

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

**António José Esteves Caldeira**

Relatório do Projeto de Dissertação  
**Engenharia Informática**  
(2<sup>o</sup> ciclo de estudos)

Orientador: Prof. Doutor Sebastião Pais

**outubro de 2025**

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Declaração de Integridade**

Eu, António José Esteves Caldeira, que abaixo assino, estudante com o número de inscrição M13748 de 2º. Ciclo 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 Integridades da Universidade da Beira Interior.

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

Universidade da Beira Interior, Covilhã 08/10/2025

*António José Esteves Caldeira*

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Agradecimentos**

Desde o meu primeiro dia no ensino superior, deparei-me com vários obstáculos, obstáculos esses que me faziam pensar em desistir da vida académica. No entanto, houve um grupo de pessoas que esteve sempre ao meu lado para conseguir ultrapassar esses desafios e de certa forma concluir uma das etapas fundamentais da minha vida académica, a dissertação de mestrado aqui redigida.

Desse grupo, agradeço em primeiro lugar aos meus pais que se sacrificaram e me ajudaram muito para cumprir com os meus objetivos académicos e no geral. Também a minha irmã que esteve sempre disponível para me ajudar no que precisasse.

Agradecer a todos os professores que, de alguma maneira, ajudaram na minha evolução como estudante e pessoa.

Em particular ao Professor Doutor Sebastião Pais, por ter aceite estar comigo como orientador nestes últimos anos e por confiar em mim não só para a realização desta dissertação mas também de trabalhar na HULTIG.

Por fim, agradeço ao Luís e ao Gonçalo que estiveram sempre ao meu lado, neste projeto e em todo o meu percurso académico.

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Resumo**

A saúde mental é um dos maiores desafios globais da atualidade, sendo a depressão e a ansiedade dois dos transtornos mais frequentes e incapacitantes. As redes sociais, pela sua natureza dinâmica e pela forma como captam estados emocionais, tornaram-se uma fonte de dados importante para a investigação em saúde mental. Neste contexto, a presente dissertação explora o uso de dados textuais provenientes de redes sociais para identificar sinais associados a transtornos mentais e a comportamentos frequentemente descritos na literatura clínica.

A metodologia seguida iniciou-se com a recolha de diferentes conjuntos de dados textuais públicos e em inglês, complementada pelo recurso ao conjunto de dados SNCrawler (Apenas os dados em inglês). Posteriormente, foram desenvolvidos modelos de classificação específicos para ansiedade e depressão, aplicados de forma cruzada para garantir maior consistência nas etiquetas. Numa fase seguinte, procedeu-se à associação de comportamentos, com base em vocabulário expandido e técnicas de *embeddings* semânticos, permitindo ter um conjunto de dados com uma nova camada de informação. A etapa final consistiu no treino de um modelo *multitask*, capaz de classificar simultaneamente transtornos e comportamentos, o qual foi aplicado ao conjunto de dados SNCrawler, resultando na criação de um recurso anotado que liga os transtornos, ansiedade e depressão, a comportamentos ligados aos mesmos descritos na literatura clínica.

Os resultados obtidos confirmam a viabilidade de recorrer a dados textuais e técnicas de *Natural Language Processing* (NLP) para apoiar a fenotipagem digital em saúde mental, mesmo perante limitações como a escassez de dados anotados e a dificuldade de explorar modalidades além do texto. Este trabalho contribui assim para a criação de novos recursos e metodologias, que poderão ser aprofundados em investigações futuras, nomeadamente através da integração multimodal e da validação clínica.

## **Palavras-chave**

Saúde Mental; Ansiedade; Depressão; *Natural Language Processing*; *Embeddings*; Fenotipagem Digital; HULTIG; Multi-task model;

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Abstract**

Mental health is one of the greatest global challenges today, with depression and anxiety being two of the most common and debilitating disorders. Social media, due to its dynamic nature and the way it captures emotional states, has become an important source of data for mental health research. In this context, this dissertation explores the use of textual data from social media to identify signs associated with mental disorders and behaviors frequently described in clinical literature.

The methodology followed began with the collection of different sets of public textual data in English, supplemented by the use of the SNCrawler dataset (English data only). Subsequently, specific classification models for anxiety and depression were developed and applied cross-cutting to ensure greater consistency in the labels. In a next phase, behaviors were associated based on expanded vocabulary and semantic embedding techniques, resulting in a dataset with a new layer of information. The final step consisted of training a multitask model capable of simultaneously classifying disorders and behaviors, which was applied to the SNCrawler dataset, resulting in the creation of an annotated resource linking disorders, anxiety, and depression to behaviors related to them described in the clinical literature.

The results obtained confirm the feasibility of using textual data and Natural Language Processing (NLP) techniques to support digital phenotyping in mental health, even in the face of limitations such as the scarcity of annotated data and the difficulty of exploring modalities beyond text. This work thus contributes to the creation of new resources and methodologies, which could be further explored in future research, particularly through multimodal integration and clinical validation.

## **Keywords**

Digital phenotyping, machine learning, anxiety, depression, computer vision, facial recognition, body gesture analysis, MENTAL AI, psychological support.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação . . . . .	1
1.2	Objetivos . . . . .	1
1.3	Organização do documento . . . . .	2
<b>2</b>	<b>Enquadramento</b>	<b>5</b>
2.1	Saúde Mental . . . . .	5
2.1.1	Transtornos Mentais: Depressão e Ansiedade . . . . .	6
2.2	Redes Sociais como Fonte de Dados . . . . .	7
2.2.1	Potencial das Redes Sociais para a Saúde Mental . . . . .	8
2.2.2	Benefícios do Uso de Redes Sociais como Fonte de Dados . . . . .	8
2.2.3	Desafios e Limitações . . . . .	8
2.2.4	Aplicações na Saúde Mental . . . . .	9
2.3	Indicadores Comportamentais no Contexto Textual . . . . .	9
2.3.1	Comportamentos Associados à Depressão . . . . .	9
2.3.2	Comportamentos Associados à Ansiedade . . . . .	10
2.3.3	Comportamentos Gerais . . . . .	10
2.4	Conclusões . . . . .	11
<b>3</b>	<b>Trabalhos Relacionados</b>	<b>13</b>
3.1	Embeddings e Representação Semântica . . . . .	13
3.1.1	Representações tradicionais de texto . . . . .	13
3.1.2	<i>Dense Word Embeddings</i> . . . . .	14
3.1.3	<i>Contextualized Embeddings</i> com <i>Transformers</i> . . . . .	14
3.1.4	Embedding Generation and Enhancement . . . . .	15
3.1.5	Aplicações na saúde mental . . . . .	15
3.2	Similaridade Semântica e Mapeamento de Comportamentos . . . . .	16
3.3	Fenotipagem Digital e Processamento de Linguagem Natural (NLP) . . . . .	17
3.3.1	Processamento de Linguagem Natural (NLP) . . . . .	19
3.3.2	Considerações éticas . . . . .	19
3.4	Estudos fundamentais em fenotipagem digital . . . . .	20
3.5	Indicadores comportamentais e modelos preditivos . . . . .	22
3.6	Padrões de saúde mental e sociabilidade dos jovens . . . . .	23
3.7	Integração de dados multimodais . . . . .	25
3.8	Dispositivos eletrónicos e doenças mentais graves . . . . .	26
3.9	Escalabilidade e aplicações ao nível da população . . . . .	28
3.10	Avanços nas tecnologias digitais de saúde mental . . . . .	29
3.11	Conclusões . . . . .	31

# Fenotipagem digital para a saúde mental: Indicadores comportamentais

<b>4</b>	<b>Abordagem / Metodologia</b>	<b>33</b>
4.1	Introdução . . . . .	33
4.2	Objetivo inicial e reformulação da abordagem . . . . .	33
4.3	Recolha de dados . . . . .	35
4.4	Fine-tuning inicial no Vertex AI . . . . .	36
4.5	Iteração cruzada e construção do conjunto de dados de transtornos . . . . .	38
4.6	Associação de comportamentos . . . . .	39
4.7	Fine-tuning <i>multitask</i> e aplicação ao conjunto de dados SNCrawler . . . . .	41
4.8	Limitações da abordagem . . . . .	43
4.9	Conclusão . . . . .	44
<b>5</b>	<b>Resultados e Discussão</b>	<b>47</b>
5.1	Fine-tuning para ansiedade e depressão . . . . .	47
5.2	Conjuntos de dados obtidos . . . . .	49
5.3	Modelo <i>multitask</i> (Transtornos + Comportamentos) . . . . .	52
5.4	Aplicação ao conjunto de dados SNCrawler . . . . .	53
5.5	Síntese dos resultados . . . . .	56
<b>6</b>	<b>Conclusões e Trabalho futuro</b>	<b>57</b>
6.1	Conclusões . . . . .	57
6.2	Trabalhos futuros . . . . .	58
	<b>Bibliografia</b>	<b>61</b>

## Lista de Figuras

4.1	Fluxograma da metodologia adotada neste estudo. . . . .	35
4.2	Número de casos acima do <i>threshold</i> de similaridade. . . . .	40
5.1	Desempenho do modelo de ansiedade: evolução da precisão ( <i>accuracy</i> ), da função de perda ( <i>loss</i> ) e do número de predições durante o treino. . . . .	48
5.2	Desempenho do modelo de depressão: evolução da precisão ( <i>accuracy</i> ), da função de perda ( <i>loss</i> ) e do número de predições durante o treino. . . . .	48
5.3	Distribuição dos transtornos no conjunto de dados inicial sem comportamentos associados. . . . .	50
5.4	Distribuição dos transtornos no conjunto de dados inicial com comportamentos, usado para fine-tuning do modelo multitask. . . . .	50
5.5	Distribuição dos comportamentos no conjunto de dados inicial com comportamentos, usado para fine-tuning do modelo multitask. . . . .	51
5.6	Desempenho do modelo <i>multitask</i> : evolução da precisão ( <i>accuracy</i> ), da função de perda ( <i>loss</i> ) e do número de predições durante o treino. . . . .	52
5.7	Distribuição dos transtornos no conjunto de dados final. . . . .	54
5.8	Distribuição dos comportamentos no conjunto de dados final. . . . .	55

