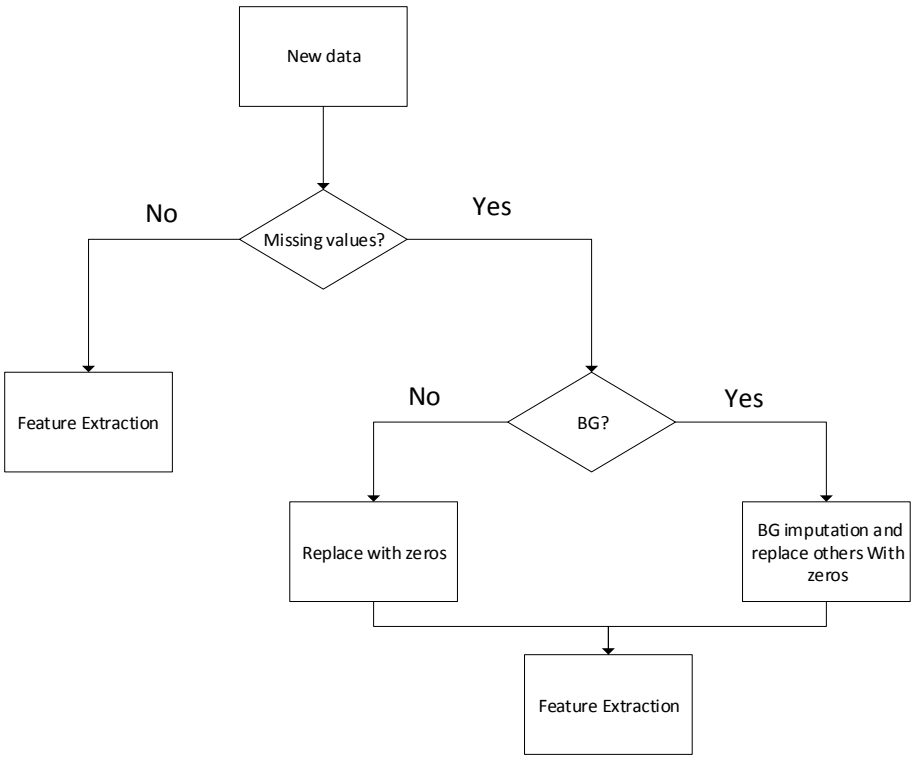

Fig. 4 – Proposed missing data imputation method.

3.4 Individual predictive models

The classifier training performs with pre-processed data (imputation and resampling). The instances of train/test set are organised as attribute vectors and for each instance containing *n* inputs attributes a *y* is defined as the target attribute or class of result. For hypoglycaemia prediction, our target vector has three classes: 0 – no risk, 1 –risk and 2 – Hypoglycaemia event (defined when the BG value is less than 70 mg/dl) [84]. The risk class (value of 1) is the risk of having a hypoglycaemia event in the next 24 hours or a BG value less than 75 mg/dl [85]. Figure 5 shows an example of how to build the target set. For the instances with a value lower than 70 mg/dl we considered all instances of the last 24h like class 1. Also, for instances with a value less or equal than 75 mg/dl we considered all instances of the last 24h like class 1 in order to mitigate the risk of a possible hypoglycaemia event. We applied a 10-fold cross validation to perform the models.

Hour	Features	BG	features	Target
17:00	…	226	…	0
22:00	…	158	…	1
08:00	…	178	…	1
12:00	…	189	…	1
19:00	…	200	…	1
22:00	…	67	…	2
08:00	…	121	…	0
12:00	…	190	…	1
17:00	…	124	…	1
22:00	…	222	…	1
08:00	…	339	…	1
11:30	…	72	…	1
17:00	…	110	…	0
22:00	…	140	…	0
07:00	…	92	…	0
12:00	…	180	…	0

Fig. 5 – Proposed target build method.

The model selection criteria are model accuracy and a good balance between sensibility and specificity. It is important because the sensitivity measures, in this approach, the ability to identify the risk of having a hypoglycaemia event in the next 24h. Likewise, the specificity measures the ability to identify the other classes. So, the good balance between these two measures is important for a good prediction of hypoglycaemia events without higher values of false positives (false alarms).

3.5 Consensus-based prediction

After building and select the individual models, an unweighted voting (also called majority voting) applied to obtain a consensus decision from different models. The main objective of the consensus decision was improving the accuracy without the increase of false alarms from the previous solution.
In this paper, we presented results for all classifier networks based on three classifiers and five classifiers (more information in section 4). The main reason of this method is to obtained suitable solutions for all patients as none individual classifiers satisfy suitable solutions for all patients.

4. Experiment design

4.1 Dataset

The design of the framework performed using a diabetes dataset [70] from the UCI Machine Learning Repository that contains datasets from 70 type-1 diabetics. Diabetes patient records from UCI dataset were obtained from two sources: an automatic electronic recording device that had an internal clock to timestamp events and paper records that only provided time-of-day related time slots such as breakfast, lunch, dinner, bedtime.

4.2 Settings

This paper used data from fifty-three patients (P1-P5, P7-P46, P50, P54-P56, P65, P67-P68, P70). The selection criteria of these patients were the number of observations and at least one hypoglycaemia event. For each patient, we built a population model with data from fifty-two patients and an external validation was performed with the data from that patient, one patient out validation.
The number of instances of the population varies between 11113 to 11738. After performing imputation of data for each patient, we applied SMOTE method to population data two times: one for class 2, with $k=6$, and other for class 1, with $k=2$. The k is the nearest neighbours in feature space. Figure 6 show an example of the distribution of the target attribute to support the SMOTE method application.

The first approach, using six (#6) features, contains BG, Insulin regular, Insulin NHP, Insulin Ultralent, Meal and Exercise. To the second approach, using twelve features (12#) we include Hour, Duration, Time slot, Diff_previous, Diff_24h and BG-1.
The classifiers used were: 1) Classification And Regression Tree (CART), 2) Support vector machine (SVM), 3) k-Nearest neighbour (kNN), 4) Adaboost+Decision Tree (AdaBTree), 5) Random Forest (RF) and 6) Subspace+ k-Nearest neighbour (SkNN). These classifiers have been chosen because our data set contained data that included some categorical and some numeric features. Table 4 shows the settings for each classifier.

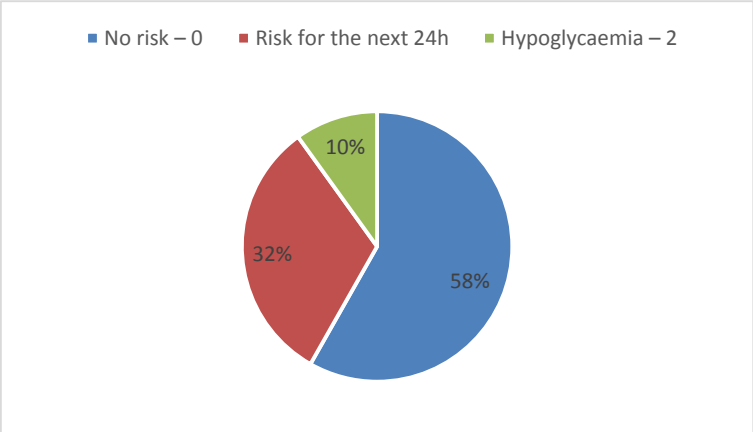

Fig. 6 Distribution of the target attribute to support the SMOTE method application.

Table 4 – Classifiers settings.

Classifier	Settings
Decision Tree (CART)	Maximum number of splits: 100 Split criterion: gini s diversity index
Support vector machine (SVM)	Kernel function: Gaussian Box constraint level: 1 Kernel scale: 0.87
k-Nearest neighbour (kNN)	Number neighbour: k=1 Distance metric: euclidian Distance weight: Equal
Adaboost+Decision Tree (AdaBTree)	Maximum number of splits: 20 Number of learners: 30 Learning rate: 0.1
Random Forest (RF)	Maximum number of splits: sample-1 Number of learners: 30
Subspace+ k-Nearest neighbour (SkNN)	Maximum number of splits: 20 Number of learners: 30 Subspace dimension: 5 (#6 features) 6 (#12 features)

For the consensus decision the individual models were combined resulting in twenty-six ensembles (twenty using three classifiers and six using five classifiers.

5. Results and Analysis

This section presents comparisons between three approaches for hypoglycaemia prediction: 1) using #6 features, 2) using #12 features, and 3) using a consensus decision. The analysis is performed to check the contribution of contextual and time-based features (#12 features) with respect to the conventional features (#6 features). This analysis also aimed to investigate the improvements achieved using a consensus decision regarding individual classifiers (#12 features models).
Figure 7 shows the accuracy (average of all patients) of train/test module for each classifier. In summary, the results show minor improvements, but in general all accuracies are higher than 80%.
The next subsections present more detailed results using one patient out validation.

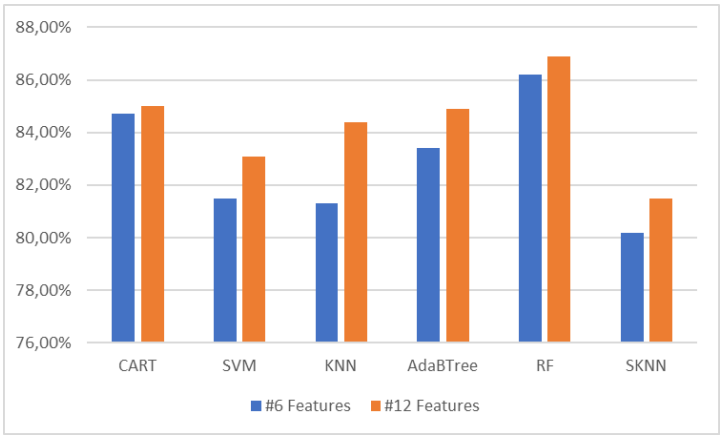

Fig. 7 – Accuracy comparison between #6 features models and #12 features models.

5.1 Integrating data context and time-based features

Table 5 shows the average results of external validation of all fifty-three patients. We call external validation because each prediction was made using data from the patient that were not used to create and train the model. In summary, the results show minor improvements, but for the #12 features approach, all individual models achieved an accuracy higher than 60%. The first column describes the different classifiers used in these experiments, the second block of columns (from column 2 to 5) represents the values for Accuracy (Acc), Sensitivity (Sen) and Specificity (Spe) and False Alarms (FA) for the approach using 6 features while the second block of columns (from column 6 to 9) show the same type of data for the approach using 12 features. Values in the table cells are percentages.

Table 5 – Performance of tested solutions.

Classifier Models	#6 features				#12 features			
	Acc	Sen	Spe	FA	Acc	Sen	Spe	FA
CART	59.9	39,9	74,3	25,8	62,3	45,9	71,8	28,4
SVM	58,4	57,1	60,0	40,1	60,8	60,0	61,2	39,8
KNN	62,8	54,6	68,0	32,2	62,0	49,6	69,2	30,9
AdaBTree	59,3	45,5	70,2	29,9	61,1	42,5	72,3	27,9
RF	64,4	43,6	77,3	22,8	67,0	50,7	75,1	25,0
SKNN	62,0	41,6	74,3	25,1	63,8	60,7	66,1	34,0

In Table 6 we present the results for the tested solutions that achieved an accuracy higher than 60% and a probability of false alarms lower than 30%. These ratios were defined after analysis of the results and are considered being adequate in terms of the expectations of the users and crucial for the purpose of valuable solution for hypoglycaemia prediction. In this table, the columns labelled as Number represent the number of solutions obtained with an accuracy higher than 60% and a probability of false alarms lower than 30%, and the columns labelled with HP refer to hypoglycaemia predicted. Values in columns Acc, FA and HP are percentages. Considering these restrictions, it can be observed that the results are better for the #12 features models. Accuracy improved between 1-10% and hypoglycaemia prediction between 3-18% (excepted for KNN and AdaBTree).

Table 6 – Solutions with an accuracy higher than 60% and false alarms lower than 30%.

Classifier Models	#6 features				#12 features			
	Number	Acc	FA	HP	Number	Acc	FA	HP
CART	23	66,0	16,3	68,6	25	67,8	20,0	72,1
SVM	11	67,2	21,1	77,5	11	77,3	17,0	81,3
KNN	21	71,0	20,5	89,0	21	72,6	18,3	82,9
AdaBTree	21	66,6	17,2	75,5	21	67,8	18,6	70,8
RF	31	69,9	16,9	76,9	33	72,3	19,7	81,2
SKNN	25	68,6	16,4	71,3	18	74,0	18,2	89,7

In Table 7, we detailed the results in terms of hypoglycaemia predicted. The approach using #6 features achieved 63,7% of solutions with a hypoglycaemia prediction above 70% and for the approach using #12 features 76,0%. In terms of improvements in patients' results, the #12 features models improved the results in 58,5% of patients.

Table 7 – Solutions results in terms of hypoglycaemia predicted.

Hypoglycaemia Predicted	#6 features		#12 features		
	Number of solutions	% of solutions	Number of solutions	% of solutions	% patients with improvements
<50%	16	12,1%	14	10,9%	**58,5%**
50%-59%	9	6,8%	8	6,2%	
60%-69%	23	17,4%	9	7,0%	
70%-79%	15	11,4%	21	16,3%	
80%-89%	19	14,4%	22	17,1%	
90%-100%	50	37,9%	55	42,6%	
Hypoglycaemia Predicted > 70%		**63,7%**		**76,0%**	

5.2 Consensus-based prediction

Table 8 shows the results for the ensembles. The results presented in Table 8 show the performance achieved for all solutions. In comparison with the results of Table 5 for #12 features models, in general, the ensembles show better results.

Table 8 – Results for the tested ensembles.

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C3.1 (CART+SVM+KNN)	64.0	56.2	66.7	33.4
C3.2 (AdaBTree+RF+SKNN)	66.8	50.9	76.6	25.5
C3.3 (CART+SVM+AdaBTree)	63.1	48.0	70.9	29.2
C3.4 (CART+KNN+ AdaBTree)	64.0	46.7	72.9	27.2
C3.5 (SVM+KNN+ AdaBTree))	64.3	55.7	67.5	32.7
C3.6 (CART+SVM+RF)	65.5	52.0	71.1	29.0
C3.7 (CART+KNN+RF)	66.9	50.8	73.9	26.2
C3.8 (CART+AdaBTree+RF)	65.5	44.6	74.6	25.5
C3.9 (CART+AdaBTree +SKNN)	63.8	46.5	73.0	27.1
C3.10 (CART+RF+SKNN)	67.2	51.8	74.6	25.5
C3.11 (CART+SVM+SKNN)	64.3	57.5	66.9	33.2
C3.12 (CART+KNN+SKNN)	65.5	54.9	70.1	30.0
C3.13 (SVM+KNN+RF)	64.8	56.6	67.5	32.6
C3.14 (SVM+KNN+SKNN)	63.7	59.8	64.4	35.7
C3.15 (KNN+AdaBTree+RF)	66.6	50.0	74.2	25.9
C3.16 (KNN+ AdaBTree+SKNN)	65.6	54.0	70.9	29.2
C3.17 (KNN+RF+SKNN)	66.2	54.3	71.2	28.9
C3.18 (SVM+ AdaBTree+RF)	65,1	50.9	71.5	28.6
C3.19 (SVM+ AdaBTree+SKNN)	64.3	55.7	67.8	32.4
C3.20 (SVM+RF+SKNN)	65.1	57.1	68.1	32.1
C5.1 (CART+SVM+KNN+AdaBTree+RF)	65.6	50.6	72.4	27.7
C5.2 (CART+SVM+KNN+AdaBTree+SKNN)	65.4	55.1	69.6	30.5
C5.3 (CART+SVM+KNN+RF+SKNN)	65.8	55.5	69.6	30.5
C5.4 (CART+SVM+AdaBTree+RF+SKNN)	66.3	51.8	72.9	27.2
C5.5 (CART+KNN+AdaBTree+RF+SKNN)	66.9	51.2	74.1	26.1
C5.6 (SVM+KNN+AdaBTree+RF+SKNN)	65.7	53.5	70.6	29.5

Also, for solutions with an accuracy higher than 60% and false alarms lower than 30% all ensembles achieved better results regarding hypoglycaemia prediction (see Table 9). For 57.4% of patients, the ensembles approaches show improvements in comparison with #12 features individual models. The results for consensus decision approach show that for 58.4% of patients, this was the best solution. Values for Acc, FA, and HP in Table 9 are percentages.

14

Table 9 – Solutions ensemble with an accuracy higher than 60% and false alarms lower than 30%.

Classifier Models	Performance				
	Acc	FA	HP	% patients with improvements	% patients with consensus as the best solution
C3.1 (CART+SVM+KNN)	73,0	21,2	84,1	57.4	57.4
C3.2 (AdaBTree+RF+SKNN)	73,5	17,9	80,2		
C3.3 (CART+SVM+AdaBTree)	67,9	19,8	76,7		
C3.4 (CART+KNN+ AdaBTree)	69,9	20,4	76,6		
C3.5 (SVM+KNN+ AdaBTree))	73,9	19,8	83,6		
C3.6 (CART+SVM+RF)	74,1	17,6	78,0		
C3.7 (CART+KNN+RF)	73,5	18,8	81,9		
C3.8 (CART+AdaBTree+RF)	69,9	20,2	81,7		
C3.9 (CART+AdaBTree +SKNN)	68,9	18,8	79,0		
C3.10 (CART+RF+SKNN)	72,7	18,7	79,3		
C3.11 (CART+SVM+SKNN)	72,3	21,1	84,2		
C3.12 (CART+KNN+SKNN)	73,4	19,5	82,4		
C3.13 (SVM+KNN+RF)	75,1	19,3	85,5		
C3.14 (SVM+KNN+SKNN)	75,1	19,4	88,0		
C3.15 (KNN+AdaBTree+RF)	73,9	18,0	80,9		
C3.16 (KNN+ AdaBTree+SKNN)	73,4	19,7	84,2		
C3.17 (KNN+RF+SKNN)	71,8	17,5	82,7		
C3.18 (SVM+ AdaBTree+RF)	74,1	17,2	78,3		
C3.19 (SVM+ AdaBTree+SKNN)	72,2	20,9	83,5		
C3.20 (SVM+RF+SKNN)	74,4	18,7	85,2		
C5.1 (CART+SVM+KNN+AdaBTree+RF)	73,0	19,4	80,1		
C5.2 (CART+SVM+KNN+AdaBTree+SKNN)	73,0	20,3	84,1		
C5.3 (CART+SVM+KNN+RF+SKNN)	75,5	19,3	86,7		
C5.4 (CART+SVM+AdaBTree+RF+SKNN)	71,7	21,3	79,2		
C5.5 (CART+KNN+AdaBTree+RF+SKNN)	73,1	18,6	80,2		
C5.6 (SVM+KNN+AdaBTree+RF+SKNN)	75,3	18,5	84,3		

To conclude the results section, Figure 8 shows the comparison of accuracy, false alarms and hypoglycaemia predicted for the solutions with accuracy higher than 60% and false alarms lower than 30%.
Comparing the three approaches: 1) #6 features, 2) #12 features and 3) consensus decision, the results show that integrating context and time-based features improve the accuracy of the solutions, and the consensus decision increases even more the accuracy with the decrease of false alarms and increase of hypoglycaemia prediction.