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## Lista de Tabelas

4.1	Número de casos retidos acima de cada <i>threshold</i> de similaridade. . . . .	40
-----	---	----

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Lista de Acrónimos**

UBI	Universidade da Beira Interior
OMS	Organização Mundial da Saúde
WHO	World Health Organization
LLM	Large Language Model
CNN	Convolutional Neural Network
IA	Inteligência Artificial
NLP	Natural Language Processing
XAI	Explainable Artificial Intelligence
TCC	Terapia Cognitivo-Comportamental
DSM-5	Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition
BoW	Bag of Words
TF-IDF	Term Frequency – Inverse Document Frequency
SVM	Support Vector Machine
SSRI	Selective Serotonin Reuptake Inhibitor
EMA	Ecological Momentary Assessment
JITAI	Just-In-Time Adaptive Intervention
EHR	Electronic Health Records
EEG	Electroencephalogram
MFCC	Mel-Frequency Cepstral Coefficients
HRV	Heart Rate Variability

# **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

# **Capítulo 1**

## **Introdução**

### **1.1 Motivação**

O interesse na saúde mental e a sua importância estão a crescer tanto a nível de investigação científica como nas políticas públicas, pois a população geral, hoje em dia, está mais exposta a situações que refletem a prevalência significativa de transtornos como a depressão e a ansiedade em diferentes grupos etários e contextos. Este interesse passa por otimizar o prognóstico clínico e fomentar intervenções personalizadas, permitindo uma identificação antecipada e monitorização destes transtornos.

Nos anos mais recentes, emergiram abordagens inovadoras que visam explorar os dados digitais produzidos no quotidiano — tais como publicações em redes sociais, padrões de comunicação ou registos de dispositivos móveis — para extrair indicadores psicológicos e comportamentais relevantes. A fenotipagem digital, um campo emergente nesta área, tem demonstrado um potencial substancial para complementar as práticas clínicas tradicionais, fornecendo sinais indiretos, embora consistentes, do estado psicológico das pessoas. Também uma hipótese ainda muito pouco explorada é através de comportamentos ditos característicos de cada transtorno conseguir detetar sinais desses transtornos e assim intervir antecipadamente, nessa linha de ideias é onde se encontra a presente dissertação.

Apesar deste potencial, o progresso na área tem sido limitado pela escassez de conjuntos de dados multimodais que permitam relacionar, de forma integrada, diferentes tipos de dados (texto, voz, imagem, fisiologia) com estados de saúde mental. A ausência destes recursos dificulta o desenvolvimento de modelos robustos e limita a validação de abordagens mais abrangentes.

Diante da limitação mencionada, a tese em questão adotou uma abordagem exclusivamente textual, para investigar a hipótese de que o discurso digital, particularmente em redes sociais, contém indícios linguísticos que permitam não só deduzir a existência de transtornos mentais, mas também os comportamentos a eles associados.

### **1.2 Objetivos**

O objetivo principal desta tese é o desenvolvimento de um conjunto de dados textuais anotados que estabeleçam uma relação entre transtornos mentais, nomeadamente a ansiedade e a depressão, e comportamentos específicos descritos como característicos desses transtornos na literatura, onde os dados são provenientes do conjunto de dados SNCrawler (conjunto de dados de redes sociais).

De forma mais detalhada, os objetivos específicos são:

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- Recolher e integrar dados textuais de diversas fontes, já anotados quanto à presença de ansiedade e depressão.
- Construir um conjunto de dados unificado de transtornos que permita a representação conjunta de ausência de transtorno, ansiedade isolada, depressão isolada e Ambos os transtornos.
- Desenvolver um vocabulário expandido de comportamentos, baseado em literatura científica e modelos de linguagem, e utilizá-lo para mapear comportamentos em texto através de técnicas de similaridade semântica.
- Produzir um conjunto de dados *multitask* com instâncias anotadas simultaneamente em termos de transtornos e comportamentos.
- Treinar modelos de aprendizagem automática para classificar transtornos e comportamentos, recorrendo a técnicas de *fine-tuning* em plataformas de NLP (como o Vertex AI da Google Cloud).
- Aplicar o modelo *multitask* ao conjunto de dados SNCrawler, criando um recurso de larga escala anotado com transtornos e comportamentos, destinado a investigações futuras em fenotipagem digital.

### 1.3 Organização do documento

O documento encontra-se organizado da seguinte forma:

- **Capítulo 2 – Enquadramento:** apresenta os conceitos fundamentais relacionados com saúde mental, de forma a permitir o panorama atual da saúde mental no mundo, o que é e também o que envolve. Este capítulo também busca de mostrar mais no que consiste a fenotipagem digital e o que envolve.
- **Capítulo 3 – Trabalhos Relacionados:** analisa estudos prévios que de alguma maneira marcaram os avanços na fenotipagem digital e também analisa estudos recentes sobre deteção de transtornos mentais em dados digitais, com foco nas abordagens e recursos utilizados.
- **Capítulo 4 – Abordagem e Metodologia:** descreve as etapas seguidas, incluindo a construção dos conjuntos de dados, o desenvolvimento do vocabulário de comportamentos, os processos de associação dos comportamentos e transtornos com os textos e o treino dos modelos.
- **Capítulo 5 – Resultados:** apresenta os resultados obtidos em cada fase do trabalho, desde o desempenho dos modelos binários e *multitask* até à aplicação ao conjunto de dados SNCrawler, incluindo uma síntese final. Aqui também é feita uma análise aos resultados.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

- **Capítulo 6 – Conclusões e Trabalho futuro:** discute as principais conclusões da tese, destacando os contributos alcançados, as limitações identificadas e possíveis direções para trabalhos futuros.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## **Capítulo 2**

### **Enquadramento**

O presente capítulo tem como objetivo contextualizar os fundamentos teóricos que sustentam o desenvolvimento desta dissertação. Parte-se de uma perspetiva abrangente sobre a saúde mental, com particular enfoque nos transtornos mais prevalentes, como a depressão e a ansiedade, que representam uma parte significativa do peso global de transtornos. Em seguida, discute-se o papel das redes sociais como fonte de dados, evidenciando o seu potencial para a investigação em saúde mental e, simultaneamente, os desafios éticos e metodológicos que o seu uso implica. Por fim, apresentam-se os principais indicadores comportamentais relevantes no contexto textual, que servem como base para a criação de um conjunto de dados anotados.

Esta estrutura fornece o enquadramento necessário para compreender a pertinência e a inovação da abordagem metodológica seguida nesta dissertação, permitindo a transição natural para a exploração de trabalhos relacionados e de técnicas avançadas de representação semântica.

#### **2.1 Saúde Mental**

A saúde mental, tal como definida pela Organização Mundial da Saúde (OMS), é um estado de bem-estar em que um indivíduo pode reconhecer as suas capacidades, lidar com as tensões normais da vida, trabalhar de forma produtiva e contribuir para a sua comunidade. Não se limita à ausência de doença, mas inclui também o bem-estar emocional e a resiliência psicológica. Os estudos mostram que os transtornos mentais são a principal causa de incapacidade em todo o mundo, contribuindo significativamente para o peso global da doença. A OMS estimava que, em 2019, mais de 970 milhões de pessoas foram afetadas por um transtorno mental, com a pandemia de COVID-19 a agravar estes números, sobretudo entre os jovens e os profissionais de saúde [1].

Os transtornos mentais mais comuns incluem a depressão e a ansiedade, que são frequentemente classificados como transtornos mentais comuns [1]. Além disso, doenças como a esquizofrenia, o transtorno bipolar e os transtornos associados ao consumo de substâncias têm uma prevalência mais baixa, mas um impacto elevado. A depressão é uma das principais causas de anos vividos com incapacidade, e o seu impacto vai para além do sofrimento individual, afetando as relações sociais, o desempenho profissional e até a economia global. De acordo com o Global Burden of Disease Study 2019, os transtornos depressivos representaram cerca de 3% do peso total da doença a nível mundial, e a prevalência dos transtornos de ansiedade foi estimada em 31,1% ao longo da vida em populações representativas a nível mundial [2].

O acesso aos cuidados de saúde mental continua a ser um desafio mundial. Segundo a OMS,

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

menos de metade das pessoas com transtornos mentais graves nos países de rendimento elevado recebem tratamento adequado. Esta taxa é ainda mais baixa nos países de baixo e médio rendimento, onde barreiras como o estigma, a falta de profissionais especializados e a fragmentação dos sistemas de saúde limitam os cuidados.

A identificação precoce dos transtornos mentais é um elemento-chave para melhorar os resultados do tratamento. Contudo, as abordagens tradicionais têm limitações significativas na deteção de sinais subtis ou emergentes. Neste contexto, os avanços tecnológicos e científicos, como a fenotipagem digital, têm merecido uma atenção crescente. Este método combina dados comportamentais e biomarcadores recolhidos por meio de dispositivos digitais, como telemóveis, para identificar padrões associados a transtornos mentais. Esta abordagem tem o potencial de identificar alterações precoces nos estados emocionais e cognitivos, permitindo intervenções personalizadas e atempadas.

Dado o impacto crescente dos transtornos mentais, a utilização de novas tecnologias para o diagnóstico, a monitorização e o acompanhamento é uma forma promissora e importante de ultrapassar as limitações tradicionais em matéria de saúde mental. Abordagens como a fenotipagem digital não só oferecem ferramentas inovadoras, como também respondem à necessidade urgente de métodos mais acessíveis, contínuos e eficazes, nomeadamente em contextos de recursos limitados.

### **2.1.1 Transtornos Mentais: Depressão e Ansiedade**

Os transtornos mentais abrangem uma ampla gama de condições que variam em gravidade, impacto e complexidade. Entre os mais prevalentes estão a depressão e a ansiedade, que juntas representam uma parcela significativa do peso global dos transtornos mentais. Estas condições são frequentemente classificadas como transtornos mentais comuns devido à sua alta incidência na população geral [3].

Além da sua prevalência, a depressão e a ansiedade frequentemente coexistem, o que aumenta a complexidade do diagnóstico e do tratamento. Estudos recentes destacam que a comorbidade entre estas condições está associada a maior gravidade dos sintomas, pior qualidade de vida e maior resistência ao tratamento [4].

#### **2.1.1.1 Depressão**

A depressão é um transtorno afetivo caracterizado por um estado persistente de tristeza, perda de interesse ou prazer em atividades diárias e uma série de sintomas emocionais, cognitivos e físicos. Conforme o *DSM-5*, os critérios diagnósticos incluem humor deprimido, alterações no apetite e no sono, fadiga, sentimentos de inutilidade ou culpa, dificuldade de concentração e pensamentos recorrentes de morte ou suicídio [5].

A depressão afeta aproximadamente 4,4% da população mundial, com variações regionais significativas. Estudos recentes indicam que a prevalência de depressão é particularmente alta em populações expostas a fatores de risco, como conflitos armados, doenças crónicas e condições socioeconómicas adversas. Por exemplo, em Gaza, 72,7% da população reportou sintomas moderados a graves de depressão após um ano de guerra [5].

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

A depressão é uma das principais causas de anos vividos com incapacidade, afetando não apenas o bem-estar individual, mas também as relações sociais, o desempenho profissional e a economia global. Em pacientes com doenças crônicas, tais como a diabetes e a hipertensão, a prevalência da depressão pode exceder 70%, intensificando os obstáculos à adesão ao tratamento e agravando os resultados clínicos [6].

O tratamento da depressão geralmente combina intervenções farmacológicas, como inibidores seletivos da recaptação de serotonina (SSRIs), e terapias psicossociais, como a terapia cognitivo-comportamental (TCC). Estão a ser exploradas abordagens inovadoras, como a imersão cognitivo-comportamental no metaverso, com vista a superar barreiras de acesso e escalabilidade. Os ensaios clínicos realizados têm produzido resultados promissores [7].

### **2.1.1.2 Ansiedade**

A ansiedade é caracterizada por sentimentos persistentes de medo ou preocupação excessiva, frequentemente acompanhados por sintomas físicos, como tensão muscular, inquietação, fadiga, dificuldade de concentração e perturbações do sono. Assim como a depressão, a ansiedade é uma condição heterogênea, com diferentes subtipos, incluindo transtorno de ansiedade generalizada (TAG), transtorno do pânico e fobias específicas [8].

Afeta cerca de 3,6% da população mundial, com taxas mais altas em mulheres e em populações expostas a fatores de risco, como estudantes de enfermagem e pacientes com doenças crônicas. Por exemplo, 47,5% dos estudantes de saúde em Karachi relataram níveis extremamente severos de ansiedade [9]. Em pacientes com cancro, a ansiedade é frequentemente moderada a severa, afetando até 68% dos indivíduos [10].

Ela está associada a uma série de consequências negativas, incluindo pior desempenho acadêmico, maior risco de doenças cardiovasculares e redução da qualidade de vida. Em populações vulneráveis, como adolescentes grávidas, a ansiedade pode agravar os desafios de saúde física e mental, destacando a necessidade de intervenções precoces [11].

O tratamento da ansiedade inclui intervenções farmacológicas, como benzodiazepinas e SSRIs, e terapias psicossociais, como a TCC. A integração de abordagens no âmbito da saúde mental com a gestão de condições subjacentes (como o abuso de substâncias) tem demonstrado eficácia no decréscimo da gravidade dos sintomas [12].

## **2.2 Redes Sociais como Fonte de Dados**

As redes sociais tornaram-se parte integrante da vida moderna, oferecendo um espaço para a autoexpressão, socialização e partilha de informações. Além de serem um meio de comunicação, estas plataformas também se tornaram uma fonte valiosa de dados para a investigação em saúde mental. Publicações, comentários e interações nas redes sociais muitas vezes refletem estados emocionais, padrões comportamentais e mudanças cognitivas, tornando-as uma ferramenta poderosa para analisar condições como a depressão e a ansiedade [13][14].

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

### 2.2.1 Potencial das Redes Sociais para a Saúde Mental

As redes sociais proporcionam uma perspectiva singular sobre os estados emocionais e cognitivos dos utilizadores, viabilizando a obtenção de um vasto conjunto de dados em tempo real. Conforme demonstrado pelos estudos realizados, é possível afirmar que:

- **Expressão emocional:** As pessoas costumam usar plataformas de redes sociais como *Twitter*, *Facebook* e *Instagram* para partilhar as suas emoções, pensamentos e experiências. Isso pode revelar sinais de depressão e ansiedade [15].
- **Padrões linguísticos:** Pessoas com depressão tendem a usar mais palavras associadas a emoções negativas, como "triste" ou "cansado", enquanto indivíduos com ansiedade costumam usar palavras relacionadas à preocupação, como "ansioso" ou "preocupado" [16].
- **Mudanças comportamentais:** Alterações no estado mental, como isolamento social ou hiperatividade, podem ser indicadas pela frequência das publicações, o tempo de atividade e a interação com outros utilizadores [17].

### 2.2.2 Benefícios do Uso de Redes Sociais como Fonte de Dados

1. **Acessibilidade e Escalabilidade:** As redes sociais facilitam a recolha de quantidades substanciais de dados de uma forma acessível e escalável, superando assim as limitações dos métodos tradicionais, como entrevistas ou questionários [14].
2. **Dados em Tempo Real:** A análise em tempo real das publicações nas redes sociais pode ajudar a identificar sinais de problemas de saúde mental numa fase inicial, permitindo uma intervenção mais rápida [15].
3. **Diversidade de Dados:** Essas plataformas oferecem uma variedade de dados, incluindo texto, imagens e vídeos. Esses dados podem ser analisados para obter uma visão mais abrangente sobre a saúde mental [18].

### 2.2.3 Desafios e Limitações

Apesar do seu potencial, utilizar as redes sociais como fonte de dados apresenta desafios significativos:

- **Privacidade e Ética:** A recolha e utilização de dados das redes sociais suscitam questões éticas, tais como a necessidade de consentimento informado e a proteção da privacidade dos utilizadores [14].
- **Ruído nos Dados:** Uma análise imprecisa pode ser resultado da presença de informações irrelevantes ou ambíguas nas publicações [13].
- **Representatividade:** Os resultados podem ser enviesados pelo facto de nem todos os grupos populacionais estarem igualmente representados nas redes sociais [18].

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

### 2.2.4 Aplicações na Saúde Mental

As redes sociais têm sido utilizadas em diversas aplicações relacionadas à saúde mental, incluindo:

1. **Deteção de Transtornos Mentais:** Modelos de aprendizagem automática, como *Random Forest* e *XGBoost*, têm sido usados para identificar sinais de depressão e ansiedade em publicações nas redes sociais com uma taxa de precisão superior a 90% [17].
2. **Monitoramento Contínuo:** A análise dos padrões de uso e linguagem permite o monitoramento contínuo dos estados mentais dos indivíduos, facilitando intervenções personalizadas [16].
3. **Promoção de Bem-Estar:** Algumas plataformas introduziram funcionalidades destinadas a promover a saúde mental, incluindo *links* para serviços de apoio e campanhas de sensibilização [14].

## 2.3 Indicadores Comportamentais no Contexto Textual

Os comportamentos associados a transtornos mentais, como depressão e ansiedade, podem ser encontrados em dados textuais e outros sinais digitais. Uma maneira de encontrar comportamentos mais físicos e não textuais é alargar o vocabulário do próprio comportamento de várias maneiras e contextos de dizer, por exemplo, a postura encurvada pode ser dita como "cabisbaixo" ou "Estou tão triste que até estou curvado, como se o peso do mundo tivesse caído sobre os meus ombros.", estas frases/expressões podem refletir o comportamento postura encurvada o que se pode deduzir existencial de depressão por exemplo.

### 2.3.1 Comportamentos Associados à Depressão

A depressão é frequentemente associada a uma redução na atividade geral e a alterações emocionais que podem ser deduzidas por meio dos seguintes comportamentos:

- **Reduced Speech Rate:** Redução na velocidade da fala, adaptada para texto como menor frequência de postagens ou textos mais curtos.
- **Reduced Pitch Variation:** Linguagem monótona, com pouca variação emocional.
- **Long Pauses:** Intervalos longos entre postagens, indicando lentificação cognitiva.
- **Reduced Facial Expression:** Linguagem neutra ou ausência de expressões emocionais.
- **Slowed Movement:** Textos que refletem lentidão cognitiva, como frases curtas e simples.
- **Hunched Posture (deduzido):** Linguagem que sugere retração ou isolamento.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- **Reduced Digital Use:** Baixa frequência de postagens ou longos períodos de inatividade.
- **Low HRV (Heart Rate Variability):** Embora não diretamente textual, pode ser correlacionado com estados emocionais de baixa energia.
- **Altered Handwriting:** No contexto textual, pode ser adaptado para padrões de escrita inconsistentes ou erros frequentes.

### 2.3.2 Comportamentos Associados à Ansiedade

A ansiedade, por sua vez, é caracterizada por estados de hiperatividade emocional e cognitiva, refletidos nos seguintes comportamentos:

- **Increased Speech Rate:** Alta frequência de postagens em curtos períodos de tempo.
- **Increased Pitch Variation:** Uso de linguagem emocionalmente carregada, com variações extremas de tom.
- **Worried Facial Expression:** Uso frequente de palavras relacionadas a preocupação, como "ansioso" ou "preocupado".
- **Physical Restlessness:** Postagens frequentes e impulsivas, com uso de linguagem repetitiva ou urgente.
- **Abnormal Eye Movement (deduzido):** Linguagem que sugere inquietação ou distração.
- **Excessive Digital Use:** Postagens excessivas, muitas vezes com conteúdo redundante ou impulsivo.
- **Altered Microexpressions (deduzido):** Linguagem que reflete mudanças emocionais rápidas ou inconsistentes.