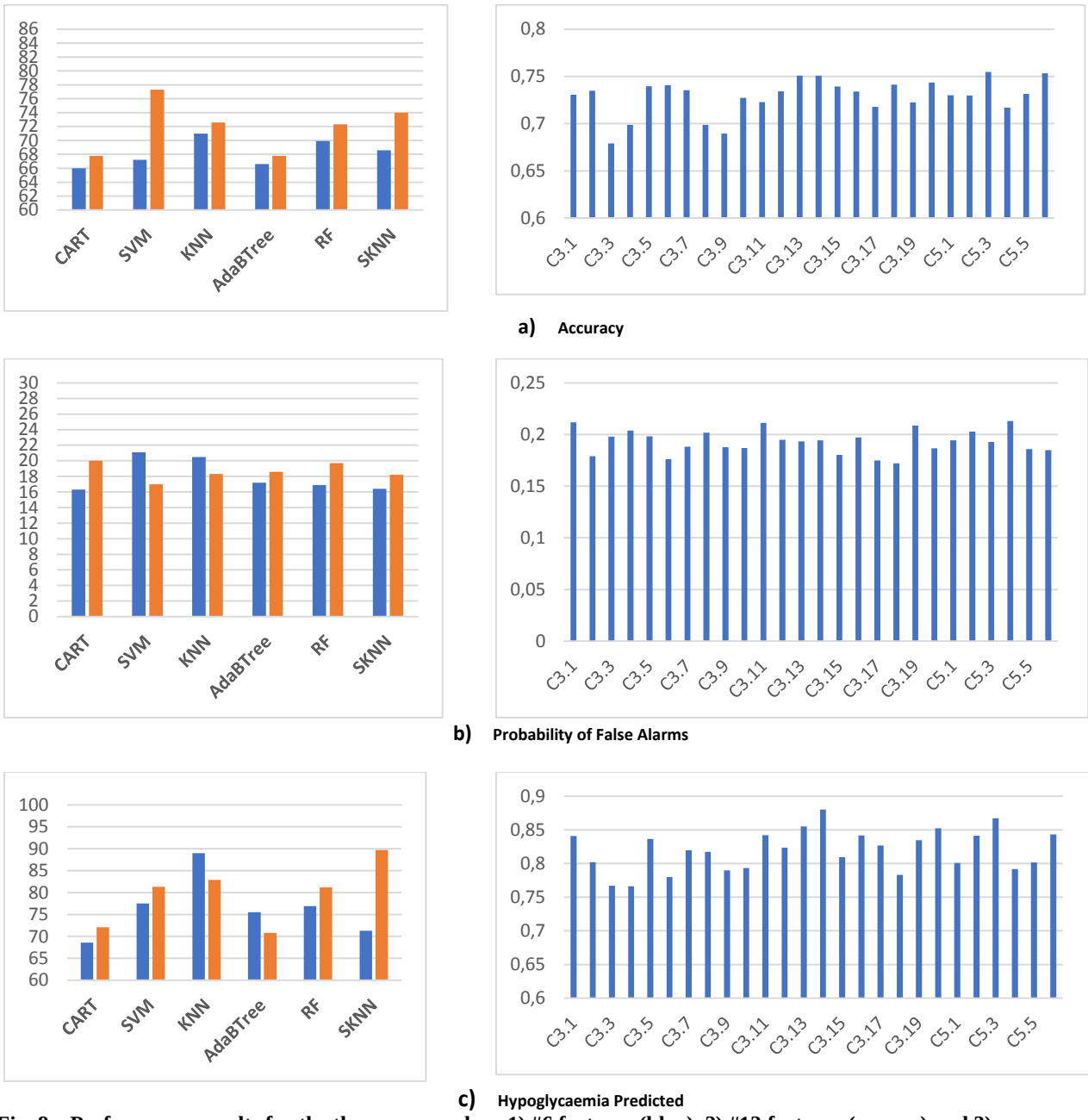

Fig. 8 – Performance results for the three approaches: 1) #6 features (blue), 2) #12 features (orange) and 3) consensus decision (blue, right). In y axis were presented a) accuracy, b) probability of false alarms and c) hypoglycaemia predicted (in percentages). In x axis were presented the classifiers and the ensembles.

5.3 Study limitations

Some limitations should be noted. First, we start this work with a manual manipulation of data from the UCI diabetes dataset due to the inadequate format of data for the classification. Second, the limited open access of data from diabetic patients with context information and hypoglycaemia events. Third, we used data from fifty-three patients, and in future research, the sample size should be increased. Fourth, the prediction horizon used was 24h and it may be decreased. Fifth, the pre-processing methods used for data imputation and resampling perform well in this approach, but other methods must be tried. For example, in SMOTE the k used for the minority's classes are fixed, and this may not be appropriate for all patients. Sixth, the threshold of 75 mg/dl used for mitigating the risk of possible occurrence of hypoglycaemia event may be optimised. Seventh, the classifiers, the settings and the ensembles proposed should be further explored.

6. Discussion

This section presents the discussion about the research work presenting the main findings, benefits, opportunities, and challenges.

Our approach supports the development of predictive solutions addressing the heterogeneity triggering of hypoglycaemia events in a scenario of self-management. The experimental design presented in the article shows that the framework supports hypoglycaemia prediction in a comprehensible step-by-step method. The features and context dimensions can properly describe patients with diabetes with different characteristics and behaviours, which is important as it allows the generalisation of the models in a self-management scenario. The context information helps in the recognition of distinct patterns associated with hypoglycaemia risk and delivering information for decision-making is a 24h window. The results about the train of individual models show an accuracy above 80% supporting that was possible performed predictions using discrete information.

The literature shows that is common the use of complementary information like as insulin units, meal, and physical activity in combination with glucose levels for diabetes predictions. However, some of these features present different dynamics for each person. Also, in the same person, there are changes over time. So, in this work, we intended the integration of contextual information to decrease the effect of inter- and intra-subject variability on prediction patterns. We added time of the day, person behaviours related to meal and physical activity and time-based glucose variability. The dimension of time is very important in creating solutions for patients with diabetes because it is that most affect variables in inter- and intra-variability. Determining the action of the patient at a specific time directly affects on the performance for a specific scenario. The time dimension operatively connects the information of the other context dimensions. For example, the dimension behaviour quantifies the meal and physical activity, but the major impact in prediction is when connected with other context dimension or variables like insulin units or blood glucose level. In this work, the major contribution of performing predictive models is the time-based features, namely time-based glucose variability. These new features show a good predictability for hypoglycaemia events, giving new knowledge about glucose profile changes in function of context.

The integration of contextual and time-based features allows the development of a greater number of solutions appropriate for an intervention (accuracy greater than 60% and false positives less than 30%) improving accuracy and hypoglycaemia predicted. These experiments show that 76% of these solutions achieved over 70% of hypoglycaemia predicted. Also, these solutions show better predictions for 58.5% of patients when compared with the predictions using conventional features. These findings suggest that contextual information and time-based features improved the predictions and allow the development of better solutions for most patients used in this study.

The context of patients with diabetes is challenging to monitor. Although mobile and wearable technology is more reachable nowadays, some commercial solutions are already in the market and in mobile app stores, however, there is not a generalised used of them yet, and the availability of their data is usually not for the public used. Hence, there is the need to create networks of monitoring devices that can provide contextual information in combination with data from personal glucometers.

We are aware that it is also expected to affect the usability of these kind of solutions as users would stop using it if they are overloaded with irrelevant predictions due to high probability of false alarms. The results from one patient out validation show that the prediction improved with the integration of contextual and time-based features, but no individual model allows a good solution for all patients and in some cases, no single solution was appropriate for a patient. So, to mitigate the problem we added to our framework a consensus decision from different individual models to improve accuracy without the increased of false alarms. We expected that with a consensus decision from different and heterogenous ensembles it is possible to decrease more the effect of inter- and intra-variability and to present to patients a better personalised solution catching appropriate patterns associated with hypoglycaemia risk. For this work, we used only the combination of three or five models for the consensus decision. The experiments results show that this approach achieved better accuracy, false positives and hypoglycaemia predicted. Also, the consensus decision improved the prediction for 57.4% of the patients offering a best solution.

The propose of personalised solutions is an important part of our approach. The creation of a population and personal databases of cases representing someone's previous patterns associated with hypoglycaemia risk brings other potential

challenge from a generalisation or personalisation perspective. So, the compromise of the user is very important because they should be aware that the more cases they create, the more accurate this kind of approach is expected to be.

7. Conclusion

Predicting future events in diabetes is a challenge for physicians and medical scientific research. Many patients with diabetes do not have access to CGM systems, so solutions using discrete data collected by *e.g.* a self-management mobile application can help to prevent some dangerous events. Our framework proposal satisfies these requirements and can be incorporated into a self-management solution.

The main findings of this work are summarized as follows:

• The new framework offers a good prediction of risk of hypoglycaemia in a 24-hour window using sparse data;

• The integration of contextual and time-based features with the conventional features improves the accuracy and the number of solutions valuable (accuracy higher than 60% and false alarms lower than 30%);

• The ensemble approaches for a consensus decision improve the performance of predictions offering good solutions for 57.4% of the patients.

However, further research is necessary to reduce the number of false alarms and to understand the complexity of these predictions due to sparse data. The present experimental work presents some limitations such as utilize all data (most patients used not provide some data about meal and exercise) and utilize data from different insulin therapeutic during the train models.

As future work it needs some improvements such as the build of models for different scenarios to offer a personalized solution taking in to account the data provided by the patient, using feature selection methods, and some similarities between patients, using clustering methods. Also, include more data from other patients from UCI diabetes dataset. The data was not very extensive, but this work shows that it is still possible to explore all combinations of classifiers in future.

Acknowledgements

This work was partially funded by FCT/MEC through national funds and co-funded by the FEDER–PT2020 partnership agreement under the project UIDB/50008/2020 (Este trabalho é financiado pela FCT/MEC através de fundos nacionais e cofinanciado pelo FEDER, no âmbito do Acordo de Parceria PT2020 no âmbito do projeto UIDB/50008/2020). This article is based on work from COST Action IC1303–AAPELE–Architectures, Algorithms and Protocols for Enhanced Living Environments and COST Action CA16226–SHELD-ON–Indoor living space improvement: Smart Habitat for the Elderly, supported by COST (European Cooperation in Science and Technology). More information at www.cost.eu.

References

[1] I. D. Federation, "IDF Diabetes Atlas 8th edition," 2017. https://diabetesatlas.org/IDF_Diabetes_Atlas_8e_interactive_EN/ (accessed Jun. 27, 2019).

[2] P. Tkachenko, G. Kriukova, M. Aleksandrova, O. Chertov, E. Renard, and S. v. Pereverzyev, "Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application," *Comput Methods Programs Biomed*, vol. 134, pp. 179–186, 2016, doi: 10.1016/j.cmpb.2016.07.003.

[3] D. Rodbard, "Continuous Glucose Monitoring: A Review of Recent Studies Demonstrating Improved Glycemic Outcomes," *Diabetes Technol Ther*, 2017, doi: 10.1089/dia.2017.0035.

[4] "Continuous Monitoring of Glucose for Type 1 Diabetes: A Health Technology Assessment," *Ont Health Technol Assess Ser*, 2018.

[5] J. C. Wong *et al.*, "Real-time continuous glucose monitoring among participants in the T1D exchange clinic registry," *Diabetes Care*, 2014, doi:

10.2337/dc14-0303.

[6] N. C. FOSTER *et al.*, "Marked Increases in CGM Use Has Not Prevented Increases in HbA1c Levels in Participants in the T1D Exchange (T1DX) Clinic Network," *Diabetes*, 2018, doi: 10.2337/db18-1689-p.

[7] H. T. A. C. Model, R. Relative, E. Assessment, E. Ja, E. Union, and H. Programme, "EUnetHTA Joint Action 3 WP4 Version," no. July 2018, pp. 1–500, 2020.

[8] T. C. Dunn, Y. Xu, G. Hayter, and R. A. Ajjan, "Real-world flash glucose monitoring patterns and associations between self-monitoring frequency and glycaemic measures: A European analysis of over 60 million glucose tests," *Diabetes Res Clin Pract*, vol. 137, pp. 37–46, 2018, doi: 10.1016/j.diabres.2017.12.015.

[9] A. Inc., "Glucose Buddy." https://www.glucosebuddy.com/ (accessed Jun. 27, 2019).

[10] D. Oy, "Dottli:A." https://dottli.com/ (accessed Jun. 27, 2019).

[11] Staywell, "Sugar Sense - Diabetes App." https://play.google.com/store/apps/details?id=org.medhelp.sugarsense&hl=pt_PT (accessed May 14, 2018).

[12] Social Diabetes, "Social Diabetes." https://www.socialdiabetes.com/ (accessed Feb. 12, 2018).

[13] Sirma Medical Systems, "Diabetes: M." https://www.diabetes-m.com/ (accessed Feb. 12, 2018).

[14] A. D. C. U. Inc., "CONTOUR®DIABETES app." https://www.contournextone.com/ (accessed Jun. 27, 2019).

[15] L. LifeScan IP Holdings, "OneTouch Reveal." https://www.onetouchreveal.com/ (accessed Jun. 27, 2019).

[16] V. Felizardo, N. M. Garcia, N. Pombo, and I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review," *Artif Intell Med*, vol. 118, no. February 2020, p. 102120, 2021, doi: 10.1016/j.artmed.2021.102120.

[17] O. Mujahid, I. Contreras, and J. Vehi, "Machine learning techniques for hypoglycemia prediction: Trends and challenges," *Sensors (Switzerland)*. 2021. doi: 10.3390/s21020546.

[18] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for T1DM patients," *Int J Numer Method Biomed Eng*, vol. 33, no. 6, pp. 1–21, 2017, doi: 10.1002/cnm.2833.

[19] A. Z. Woldaregay *et al.*, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," *Artificial Intelligence in Medicine*. 2019. doi: 10.1016/j.artmed.2019.07.007.

[20] R. H. Botwey, E. Daskalaki, P. Diem, and S. G. Mougiakakou, "Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 4843–4846. doi: 10.1109/EMBC.2014.6944708.

[21] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, "A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management," *Modern Artifical Intelligence for Health Analytics*, pp. 35–39, 2014.

[22] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, "Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning," *Health Informatics J*, 2019, doi: 10.1177/1460458219850682.

[23] F. Stahl, R. Johansson, and M. L. Olsson, "Predicting Nocturnal Hypoglycemia Using a Non-parametric Insulin Action Model," *Proceedings - 2015 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2015*, pp. 1583–1588, 2016, doi: 10.1109/SMC.2015.280.

[24] C. Zecchin, A. Facchinetti, G. Sparacino, and C. Cobelli, "How much is short-term glucose prediction in type 1 diabetes improved by adding insulin delivery and meal content information to CGM data? A proof-of-concept study," *J Diabetes Sci Technol*, vol. 10, no. 5, pp. 1149–1160, 2016, doi: 10.1177/1932296816654161.

[25] T. Hamdi *et al.*, "Artificial neural network for blood glucose level prediction," *2017 International Conference on Smart, Monitored and Controlled Cities, SM2C 2017*, pp. 91–95, 2017, doi: 10.1109/SM2C.2017.8071825.

[26] F. Dubosson, N. Mordvanyuk, B. López Ibáñez, and M. Schumacher, "Negative results for the prediction of postprandial hypoglycemias from insulin intakes and carbohydrates: analysis and comparison with simulated data," pp. 25–29, 2017, [Online]. Available: https://dugi-doc.udg.edu/handle/10256/14859

[27] H. N. Mhaskar, S. v. Pereverzyev, and M. D. van der Walt, "A Deep Learning Approach to Diabetic Blood Glucose Prediction," *Front Appl Math Stat*, vol. 3, no. July, pp. 1–11, 2017, doi: 10.3389/fams.2017.00014.

[28] J. Ben Ali, T. Hamdi, N. Fnaiech, V. Di Costanzo, F. Fnaiech, and J. M. Ginoux, "Continuous blood glucose level prediction of Type 1 Diabetes based on Artificial Neural Network," *Biocybern Biomed Eng*, vol. 38, no. 4, pp. 828–840, 2018, doi: 10.1016/j.bbe.2018.06.005.

[29] T. Hamdi, J. Ben Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, and J. M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybern Biomed Eng*, vol. 38, no. 2, pp. 362–372, 2018, doi: 10.1016/j.bbe.2018.02.005.

[30] N. Li, J. Tuo, and Y. Wang, "Chaotic time series analysis approach for prediction blood glucose concentration based on echo state networks," in *2018 Chinese Control And Decision Conference (CCDC)*, 2018, pp. 2017–2022. doi: 10.1109/CCDC.2018.8407457.

[31] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting learning to accurately predict blood glucose trajectories," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1387–1395, 2018, doi: 10.1145/3219819.3220102.

[32] C. Midroni, P. J. Leimbigler, G. Baruah, M. Kolla, A. J. Whitehead, and Y. Fossat, "Predicting glycemia in type 1 diabetes patients: Experiments with XGBoost," *CEUR Workshop Proc*, vol. 2148, pp. 79–84, 2018.

[33] T. El Idrissi, A. Idri, I. Abnane, and Z. Bakkoury, "Predicting blood glucose using an LSTM Neural Network," *Proceedings of the 2019 Federated Conference on Computer Science and Information Systems*, vol. 18, pp. 35–41, 2019, doi: 10.15439/2019f159.

[34] T. Koutny, A. Della Cioppa, I. De Falco, E. Tarantino, U. Scafuri, and M. Krcma, "De�randomized Meta-Differential Evolution for Calculating and Predicting Glucose Levels," in *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, 2019, pp. 269–274. doi: 10.1109/CBMS.2019.00064.