### 2.3.3 Comportamentos Gerais

Além dos comportamentos específicos para depressão e ansiedade, alguns indicadores podem ser comuns a ambos os transtornos, dependendo do contexto:

- **Altered MFCC (Mel-Frequency Cepstral Coefficients):** Embora mais aplicável a dados de áudio, pode ser adaptado para padrões textuais que refletem alterações na entonação ou ritmo.
- **Jitter/Shimmer:** No contexto textual, pode ser interpretado como inconsistências na estrutura ou no tom das postagens.

Esses comportamentos foram mapeados para textos utilizando técnicas de **Processamento de Linguagem Natural (NLP)**, como *embeddings* semânticos e similaridade do cosseno, isto é, foi aumentado o vocabulário de cada comportamento e depois sim foram usadas as técnicas anteriores. A abordagem permitiu associar cada entrada textual a um comportamento, contribuindo para a criação de um conjunto de dados com anotações comportamentais.

### 2.4 Conclusões

Este capítulo apresentou o enquadramento fundamental da investigação. Foram discutidas a relevância da saúde mental e o impacto dos transtornos mais prevalentes, como a depressão e a ansiedade, destacando não só a sua elevada incidência, mas também as implicações sociais e clínicas associadas. Explorou-se igualmente o potencial das redes sociais como fonte de dados para o estudo da saúde mental, realçando tanto as oportunidades que oferecem em termos de acessibilidade e diversidade de informação, como os desafios éticos e metodológicos que colocam.

Adicionalmente, foram sistematizados os principais indicadores comportamentais que podem ser encontrados a partir de dados textuais, permitindo a sua adaptação para tarefas de análise automática através de técnicas de *NLP*. Estes indicadores constituem uma base essencial para a construção de conjunto de dados anotados que possibilitam a identificação de padrões associados a transtornos mentais.

Assim, o enquadramento teórico aqui desenvolvido fornece as bases conceptuais para os capítulos seguintes, onde são analisados trabalhos relacionados e exploradas metodologias que procuram ultrapassar as limitações atuais na deteção e monitorização de transtornos mentais através de dados digitais.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## Capítulo 3

### Trabalhos Relacionados

O presente capítulo tem como objetivo apresentar os principais trabalhos relacionados com a investigação desenvolvida nesta dissertação. São analisados contributos relevantes em três eixos centrais: (i) a utilização de *embeddings* e técnicas de representação semântica no processamento de linguagem natural (NLP); (ii) o desenvolvimento e aplicação de modelos preditivos no domínio da saúde mental, com destaque para a depressão e a ansiedade; e (iii) o papel da fenotipagem digital e dos indicadores comportamentais na construção de perfis psicológicos a partir de dados digitais.

A revisão da literatura permite identificar não apenas os avanços alcançados, mas também as limitações e lacunas que justificam a abordagem seguida neste trabalho. Além disso, esta análise evidencia como diferentes técnicas, desde modelos clássicos até *large language models* (LLMs), têm sido aplicadas em contextos clínicos e em dados provenientes de redes sociais, estabelecendo o enquadramento científico para a metodologia proposta nos capítulos seguintes.

#### 3.1 Embeddings e Representação Semântica

A representação semântica e os *embeddings* são conceitos fundamentais no *Natural Language Processing* (NLP), permitindo que as máquinas compreendam e processem a linguagem humana de forma mais eficaz. A insuficiência dos métodos tradicionais, como o *Bag-of-Words* (BoW) ou o *TF-IDF*, na captura do contexto e significado das palavras, é compensada pelos *embeddings* modernos, que oferecem representações ricas e matizadas do texto, devido à sua densidade e contextualização.

##### 3.1.1 Representações tradicionais de texto

Os métodos tradicionais de representação de texto, como *Bag-of-Words* (BoW) e *Term Frequency–Inverse Document Frequency* (TF-IDF), desempenharam um papel importante nas fases iniciais do NLP, mas apresentam limitações significativas na captura de nuances semânticas e contextuais.

O BoW representa o texto como um vetor esparsos de frequências de palavras, ignorando a ordem e o contexto em que estas aparecem. O TF-IDF expande esta ideia, ponderando os termos com base na sua frequência num documento (*Term Frequency*) e na sua raridade no corpus (*Inverse Document Frequency*). Este mecanismo reduz a influência de palavras comuns e destaca os termos mais informativos [19][20].

Apesar de úteis em tarefas simples, como classificação de documentos ou extração de palavras-chave, estes métodos não conseguem capturar polissemia, sinonímia ou relações contextuais — aspetos críticos para a análise de saúde mental, onde a linguagem tende a ser

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

subtil e carregada de significados implícitos. Estudos demonstram que BoW e TF-IDF apresentam desempenho inferior a modelos baseados em *embeddings*, como o BERT, em tarefas que exigem compreensão semântica, incluindo detecção de stress ou análise de sentimentos [21][22]. Ainda assim, pela sua simplicidade e baixo custo computacional, continuam a ser usados como linhas de base ou em cenários com recursos limitados [23].

### 3.1.2 Dense Word Embeddings

Modelos como o *Word2Vec* e o *GloVe* representam um avanço significativo em relação às representações esparsas.

O *Word2Vec* introduziu dois objetivos principais de treino: o *Skip-gram*, que prevê palavras de contexto a partir de uma palavra-alvo, e o *Continuous Bag-of-Words (CBOW)*, que prevê a palavra-alvo com base no contexto. Ambos os métodos otimizam os *embeddings* através de estratégias como *negative sampling* ou *hierarchical softmax*, reduzindo a complexidade computacional sem sacrificar a qualidade das representações [24][25].

O *GloVe (Global Vectors for Word Representation)* utiliza uma abordagem baseada em estatísticas globais de coocorrência. Ao fatorizar matrizes de coocorrência palavra-palavra, este modelo produz *embeddings* capazes de codificar simultaneamente informações semânticas locais e globais, oferecendo representações que refletem tanto o uso imediato das palavras como a sua distribuição geral no corpus [26][27].

Uma das propriedades mais notáveis dos *embeddings* densos é a sua capacidade de capturar relações semânticas através de operações vetoriais. Um exemplo clássico é a equação **rei – homem + mulher ≈ rainha**, que demonstra como estas representações conseguem codificar relações analógicas e estruturais de forma consistente [26] [28].

Graças a estas características, os *embeddings* densos tornaram-se fundamentais para diversas tarefas de *NLP*. Eles oferecem representações compactas, ricas em significado e muito mais eficazes do que os métodos esparsos tradicionais, tendo permitido avanços notáveis em áreas como análise de sentimentos, tradução automática e detecção de padrões linguísticos associados a transtornos mentais [24][26].

### 3.1.3 Contextualized Embeddings com Transformers

Os *embeddings contextualizados* revolucionaram a *NLP* ao atribuir significados dinâmicos às palavras conforme o contexto. Modelos baseados em *ELMo* e, sobretudo, em *Transformers*, como o *BERT* e o *GPT*, são os principais responsáveis por essa mudança.

O *ELMo* utiliza uma arquitetura *BiLSTM* para gerar *embeddings* dependentes do contexto anterior e posterior de uma palavra. Apesar da sua eficácia na modelação da polissemia e da sintaxe, o processamento sequencial limita a escalabilidade [29] [30]. Já os *Transformers*, introduzidos em Attention is All You Need [31], substituem a recorrência pelo mecanismo de *self-attention*, permitindo paralelização e melhor escalabilidade.

- O *BERT* é pré-treinado com *Masked Language Modeling (MLM)*, destacando-se em tarefas de compreensão semântica [32].

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- O *GPT*, por outro lado, usa modelação autorregressiva (*next token prediction*), tornando-se mais eficiente em geração de texto [33].

Estes modelos impulsionaram enormemente o desempenho em NLP, estabelecendo novos padrões de qualidade em tarefas como resposta automática, análise de sentimentos e geração de linguagem natural.

### 3.1.4 Embedding Generation and Enhancement

O *Ollama* oferece uma estrutura robusta para a geração e o refinamento de *embeddings*, permitindo a sua adaptação a aplicações específicas do domínio, como a análise de saúde mental.

O processo inicia-se com a preparação do conjunto de dados, selecionando textos relevantes para o domínio, de forma a garantir que os *embeddings* capturam as nuances semânticas necessárias. O uso de conjuntos de dados clínicos ou provenientes de redes sociais melhora significativamente a qualidade das representações [34][35].

A *prompt engineering* é uma etapa crítica, consistindo na criação de instruções otimizadas para guiar o modelo na geração de frases representativas de cada comportamento. Técnicas como *few-shot prompting* ou a integração de pistas emocionais permitem resultados mais contextualmente adequados, alinhando-se com critérios clínicos como os estabelecidos no DSM-5 [36][37].

A filtragem de qualidade assegura que os resultados sejam representativos e consistentes. Esta etapa inclui a comparação com referências (por exemplo, via similaridade cosseno) e a eliminação de ruído [38]. Termos novos podem ser incorporados aos vocabulários já existentes através de *transfer learning* ou *fine-tuning*, garantindo compatibilidade retroativa [39].

Finalmente, o desempenho dos *embeddings* é avaliado com métricas como *F1-score* e validações humanas, assegurando que atendem aos requisitos específicos da tarefa [40][41]. Este fluxo de trabalho mostra como o *Ollama* é capaz de fazer representações de qualidade, ajustadas a domínios especializados.

### 3.1.5 Aplicações na saúde mental

Os *embeddings* são fundamentais para otimizar a análise da saúde mental, ao facilitarem a identificação de marcadores linguísticos, medir a semelhança semântica, construir conjuntos de dados e detetar preconceitos.

1. **Identificação de marcadores linguísticos:** Embeddings, particularmente aqueles derivados de modelos como BERT e RoBERTa, são essenciais para identificar padrões linguísticos associados a condições de saúde mental, como depressão e ansiedade. Estudos mostraram, por exemplo, que *embeddings* podem capturar marcadores subtis, como autoaversão, solidão e distúrbios alimentares, com alta precisão [42] [43].
2. **Similaridade semântica para classificação:** Medidas de similaridade semântica, como a similaridade cosseno, melhoram a classificação de textos ao identificar relações

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

entre textos relacionados à saúde mental. Essas medidas são particularmente eficazes para distinguir entre condições sobrepostas, como depressão e ansiedade, conforme demonstrado em tarefas de classificação com várias etiquetas [44] [45].