[35] S. Mirshekarian, H. Shen, R. Bunescu, and C. Marling, "LSTMs and Neural Attention Models for Blood Glucose Prediction: Comparative Experiments on Real and Synthetic Data," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 706–712. doi: 10.1109/EMBC.2019.8856940.

[36] J. I. Hidalgo, J. M. Colmenar, G. Kronberger, S. M. Winkler, O. Garnica, and J. Lanchares, "Data Based Prediction of Blood Glucose Concentrations Using Evolutionary Methods," *J Med Syst*, vol. 41, no. 9, 2017, doi: 10.1007/s10916-017-0788-2.

[37] B. Sudharsan, M. Peeples, and M. Shomali, "Hypoglycemia Prediction Using Machine Learning Models for Patients With Type 2 Diabetes," *J Diabetes Sci Technol*, vol. 9, no. 1, pp. 86–90, 2014, doi: 10.1177/1932296814554260.

[38] J. Li and C. Fernando, "Smartphone-based personalized blood glucose prediction," *ICT Express*, vol. 2, no. 4, pp. 150–154, 2016, doi: 10.1016/j.icte.2016.10.001.

[39] K. Zarkogianni *et al.*, "Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring," *Med Biol Eng Comput*, vol. 53, no. 12, pp. 1333–1343, 2015, doi: 10.1007/s11517-015-1320-9.

[40] K. Li, C. Liu, T. Zhu, P. Herrero, and P. Georgiou, "GluNet: A Deep Learning Framework For Accurate Glucose Forecasting," *IEEE J Biomed Health Inform*, vol. PP, no. c, pp. 1–1, 2019, doi: 10.1109/jbhi.2019.2931842.

[41] M. Gadaleta, A. Facchinetti, E. Grisan, and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," *IEEE J Biomed Health Inform*, vol. 23, no. 2, pp. 650–659, 2019, doi: 10.1109/JBHI.2018.2823763.

[42] S. Mirshekarian, R. Bunescu, C. Marling, and F. Schwartz, "Using LSTMs to learn physiological models of blood glucose behavior," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 2887–2891. doi: 10.1109/EMBC.2017.8037460.

[43] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, "Online prediction of glucose concentration in type 1 diabetes using extreme learning machines," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2015-Novem, pp. 3262–3265, 2015, doi: 10.1109/EMBC.2015.7319088.

[44] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, "Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models," *Med Biol Eng Comput*, vol. 53, no. 12, pp. 1305–1318, 2015, doi: 10.1007/s11517-015-1263-1.

[45] T. Zhu, K. Li, P. Herrero, J. Chen, and P. Georgiou, "A deep learning algorithm for personalized blood glucose prediction," *CEUR Workshop Proc*, vol. 2148, pp. 64–78, 2018.

[46] J. Chen, K. Li, P. Herrero, T. Zhu, and P. Georgiou, "Dilated recurrent neural network for short-time prediction of glucose concentration," *CEUR Workshop Proc*, vol. 2148, pp. 69–73, 2018.

[47] L. O. Griva and M. S. Basualdo, "Evaluating Clinical Accuracy of Models for Predicting Glycemic Behavior for Diabetes Care," in *2018 Argentine Conference on Automatic Control (AADECA)*, Nov. 2018, pp. 1–6. doi: 10.23919/AADECA.2018.8577323.

[48] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," *Int J Med Inform*, vol. 126, no. March, pp. 1–8, 2019, doi: 10.1016/j.ijmedinf.2019.03.008.

[49] S. Faccioli, B. Ozaslan, J. F. Garcia-Tirado, M. Breton, and S. Del Favero, "Black-box Model Identification of Physical Activity in Type-l Diabetes Patients," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 3910–3913. doi: 10.1109/EMBC.2018.8513378.

[50] C. Mosquera-Lopez, R. Dodier, N. Tyler, N. Resalat, and P. Jacobs, "Leveraging a Big Dataset to Develop a Recurrent Neural Network to Predict Adverse Glycemic Events in Type 1 Diabetes," *IEEE J Biomed Health Inform*, pp. 1–1, 2019, doi: 10.1109/jbhi.2019.2911701.

[51] M. Sarwar and T. R. Soomro, "Impact of Smartphone ' s on Soci ety," *European Journal of Scientific Research*, vol. 98, no. 2, pp. 216–226, 2013.

[52] P. Vallejo-Correa, J. Monsalve-Pulido, and M. Tabares-Betancur, "Systematic mapping review of context-aware analysis and its approach to mobile learning and ubiquitous learning processe," *Comput Sci Rev*, vol. 39, p. 100335, 2021, doi: 10.1016/j.cosrev.2020.100335.

[53] S. Zammali and S. Ben Yahia, "How to select and weight context dimensions conditions for context-aware recommendation?," *Expert Syst Appl*, vol. 182, no. June, p. 115176, 2021, doi: 10.1016/j.eswa.2021.115176.

[54] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles, "Towards a better understanding of context and context-awareness," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1999. doi: 10.1007/3-540-48157-5_29.

[55] A. R. M. Forkan and W. Hu, "A context-aware, predictive and protective approach for wellness monitoring of cardiac patients," *2016 Computing in Cardiology Conference (CinC)*, vol. 43, pp. 369–372, 2016, doi: 10.22489/CinC.2016.108-324.

[56] A. R. M. Forkan, I. Khalil, Z. Tari, S. Foufou, and A. Bouras, "A context-aware approach for long-term behavioural change detection and abnormality prediction in ambient assisted living," *Pattern Recognit*, vol. 48, no. 3, pp. 628–641, 2015, doi: 10.1016/j.patcog.2014.07.007.

[57] H. D. Vianna and J. L. V. Barbosa, "A scalable model for building context-aware applications for noncommunicable diseases prevention," *Inf Process Lett*, vol. 148, pp. 1–6, 2019, doi: 10.1016/j.ipl.2019.03.010.

[58] M. Quinde, J. C. Augusto, N. Khan, and A. van Wyk, "ADAPT: Approach to Develop context-Aware solutions for Personalised asthma managemenT," *J Biomed Inform*, vol. 111, no. October, p. 103586, 2020, doi: 10.1016/j.jbi.2020.103586.

[59] E. Zdravevski *et al.*, "Improving Activity Recognition Accuracy in Ambient-Assisted Living Systems by Automated Feature Engineering," vol.

5, 2017.

[60] J. Yang, L. Li, Y. Shi, and X. Xie, "An ARIMA Model with Adaptive Orders for Predicting Blood Glucose Concentrations and Hypoglycemia," *IEEE J Biomed Health Inform*, vol. 23, no. 3, pp. 1251–1260, 2019, doi: 10.1109/JBHI.2018.2840690.

[61] A. Aliberti *et al.*, "A Multi-Patient Data-Driven Approach to Blood Glucose Prediction," *IEEE Access*, vol. 7, pp. 69311–69325, 2019, doi: 10.1109/ACCESS.2019.2919184.

[62] Y. Dong, R. Wen, Z. Li, K. Zhang, and L. Zhang, "Clu-RNN: A New RNN Based Approach to Diabetic Blood Glucose Prediction," in *2019 IEEE 7th International Conference on Bioinformatics and Computational Biology (ICBCB)*, 2019, pp. 50–55. doi: 10.1109/ICBCB.2019.8854670.

[63] X. Chen, J. Tuo, and Y. Wang, "A prediction method for blood glucose based on grey wolf optimization evolving kernel extreme learning machine," *Chinese Control Conference, CCC*, vol. 2019-July, pp. 3000–3005, 2019, doi: 10.23919/ChiCC.2019.8866210.

[64] J. He and Y. Wang, "Blood glucose concentration prediction based on kernel canonical correlation analysis with particle swarm optimization and error compensation," *Comput Methods Programs Biomed*, vol. 196, p. 105574, 2020, doi: 10.1016/j.cmpb.2020.105574.

[65] G. Kriukova, N. Shvai, and S. v. Pereverzyev, "Application of regularized ranking and collaborative filtering in predictive alarm algorithm for nocturnal hypoglycemia prevention," *Proceedings of the 2017 IEEE 9th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2017*, vol. 2, no. September, pp. 634–638, 2017, doi: 10.1109/IDAACS.2017.8095169.

[66] B. P. Kovatchev, E. Otto, D. Cox, L. Gonder-Frederick, and W. Clarke, "Evaluation of a new measure of blood glucose variability in diabetes," *Diabetes Care*, 2006, doi: 10.2337/dc06-1085.

[67] L. Monnier, A. Wojtusciszyn, C. Colette, and D. Owens, "The Contribution of Glucose Variability to Asymptomatic Hypoglycemia in Persons with Type 2 Diabetes," *Diabetes Technol Ther*, 2011, doi: 10.1089/dia.2011.0049.

[68] V. V. Klimontov and N. E. Myakina, "Glucose variability indices predict the episodes of nocturnal hypoglycemia in elderly type 2 diabetic patients treated with insulin," *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, vol. 11, no. 2, pp. 119–124, 2017, doi: 10.1016/j.dsx.2016.08.023.

[69] S. Whalen and G. Pandey, "A comparative analysis of ensemble classifiers: Case studies in genomics," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2013. doi: 10.1109/ICDM.2013.21.

[70] U. Machine Learning Repository, "Diabetes Data Set." https://archive.ics.uci.edu/ml/datasets/diabetes (accessed Jul. 20, 2005).

[71] E. B. Schroeder, S. Xu, G. K. Goodrich, G. A. Nichols, P. J. O'Connor, and J. F. Steiner, "Predicting the 6-month risk of severe hypoglycemia among adults with diabetes: Development and external validation of a prediction model," *J Diabetes Complications*, vol. 31, no. 7, pp. 1158–1163, 2017, doi: 10.1016/j.jdiacomp.2017.04.004.

[72] T. Khan, M. A. Masud, and K. A. Mamun, "Methods to predict blood glucose level for type 2 diabetes patients," *5th IEEE Region 10 Humanitarian Technology Conference 2017, R10-HTC 2017*, vol. 2018-Janua, pp. 392–395, 2018, doi: 10.1109/R10-HTC.2017.8288982.

[73] G. Roffo, S. Melzi, and M. Cristani, "Infinite feature selection," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015. doi: 10.1109/ICCV.2015.478.

[74] G. Roffo, "Ranking to Learn and Learning to Rank: On the Role of Ranking in Pattern Recognition Applications," 2017.

[75] G. Roffo and S. Melzi, "Ranking to learn: Feature ranking and selection via eigenvector centrality," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017. doi: 10.1007/978-3-319-61461-8_2.

[76] G. Roffo, S. Melzi, U. Castellani, and A. Vinciarelli, "Infinite Latent Feature Selection: A Probabilistic Latent Graph-Based Ranking Approach," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017. doi: 10.1109/ICCV.2017.156.

[77] X. He, D. Cai, and P. Niyogi, "Laplacian Score for Feature Selection," in *Proceedings of Advances in Neural Information Processing Systems*, 2005. doi: http://books.nips.cc/papers/files/nips18/NIPS2005_0149.pdf.

[78] Y. Yang, H. T. Shen, Z. Ma, Z. Huang, and X. Zhou, "ℓ2,1-Norm regularized discriminative feature selection for unsupervised learning," in *IJCAI International Joint Conference on Artificial Intelligence*, 2011. doi: 10.5591/978-1-57735-516-8/IJCAI11-267.

[79] H. Zeng and Y. M. Cheung, "Feature selection and kernel learning for local learning-based clustering," *IEEE Trans Pattern Anal Mach Intell*, 2011, doi: 10.1109/TPAMI.2010.215.

[80] A. Folch-Fortuny, F. Arteaga, and A. Ferrer, "Missing Data Imputation Toolbox for MATLAB," *Chemometrics and Intelligent Laboratory Systems*, vol. 154, pp. 93–100, 2016, doi: 10.1016/j.chemolab.2016.03.019.

[81] B. Walczak and D. L. Massart, "Dealing with missing data: Part II," *Chemometrics and Intelligent Laboratory Systems*, 2001, doi: 10.1016/S0169-7439(01)00132-0.

[82] S. Gao, "Analysis of incomplete multivariate data," *Statistical Methods in Medical Research*. 1999. doi: 10.1177/096228029900800109.

[83] N. v. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, 2002.

[84] M. B. Abraham *et al.*, "ISPAD CLINICAL PRACTICE CONSENSUS GUIDELINES ISPAD Clinical Practice Consensus Guidelines 2018 : Assessment and management of hypoglycemia in children and adolescents with diabetes," vol. 19, no. March, pp. 178–192, 2018, doi: 10.1111/pedi.12698.

[85] F. Cameron, G. Niemeyer, D. Ph, K. Gundy-burlet, D. Ph, and B. Buckingham, "Statistical Hypoglycemia Prediction," vol. 2, no. 4, pp. 612–621, 2008.

Chapter 4

Framework evaluation

The chapter presents and discusses the evaluation of the hypoglycaemia predictive framework and is presented as the research paper intitled "Hypoglycaemia prediction using information fusion and classifiers consensus.".

4.1 Hypoglycaemia prediction using information fusion and classifiers consensus.

Virginie Felizardo, Nuno Garcia, Imen Megdiche, Nuno Pombo, Miguel Sousa and František Babič.

Hypoglycaemia prediction using information fusion and classifiers consensus.

submitted to a Computer Science journal, Q1, by Scimago Journal & Country Rank.

This research paper presents the evaluation of the framework for hypoglycaemia prediction, performing personalised predictions for real patients. The predictive algorithm offers a good prediction of the risk of hypoglycaemia in a 24-hour time frame using sparse data.

Hypoglycaemia prediction using information fusion and classifiers consensus

Virginie Felizardo[a,b]*, Nuno Garcia[a,b], Imen Megdiche[c], Nuno Pombo[a,b], Miguel Sousa [a], František Babič[d]

[a]*Universidade da Beira Interior, Covilhã, Portugal*
[b]*Instituto de Telecomunicações, Covilhã, Portugal*
[c] *Institut National Universitaire JF Champollion - Ecole d'ingénieur ISIS , Castres, France*
[d] *Department of Cybernetics and Artificial Intelligence of Technical University of Kosice, Košice, Slovakia*

virginie@it.ubi.pt, ngarcia@di.ubi.pt, imen.megdiche@univ-jfc.fr, ngpombo@di.ubi.pt, mcbranco@ubi.pt, frantisek.babic@tuke.sk

Abstract

The recommendation that there must be a balance between insulin, food, and exercise to keep diabetes under control provides an opportunity for developing mobile applications for real predictions considering context. These predictions can improve the quality of patients' lives by avoiding unwanted events, namely, hypoglycaemia. We propose a hypoglycaemia prediction approach combining information fusion and classifiers consensus. The approach fuses information on two processes, Learning Process and Application Process, to predict the risk of hypoglycaemia in a 24-hour window. The first process learns multi-classifiers from different sources of different patients. The process integrates contextual and time-based features with the conventional inputs. In the second one, the data are collected from a unique patient; after extracting and imputing the features, we use the classifiers resulting from the learning process in order to get the classifiers consensus decision. So that we fuse the evaluation information from different classifiers. This is used to evaluate the effectiveness of the method. We performed experiments using data from 54 patients extracted from the University of California Irvine diabetes dataset. The results from classifiers consensus decision provide very promising results, which are acceptable considering that we used discrete data which are so far not studied in the literature.

Keywords: data fusion, information fusion, hypoglycaemia prediction, classifiers consensus.

1. Introduction

Diabetes mellitus (DM) is a chronic disease characterised by the difficulty to control blood glucose (BG) due to the dysfunction of insulin production by the pancreas. In 2021, approximately 537 million adults around the world are affected by this disease [1] which prevalence is estimated to reach 783 million adults in 2045 [1].

The poor control of BG results in some short- and long-term complications. These complications can seriously decrease the quality of life of diabetic patients; they can lead to blindness, amputation, or heart problems. Hypoglycaemia (or hypoglycemia in US English) is one of the most problematic short-term issue of a bad glycaemia control. Hypoglycaemia is the drop in BG

below 70 mg/dL or 3.9 mmol/L. This threshold is the most accepted in medical circles [2], [3]. The hypoglycaemia is commonly feared because it arrives unannounced and is essentially asymptomatic. It may lead to loss of consciousness, confusion, seizures, blurred vision, erratic behaviour, garbled speech, clumsiness and, in the worst cases, it can cause death [4].

Currently, the trend for glycaemia control is the use of continuous glucose monitoring sensors, often coupled with an insulin pump, named sensor-augmented pump therapy. Despite its increasingly adoption by patients and practitioners, several limitations were observed due to its: (1) inaccuracy [5], [6], (2) reimbursement [6], (3) cost [5], [7] and (4) frequent annoying false alarms [5]. In line with this, the invasive method based on finger sticks is broadly used for the blood BG monitoring. A complementary solution based on technology is available in online web stores named the self-management applications (apps), such as Glucose Buddy [8], Dottli: A [9], Sugar Sense [10], Social Diabetes [11], and Diabetes M [12] just to mention a few. These applications allow a continuous recording of patient data, but they are sometimes not supported by medical evidence. Furthermore, there are similar products oriented to specific glucose meters devices, e.g., Contour®Diabetes app [13] and OneTouch Reveal® [14].

Practitioners and researchers proposed solutions to promote the patients' insight through embedding expert knowledge within the glucose monitoring system. Extensive work on predictive models has been done using data from continuous glucose monitoring [15]–[27]. Hypoglycaemia prediction using discrete data, to the best of our knowledge, lacks for evidence and/or is understudied, despite being the most common scenario for patients who try to control their BG levels. The major brake on this approach relies on the data sparsity [28], [29] and complexity. In addition, this method is not applied to patients using BG monitoring performed using finger sticks.

These limitations associated with sensors, combined with the lack of contextual open resources, noncompliance of patients with the study protocols and the use of discrete information in predictive works, challenges for innovative and disruptive solutions. Introducing predictive algorithms for the discovery of personalised knowledge may lead to an earlier detection of changes in patients' health status and to expect activities to mitigate risks. The cornerstone of predictive models is its ability to "produce" knowledge. For that reason, machine learning (ML) algorithms are implemented to improve the accuracy of prediction over the use of physiological models by capturing complex, nonlinear relationships in the data [30], [31]. Also, there is still a need for fuses different sources of information and fuses different decisions to make the predictions more useful and accurate for the patients with diabetes.

In the present research, we propose a hypoglycaemia prediction approach using information fusion and classifiers consensus. We summarise the major contributions of this paper as follows:

- A method for hypoglycaemia prediction using sparse data, information fusion and classifiers consensus decision in a 24h hour timeframe;

- An experiment's evaluation presenting personalised predictions using 54 patients from diabetes dataset of UCI Machine Learning Repository;

- A contribution to the state-of-the-art including a benchmark to determine which classifier combinations perform better.

We organise the rest of the paper as follows. Section 2 and 3 present the problem definition and related work that inspired our approach. In section 4, we introduce the method of the predictive algorithm framework. Section 5 presents the results from the evaluation experiments. We present the discussion of experiments in Section 6. Section 7 presents the conclusion and a proposal for future research work.

2. Problem definition

An extensive work [15], [16], [30], [32]–[40], [18], [41], [19], [20], [23]–[27] has been done on BG prediction using data collected from continuous glucose monitoring (CGM) systems. In the literature, we observe that real-time CGM and the hypoglycaemic alarm contribute to reduce the incidence of hypoglycaemia [6] and time spent when hypoglycaemia [6] occurs. However, these results are not extensible to patients who use finger sticks to monitor BG. The sparsity of data from this type of BG measurements leads to discrete data, which are not yet explored in the literature [42]. However, most medical experts agree that self-monitored blood glucose (SMBG) performed using finger sticks is essential to manage hypoglycaemia [43].

In order to address this issue, we have proved in our previous work [44] that integrating contextual and time-based features with the meal, exercise and therapy inputs improves the accuracy of the predictions. Our framework [44] offers a good prediction of risk of hypoglycaemia in a 24-hour window using sparse data. In order to produce as accurate predictions as possible for each patient, even if they do not have access to CGM systems, we propose an extension of our work in which we explore all combinations of classifiers to generate a consensus prediction. Figure 1 shows the proposed framework that comprises information fusion, given by a machine learning system composed of a multi-classifier learning, multi-classifiers decision and the classifiers consensus decision by majority vote. The extension of our previous work is focused on classifiers consensus decision, exploring all classifiers' combination in order to provide some insights into which combinations present suitable predictions and which are their characteristics.

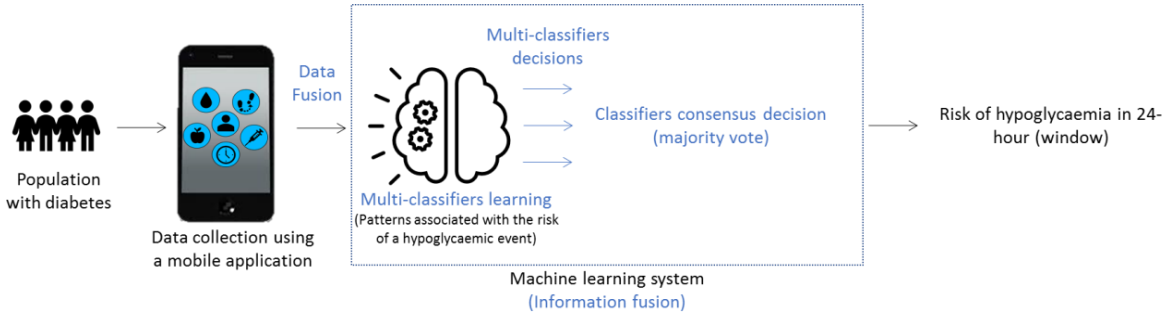

Figure 1 – Proposed framework.

3. Related work

Extensive work has been done on blood glucose prediction [4], [42], [45]–[47] but just some work [21], [24], [27], [36], [37] address the prediction of events in diabetes, such as hyperglycaemia or hypoglycaemia. Recent reviews [42], [45] of data-based techniques and approaches based on data fusion for blood glucose and hypoglycaemia prediction highlight the following topics:

1) The hypoglycaemia prediction has been addressed as short-term glucose predictions based on a prediction horizon between 5 and 180 minutes.

2) The inter-variability in daytime of meals, medication and exercise motivated some studies to use only glucose data as predictor even if it is common to combine with other predictors like meal, insulin, and physical activity. Some studies [16], [18]–[20], [23], [24] used feature selection and tested different combinations of features, or showed the importance of some features relative to others.

The selection of proper predictors is crucial to create models that perform good results for events predictions [45], [48]. Some works showed the importance of data fusion using different sources of features. Georga *et al.* [18] show that the fusion of features like meal intake, insulin and exercise are small but not irrelevant. They also show that exercise features based on time are more significant in short-term predictions than in long-term predictions. Zarkogianni *et al.* [19] show the importance of the fusion of exercise features with glucose features and the time of the day, including some features connected to glucose dynamics. Also, He *et al.* [23] show that the fusion of information on meals, insulin and time provide dominant causal correlations. These findings support that data fusion can be appropriate for diabetes future events prediction. In fact, as evidenced in our previous work [44], integrating contextual and time-based features with the conventional features improves the prediction accuracy. In our work, the context is provided by three dimensions: time, behaviour, and time-based features, giving supplementary information to bridge the discrete data and allow the recognition of distinct patterns associated with hypoglycaemia risk.

Fewer studies in literature [28], [29], [44] explored the idea of creating a glucose level prediction model using discrete data from a population of diabetics to provide a personalised prediction.

Sudharsan *et al.* [28] proposed a classification approach using a mixed measure dataset (with numerical and categorical inputs) and a binary target class. The authors proposed a 24-hour prediction window, like time frame, for delivering an intervention. Li and Fernando [29] proposed an approach based on smartphone-collected data using pooled population data to deliver a personalised BG prediction without window time.

In our previous work [44], we proposed a novel framework prediction of risk of hypoglycaemia in a 24-hour window using a discrete data fusion and a multi-classifier decision, that we named consensus. These work results showed the potentialities of data and information fusion of combinations of three or five classifiers to improve the performance of predictions.

Scientific community proposed the combination of predictors and techniques hoping to represent spatial and temporal input and output dependencies. However, no specific method

provides 100% reliable event predictions. Because of individual inter- and intra-variability, no specific classifiers provide valuable solutions for all patients [49].

Some work in several domains, namely bioinformatics [50]–[52], Alzheimer's disease [53], disease prediction [54]–[56] and for fatigue recognition [57] resort to ensemble methods to achieve accurate predictions for several complex tasks. So, a way to get a successful prediction model is to use ensemble methods, combining the output of individual classifiers [50]. Bashir et al. [54], [55] stated that there are two categories of ensemble frameworks: homogeneous ensemble frameworks and heterogeneous ensembles frameworks. The homogeneous ensembles used the same base classifiers, whereas the heterogeneous ensembles used different base classifiers. Our approach uses single classifiers but also ensembles. So, by exploring all combinations, we can get heterogeneous ensembles or ensembles of ensembles decisions resulting from the fusion of the different models. The ensemble of ensembles approach is already used in literature [58]. In our perspective, fusing decisions from heterogenous ensembles or ensembles of ensembles can be solve some problems between them such as complexity and variability of data from different person.

4. Method

The proposed framework is based on the rationale of data and information fusion applied in distinct steps for hypoglycaemia prediction. Figure 2 explains the method aimed at predicting hypoglycaemia events, which comprises two processes, learning process and application process. The purpose of the learning process is to build and train the models using data from a population of patients and the application process purpose is the prediction of hypoglycaemia using data from a never-before-seen patient.

4.1 Learning Process

The learning process comprises three different levels: 1) Data collection using a mobile application, 2) Pre-processing, and 3) multi-classifiers learning.

Data collection may be performed via a mobile application and organised to allow that user can register their data: BG level, type and quantity of insulin, information about meal, information about physical activity and contextual information (Figure 2, A). For the framework evaluation, we used the diabetes dataset [59] from the UCI Machine Learning Repository that contains datasets from 70 type-1 diabetics. This paper used data from fifty-two patients from fifty-three (P1-P5, P7-P46, P50, P54-P56, P65, P67-P68, P70) for the learning process, being that one patient is always left out for the application process. In application process, we apply the principle of leaving one patient out validation for all patients, including now P6, patient without hypoglycaemia event record not used in the learning process.

Pre-processing includes data imputation comprising blood glucose imputation when the user does not provide this information, feature extraction that comprises time-based features, and re-sampling to balance the data because hypoglycaemia events are always a minority class in

comparison with the totality of the data (Figure 2, B). The methods of feature extraction and data imputation are performed in each patient data to impute more realistic data and re-sampling in population data to solve unbalanced data.

The features applied on this approach that can be obtained directly from the dataset were: BG, Insulin regular, Insulin NHP, Insulin Ultralent, Meal, Exercise, Hour of the day and Time slot. Some of these features belong to context dimensions: 1) time, allowing information deliver about when some actions were performed and 2) behaviour, is given by information about meal or physical activity, the behaviour feature answers the question 'how these actions were performed?', 3) For the time-based features which are extracted from the original features, we find 3.1) the duration between consecutive records, 3.2) the difference between actual Blood Glucose record and previous one, 3.3) the difference between actual Blood Glucose record and previous 24h record and 3.4) the Blood Glucose of the previous record. The glucose time-based features, specifically Difference between actual Blood Glucose and previous record and Difference between actual Blood Glucose and previous 24h, constitute the link between the time feature and glucose level feature, summarising how the glucose level varies over time.

The multi-classifiers learning comprises building different models using population data to learn (Figure 2, C). The single classifiers used were 1) Classification and Regression Tree (CART), 2) Support vector machine (SVM), 3) k-Nearest neighbour (kNN), and the ensemble classifiers used were 4) Adaboost Decision Tree (AdaBTree), 5) Random Forest (RF) and 6) Subspace k-Nearest neighbour (SkNN).

4.2 Application Process

Application process comprises four different levels: 1) Data collection using a mobile application, 2) pre-processing, 3) prediction and 4) consensus decision.

The procedure for data collection is the same as that for the learning process (Figure 2, D). In data pre-processing, only feature extraction and data imputation are used (Figure 2, E). Re-sampling is not used because the aim is the prediction of hypoglycaemia using the models developed in the learning process. So, using the models developed in the learning process to perform the multi-classifiers decision, getting the decision from each model for the data provided by the user (Figure 2, F). After the fusion of different models' decision by majority vote, we get the classifiers consensus decision (Figure 2, G).

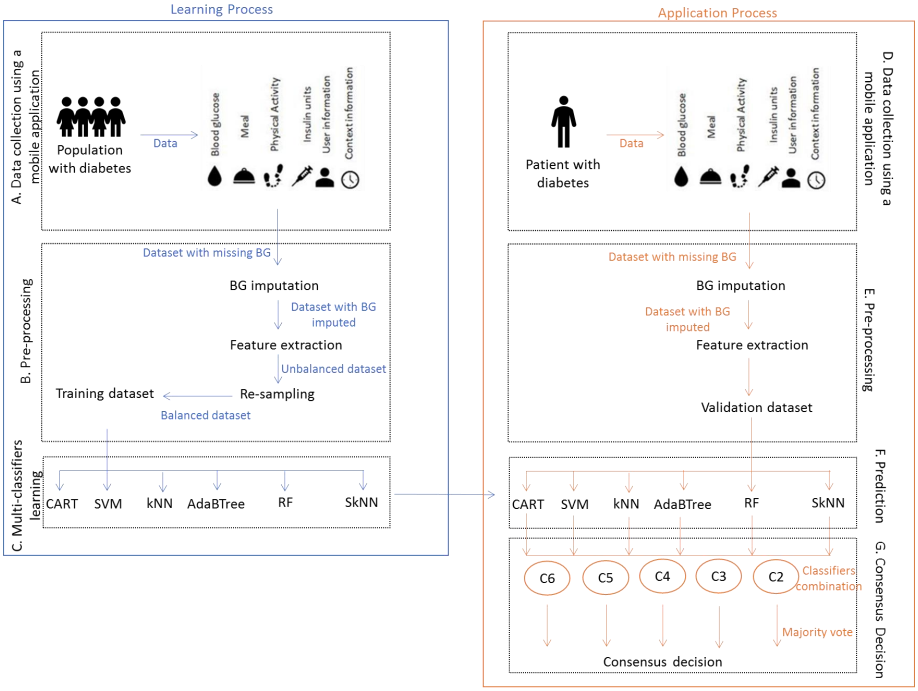

Figure 2 - The method for Hypoglycaemia prediction.

5. Evaluation

The method described in the previous section was performed for fifty-four patients (P1-P5, P6-P46, P50, P54-P56, P65, P67-P68, P70) in order to get a personalised consensus decision. First, for each patient left out, the models are trained using data from fifty-two patients (without P6, only used for predictions), that is, six models are built each time. Second, the models got are used for the predictions performing six decisions that are combined to get the consensus decision. Figure 3 shows an example of how the method is applied.

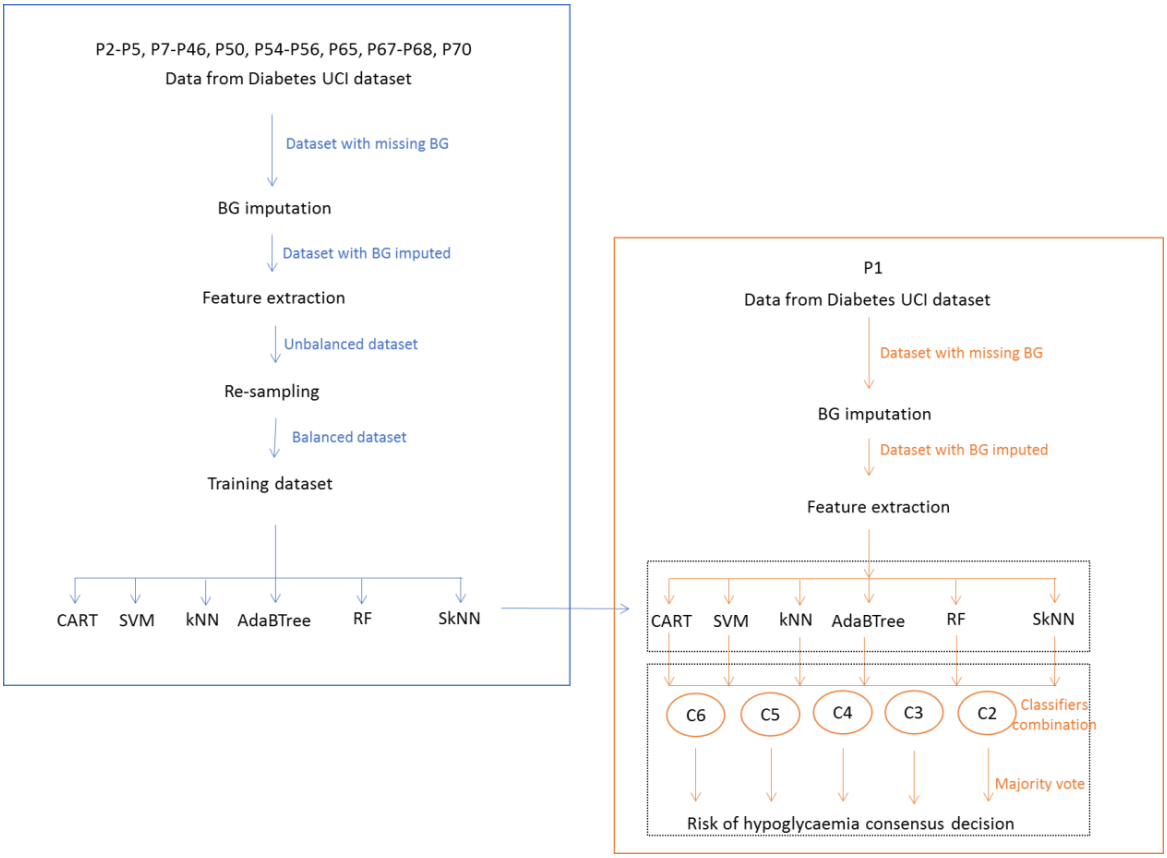

Figure 3–Example of how the method applies to get P1 personalised prediction.

5.1 Dataset and settings

According to the patients that we perform their prediction, the number of instances of the population varies between 11113 to 11738. After performing imputation of data for each patient, we applied SMOTE method to population data two times: one for class 2, with k=6, and other for class 1, with k=2. The k is the nearest neighbours in feature space and this value was fixed for all patients. Table 1 shows the settings for each classifier used and fixed for all patients.

Table 1 - Classifier's settings.

Classifier	**Settings**
Decision Tree (CART)	Maximum number of splits: 100 Split criterion: gini s diversity index
Support vector machine (SVM)	Kernel function: Gaussian Box constraint level: 1 Kernel scale: 0.87
k-Nearest neighbour (kNN)	Number neighbour: k=1 Distance metric: euclidian Distance weight: Equal
Adaboost+Decision Tree (AdaBTree)	Maximum number of splits: 20 Number of learners: 30 Learning rate: 0.1
Random Forest (RF)	Maximum number of splits: sample-1 Number of learners: 30
Subspace+ k-Nearest neighbour (SkNN)	Maximum number of splits: 20 Number of learners: 30 Subspace dimension: 6

5.2 Classifier's consensus decision

By applying the data of each patient to the six models previously cited, we obtain six individual decisions for each patient. Each classifier consensus decision resulted from the different combinations of classifiers' decision aggregated by majority vote. We tested different combinations: 1) combination of two classifiers (C2), 2) combination of three classifiers (C3), 3) combination of four classifiers (C4), 4) combination of five classifiers (C5) and 5) a unique combination of six classifiers (C6).

Table 2, 3, 4 and 5, in Annex, show the average results for each classifiers combination. In general, the results are similar for the different combinations. For C2 and C4, the results show less false alarms (FA), but the sensitivity results suggest a reduction of predicted events. Contrariwise, for C3, the results show better sensibility but more false alarms. As through this first analysis, it is not possible to find out major differences, so we will analyse in detail other aspects.

Figures 4 and 5 show a comparative analysis between all classifiers consensus that achieved an accuracy higher than 60% and false alarms less than 30%. These ratios were defined after analysis of the results presented in Table 2, 3, 4 and 5, and are considered being adequate in

terms of the expectations of the users and crucial for a valuable solution for hypoglycaemia prediction.

In Figure 4 (a) and (b), is presented the accuracy results and false alarms results in percentage for each combination of classifiers. The results show that the average accuracy is 71.8% and the average of false alarms is 17.2%.