3. **Construção de conjuntos de dados:** Embeddings facilitam a criação de conjuntos de dados anotados, agrupando publicações com base em características linguísticas, permitindo assim o desenvolvimento de modelos preditivos. Por exemplo, conjuntos de dados baseados no Reddit têm sido empregados para treinar modelos para a identificação precoce de problemas de saúde mental [46] [47].
4. **Análise de preconceitos e dados demográficos:** Os Embeddings também podem ajudar a revelar preconceitos em modelos de saúde mental, como disparidades culturais ou linguísticas. Estudos interculturais enfatizam a importância de criar embeddings adaptados a dados demográficos específicos, a fim de garantir que as intervenções em saúde mental sejam equitativas [48] [49].

O potencial transformador do *embedding* na investigação em saúde mental e na prática clínica é demonstrado por estas aplicações. Estando cada vez mais próximos de um diagnóstico mais real e fidedigno.

### 3.2 Similaridade Semântica e Mapeamento de Comportamentos

A *similaridade semântica* é um conceito central no *Natural Language Processing (NLP)*, utilizado para medir a proximidade entre palavras, frases ou documentos com base no seu significado. Esta técnica é fundamental em tarefas como recuperação de informação, classificação de textos e análise de sentimentos. No contexto desta tese, a *similaridade semântica* é aplicada para mapear comportamentos humanos, associando padrões linguísticos a categorias comportamentais específicas, como *reduced speech rate* ou *worried facial expression*.

#### O Conceito de Similaridade Semântica

O princípio da *similaridade semântica* assenta na hipótese distributiva: palavras que ocorrem em contextos semelhantes tendem a ter significados próximos. Para operacionalizar esta ideia, o texto é convertido em representações vetoriais e são aplicadas métricas matemáticas que calculam a proximidade entre vetores.

Os métodos mais comuns para medir a similaridade semântica incluem:

- **Similaridade cosseno:** mede o ângulo entre dois vetores e é amplamente utilizada para comparar *embeddings* de palavras ou frases [50].
- **Distância euclidiana:** calcula a distância linear entre vetores, útil para avaliar diferenças absolutas.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- **Métodos baseados em *kernel*:** permitem medir similaridade em espaços de alta dimensão, sendo aplicados em modelos como *Support Vector Machines* [51].

Estas técnicas são geralmente aplicadas em conjunto com *embeddings*, que transformam palavras e frases em vetores densos e informativos. Modelos como *Word2Vec*, *GloVe* ou *BERT* são amplamente usados neste contexto, por captarem relações semânticas e contextuais [24][26][32].

No âmbito desta tese, a *similaridade semântica* foi usada para mapear textos de redes sociais a comportamentos humanos relacionados com transtornos mentais. Para tal, o vocabulário comportamental foi expandido com o auxílio do modelo *Ollama*, que gerou frases adicionais para cada categoria. Essas frases foram depois integradas no processo de mapeamento.

### Mapeamento de Comportamentos

O mapeamento comportamental consiste em associar textos a categorias pré-definidas de comportamento, com base no cálculo da sua proximidade semântica. Esta técnica é particularmente relevante em saúde mental, dado que expressões linguísticas podem refletir estados emocionais ou cognitivos específicos. Por exemplo, frases como “falo muito devagar” ou “tenho dificuldade em expressar-me” podem ser associadas ao comportamento *reduced speech rate*. Da mesma forma, expressões como “não consigo ficar quieto” ou “estou sempre a balançar a perna” podem ser mapeadas para *physical restlessness*.

As etapas a seguir geralmente estão envolvidas no processo de mapeamento:

- **Representação vetorial:** tanto os textos como o vocabulário comportamental são convertidos em *embeddings*, obtidos a partir de modelos pré-treinados ou ajustados ao domínio.
- **Cálculo de similaridade:** métricas como a similaridade cosseno são aplicadas para determinar a proximidade entre os vetores das frases e das categorias comportamentais.
- **Classificação:** cada texto é atribuído ao comportamento mais próximo, de acordo com os valores de similaridade obtidos.

Este método permite a análise eficiente de grandes volumes de dados textuais, como publicações em redes sociais, identificando padrões comportamentais que de outra forma poderiam passar despercebidos. No contexto da saúde mental, representa uma abordagem para explorar sinais precoces de depressão e ansiedade, ou até mesmo dos transtornos restantes.

### 3.3 Fenotipagem Digital e Processamento de Linguagem Natural (NLP)

A fenotipagem digital consiste na recolha e análise de dados digitais gerados no dia a dia das pessoas, através do uso de *smartphones*, aplicações móveis e outros dispositivos. Surge

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

como uma alternativa eficaz e acessível aos métodos clínicos tradicionais, permitindo uma monitorização contínua, objetiva e em larga escala da saúde mental. Enquanto os métodos convencionais dependem de relatórios pontuais e percepções subjetivas, a fenotipagem digital capta sinais comportamentais subtis e contextuais em tempo real [52][53].

Esta abordagem associa variáveis comportamentais diárias — como atividade física, interações sociais e padrões de comunicação — a indicadores de estados emocionais e psicológicos. Por exemplo, mudanças na mobilidade ou na frequência de interações online podem sinalizar sintomas de depressão, ou ansiedade. Assim, a fenotipagem digital contribui para a deteção precoce e para o desenho de intervenções mais direcionadas e personalizadas [54][55].

Os objetivos desta abordagem vão além do diagnóstico, incluindo a monitorização de padrões recorrentes que auxiliam no desenvolvimento de intervenções adaptativas. As suas principais vantagens incluem:

- **Identificação Precoce:** permite detetar alterações comportamentais antes de os sintomas evoluírem para estados graves, como evidenciado em estudos que correlacionam padrões de mobilidade com episódios depressivos [56][57].
- **Monitorização contínua:** ao contrário de consultas pontuais, fornece uma visão longitudinal e dinâmica do estado mental dos indivíduos [58][59].
- **Intervenções Personalizadas:** os dados recolhidos servem de base para planos terapêuticos adaptados ao perfil de cada utilizador, aumentando a eficácia do tratamento [60].
- **Acessibilidade:** a ubiquidade dos dispositivos móveis democratiza o acesso a cuidados de saúde mental, especialmente em contextos com recursos limitados [61].

A recolha de dados em fenotipagem digital é um dos passos fundamentais na implementação de soluções baseadas em Inteligência Artificial (IA). Estes dados, recolhidos a partir de múltiplas fontes, fornecem uma visão abrangente dos comportamentos e estados emocionais de um indivíduo, criando a base para modelos preditivos. Dependendo da forma como os dados são recolhidos, podem ser divididos em duas categorias principais:

### 1. Dados Passivos

Os dados passivos são recolhidos automaticamente, sem intervenção direta do utilizador, e captam as interações diárias com dispositivos e informações ambientais. Estas fontes incluem:

- **Geolocalização:** Os indicadores das distâncias percorridas e das deslocações efetuadas ajudam a compreender os padrões de mobilidade, que são frequentemente reduzidos nos doentes com depressão moderada ou grave. [62].
- **Sensores Biométricos:** Os dispositivos que podem medir dados como o ritmo cardíaco, os padrões de sono e os movimentos diários são utilizados para detetar alterações que correspondem a episódios de ansiedade [63] [64].

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- **Padrões de Comunicação:** As mensagens de texto e as chamadas, bem como os padrões de utilização de aplicações, podem ser utilizados para identificar o afastamento social ou o aumento da atividade em situações de stress [65].

### 2. Dados Ativos

Os dados ativos são fornecidos voluntariamente pelos utilizadores, normalmente através de tarefas específicas concebidas para captar informações relevantes sobre o seu estado emocional e cognitivo:

- **Relatórios Subjetivos:** Os questionários diários ou semanais permitem captar as mudanças que os indivíduos percebem em si próprios. Esta informação é particularmente útil quando integrada com dados passivos. [66].
- **Testes Cognitivos e Jogos:** As tarefas de memória, tempo de reação e coordenação avaliam funções cognitivas frequentemente associadas a condições como a ansiedade e o esgotamento. [67].
- **Análise de Linguagem:** Estudos recentes mostraram que o tom, a escolha de palavras e a frequência da comunicação em mensagens escritas podem ser indicadores úteis para monitorizar alterações emocionais [68].

### 3.3.1 Processamento de Linguagem Natural (NLP)

A integração de *Large Language Models (LLMs)* na fenotipagem digital tem revelado um enorme potencial para avaliação psicológica. Modelos como *BERT* e *GPT* conseguem analisar dados textuais e transcrições com elevada sensibilidade contextual, captando padrões linguísticos associados a estados de saúde mental.

Indivíduos com depressão tendem a utilizar palavras com conotações negativas ou egocêntricas, enquanto aqueles com ansiedade recorrem a construções mais repetitivas e fragmentadas. Estudos recentes demonstram que o *BERT*, por exemplo, pode identificar linguagem melancólica ou menções de culpa em redes sociais [69].

Além da análise textual, os LLMs também têm utilidade na modelação temporal, permitindo acompanhar a progressão de estados emocionais ao longo do tempo. Essa capacidade é particularmente relevante em contextos clínicos, como análise de sessões terapêuticas ou suporte digital via *chatbots* com modelos de linguagem, que fornecem apoio emocional em tempo real [70].

### 3.3.2 Considerações éticas

A fenotipagem digital é uma abordagem promissora para avaliar a saúde mental, mas a sua utilização levanta questões éticas importantes que não podem ser ignoradas. Entre elas, a privacidade do utilizador é uma preocupação fundamental. A recolha contínua de dados através de dispositivos digitais, como telemóveis, permite uma análise comportamental rica, mas também expõe informações altamente sensíveis. Mesmo quando os dados são anonimizados, a possibilidade de identificação permanece, especialmente com pequenos conjuntos

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

de dados. Para mitigar este risco, o Regulamento Geral sobre a Proteção de Dados (RGPD) da União Europeia estabelece normas rigorosas para proteger a privacidade, como a exigência de consentimento informado, que deve ser informado, voluntário e revogável em qualquer altura [71].