In Figure 5 (a) and (b), is presented the hypoglycaemia predicted results and patients with adequate solutions in percentage for each combination of classifiers. The results show that the average hypoglycaemia predicted is 77.3% and the average of patients with adequate solutions is 56.6%. Regarding these results, we highlight the consensus decision using four classifiers: C4.7, C4.8 and C4.10. These solutions presented over 70% of hypoglycaemia predicted for over 70% of patients with adequate solutions (accuracy>60% and FA<30%).

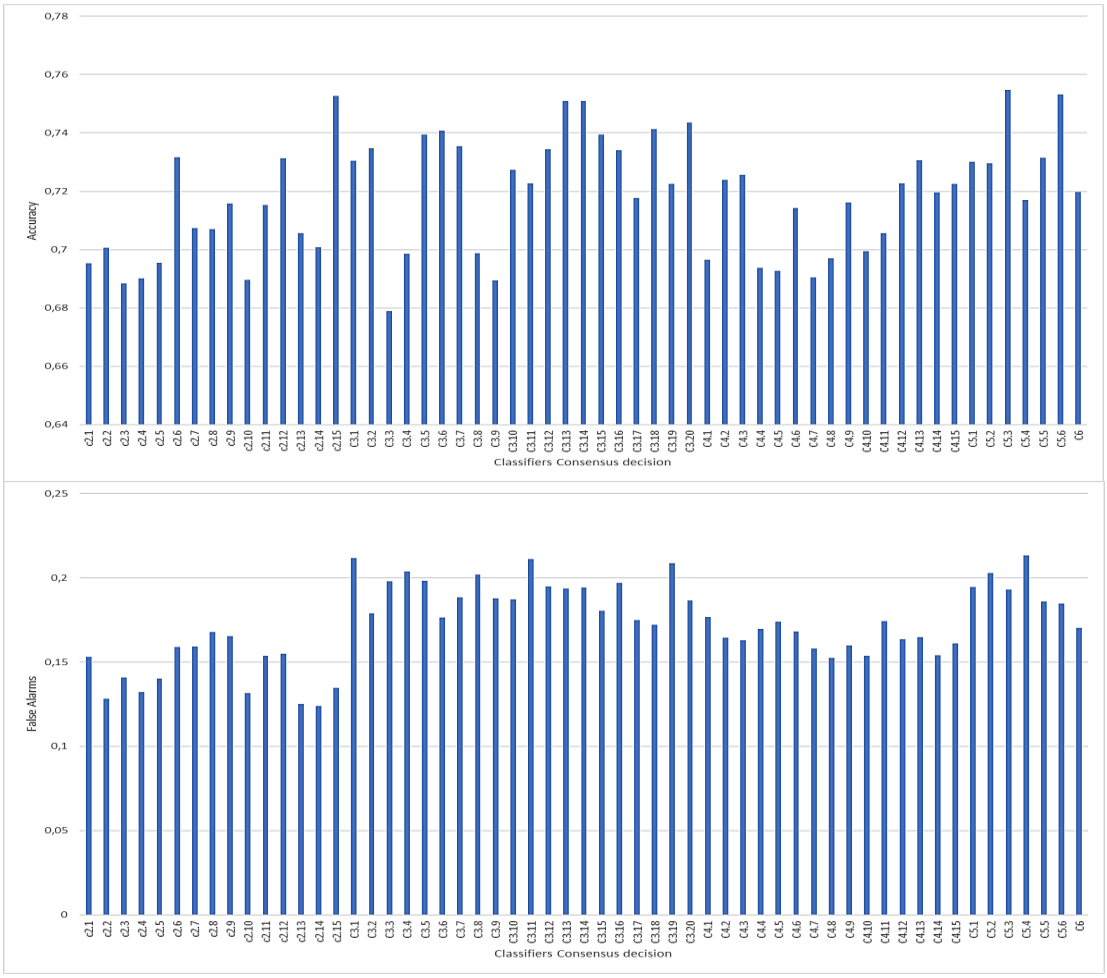

Figure 4–Accuracy (top) and false alarms (bottom) results for each combination of classifiers.

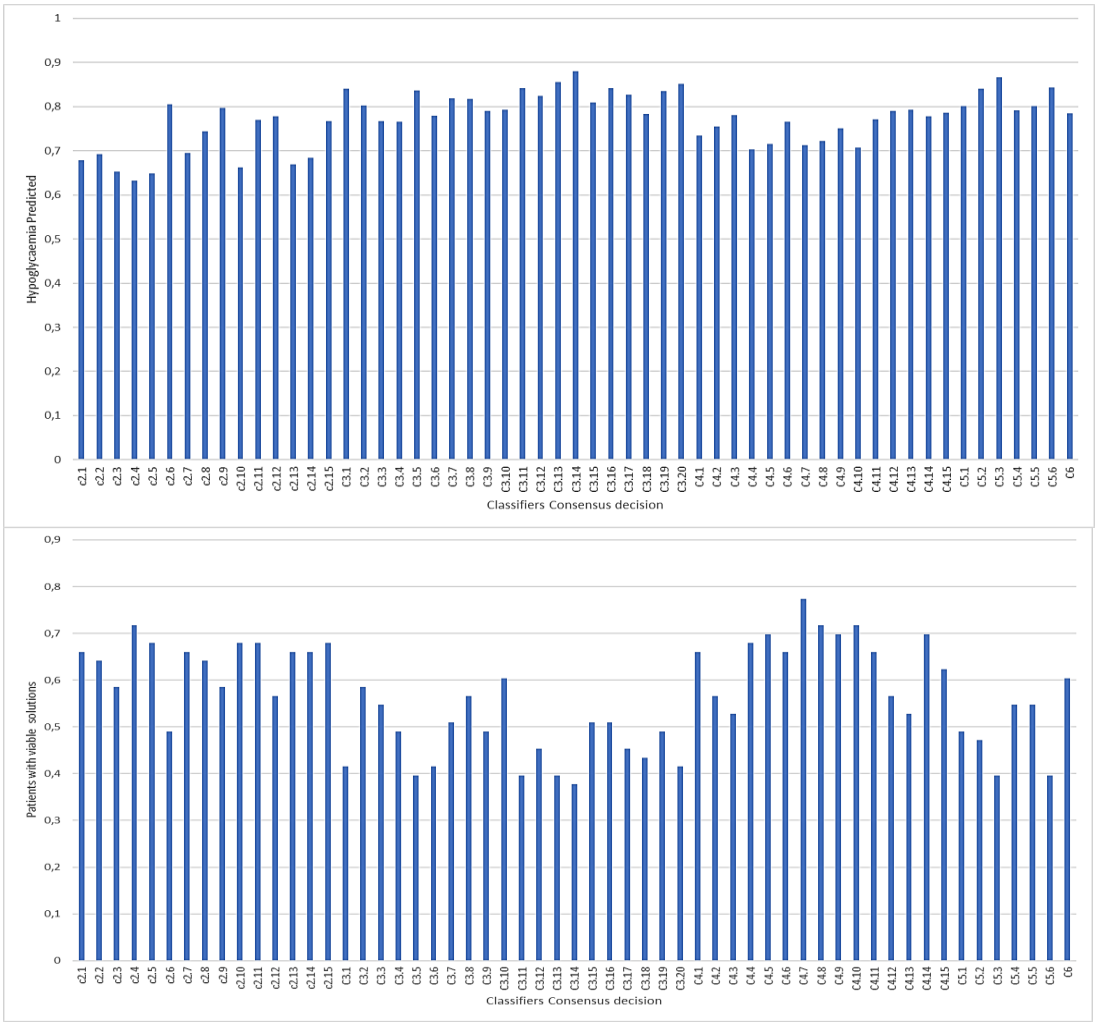

Figure 5 - Hypoglycaemia predicted (top) and patients with adequate solutions (bottom) results for each combination of classifiers.

In terms of better patients' results, Table 6 shows four perspectives: 1) individual classifier vs consensus decision with combinations of classifiers, 3) Hypoglycaemia prediction and 4) False alarms. This Table shows the representativity of patients between each perspective.

Table 6 – Representativeness of each aspect: 1) patients with best solution achieved with an individual classifier or classifier consensus decision with discriminated combinations of classifiers Decision (CCD); 2) Hypoglycaemia prediction; and 3) False alarms.

Perspectives		**Number of Patients (% of patients)**	**Patients**
Consensus Classifier Decision (CCD) Vs Individual Classifier (54 patients)	**C1**	10 (18.5)	P7, P13, P14, P17, P24, P25, P27, P30, P35, P39
	C2	18 (33.3)	P3, P4, P10, P15, P16, P18, P19, P32, P33, P34, P36, P37, P40, P43, P46, P68, P70, P6
	C3	17 (31.5)	P2, P5, P8, P12, P19, P20, P21, P23, P26, P28, P29, P31, P40, P41, P42, P44, P50, P55, P67
	C4	14 (25.9)	P1, P9, P11, P12, P19, P22, P38, P40, P45, P54, P56, P65
	C5	1 (1.9)	P19
	C6	1 (1.9)	P11
Hypoglycaemia Prediction (53 patients)	**100%**	22 (41.5)	P3, P4, P5, P7, P9, P10, P11, P12, P13, P17, P18, P19, P30, P31, P32, P33, P35, P36, P38, P40, P44, p45, P67
	90%-100%	13 (24.5)	P15, P23, P24, P26, P28, P34, P39, P42, P46, P50, P55, P56, P65
	80%-90%	6 (11.3)	P14, P20, P25, P27, P54, P68
	70%-80%	10 (18.9)	P1, P2, P8, P16, P22, P29, P37, P41, P43
	60-70%	1 (1.9)	P21
	<60%	1 (1.9)	P70
False Alarms (54 patients)	**>30**	2 (3.7)	P43,P46
	25-30	18 (33.3)	P1, P2, P3, P9, P13, P14, P20, P22, P25, P28, P29, P37, P38, P40, P41, P42, P54, P56, P67
	20-25	16 (29.6)	P7, P10, P21, P24, P26, P27, P30, P31, P33, P35, P45, P55, P65, P68, P70
	15-20	7 (13.0)	P8, P12, P15, P16, P23, P39, P44
	10-15	5 (9.3)	P17, P32, P34, P36, P50, P6
	5-10	4 (7.4)	P5, P11, P18
	<5	2 (3.7)	P4, P19

Analysing the Table 6, the first perspective shows the representativity of patients with best solution achieved with individual classifier or classifier consensus decision. So, most of the patients (81.5%) achieved best solution with the classifier's consensus decision. Between the CCD, the combinations with two and three classifiers have a greater representativeness (33.3% and 31.5%, respectively), followed by combinations with four classifiers (25.9%), combinations

with five and six classifiers show fewer representativeness (1.9%). For 18.5% of the patients, the best solution is achieved with an individual classifier (C1). In addition, 41.5% of the patient's best solution achieved 100% of hypoglycaemia prediction, followed by 24.5% that achieved between 90-100%, 18.9% between 70-80%, 11.3% between 80-90% and 3.8% achieved less than 70%. Finally, 33.3% of the best solution achieved false alarms between 25-30%, 29.6% between 20-25%, 13% between 15-20%, 9.3% between 10-15%, 7.4% between 5-10%, 3.7% with less than 5% and 3.7% with over 30% of false alarms.

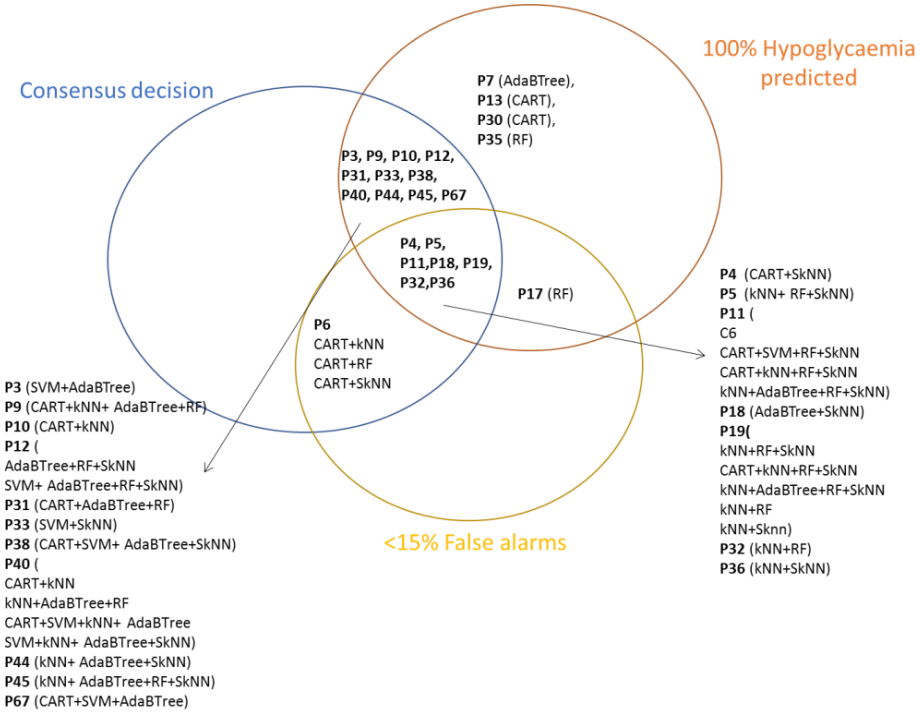

Figure 6–Best patients solution representativity from three perspectives: 1) consensus decision (blue circle), 2) 100% of hypoglycaemia predicted (orange circle) and 3) less than 15% of false alarms (yellow circle).

In Figure 6 are presented, the patients and the respective best solution distributed between three perspectives: 1) consensus decision (blue circle), 2) 100% of hypoglycaemia predicted (orange circle) and 3) less than 15% of false alarms (yellow circle). This figure aims to show the representativity of best patient solutions from these three perspectives. This information shows that to achieve 100% of hypoglycaemia predicted, each patient's best solution is different, with only two combinations providing best solutions for P10, P40 and P6 with C2.2 and for P11, P19 and P45 with C4.14. These results show that the influence of the features dimensionality and inter- and intra-variability and the need for further work in terms of dimensionality reduction and clustering. However, some interesting aspects can be extracted from these results, e.g., for the patient's best solution that achieved 100% of hypoglycaemia predicted with an individual classifier, the tree-based solutions (CART or RF or AdaBTree) are presented for all of them. For the solutions that use the consensus decision and achieved 100% of hypoglycaemia predicted,

we note that most solutions combined tree-based classifiers and kNN-based classifiers (13 patients from 18).

Next, the Figures 7 and 8 show four individual BG profiles (patients 3, 5, 19 and 6, respectively) with the real and predicted alarms. The predicted alarms presented in figures were achieved with the better solution for each of the patients. According to Figure 6, the patient 3 belongs to the group that achieved 100% of Hypoglycaemia with the consensus decision and over 15% of false alarms while patient 5 belongs to the group that achieved 100% of Hypoglycaemia with the consensus decision and less than 15% of false alarms. The principal difference in the data of these two patients is that P5 collected the data four times per day and the data from P3 present missing slots. So, our method performs well when the data from time has no missing slots making fewer mistakes, 26.9% of false alarms for P3 (Figure 7, top) and 7.6% for P5 (Figure 7, bottom). In Figure 8, the patient 19 belongs to the same group as patient 5, but in this case our method does not present false alarms (Figure 8, top). As for patient 5, the data of patient 19 has no missing slots but the data are collected regularly, this is always in the same hour. Patient 6 was used because he had no hypoglycaemic events and we thought it would be a good scenario to evaluate our method. We can see that our method only failed 10% and that the alarm appears following sudden drops in BG data (Figure 8, bottom). In addition, regularity at the time of data collection is also verified.

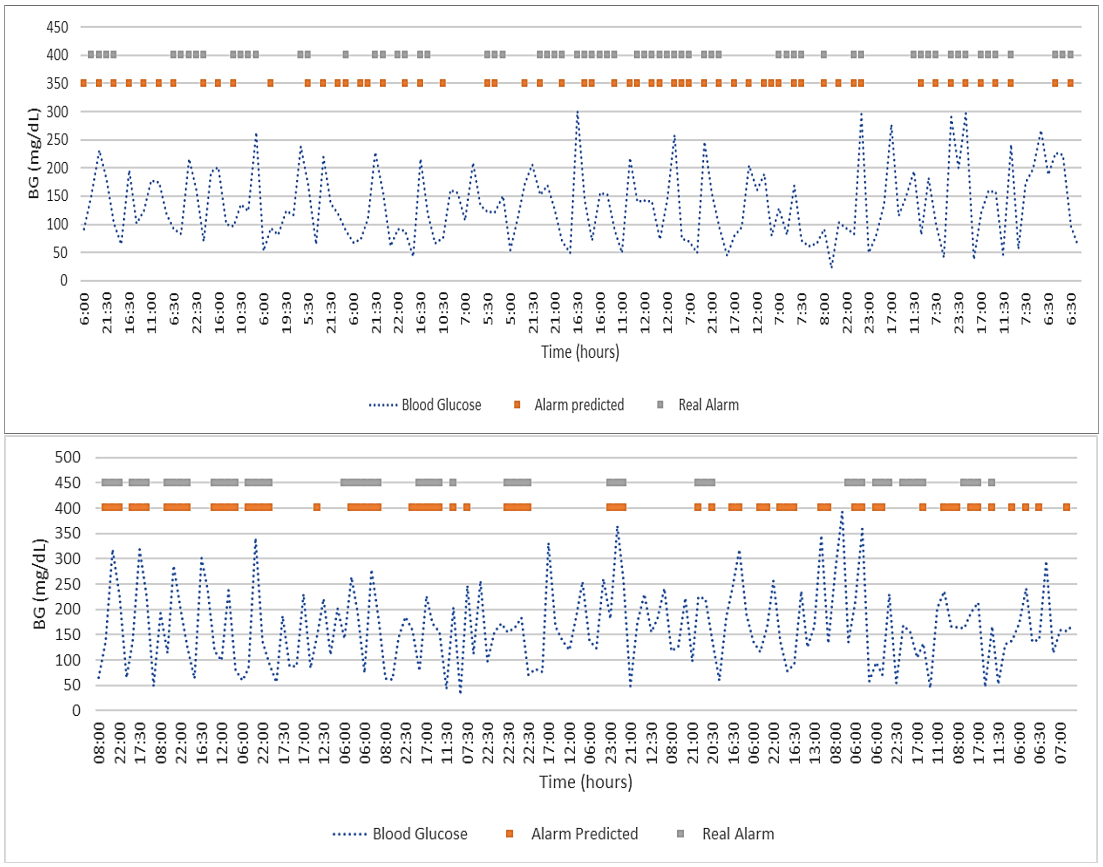

Figure 7 - Individual BG profile Patient 3 (top) and Patient 5 (bottom).