Apesar do seu potencial, a fenotipagem digital levanta questões éticas críticas. A recolha contínua de dados expõe informações altamente sensíveis, mesmo quando anonimizadas. O Regulamento Geral sobre a Proteção de Dados (RGPD) da União Europeia exige consentimento informado, voluntário e revogável, reforçando a proteção da privacidade [71].

Outro risco é o impacto negativo de alertas baseados em algoritmos, que podem induzir ansiedade ou diagnósticos incorretos. Para mitigar esses problemas, é essencial garantir transparência, revisão contínua e acompanhamento clínico adequado [72].

Além disso, o uso secundário dos dados — como marketing ou vigilância — compromete a confiança pública. Por isso, é imprescindível restringir a utilização aos fins explicitamente autorizados, respeitando a dignidade e os direitos dos participantes [72].

### 3.4 Estudos fundamentais em fenotipagem digital

A fenotipagem digital tem emergido como uma abordagem transformadora para compreender a saúde mental, aproveitando dados recolhidos em tempo real a partir de dispositivos digitais, de forma a quantificar comportamentos e estados mentais. O termo *fenótipo digital* foi introduzido pela primeira vez por Jain (2015), que o descreveu como “a quantificação momento a momento do fenótipo humano ao nível individual *in situ*, utilizando dados de dispositivos digitais pessoais” [73]. Inspirando-se no conceito de *fenótipo alargado* proposto por Richard Dawkins, esta definição destacou o potencial dos traços digitais para revelar padrões comportamentais e de saúde para além dos contextos clínicos tradicionais. O trabalho de Jain estabeleceu as bases teóricas do campo, demonstrando a capacidade de *smartphones* e dispositivos vestíveis em capturar medidas contínuas e ecológicas do comportamento humano.

A partir desta proposta, Insel (2017) enquadrou a fenotipagem digital como uma “nova ciência do comportamento”, ao integrar dados biológicos, comportamentais e sociais para promover a psiquiatria de precisão [74]. O autor salientou que a falta de medições objetivas na psiquiatria constituía, há décadas, um obstáculo ao progresso no diagnóstico e tratamento. Argumentou, assim, que a fenotipagem digital poderia colmatar esta lacuna, ao fornecer dados contínuos e contextualizados sobre as experiências dos pacientes. Alterações nos padrões de mobilidade ou nas características da fala, por exemplo, poderiam servir como indicadores precoces de deterioração da saúde mental, possibilitando intervenções atempadas.

Um contributo marcante foi a criação da plataforma *Beiwe*, desenvolvida por Onnela et al. (2016), que operacionalizou a fenotipagem digital em contexto de investigação. Esta ferramenta permitia a recolha de dados passivos de *smartphones*, como leituras de GPS, dados de acelerómetros e padrões de utilização do telefone, para o estudo de condições psiquiátricas como a esquizofrenia [75]. Ao demonstrar a viabilidade de capturar manifestações comportamentais de doenças mentais em ambientes naturalistas, a plataforma abriu caminho a

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

estudos subsequentes. Os resultados iniciais evidenciaram que dados passivos conseguiam revelar alterações sociais e comportamentais associadas a condições de saúde mental, oferecendo um método escalável e objetivo de monitorização de pacientes.

Num plano mais alargado, Insel (2018) destacou as implicações globais da fenotipagem digital, em particular para colmatar lacunas nos cuidados de saúde mental em contextos com poucos recursos [76]. A ubiquidade dos *smartphones* e a riqueza dos fluxos de dados por eles gerados foram apontadas como oportunidades para transformar a psiquiatria, ao proporcionar medidas objetivas e ecológicas de saúde mental. Estudos revelaram, por exemplo, que métricas de interação humano-computador — como velocidade de digitação e padrões de rolagem — se correlacionam com estados cognitivos e afetivos. Estas descobertas sublinham o potencial da fenotipagem digital como um “alarme de fumo” da saúde mental, permitindo sinalizar precocemente a necessidade de intervenção antes do agravamento dos sintomas.

Revisões recentes consolidaram ainda mais o papel desta abordagem na investigação em saúde mental. Dlima et al. (2022), por exemplo, realizaram uma revisão exploratória de 46 estudos e destacaram a diversidade de fontes de dados e métodos analíticos aplicados no domínio [77]. Os autores observaram que a maioria dos trabalhos se centrava em transtornos psiquiátricos e recorria a algoritmos de *machine learning*, como florestas aleatórias e máquinas de vetores de suporte, para análise dos dados. Foi ainda salientada a necessidade de estudos longitudinais e de conjuntos de dados mais robustos para validar a utilidade da fenotipagem digital em contextos clínicos.

Outro contributo relevante é o de Bernardos et al. (2019), que propuseram uma arquitetura geral para plataformas de fenotipagem digital e identificaram barreiras centrais à sua adoção [78]. Entre estas destacavam-se a necessidade de validação extensiva junto dos utilizadores, preocupações éticas relacionadas com a privacidade dos dados e desafios na integração destas ferramentas nos sistemas de saúde existentes. Os autores defenderam que a superação destes obstáculos é essencial para que a fenotipagem digital realize o seu potencial como pedra angular dos cuidados de saúde mental personalizados.

As dimensões éticas têm assumido crescente relevância neste campo. Martinez-Martin (2018) explorou as implicações éticas, legais e sociais da fenotipagem digital, chamando a atenção para a importância da transparência, do consentimento informado e da responsabilidade [79]. O estudo sublinhou os riscos do uso indevido de dados e a insuficiência das estruturas regulatórias existentes para proteger informação pessoal sensível. Estas preocupações levaram a apelos para a criação de diretrizes éticas robustas, capazes de regular o uso da fenotipagem digital tanto em investigação como na prática clínica.

Em síntese, os estudos fundamentais estabeleceram os alicerces teóricos e práticos da fenotipagem digital, evidenciando o seu potencial para transformar os cuidados de saúde mental. Desde a conceção inicial de Jain (2015), passando pela operacionalização de Onnela et al. (2016) e pelas reflexões éticas de Martinez-Martin (2018), estes trabalhos lançaram as bases de uma nova era da psiquiatria de precisão. Contudo, permanecem desafios críticos, como a validação em larga escala, a proteção ética e a integração nos sistemas de saúde, que constituem áreas prioritárias para investigação futura.

### 3.5 Indicadores comportamentais e modelos preditivos

Os indicadores comportamentais constituem uma via promissora para compreender a saúde mental, captando mudanças subtis e frequentemente imperceptíveis que ocorrem no cotidiano dos indivíduos. Estes indicadores permitem monitorizar, de forma contínua e não invasiva, variações de comportamento que podem sinalizar o surgimento ou agravamento de transtornos mentais, como ansiedade e depressão.

Um exemplo comum é o uso de dados de mobilidade obtidos através de GPS. A variação de localização, a frequência de deslocações ou o tempo passado em determinados locais são alguns dos padrões que têm sido associados a estados depressivos. Os indivíduos com mobilidade reduzida ou que visitam menos locais tendem a apresentar níveis mais baixos de humor, o que reforça a relação entre o movimento físico e o bem-estar emocional [80, 81]. Outro indicador central é o sono, cuja duração, regularidade e variabilidade, frequentemente retiradas a partir de acelerómetros, correlacionam-se com sintomas de ansiedade e depressão.

As interações sociais também emergem como indicadores relevantes. A análise de registos de chamadas, mensagens de texto ou proximidade via Bluetooth permite avaliar o nível de envolvimento social de um indivíduo. Reduções consistentes nestas interações têm sido associadas a sentimentos de solidão ou a estados ansiosos [82]. Para além disso, sinais fisiológicos como a variabilidade da frequência cardíaca e a atividade eletrodérmica acrescentam camadas adicionais de informação, ao relacionar respostas corporais diretas com o bem-estar emocional [80].

Mais recentemente, o uso de indicadores linguísticos tem ganho relevo. O recurso a técnicas de *Natural Language Processing (NLP)* aplicadas a mensagens de texto e publicações em redes sociais permite identificar sentimentos, temas recorrentes e pistas semânticas associadas a estados psicológicos. Estas análises revelam padrões linguísticos que funcionam como marcadores de ansiedade, depressão ou stress, reforçando o valor do texto digital como fonte de dados comportamentais [83].

Os modelos preditivos desempenham um papel fundamental na interpretação destes indicadores. Técnicas de *machine learning*, como florestas aleatórias, máquinas de vetores de suporte e redes neuronais, são amplamente utilizadas para identificar padrões complexos e prever desfechos de saúde mental. Zhang et al. (2024), por exemplo, recorreram a modelos *XGBoost* para prever a gravidade da depressão e ansiedade com base em dados recolhidos por dispositivos vestíveis, alcançando elevados níveis de precisão [81].

O advento de *large language models (LLMs)* ampliou significativamente estas capacidades. Yuan et al. (2025) demonstraram que *LLMs* conseguem processar e integrar dados comportamentais heterogêneos de forma mais eficaz do que abordagens tradicionais de *machine learning*, oferecendo previsões mais robustas sobre a evolução da depressão [83].

Abordagens multimodais, que combinam dados de várias fontes, também demonstraram desempenho superior. Beames et al. (2025), por exemplo, exploraram a integração de dados de fenotipagem digital com conjuntos de dados clínicos e genéticos, demonstrando o potencial dos modelos multimodais na criação de perfis abrangentes de saúde mental [84]. Da mesma forma, Kadirvelu et al. (2025) demonstraram a utilidade de combinar dados ativos e passivos de smartphones para prever riscos de saúde mental em adolescentes, alcançando

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

precisões de até 77% na previsão de ideação suicida [85].

As aplicações práticas destes modelos já se fazem notar em contextos diversos. Melcher et al. (2020), por exemplo, estudaram estudantes universitários e verificaram que variáveis relacionadas com o sono, a atividade física e as interações sociais permitiam prever alterações de humor, stress e ansiedade com elevada fiabilidade [86]. Outros trabalhos destacaram o potencial de prever recaídas em transtornos de humor e ansiedade, reforçando a utilidade da fenotipagem digital na gestão de crises [80].