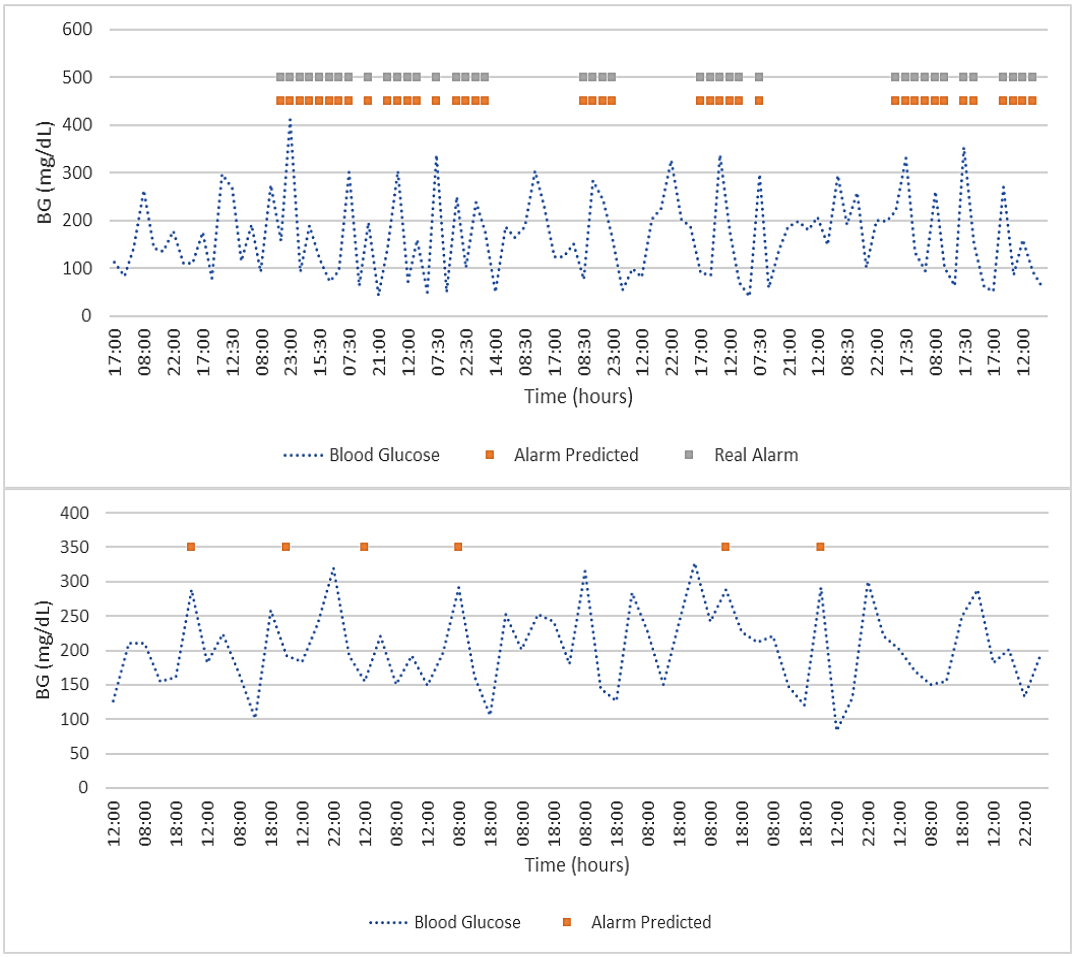

Figure 8 - Individual BG profile Patient 19 (top) and Patient 6 (bottom).

6. Discussion

The principal goal of this work was to show the efficiency of hypoglycaemia prediction in a 24h hour timeframe, using sparse data, information fusion and classifiers' consensus decision. In the following paragraphs, we discuss the present research work presenting the main findings, benefits, opportunities, and challenges.

Our approach supports the development of predictive solutions addressing **data and information fusion** in order to make their use viable in a self-management scenario where diabetics are outpatients performing their self-care. Hypoglycaemia may be several triggering, *e.g.* an overdose of insulin, a low intake of carbohydrates, an excess of physical activity or use of some medicines. This is one reason for the use of data fusion that will allow to describe the various scenarios of a patient with diabetes. In our previous work [44], we show that features and context dimensions, also used in this work, can describe different characteristics and behaviours of patients with diabetes and improve prediction accuracy. This is important for the

generalisation of the models and for catching several patterns associated with hypoglycaemia risk. In addition, the predictive framework supports hypoglycaemia prediction with sequential information fusion. First, the data collected was used for extraction of features where a fusion of features and context dimension were performed. The fusion of the dimension time with blood glucose gives a new knowledge about glucose profile changes in function of time. Second, imputation and re-sampling methods also contribute to the addition of information. Third, we used a multi-classifier learning that is a promising means to fuse information from data influenced by different person dynamics, inter- and intra-variability and catching appropriate patterns associated with hypoglycaemia prediction.

Data fusion of patients with diabetes can be challenging, but nowadays, some mobile and wearable technology allows monitoring different source of information. However, there is the need to create networks of monitoring devices that combine physiological or clinical data with different context dimensions.

Our approach uses information fusion of multi-models to give a **classifiers consensus decision** to predict the risk of hypoglycaemia for a specific patient. Considering that hypoglycaemia events are commonly feared in patients with insulin therapy and that can cause patient loses consciousness, we favour solutions that have a higher value of predicted events. So, the combinations of three or five classifiers show achievements of over 80% of hypoglycaemia predictions. However, we know that this kind of solution need acceptance from the users, so, it is crucial to consider the false alarms. The average result of false alarms shows that the combinations of three or five achieved FA between 25-30%, a combination of four and six classifiers percentage values between 20-25% and combinations of two classifiers between 15-20%. Although the results for the combination of two classifiers show less FA, its predictive ability is also smaller. For these experiments, the consensus decision achieved a better solution for 81.5% of patients. Only for ten patients (P7, P13, P14, P17, P24, P25, P27, P30, P35 and P39) the best solution was achieved with an individual classifier. However, we did not see any significant pattern between the combinations that provide better solutions. This is explained by inter- and intra-variability of patients. The highest percentage of patients' best solution was achieved with combinations of two and three classifiers. As we mentioned earlier, the solution that provides better results varies from patient to patient, but we highlight three combinations that provide better results for four patients each. The combination C2.2 (CART+KNN) provides better results for P6, P10, P40 and P43; the combination C2.5 (CART+SKNN) for P4, P6, P37 and P70; and combination C2.12 (KNN+SKNN) for P19, P34, P36 and P46. From these results, we can conclude that KNN-based ensembles show some potentialities for some clusters of patients. Also, considering the solutions that use the consensus decision and achieved 100% of hypoglycaemia predicted, most solutions combined tree-based classifiers and KNN-based classifiers.

Hypoglycaemia prediction is done by the decision of the multi-models instead of individual model decision. Regarding the best results with an individual model, the solution achieved over 90% of hypoglycaemia predicted for 54.7% of the patients (30.2% with 100% of hypoglycaemia predicted). Using the classifiers consensus decision 66% of the patients have over 90% of hypoglycaemia predicted (37.7% with 100% of hypoglycaemia predicted).

Considering both approaches, the hypoglycaemia predicted was achieved with over 90% for 66% of patients' best solutions (41.5% of patients best solutions with 100% of hypoglycaemia predicted). From these solutions, 69.8% are KNN-based. Considering the results with better individual classifier, the results show false alarms are between 20-30% for 56.6% of the patients. Using the classifiers consensus, the interval of false alarms 20-30% is the one with the greatest representativeness, 62.3% of the patients. Considering both approaches, for 63% patients, the false alarms are between 20-30%. In addition, the findings show that our method performs better in patients with regular time slots and time presenting less false alarms.

To conclude, this work shows the importance of data fusion and consensus decision to catch the patterns associated with hypoglycaemia risk and its prediction. Data fusion from different sources and context can represent patients with diabetes in different scenarios. A sequential information fusion can help to understand the complexity of this problem and the classifiers consensus decision can allow the proposal of a better solution for a specific patient, decreasing the effects of inter- and intra-variability.

7. Conclusion

Data and information fusion for hypoglycaemia events prediction is a challenge for physicians and medical scientific research. Many patients with diabetes are out-patients doing our self-monitoring and do not have access to CGM systems, so solutions using discrete data collected by, *e.g.* a self-management mobile application can help to prevent this kind of dangerous events. Our predictive algorithm satisfies these requirements, and the results achieved in this work show the potentialities for a future incorporation into a self-management solution. Also, data and information fusion can give a good contribution to represent patients with diabetes in different scenarios and giving new knowledge about the complexity of patterns associated with risk of hypoglycaemia.

The main findings of this work are summarised as follows:

• The predictive algorithm offers a good prediction of risk of hypoglycaemia in a 24-hour timeframe using sparse data;

• The classifiers consensus decision using data and information fusion showed potentialities for the predictions;

• The approaches improve the performance of predictions offering suitable solutions for 81.5% of the patients;

•The consensus decision achieved over 90% of hypoglycaemia predicted for 66% of patients.

These results leave open a set of opportunities for technological solutions centred on self-care patients. However, further research is necessary to reduce the number of false alarms and to understand the complexity of these predictions because of sparse data. Also, we noticed that a high percentage of best consensus is KNN-based. So, this finding will be one target for future work. The data was not very extensive, but this work shows that it is still possible to explore

more of the topic in the future. As future work it needs some improvements, such as the build of models considering data from the patient to offer a personalised solution considering patient similarities networks. Also, taking into account the specifies of data provided by the patient, using feature selection methods.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgment

This work is funded by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/EEA/50008/2020 (Este trabalho é financiado pela FCT/MCTES através de fundos nacionais e quando aplicável cofinanciado por fundos comunitários no âmbito do projeto UIDB/EEA/50008/2020).

This article/publication is based on work from COST Action IC1303 - AAPELE - Architectures. Algorithms and Protocols for Enhanced Living Environments and COST Action CA16226 - SHELD-ON - Indoor living space improvement: Smart Habitat for the Elderly. supported by COST (European Cooperation in Science and Technology). More information in www.cost.eu.

References

[1] I. D. Federation, “IDF Diabetes Atlas 10th edition 2021,” 2021.

[2] NICE Guideline 18, “Diabetes (type 1 and type 2) in children and young people: diagnosis and management,” *Diabetes UK*, 2015.

[3] NICE, “Type 1 diabetes in adults: diagnosis and management (NICE guideline).,” *NICE Guidel. [NG17]*, 2015.

[4] O. Mujahid, I. Contreras, and J. Vehi, “Machine learning techniques for hypoglycemia prediction: Trends and challenges,” *Sensors (Switzerland)*. 2021.

[5] P. Tkachenko, G. Kriukova, M. Aleksandrova, O. Chertov, E. Renard, and S. V. Pereverzyev, “Prediction of nocturnal hypoglycemia by an aggregation of previously known prediction approaches: proof of concept for clinical application,” *Comput. Methods Programs Biomed.*, vol. 134, pp. 179–186, 2016.

[6] D. Rodbard, “Continuous Glucose Monitoring: A Review of Recent Studies Demonstrating Improved Glycemic Outcomes,” *Diabetes Technol. Ther.*, 2017.

[7] Health Quality Ontario, “Continuous Monitoring of Glucose for Type 1 Diabetes: A Health Technology Assessment,” *Ont. Health Technol. Assess. Ser.*, 2018.

[8] A. Inc., “Glucose Buddy.” [Online]. Available: https://www.glucosebuddy.com/. [Accessed: 27-Jun-2019].

[9] D. Oy, “Dottli:A.” [Online]. Available: https://dottli.com/. [Accessed: 27-Jun-2019].

[10] Staywell, “Sugar Sense - Diabetes App.” [Online]. Available: https://play.google.com/store/apps/details?id=org.medhelp.sugarsense&hl=pt_PT. [Accessed: 14-May-2018].

[11] Social Diabetes, “Social Diabetes.” [Online]. Available: https://www.socialdiabetes.com/. [Accessed: 12-Feb-2018].

[12] Sirma Medical Systems, “Diabetes: M.” [Online]. Available: https://www.diabetes-m.com/. [Accessed: 12-Feb-2018].

[13] A. D. C. U. Inc., “CONTOUR®DIABETES app.” [Online]. Available: https://www.contournextone.com/. [Accessed: 27-Jun-2019].

[14] L. LifeScan IP Holdings, “OneTouch Reveal.” [Online]. Available: https://www.onetouchreveal.com/. [Accessed: 27-Jun-2019].

[15] K. Plis, R. Bunescu, C. Marling, J. Shubrook, and F. Schwartz, “A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management,” *Mod. Artifical Intell. Heal. Anal.*, pp. 35–39, 2014.

[16] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, “Online prediction of glucose concentration in type 1 diabetes using extreme learning machines,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 3262–3265.

[17] S. Kriventsov, A. Lindsey, and A. Hayeri, “The diabits app for smartphone-assisted predictive monitoring of glycemia in patients with diabetes: Retrospective observational study,” *JMIR Diabetes*, vol. 5, no. 3, pp. 1–14, 2020.

[18] E. I. Georga, V. C. Protopappas, D. Polyzos, and D. I. Fotiadis, “Evaluation of short-term predictors of glucose concentration in type 1 diabetes combining feature ranking with regression models,” *Med. Biol. Eng. Comput.*, vol. 53, no. 12, pp. 1305–1318, 2015.

[19] K. Zarkogianni *et al.*, “Comparative assessment of glucose prediction models for patients with type 1 diabetes mellitus applying sensors for glucose and physical activity monitoring,” *Med. Biol. Eng. Comput.*, vol. 53, no. 12, pp. 1333–1343, 2015.

[20] C. Zecchin, A. Facchinetti, G. Sparacino, and C. Cobelli, “How much is short-term glucose prediction in type 1 diabetes improved by adding insulin delivery and meal content information to CGM data? A proof-of-concept study,” *J. Diabetes Sci. Technol.*, vol. 10, no. 5, pp. 1149–1160, 2016.

[21] J. Vehí, I. Contreras, S. Oviedo, L. Biagi, and A. Bertachi, “Prediction and prevention of hypoglycaemic events in type-1 diabetic patients using machine learning,” *Health Informatics J.*, 2019.

[22] S. Mirshekarian, R. Bunescu, C. Marling, and F. Schwartz, “Using LSTMs to learn physiological models of blood glucose behavior,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 2887–2891.

[23] M. He, W. Gu, Y. Kong, L. Zhang, C. J. Spanos, and K. M. Mosalam, “CausalBG: Causal Recurrent Neural Network for the Blood Glucose Inference with IoT Platform,” *IEEE*

Internet Things J., p. 1, 2019.

[24] M. H. Jensen, C. Dethlefsen, P. Vestergaard, and O. Hejlesen, "Prediction of Nocturnal Hypoglycemia From Continuous Glucose Monitoring Data in People With Type 1 Diabetes: A Proof-of-Concept Study," *J. Diabetes Sci. Technol.*, vol. 14, no. 2, pp. 250–256, 2019.

[25] J. Pavan *et al.*, "Personalized machine learning algorithm based on shallow network and error imputation module for an improved blood glucose prediction," *CEUR Workshop Proc.*, vol. 2675, pp. 95–99, 2020.

[26] G. Cappon *et al.*, "A personalized and interpretable deep learning based approach to predict blood glucose concentration in type 1 diabetes," *CEUR Workshop Proc.*, vol. 2675, pp. 75–79, 2020.

[27] D. Dave *et al.*, "Feature-Based Machine Learning Model for Real-Time Hypoglycemia Prediction," *J. Diabetes Sci. Technol.*, 2020.

[28] B. Sudharsan, M. Peeples, and M. Shomali, "Hypoglycemia prediction using machine learning models for patients with type 2 diabetes," *J. Diabetes Sci. Technol.*, vol. 9, no. 1, pp. 86–90, 2015.

[29] J. Li and C. Fernando, "Smartphone-based personalized blood glucose prediction," *ICT Express*, vol. 2, no. 4, pp. 150–154, 2016.

[30] J. I. Hidalgo, J. M. Colmenar, G. Kronberger, S. M. Winkler, O. Garnica, and J. Lanchares, "Data Based Prediction of Blood Glucose Concentrations Using Evolutionary Methods," *J. Med. Syst.*, vol. 41, no. 9, 2017.

[31] C. Novara, N. M. Pour, T. Vincent, and G. Grassi, "A Nonlinear Blind Identification Approach to Modeling of Diabetic Patients," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 3, pp. 1092–1100, 2016.

[32] G. Alfian *et al.*, "Blood glucose prediction model for type 1 diabetes based on artificial neural network with time-domain features," *Biocybern. Biomed. Eng.*, vol. 40, no. 4, pp. 1586–1599, 2020.

[33] R. H. Botwey, E. Daskalaki, P. Diem, and S. G. Mougiakakou, "Multi-model data fusion to improve an early warning system for hypo-/hyperglycemic events," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 4843–4846.

[34] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting learning to accurately predict blood glucose trajectories," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 1387–1395, 2018.

[35] H. N. Mhaskar, S. V. Pereverzyev, and M. D. van der Walt, "A Deep Learning Approach to Diabetic Blood Glucose Prediction," *Front. Appl. Math. Stat.*, vol. 3, no. July, pp. 1–11, 2017.

[36] A. Bertachi, L. Biagi, N. Luo, and J. Veh, "Prediction of blood glucose levels and nocturnal hypoglycemia using physiological models and artificial neural networks," 2018.

[37] R. Reddy, N. Resalat, L. M. Wilson, J. R. Castle, J. El Youssef, and P. G. Jacobs,

"Prediction of Hypoglycemia During Aerobic Exercise in Adults With Type 1 Diabetes," *J. Diabetes Sci. Technol.*, vol. 13, no. 5, pp. 919–927, 2019.

[38] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," *Int. J. Med. Inform.*, vol. 126, no. March, pp. 1–8, 2019.

[39] S. Mirshekarian, H. Shen, R. Bunescu, and C. Marling, "LSTMs and Neural Attention Models for Blood Glucose Prediction: Comparative Experiments on Real and Synthetic Data," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 706–712.

[40] A. Aliberti *et al.*, "A Multi-Patient Data-Driven Approach to Blood Glucose Prediction," *IEEE Access*, vol. 7, pp. 69311–69325, 2019.

[41] A. Aliberti, A. Bagatin, A. Acquaviva, E. Macii, and E. Patti, "Data driven patient-specialized neural networks for blood glucose prediction," *2020 IEEE Int. Conf. Multimed. Expo Work. ICMEW 2020*, pp. 1–6, 2020.

[42] V. Felizardo, N. M. Garcia, N. Pombo, and I. Megdiche, "Data-based algorithms and models using diabetics real data for blood glucose and hypoglycaemia prediction – A systematic literature review," *Artif. Intell. Med.*, vol. 118, no. February 2020, p. 102120, 2021.

[43] E. M. Benjamin, "Self-Monitoring of Blood Glucose: The Basics," *Clin. Diabetes*, 2002.

[44] V. Felizardo, N. M. Garcia, I. Megdiche, N. Pombo, and F. Babič, "A novel context-aware framework for Hypoglycaemia Events Prediction," *Under Rev. Publ. Pat. being Submitt.*, 2022.

[45] V. Felizardo, D. Machado, N. Garcia, N. Pombo, and P. Brandão, "Hypoglycaemia prediction models with auto explanation," *IEEE Access*, 2021.

[46] S. Oviedo, J. Vehí, R. Calm, and J. Armengol, "A review of personalized blood glucose prediction strategies for T1DM patients," *Int. j. numer. method. biomed. eng.*, vol. 33, no. 6, pp. 1–21, 2017.

[47] A. Z. Woldaregay *et al.*, "Data-driven modeling and prediction of blood glucose dynamics: Machine learning applications in type 1 diabetes," *Artificial Intelligence in Medicine*. 2019.

[48] E. Zdravevski *et al.*, "Improving Activity Recognition Accuracy in Ambient-Assisted Living Systems by Automated Feature Engineering," vol. 5, 2017.

[49] C. Mosquera-Lopez, R. Dodier, N. Tyler, N. Resalat, and P. Jacobs, "Leveraging a Big Dataset to Develop a Recurrent Neural Network to Predict Adverse Glycemic Events in Type 1 Diabetes," *IEEE J. Biomed. Heal. Informatics*, pp. 1–1, 2019.

[50] S. Whalen and G. Pandey, "A comparative analysis of ensemble classifiers: Case studies in genomics," in *Proceedings - IEEE International Conference on Data Mining, ICDM*, 2013.

[51] P. Yang, Y. Hwa Yang, B. B. Zhou, and A. Y. Zomaya, "A Review of Ensemble Methods in Bioinformatics," *Curr. Bioinform.*, vol. 5, no. 4, pp. 296–308, 2010.

[52] A. Altmann *et al.*, "Comparison of classifier fusion methods for predicting response to anti HIV-1 therapy," *PLoS One*, vol. 3, no. 10, 2008.

[53] M. Liu, D. Zhang, and D. Shen, "Ensemble sparse classification of Alzheimer's disease," *Neuroimage*, vol. 60, no. 2, pp. 1106–1116, 2012.

[54] S. Bashir, U. Qamar, and F. H. Khan, "IntelliHealth: A medical decision support application using a novel weighted multi-layer classifier ensemble framework," *J. Biomed. Inform.*, vol. 59, pp. 185–200, 2016.

[55] S. Bashir, U. Qamar, F. H. Khan, and L. Naseem, "HMV: A medical decision support framework using multi-layer classifiers for disease prediction," *J. Comput. Sci.*, vol. 13, pp. 10–25, 2016.

[56] Q. He *et al.*, "Feasibility study of a multi-criteria decision-making based hierarchical model for multi-modality feature and multi-classifier fusion: Applications in medical prognosis prediction," *Inf. Fusion*, vol. 55, no. July 2019, pp. 207–219, 2020.

[57] S. Chen *et al.*, "Information fusion and multi-classifier system for miner fatigue recognition in plateau environments based on electrocardiography and electromyography signals," *Comput. Methods Programs Biomed.*, vol. 211, p. 106451, 2021.

[58] X. C. Yin, K. Huang, and H. W. Hao, "DE2: Dynamic ensemble of ensembles for learning nonstationary data," *Neurocomputing*, vol. 165, pp. 14–22, 2015.

[59] U. Machine Learning Repository, "Diabetes Data Set." [Online]. Available: https://archive.ics.uci.edu/ml/datasets/diabetes. [Accessed: 20-Jul-2005].

Table 2 – Consensus decision of two classifiers (C2).

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C2.1 (CART+SVM)	0.64651	0.345751	0.817245	0.183756
C2.2 (CART+KNN)*	0.658682	0.309507	0.856854	0.14438
C2.3 (CART+AdaBTree)	0.633003	0.323455	0.809519	0.191481
C2.4 (CART+RF)	0.654882	0.31246	0.849054	0.151947
C2.5 (CART+SKNN)*	0.656279	0.320456	0.851895	0.149339
C2.6 (SVM+KNN)	0.653577	0.477143	0.740847	0.260154
C2.7 (SVM+ AdaBTree)*	0.640084	0.340516	0.809491	0.191509
C2.8 (SVM+RF)	0.664469	0.417536	0.793057	0.207943
C2.9 (SVM+SKNN)*	0.659084	0.44042	0.777724	0.223276
C2.10 (KNN+AdaBTree)	0.653194	0.305198	0.849483	0.151751
C2.11 (KNN+RF)*	0.674764	0.389729	0.82614	0.175095
C2.12 (KNN+SKNN)*	0.668696	0.407374	0.813128	0.188107
C2.13 (AdaBTree+RF)	0.662445	0.320742	0.851357	0.149877
C2.14 (AdaBTree+SKNN)*	0.654606	0.328569	0.844988	0.156246
C2.15 (RF+SKNN)	0.684779	0.385665	0.842061	0.159174

*achieved 100% of hypoglycaemia predicted: C2.2 for P10 and P40; C2.5 for P4; C2.7 for P3; C2.9 for P33; C2.11 for P32; C2.12 for P36; and C2.14 for P18.

Table 3– Consensus decision of three classifiers (C3).

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C3.1 (CART+SVM+KNN)	0.640219	0.561961	0.667401	0.333599
C3.2 (AdaBTree+RF+SKNN)*	0.668405	0.508653	0.746254	0.254746
C3.3 (CART+SVM+AdaBTree)*	0.631186	0.480275	0.709207	0.291793
C3.4 (CART+KNN+ AdaBTree)	0.639595	0.467309	0.729143	0.271857
C3.5 (SVM+KNN+ AdaBTree))	0.643486	0.556533	0.674455	0.326545
C3.6 (CART+SVM+RF)	0.655237	0.520432	0.710844	0.290156
C3.7 (CART+KNN+RF)*	0.66904	0.5078	0.739078	0.261922
C3.8 (CART+AdaBTree+RF)*	0.644615	0.44558	0.746093	0.254907
C3.9 (CART+AdaBTree +SKNN)	0.638379	0.464717	0.730118	0.270883
C3.10 (CART+RF+SKNN)*	0.671543	0.518042	0.745992	0.255008
C3.11 (CART+SVM+SKNN)	0.643403	0.574561	0.668711	0.33229
C3.12 (CART+KNN+SKNN)*	0.655091	0.549152	0.701376	0.299625
C3.13 (SVM+KNN+RF)	0.647559	0.566185	0.674975	0.326025
C3.14 (SVM+KNN+SKNN)*	0.6367	0.597689	0.644434	0.356567

Table 4 – Consensus decision of three classifiers (C3) (continuation).

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C3.15 (KNN+AdaBTree+RF)*	0.665807	0.499118	0.741678	0.259322
C3.16 (KNN+ AdaBTree+SKNN)*	0.655724	0.540373	0.708859	0.292142
C3.17 (KNN+RF+SKNN)*	0.661605	0.543407	0.711805	0.289195
C3.18 (SVM+ AdaBTree+RF)	0.650847	0.508639	0.715086	0.285914
C3.19 (SVM+ AdaBTree+SKNN)	0.642718	0.55706	0.677519	0.323481
C3.20 (SVM+RF+SKNN)	0.65124	0.571434	0.68019	0.32081

*achieved 100% of hypoglycaemia predicted: C3.2 for P12; C3.3 for P67; C3.7 for P19; C3.8 for P31; C3.10, C3.12, C3.14, C3.16 for P19; and 3.17 for P5 and P19.

Table 5– Consensus decision of four classifiers (C4).

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C4.1 (CART+SVM+KNN+ AdaBTree)*	0.651505	0.399012	0.789709	0.211292
C4.2 (CART+SVM+KNN+RF)	0.667937	0.447521	0.776849	0.224151
C4.3 (CART+SVM+KNN+SKNN)	0.661634	0.470555	0.758062	0.242938
C4.4 (CART+SVM+ AdaBTree+RF)	0.652864	0.387046	0.797317	0.203683

Table 5– Consensus decision of four classifiers (C4) (continuation).

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C4.5 (CART+SVM+ AdaBTree+SKNN)*	0.649603	0.402479	0.787399	0.213601
C4.6 (CART+SVM+RF+SKNN)*	0.667196	0.442769	0.783808	0.217192
C4.7 (CART+KNN+ AdaBTree+RF)*	0.663654	0.376693	0.817656	0.183344
C4.8 (CART+KNN+ AdaBTree+SKNN)	0.657466	0.383077	0.81116	0.18984
C4.9 (CART+KNN+RF+SKNN)*	0.67676	0.427634	0.805424	0.195576
C4.10 (CART+ AdaBTree+RF+SKNN)	0.660369	0.38089	0.813373	0.187628
C4.11 (SVM+KNN+ AdaBTree+RF)	0.66316	0.438523	0.777689	0.223311
C4.12 (SVM+KNN+ AdaBTree+SKNN)*	0.660537	0.455002	0.762677	0.238323
C4.13 (SVM+KNN+RF+SKNN)	0.661422	0.475331	0.753407	0.247593
C4.14 (KNN+ AdaBTree+RF+SKNN)*	0.679032	0.429622	0.804963	0.196037
C4.15 (SVM+ AdaBTree+RF+SKNN)*	0.667593	0.441939	0.783013	0.217987

*achieved 100% of hypoglycaemia predicted: C4.1, C4.12 for P40; C4.5 for P38; C4.6 and C4.9 for P11; C4.7 for P9; C4.14 for P11, P44 and P45; and C4.15 for P12.

Table 6*– Consensus decision of five and six classifiers (C5 and C6)

Classifier Models	Performance			
	Acc	Sen	Spe	FA
C5.1 (CART+SVM+KNN+AdaBTree+RF)	0,656401	0,505518	0,724276	0,276724
C5.2 (CART+SVM+KNN+AdaBTree+SKNN)	0,653925	0,551251	0,695939	0,305062
C5.3 (CART+SVM+KNN+RF+SKNN)	0,658227	0,554769	0,695697	0,305303
C5.4 (CART+SVM+AdaBTree+RF+SKNN)	0,663255	0,517665	0,728942	0,272058
C5.5 (CART+KNN+AdaBTree+RF+SKNN)	0,669123	0,512296	0,740557	0,260443
C5.6 (SVM+KNN+AdaBTree+RF+SKNN)	0,656447	0,535335	0,70614	0,29486
C6 (all classifiers)	0,666147	0,454589	0,771653	0,229347

Chapter 5

Conclusion and future work

This chapter presents the main conclusions drawn from the research work described in this thesis and also discusses some research topics that may be addressed as a continuation or a complement of the work developed in this doctoral program.

5.1 Conclusion

This thesis proposes a new context-aware framework for hypoglycemia prediction, that comprises data fusion and classifiers consensus decision in a 24h hour time frame. The literature shows an extensively studied approach for glycaemia prediction using data from continuous glucose monitoring systems, but these approaches are unapproachable to patients performing their self-management of diabetes. The development of self-management solutions with prediction purposes based on discrete information is a challenge. However, integrating additional context information with the conventional features can contribute in a way to complement the information, reducing inter- and intra-patient variability. The definition of a prediction framework that comprises data fusion and multi-classifiers decision significantly contributes to the creation of suitable and generalised solutions for hypoglycaemia prevention to assist patients and caregivers in managing diabetes. To support this thesis, four main objectives were achieved: **1)** the identification of the existing predictive algorithms for hypoglycaemia and identification of possible open challenges; **2)** the design of a novel context-aware framework for hypoglycaemia prediction; **3)** the development of a context-aware predictive algorithm using discrete data,and, **4)** the application of the framework using real patient data to evaluate the data fusion and the multi-classifiers decision. Following, the research contributions regarding these objectives are summarized.

The first objective was presented in chapter 2, consists in a systematic literature review addressing the existing data-based models using real data for hypoglycemia prediction. This review analyzes sixty-three studies in which the models have been addressed massively, as short-term blood glucose prediction presented on a prediction horizon between 5 min and 180 min according to the time of day, *e.g.* nocturnal or postprandial predictions. The predictive models resort to different inputs as meal, insulin and physical activity data combined with glucose data, but some studies use only glucose data due to inter-variability in day-time meals, medication, and exercise. The studies used different methods and techniques showing good predictive results. There are some trends related to the datasets used. There are some datasets available, nevertheless only OhioT1DM Dataset allows a perfect replicability of the methods.

The second and third objectives were presented in chapter 3, consist in the design and development of the context-aware framework for hypoglycaemia prediction. The framework comprises data fusion and multi-classifiers decision providing a good prediction of risk of hypoglycaemia in a 24-hour window using discrete data. With the achievement of these objectives, we show that integrating contextual and time-based features with the conventional features improves the accuracy and the number of solutions valuable (accuracy higher than 60% and false alarms lower than 30%). In addition, the consensus decision improves the performance of predictions offering good solutions for 57.4% of the patients.

The fourth objective was presented in chapter 4, present the framework evaluation using real patient data. From the framework application, the predictive algorithm offers a good prediction of risk of hypoglycaemia in a 24-hour time frame using discrete data. The classifiers consensus decision using data and information fusion showed potentialities for the predictions offering suitable solutions for 81.5% of the patients achieving over 90% of hypoglycaemia predicted for 66% of patients.

As presented in the Introduction there are several open research questions involved in the development of the proposed framework. Following, we discuss how we have addressed and answered the questions stated in the Introduction: **1)** Question 1 - How to convert collected data into useful context-knowledge?; **2)** Question 2 - Which pre-processing techniques are most suitable for our scenario?; **3)** Question 3 - How to prepare the target (output) for the learning process?, and, **4)** Question 4 - Which classifiers are most suitable for the proposed scenario?.

We answer Question 1 in Chapter 2, where we analyse the input attributes of each study identified in the systematic literature review. The predictive algorithms are mostly based on data from continuous glucose monitoring systems used alone or others used these data in combination with information about meals, therapy, and exercise. For our approach, focusing on the solution for diabetes self-management, we opted for the collection of discrete data through a mobile application. Therefore, the data source is similar to that found in the literature. In Chapter 3, the framework design is presented in detail, the main contribution being the proposal of input features as contextual features. As hypoglycemic events can have different triggers, the contextual information can be useful in defining and support the decision process and it is given by three context dimensions, such as 1) time, 2) behaviour, and 3) time-based features. The dimension of time is given by the feature hour of the day or time of register, allowing deliver information about when some actions were performed. The behaviour dimension is given by information about meal or physical activity and allow answer to how these actions were performed. The time-based features are the link between the time and glucose level feature, giving a notion of how the glucose level varies over time. So, the context works like as supplementary information to bridge the discrete data. The results presented in chapter 3 show the contribution of fusing contextual information with conventional features like glucose level and insulin therapy for the hypoglycaemia predictions.

To answer question 2, we present in chapter 3 the framework for hypoglycaemia prediction that comprises two steps of pre-processing: 1) Imputation and 2) Re-sampling. Data imputation is necessary due to the impossibility of replacing the blood glucose missing data with zero. The other option could be to delete the instance with the missing value, but this could result in losing relevant information. Because hypoglycaemia events are always a minority class in comparison with the totality of the data, re-sampling methods are crucial for unbalanced data. So, for our approach, we opted for the Synthetic Minority Over-sampling Technique (SMOTE). These options are not new, but they show that when we propose a new framework, it is crucial focuses on the problem as stated and try to synthesize information and knowledge to achieve a solution.

In order to answer the question 3, in chapter 3, the target (output) for the learning process has three classes: no risk (class 0), hypoglycaemia risk (class 1) and hypoglycaemia event (class 2). The risk class is the risk of having a hypoglycaemia event in the next 24 hours or a BG value less than 75 mg/dl. So, for the instances with a value lower than 70 mg/dl we considered all instances of the last 24h like class 1. Also, for instances with a value less or equal than 75 mg/dl we considered all instances of the last 24h like class 1 in order to mitigate the risk of a possible hypoglycaemia event.

In order to answer Question 4 in Chapter 3, we proposed a consensus decision from different individual models to mitigate the fact that no individual models allow a good solution for all patients. The results presented in chapter 3 show that this approach achieved better accuracy, false positives and hypoglycaemia predicted. In chapter 4, we present the framework evaluation using real patient data wherein the consensus decision offers a good prediction of risk of hypoglycaemia in a 24-hour time frame using discrete data.

5.2 Future work

On top of the general ideas for further work on context-aware hypoglycaemia prediction, we describe several open issues that raise from the research presented in this thesis, and the corresponding lines of future work that could address them. In particular, the research introduced in this thesis opens several paths to future developments. We organise these ideas around the data fusion, model decision fusion and model interpretability.

In this thesis, when we identified several existing predictive algorithms for hypoglycaemia and we highlight the lack of available and public access datasets. First, algorithms' development and validation would benefit from public datasets. Second, the contextual information recordings are of utmost interest, as they provide a better perspective on the patient's profile and behaviour. Third, the Diabetes UCI contained a population with frequent hypoglycemic events. It would be interesting to test the algorithm on a population with fewer events. Fourth, it would be interesting to explore different data sources in order to define different context dimensions and time-based features.

In chapter 4, the framework was applied providing personalised predictions for each patient. Although promising results have been obtained, we must reduce false alarms and thus improve the accuracy. First, there are different classifiers that have not been tried. Second, KNN-based ensembles presented some potentialities for some clusters of patients. Third, integrating patient past data in a population train dataset.

The purpose of this thesis was not to give interpretability to the developed algorithm, however, it is important to emphasize the importance of this feature for this approach. First, it would be interesting to explore personalized models considering patient similarities networks. Second, to explore and implement a recommender system to provide proper information to patient and caregivers.