Contudo, persistem desafios. A heterogeneidade dos dados recolhidos, as limitações metodológicas e as preocupações com privacidade e consentimento informado constituem barreiras relevantes. Martinez-Martin (2021) enfatizou ainda a necessidade de garantir a equidade algorítmica, de forma a que os modelos não reproduzam preconceitos culturais ou demográficos [87].

Para o futuro, a investigação deverá privilegiar o aumento da precisão e escalabilidade dos modelos, recorrendo à integração de dados multimodais e a técnicas emergentes como a *learning federada* e a *explainable AI (XAI)*. Estas abordagens podem não só aumentar a capacidade preditiva dos modelos, mas também responder às exigências de privacidade e interpretabilidade, aspetos fundamentais para a sua adoção em contextos clínicos.

Em suma, os indicadores comportamentais e os modelos preditivos são pilares da fenotipagem digital, ao possibilitarem uma abordagem contínua, objetiva e personalizada da saúde mental. Apesar dos desafios ainda existentes, os avanços tecnológicos e metodológicos apontam para um futuro em que a deteção precoce e a intervenção personalizada serão cada vez mais viáveis.

### **3.6 Padrões de saúde mental e sociabilidade dos jovens**

A fenotipagem digital tem-se mostrado bastante promissora no tratamento de problemas de saúde mental entre os jovens, aproveitando indicadores comportamentais como mobilidade, padrões de sono e interação social. A adolescência e a idade adulta jovem são períodos críticos para a saúde mental, pois são marcados por um desenvolvimento emocional, social e cognitivo significativo. Durante este período, os indivíduos são particularmente vulneráveis a condições como depressão e ansiedade, que podem ter consequências a longo prazo se não forem tratadas. A fenotipagem digital oferece uma abordagem escalável e não invasiva para monitorizar e prever os estados de saúde mental nesta população, utilizando dados recolhidos de *smartphones* e dispositivos vestíveis.

Estudos recentes destacaram a eficácia da fenotipagem digital neste campo. Beames et al. (2024) realizaram uma revisão de 35 estudos com foco em jovens de 12 a 25 anos, examinando o uso de sensores passivos e dados de uso eletrónico para detetar e prever depressão e ansiedade. A revisão constatou que os padrões de mobilidade e de sociabilidade foram os indicadores mais frequentemente explorados, com características como frequência de transição, tempo passado em locais específicos e uniformidade de movimento a revelarem-se particularmente informativas. Os autores também observaram que modelos que combinavam múltiplas características — como mobilidade, sociabilidade e sono — eram mais eficazes

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

do que aqueles que dependiam de uma única característica, enfatizando o valor de uma abordagem multimodal [88].

De forma semelhante, Melcher et al. (2020) realizaram uma revisão de 25 estudos sobre fenotipagem digital em estudantes universitários. A revisão centrou-se na utilização de fluxos de dados baseados em sensores, tais como leituras de localização e acelerómetro, bem como informações sociais, para encontrar comportamentos como sono, exercício físico e interações sociais. Os autores destacaram a estreita ligação entre estas características e os resultados de saúde mental, como humor, ansiedade e *stress*, bem como a viabilidade do uso da fenotipagem digital para personalizar o acompanhamento dos estudantes, especialmente em ambientes de aprendizagem remota onde os serviços tradicionais de saúde mental podem ser menos acessíveis [86].

O uso de *Ecological Momentary Assessments* (EMAs), em conjunto com sensores passivos, também foi explorado como um meio de aumentar a precisão do monitoramento da saúde mental em jovens. Marciano et al. (2023) realizaram uma revisão abrangente de estudos que utilizam EMAs e dados de rastreamento para monitorizar a saúde mental nesta população. Verificou-se que as EMAs, que envolvem autoavaliações em tempo real do humor e do comportamento, complementam os dados passivos ao fornecer um contexto subjetivo. Esta combinação oferece uma compreensão mais holística dos estados de saúde mental e tem demonstrado melhorar a deteção de condições como depressão e ansiedade [89].

Os padrões de interação social têm sido identificados como indicadores críticos da saúde mental dos jovens. Alterações na frequência de comunicação, atividade nas redes sociais e proximidade com outras pessoas — medidas através de *Bluetooth* e registos de chamadas — têm sido associadas ao bem-estar emocional. A redução das interações sociais, por exemplo, está frequentemente correlacionada com solidão e ansiedade, enquanto o aumento do uso das redes sociais tem sido associado a resultados tanto positivos como negativos para a saúde mental, dependendo do contexto [90] [91].

Apesar destes resultados promissores, persistem desafios na aplicação da fenotipagem digital à saúde mental dos jovens. Beames et al. (2024) salientaram a inconsistência na comunicação de variáveis importantes, como idade, género e características do dispositivo, entre os estudos. Identificaram ainda um risco moderado a elevado de viés nas metodologias utilizadas, o que limita a generalização dos resultados e evidencia a necessidade de estruturas de comunicação padronizadas que aumentem a transparência e a replicabilidade. Além disso, as preocupações éticas relacionadas com privacidade, consentimento e propriedade dos dados são particularmente relevantes quando se trabalha com jovens, como enfatizado por Martinez-Martin (2021) [87].

Pesquisas futuras devem concentrar-se em superar estes desafios, desenvolvendo metodologias padronizadas e diretrizes éticas adaptadas às necessidades dos jovens. Há também necessidade de estudos longitudinais de maior dimensão que validem a utilidade da fenotipagem digital em diferentes contextos culturais e socioeconómicos. À medida que esta abordagem continua a evoluir, apresenta um forte potencial para transformar os cuidados de saúde mental dos jovens, permitindo a deteção precoce, intervenções personalizadas e um melhor acesso a sistemas de apoio.

### 3.7 Integração de dados multimodais

A integração de dados multimodais é um avanço crítico na fenotipagem digital, permitindo que diversas fontes de informação sejam combinadas para criar perfis abrangentes de saúde mental. Ao integrar dados de *smartphones*, dispositivos vestíveis, registros eletrônicos de saúde (*EHRs*), informações genéticas e outras fontes, os investigadores conseguem capturar a natureza multifacetada das condições de saúde mental. Embora esta abordagem aumente a precisão preditiva e proporcione uma compreensão mais holística, apresenta também desafios técnicos, éticos e processuais significativos.

#### O papel dos dados multimodais na fenotipagem digital

A integração de dados multimodais combina os pontos fortes de diferentes fluxos de dados para construir uma visão mais completa da saúde mental. Por exemplo, dados recolhidos por *smartphones* — como localização GPS, leituras de acelerómetro e padrões de utilização de aplicações — podem ser complementados por sinais fisiológicos captados por dispositivos vestíveis, como variabilidade da frequência cardíaca e padrões de sono, para monitorizar estados comportamentais e emocionais. Beames et al. (2025) demonstraram a viabilidade de ligar dados de fenotipagem digital a conjuntos de dados genéticos e clínicos no estudo piloto *Mobigene*. O estudo, que envolveu participantes com histórico de depressão, mostrou que a integração de dados multimodais podia não só melhorar as taxas de adesão e envolvimento, mas também fornecer informações clínicas valiosas sobre a saúde mental [84].

De forma complementar, Yuan et al. (2025) aplicaram *Large Language Models* (LLMs) à análise de dados heterogêneos de fenotipagem digital, incluindo sinais comportamentais e fisiológicos. As suas descobertas destacaram o potencial dos LLMs para generalizar em diferentes tarefas e melhorar a previsão de alterações na gravidade da depressão, reforçando a importância de técnicas avançadas de *machine learning* para a integração eficaz de dados multimodais [83].

#### Aplicações da integração de dados multimodais

A integração multimodal já tem sido aplicada em diversos contextos com resultados promissores. Ali et al. (2025), por exemplo, propuseram uma abordagem *trimodal* para análise de dados de fala no contexto da deteção de depressão. O modelo combinava texto, características acústicas e biomarcadores vocais, alcançando maior precisão que os métodos unimodais. A análise longitudinal integrada permitiu ainda acompanhar variações temporais na saúde mental, evidenciando o valor da dimensão temporal na modelação multimodal [92].

Noutro exemplo, Mendes et al. (2021) analisaram conjuntos de dados públicos usados em fenotipagem digital, identificando o potencial dos dados multimodais para modelar comportamentos como sociabilidade, humor e atividade física. Contudo, os autores observaram uma escassez de conjuntos de dados públicos de alta qualidade, o que limita tanto a escalabilidade como a generalização das abordagens multimodais [93].

### Desafios na integração de dados multimodais

Apesar do seu potencial, a integração de dados multimodais enfrenta diversos obstáculos. Entre os principais desafios técnicos estão a heterogeneidade dos dados, o desequilíbrio entre modalidades e a complexidade inerente à fusão de fluxos de informação distintos. Por exemplo, fatores demográficos e contextuais podem afetar a fiabilidade de determinadas modalidades, tornando necessárias estratégias robustas de alinhamento intermodal. Hu et al. (2024) mostraram as vantagens da integração de dados de *electroencephalogram* (EEG) com áudio e expressões faciais para a avaliação de saúde mental, mas também apontaram as dificuldades de sincronização entre estas modalidades [94].

Os desafios éticos e processuais são igualmente relevantes. A recolha e integração de dados altamente sensíveis — como informação genética e registos de saúde eletrónicos — levantam preocupações relacionadas com privacidade, consentimento e propriedade dos dados. Martinez-Martin (2021) destacou a necessidade de regras claras de transparência e responsabilidade, sublinhando que as estruturas regulatórias atuais são insuficientes para lidar com a complexidade e sensibilidade dos dados multimodais [87].

### Direções futuras

O futuro da integração multimodal passa pelo desenvolvimento de protocolos padronizados para recolha, alinhamento e análise de dados. Avanços em áreas como *federated learning* e *explainable AI* oferecem soluções promissoras para melhorar a escalabilidade, a privacidade e a interpretabilidade dos modelos. Além disso, a criação de conjuntos de dados de alta qualidade e abertos à comunidade científica será essencial para validar e generalizar abordagens multimodais em contextos clínicos e populacionais.