Appendix A

Appendix

A.1 Patent

In this Appendix we present the patent application titled "*Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos*" (in English, "Method and machine for hypoglycaemia events prediction using data information and a consensus of predictive models"). In short, the present invention relates to the hypoglycaemia events prediction using discrete information fusion and a consensus of predictive models. The information used by the predictive models are glucose levels, therapy used, meal, physical exercise and context information, related to time-based information, for example, time of data collection, type of meal and variability of glucose levels in the time. Predictive models are trained and tested using machine learning techniques. The consensus decision of the predictive models is given in a personalised way to the patient, indicating the risk of a hypoglycemia event in the next 24 hours.

RESUMO

"Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos"

A presente invenção está relacionada com a predição de eventos de hipoglicemia usando a fusão de informação discreta e recorrendo ao consenso de modelos preditivos.

Desta forma a presente invenção é útil para prever eventos de hipoglicemia em pacientes com diabetes, que não tendo acesso a sistemas contínuos de monitorização de glicose, realizem a sua auto monitorização recorrendo ao seu glicosímetro, podendo ser combinado com tecnologia móvel adicional, nomeadamente sistemas de apoio à decisão, de forma a garantir a transferência de dados médicos para o centro de saúde remoto e/ou cuidadores.

DESCRIÇÃO

"Método e máquina para predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos"

Domínio técnico da invenção

A presente invenção está relacionada com a predição de eventos de hipoglicemia usando a fusão de informação discreta e recorrendo ao consenso de modelos preditivos. A informação utilizada pelos modelos preditivos são os níveis de glicose, terapia utilizada, refeição, exercício físico e informação de contexto, relacionada com informação dependente do tempo, por exemplo, hora de recolha de dados, tipo de refeição e variabilidade dos níveis de glucose no tempo. Os modelos preditivos são treinados e testados recorrendo a técnicas de aprendizagem por máquina. A decisão de consenso dos modelos preditivos é dada de forma personalizada ao paciente indicando o risco de ocorrer um evento de hipoglicemia nas próximas 24 horas.

Sumário da invenção

É objetivo da presente invenção a predição de risco de hipoglicemia nas próximas 24 horas usando uma fusão de dados discretos e uma decisão de consenso de vários modelos preditivos.

A presente invenção é útil para prever eventos de hipoglicemia em pacientes com diabetes, que não tendo acesso a sistemas contínuos de monitorização de glicose, realizem a sua auto monitorização recorrendo ao seu glicosímetro, podendo ser combinado com tecnologia móvel adicional, nomeadamente

sistemas de apoio à decisão, de forma a garantir a transferência de dados médicos para o centro de saúde remoto e/ou cuidadores.

Estado da técnica

Nos últimos anos, *telemóveis* e dispositivos inteligentes têm mostrado um papel importante no que diz respeito à gestão da saúde e comunicação entre profissionais de saúde, cuidadores e pacientes. Além disso, as aplicações móveis médicas têm permitido a autogestão de doenças crónicas, dando autonomia aos pacientes para gerir as condições associadas à sua doença. O uso de aplicações de autogestão de doenças crónicas está em crescimento e, em áreas como diabetes, algumas soluções já estão disponíveis em lojas da *internet* [1], [2]. Em relação aos pacientes com diabetes, dependentes de insulina, doses excessivas ou insuficientes de insulina causam um mau controlo glicémico, resultando em algumas complicações de longo e curto prazo, como hipoglicemia (baixo nível de glicose no sangue). Geralmente, os pacientes temem estes eventos porque, em alguns casos graves, é necessária assistência médica imediata. A literatura mostra um extenso trabalho na predição dos níveis de glicose no sangue e na predição de eventos futuros para apoio a pacientes com diabetes, sendo que a maioria desses trabalhos se concentrou em predições usando dados provenientes deuma monitorização contínua da glicemia [3] - [8]. Os algoritmos de predição usando informações discretas são apresentados com uma baixa expressão [9], [10]. A presente invenção permite a predição de risco de hipoglicemia nas próximas 24 horas usando uma fusão de dados discretos e uma decisão de consenso de vários modelos preditivos, o que permite estar acessível a pacientes com diabetes que não têm acesso a sistemas de monitorização contínua de glicemia. Alguns modelos utilizam mais informações que doo nível de

glicose no sangue, como informações sobre o nível insulina, refeições e exercício físico. A presente invenção propõe a integração de informação contextual e informação dependente do tempo com a informação convencional, o que permite melhorar a predição.
A literatura mostra que não existe um método que apresente 100% de acerto das predições e, devido à inter- e intra-variabilidade individual, nenhum classificador apresenta soluções viáveis para todos os pacientes. A presente invenção propõe uma decisão baseada no consenso de vários classificadores, o que permite determinar, a médio prazo, a melhor solução personalizada para um determinado paciente.

A maioria dos documentos de patente [11]-[13] apresenta sistemas e métodos que recorrem a dados contínuos de glicemia por via de sistemas de monitorização contínua (SMC). A presente invenção usa apenas dados discretos, o que facilita o registo de dados pelo paciente no seu quotidiano, através de um telemóvel.
O documento CN111755122A [14] propõe um método que pode usar dados discretos, mas apenas os dados dos níveis de glicose são usados tal como no método reivindicado no presente documento. A presente invenção utiliza dados contextuais e variáveis dependentes do tempo, de forma a complementar a informação das restantes variáveis (níveis de glucose, unidades de insulina, informação acerca da dieta e atividade física).
As patentes já referidas acima [11]-[14] usam técnicas de aprendizagem para a criação dos modelos preditivos. No entanto, propõem apenas um modelo preditivo baseado numa técnica específica.
O documento CN111631730A [15] refere fusão de modelos, embora cada modelo seja treinado com variáveis diferentes (por exemplo um modelo relacionado com a dieta, um modelo relacionado com atividade física, etc…). O mesmo se verifica

do documento US2010174228A1 [16], que usa a fusão de um modelo treinado com dados obtidos por via de SMC e outro com dados dos níveis de insulina. A presente invenção propõe a construção de diferentes modelos preditivos utilizando as mesmas variáveis para todos modelos, alterando o tipo de classificador. Esta solução tem o intuito de captar diferentes padrões de acordo com o tipo de classificador e diminuir a influência da inter- e intra- variabilidade existente entre pacientes. A presente invenção propõe uma decisão baseada no consenso de vários classificadores, o que permite determinar, a médio prazo, a melhor solução personalizada para um determinado paciente.

Descrição geral da invenção

A recomendação de que deve haver um equilíbrio entre insulina, alimentação e exercício físico para manter a diabetes controlada oferece uma oportunidade para o desenvolvimento de aplicações móveis para previsões reais. Essas previsões podem ser usadas para melhorar a qualidade de vida de pacientes com eventos indesejados, como a hipoglicemia.

A presente invenção diz respeito a uma abordagem de predição de hipoglicemia usando fusão de informações e consenso de modelos preditivos. A abordagem combina informações em dois processos, Processo de Aprendizagem e Processo de Aplicação. O primeiro é composto por 1) coleta de dados, onde são coletados dados de diferentes fontes e diferentes pacientes; 2) extração de variáveis, em que as variáveis baseadas no tempo são extraídas dos dados coletados; 3) imputação de glicose no sangue, se um dado ausente for verificado; 4) reamostragem, para obter um conjunto de dados balanceado e 5) aprendizagem de classificadores, onde diferentes classificadores aprendem a partir dos dados. No segundo, os dados são coletados de um único paciente, as características são extraídas e imputadas e, usando os modelos preditivos obtidos no processo de

aprendizagem, obtém-se a decisão de consenso dos modelos preditivos, fundindo as informações de avaliação de cada modelo preditivo. A presente invenção é usada para prever o risco de hipoglicemia numa janela de 24 horas.

Descrição detalhada da invenção

A presente invenção envolve a aquisição de dados, pré-processamento, treino/teste de múltiplos classificadores e a decisão do consenso dos modelos preditivos do risco de ter um evento de hipoglicemia nas próximas 24 horas. A abordagem combina informações em dois processos, Processo de Aprendizagem e Processo de Aplicação.

Processo de Aprendizagem

O Processo de Aprendizagem é composto por 1) coleta de dados, onde são coletados dados de diferentes fontes e diferentes pacientes; 2) extração de variáveis, em que variáveis baseados no tempo são extraídos dos dados coletados; 3) imputação de glicose no sangue, se um dado ausente for verificado; 4) reamostragem, para obter um conjunto de dados balanceado e 5) aprendizagem de classificadores, onde diferentes classificadores aprendem a partir dos dados.

1) Coleta de dados e extração de variáveis

A aquisição de dados pode ser realizada por meio de uma aplicação móvel para autogestão da diabetes e organizada para permitir que o utilizador possa registar o seu nível de glicose, dando um contexto para esta medida, por exemplo, tipo de refeição (pré ou pós pequeno-almoço, pré ou pós almoço, pré ou pós jantar), informações sobre a refeição (mais, menos ou típica), informações sobre exercício físico (mais, menos ou típico) e tipo e quantidade de insulina.

As informações de contexto e a hora de registo dão a esta abordagem uma nova contribuição usando essas informações para o cálculo de três variáveis baseadas no tempo. Essas três

variáveis são: a diferença entre o nível de glicose atual e o registo anterior, a diferença entre o nível de glicose atual e o registo das 24 horas anteriores que fornecem algumas informações sobre a variabilidade dos níveis de glicose no tempo.

2) Imputação de dados e reamostragem

Com a impossibilidade de substituir os dados ausentes do nível de glicose por zero, podemos substituir os dados ausentes usando técnicas de imputação. A outra opção pode ser excluir a observação com o valor ausente, mas isso pode resultar na perda de informações relevantes. A imputação de dados foi realizada para os dados de cada paciente, de forma a realizar imputações mais realistas utilizando a técnica regressão de Ridge proposta em [17].

Após a imputação de dados, a proposta de método apresenta no presente documento usa a técnica de sobreamostragem de minoria sintética [18] para reamostrar o conjunto de dados. Como os eventos de hipoglicemia são sempre uma classe minoritária em comparação com a totalidade dos dados, os métodos de reamostragem são cruciais para dados não balanceados.

3) Modelos preditivos

O treino dos classificadores é executado com dados pré-processados (imputação de dados e reamostragem). As observações do conjunto de treino/teste são organizadas como vetores de atributos e, para cada observação contendo n atributos de entrada, um y é definido como o atributo alvo ou classe. Para a previsão de hipoglicemia, o nosso vetor-alvo tem três classes: 0 - sem risco, 1 - risco e 2 - evento de hipoglicemia (definido quando o valor do nível de glicose é inferior a 70 mg/dl) [19]. A classe de risco (valor 1) é o risco de ocorrer um evento de hipoglicemia nas próximas 24 horas ou um valor de glicemia inferior a 75 mg/dl [20]. Para

as observações com um valor inferior a 70 mg/dl, consideramos todas as observações das últimas 24h como classe 1. Além disso, para as observações com um valor menor ou igual a 75 mg/dl, consideramos todas as observações das últimas 24 horas como classe 1, a fim de reduzir o risco de um possível evento de hipoglicemia. Aplicamos uma validação cruzada (repetida 10 vezes) para testar os modelos.
Os critérios de seleção do modelo são a precisão do modelo e um bom equilíbrio entre sensibilidade e especificidade. Tal é importante porque a sensibilidade mede, nesta abordagem, a capacidade de identificar corretamente o risco de ocorrer um evento de hipoglicemia nas próximas 24 horas. Da mesma forma, a especificidade mede a capacidade de identificar corretamente as outras classes. Portanto, o bom equilíbrio entre essas duas medidas é importante para uma boa predição de eventos de hipoglicemia sem número elevado de falsos positivos (falsos alarmes).

Processo de Aplicação

No Processo de Aplicação, os dados são coletados de um único paciente, as características são extraídas e imputadas da mesma forma como descrito anteriormente e, recorrendo aos modelos obtidos no processo de aprendizagem, obtemos a decisão de consenso dos modelos, fundindo as informações de avaliação de diferentes modelos. O consenso de modelos, podendo ser consenso de dois ou mais modelos, é usado para prever o risco de hipoglicemia numa janela de 24 horas. A técnica utilizada para obter a decisão é denominada de voto maioritário. A médio prazo é possível obter a melhor combinação de modelos para um determinado paciente.

Exemplos de aplicação

O método proposto poderá ser embutido num equipamento de monitorização que, combinado com tecnologia móvel (por

exemplo, uma aplicação de autogestão) poderá auxiliar o paciente e cuidadores na gestão da diabetes, podendo evitar eventos de hipoglicemia. Ou ainda, embutido num sistema de apoio à decisão, de forma a garantir a transferência de dados médicos para o centro de saúde remoto e/ou cuidadores.

REFERÊNCIAS

[1] Google, "Diabetes Android Applications." [Online]. Available: https://play.google.com/store/search?q=diabetes&c=apps&hl=pt_PT. [Accessed: 25-Sep-2020].

[2] S. Veazie et al., "Rapid Evidence Review of Mobile Applications for Self-management of Diabetes," Journal of General Internal Medicine. 2018.

[3] C. Tronstad et al., "Non-invasive prediction of blood glucose trends during hypoglycemia," Anal. Chim. Acta, vol. 1052, pp. 37–48, 2019.

[4] S. Oviedo, I. Contreras, C. Quirós, M. Giménez, I. Conget, and J. Vehi, "Risk-based postprandial hypoglycemia forecasting using supervised learning," Int. J. Med. Inform., vol. 126, no. March, pp. 1–8, 2019.

[5] I. De Falco et al., "A genetic programming-based regression for extrapolating a blood glucosedynamics model from interstitial glucose measurements and their first derivatives," Appl. Soft Comput. J., vol. 77, pp. 316–328, 2019.

[6] H. M. Romero-Ugalde et al., "ARX model for interstitial glucose prediction during and after physical activities," Control Eng. Pract., vol. 90, no. July, pp. 321–330, 2019.

[7] K. Li, C. Liu, T. Zhu, P. Herrero, and P. Georgiou, "GluNet: A Deep Learning Framework For Accurate Glucose

Forecasting," IEEE J. Biomed. Heal. Informatics, vol. PP, no. c, pp. 1–1, 2019.
[8] M. Gadaleta, A. Facchinetti, E. Grisan, and M. Rossi, "Prediction of Adverse Glycemic Events From Continuous Glucose Monitoring Signal," IEEE J. Biomed. Heal. Informatics, vol. 23, no. 2, pp. 650–659, 2019.
[9] K. Sakurai, Y. Kawai, M. Yamazaki, and M. Komatsu, "Prediction of lowest nocturnal blood glucose level based on self-monitoring of blood glucose in Japanese patients with type 2 diabetes," J. Diabetes Complications, vol. 32, no. 12, pp. 1118–1123, 2018.
[10] B. R. Shah, S. Walji, A. Kiss, J. E. James, and J. M. Lowe, "Derivation and Validation of a RiskPrediction Tool for Hypoglycemia in Hospitalized Adults With Diabetes: The Hypoglycemia During Hospitalization (HyDHo) Score," Can. J. Diabetes, pp. 1–5, 2018.
[11] Breton, Marc D,; Hughes, Jonathan,; Andreson, Stacey; " System and method for online domain adaptation of models for hypoglycemia prediction in type 1 diabetes."; WO2021007485 (A1), issued January 14, 2021.
[12] Zedong, Nie; Jingzhen, Li; Yuhang, Liu; " Hypoglycemia early warning method and hypoglycemia early warning device."; CN111128327 (A), issued May 8, 2020.
[13] Imanbayev, Anuar; "Blod glucose data set optimization for improved hypoglycemia prediction based on machine learning implementation ingestion."; WO2020120571 (A1), issued June 18, 2020.
[14] Junyan, Du; Guiyong, Li; Yaping, Fu; Baohua, Liang; Xiaoqi, Miao; Meng, Wen; Naidan, Yan; Yaya, Wang; "Diabetes blood glucose prediction system and method based on CNN and model fusion."; CN111755122 (A), issued October 9, 2020.
[15] Zedong, Nie; Jingzhen, Li; Yuhang, Liu; " Hypoglycemia early warning method and system based on sensing data and

physiological information fusion."; CN111631730 (A), issued September 8, 2020.
[16] Buckingham, Bruce; Guenter, Niemeyer; Darrell M., Wilson; Fraser, Cameron; Peter, Chase H.; Eyal, Dassau; Hyunjin, Lee; Wayne, Bequette B.; Francis J., Doyle III; " Hypoglycemia prediction and control."; US2010174228 (A1), issued july 8, 2010.
[17] A. Folch-Fortuny, F. Arteaga, and A. Ferrer, "Missing Data Imputation Toolbox for MATLAB," Chemom. Intell. Lab. Syst., vol. 154, pp. 93–100, 2016.
[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., 2002.
[19] M. B. Abraham et al., "ISPAD CLINICAL PRACTICE CONSENSUS GUIDELINES ISPAD Clinical Practice Consensus Guidelines 2018: Assessment and management of hypoglycemia in children and adolescents with diabetes," vol. 19, no. March, pp. 178–192, 2018.
[20] F. Cameron, G. Niemeyer, D. Ph, K. Gundy-burlet, D. Ph, and B. Buckingham, "Statistical Hypoglycemia Prediction," vol. 2, no. 4, pp. 612–621, 2008.

Covilhã, 22 de junho de 2021

REIVINDICAÇÕES

1. Método de predição de eventos de hipoglicemia usando fusão de informação e consenso de modelos preditivos caracterizado por utilizar uma fusão de dados discretos e uma decisão de consenso de vários modelos preditivos.

2. Método de predição de eventos de hipoglicemia, de acordo com a reivindicação 1, caracterizado por utilizar múltiplos classificadores.

3. Método de predição de eventos de hipoglicemia, de acordo com a reivindicação 1, caracterizado pelos modelos preditivos serem testados recorrendo a técnicas de aprendizagem.

4. Método de predição de eventos de hipoglicemia, de acordo com a reivindicação 1, caracterizado pela integração de dados contextuais e variáveis dependentes do tempo em combinação com dados discretos de níveis de glicose no sangue, unidades de insulina, informação acerca da dieta do paciente e sua atividade física.

5. Método de predição de eventos de hipoglicemia, de acordo com a reivindicação 1, caracterizado pela predição do risco de ocorrer um evento de hipoglicemia nas próximas 24 horas.

Covilhã, 22 de junho de 2021