Em suma, a integração de dados multimodais representa um passo decisivo na evolução da fenotipagem digital, permitindo a construção de perfis de saúde mental mais precisos e abrangentes. Superar os desafios técnicos e éticos associados aumentará o poder preditivo e a utilidade clínica destas abordagens, consolidando o papel da fenotipagem digital como ferramenta-chave nos cuidados de saúde mental personalizados.

## 3.8 Dispositivos eletrónicos e doenças mentais graves

Os dispositivos vestíveis, como *smartwatches* e monitores de fitness, estão equipados com sensores capazes de medir sinais fisiológicos, como variabilidade da frequência cardíaca, condutância da pele e padrões de sono. Estes sinais têm sido amplamente associados a estados de saúde mental, funcionando como biomarcadores valiosos para o monitoramento de doenças mentais graves (*Severe Mental Illnesses – SMIs*). Lanatà et al. (2015) mostraram que a variabilidade da frequência cardíaca e a atividade eletrodérmica podiam prever estados de humor em pessoas com transtorno bipolar. Estes sinais atuaram como marcadores de alerta precoce para episódios maníacos ou depressivos, permitindo intervenções mais oportunas [95].

De forma semelhante, Valenza et al. (2014) investigaram o potencial de sensores vestíveis

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

para monitorizar a atividade do sistema nervoso autónomo em pacientes com esquizofrenia. Verificaram que alterações na frequência cardíaca e na condutância da pele estavam associadas à gravidade dos sintomas, reforçando o potencial de dispositivos vestíveis para fornecer medidas objetivas de saúde mental [96].

Além dos dispositivos vestíveis, os *smartphones* desempenham um papel igualmente relevante no monitoramento de SMIs através da deteção passiva de padrões de mobilidade, comunicação e uso de aplicações. Wang et al. (2016) desenvolveram o sistema *CrossCheck*, que utiliza sensores de *smartphones* para monitorizar alterações comportamentais em pacientes com esquizofrenia. O estudo mostrou que mobilidade reduzida e padrões de comunicação irregulares estavam associados à exacerbação dos sintomas, demonstrando a viabilidade do uso de *smartphones* como ferramenta de monitorização em tempo real [97].

Investigações recentes exploraram a integração de dados recolhidos por dispositivos vestíveis e *smartphones* para aumentar a precisão preditiva. Hassan et al. (2025), numa revisão sobre o uso de dispositivos de consumo em SMIs, destacaram o potencial de combinar dados fisiológicos e comportamentais. Apesar disso, observaram grande variabilidade nos tipos de dispositivos e metodologias de recolha, o que limita a generalização dos resultados [98].

Estudos também evidenciam o uso destas tecnologias na previsão de recaídas e no acompanhamento de respostas ao tratamento. Barnett et al. (2018) utilizaram fenotipagem digital baseada em *smartphones* para prever recaídas em pacientes com esquizofrenia, identificando mobilidade reduzida e padrões de sono irregulares como sinais de alerta precoce [99]. Da mesma forma, Faurholt-Jepsen et al. (2019) mostraram que níveis de atividade medidos por acelerómetros estavam correlacionados com estados de humor em pacientes com transtorno bipolar, fornecendo indicadores objetivos para avaliar a eficácia do tratamento e apoiar intervenções personalizadas [100].

Apesar dos avanços, a utilização de dispositivos vestíveis e *smartphones* enfrenta barreiras importantes. Entre os desafios técnicos destacam-se a heterogeneidade dos dados, a falta de normalização e os problemas de interoperabilidade entre dispositivos. Sheikh et al. (2021) enfatizaram a necessidade de protocolos padronizados para assegurar consistência e fiabilidade na recolha e análise de dados [101].

As barreiras clínicas também são significativas. Baixas taxas de adesão ao uso de dispositivos, dificuldades de pacientes com SMIs em manter rotinas consistentes e a complexidade da interpretação de dados passivos representam obstáculos para a implementação em larga escala. A análise destes dados exige ferramentas avançadas e equipas multidisciplinares, que muitas vezes não estão disponíveis em contextos clínicos [102].

Questões éticas assumem ainda maior relevância neste contexto. A recolha de dados sensíveis levanta preocupações relacionadas com privacidade, consentimento informado e segurança. Martinez-Martin (2021) salientou os riscos de uso indevido dos dados e defendeu a necessidade de estruturas éticas robustas, particularmente quando se trata de populações vulneráveis, como pacientes com SMIs [87].

### 3.9 Escalabilidade e aplicações ao nível da população

A escalabilidade da fenotipagem digital abriu novas possibilidades para monitorizar a saúde mental em larga escala, permitindo a recolha e análise de dados comportamentais em populações amplas e diversificadas. Ao explorar a ubiquidade de *smartphones* e dispositivos vestíveis, esta abordagem oferece uma solução económica e escalável para compreender tendências populacionais em saúde mental. O seu impacto é particularmente relevante em contextos de saúde pública, desde a identificação de grupos de risco até à monitorização em períodos de crise e ao suporte à formulação de políticas.

#### Estudos em grande escala

Diversos estudos demonstraram a viabilidade e utilidade da fenotipagem digital em contextos populacionais. Zhang et al. (2024) realizaram um estudo com mais de 10 000 participantes no Reino Unido, integrando dados de dispositivos vestíveis Fitbit com questionários de saúde mental auto-relatados. Os resultados mostraram correlações significativas entre a gravidade da depressão e da ansiedade e indicadores comportamentais como atividade física, padrões de sono e variabilidade da frequência cardíaca. Este estudo demonstrou o potencial da fenotipagem digital como ferramenta de rastreio em larga escala [81].

De forma complementar, Torous et al. (2025) destacaram o papel das tecnologias digitais de saúde mental na escalabilidade, especialmente ao integrar a fenotipagem digital com intervenções adaptativas *just-in-time* (JITAI). Estas intervenções utilizam dados em tempo real para fornecer suporte personalizado e sensível ao contexto, revelando-se eficazes no reforço da saúde mental durante crises como a pandemia, quando os serviços clínicos tradicionais foram interrompidos [103].

#### Aplicações na saúde pública

A fenotipagem digital tem sido aplicada para identificar tendências populacionais e apoiar estratégias de saúde pública. Durante a pandemia, por exemplo, foi utilizada para monitorizar alterações na mobilidade, nos padrões de sono e nas interações sociais, fatores associados ao aumento de ansiedade, depressão e stress. Huckins et al. (2020) analisaram dados de estudantes universitários recolhidos por *smartphones* e identificaram perturbações significativas nestes comportamentos, que se correlacionaram com um agravamento da saúde mental auto-relatada [100].

Além disso, esta abordagem tem demonstrado utilidade na identificação de grupos de risco. Perlmutter et al. (2024), por exemplo, aplicaram técnicas de fenotipagem digital para identificar marcadores comportamentais de ideação suicida em adolescentes. Através da análise de padrões de mobilidade e comunicação, conseguiram prever ideação suicida com elevada precisão, mostrando o potencial destas ferramentas em intervenções preventivas e direcionadas [90].

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

### Desafios na escalabilidade

Apesar do potencial, a aplicação em larga escala enfrenta vários obstáculos. A qualidade dos dados é um dos principais desafios: estudos populacionais registam frequentemente elevadas taxas de dados ausentes ou incompletos, devido a falhas técnicas nos dispositivos ou inconsistência no uso por parte dos participantes. Currey et al. (2023) sublinharam a importância de estratégias de monitorização em tempo real e períodos de adaptação para melhorar a qualidade dos dados em estudos de grande escala [104].

Outro desafio é a representatividade das amostras. Grande parte dos estudos recorre a amostras de conveniência (como estudantes universitários), que não refletem a diversidade da população em geral. Mokadam et al. (2025) chamaram a atenção para as divisões digitais, especialmente em comunidades marginalizadas com acesso limitado a *smartphones* ou dispositivos vestíveis, e defenderam o desenvolvimento de dispositivos de baixo custo e estratégias de recrutamento inclusivas [105].

Questões éticas e de privacidade são igualmente críticas. A recolha de dados comportamentais em larga escala levanta preocupações sobre consentimento, propriedade dos dados e risco de uso indevido. Martinez-Martin (2021) propôs uma estrutura ética para a administração destes dados, assente na transparência, responsabilidade e proteção dos utilizadores, sublinhando a urgência de regulamentação robusta para estudos de grande escala [87].

### Direções futuras

Pesquisas futuras devem concentrar-se no desenvolvimento de protocolos padronizados para recolha e análise de dados em grande escala. Avanços em aprendizagem automática, como *federated learning* e técnicas de privacidade diferencial, oferecem soluções promissoras para aumentar a segurança e a escalabilidade. Além disso, a criação de conjuntos de dados públicos e de alta qualidade será crucial para validar resultados e generalizar descobertas em diferentes contextos culturais e socioeconómicos.

Em conclusão, a escalabilidade da fenotipagem digital tem o potencial de transformar a monitorização da saúde mental a nível populacional, fornecendo informações valiosas sobre tendências comportamentais e orientando políticas públicas. Superar os desafios atuais será essencial para que esta abordagem se consolide como uma ferramenta central na promoção da saúde mental em larga escala.

## 3.10 Avanços nas tecnologias digitais de saúde mental

Os avanços recentes nas tecnologias digitais de saúde mental têm expandido significativamente as capacidades da fenotipagem digital, permitindo intervenções mais personalizadas, adaptativas e escaláveis. Estas tecnologias integram dados comportamentais recolhidos a partir de *smartphones*, dispositivos vestíveis e outras fontes digitais com algoritmos de *machine learning*, viabilizando o monitoramento, a previsão e a melhoria dos resultados de saúde mental. A incorporação de inovações como intervenções adaptativas *just-in-time* (JITs), modelos de *inteligência artificial* (IA) e técnicas de fusão de dados multimodais está

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

a redefinir o panorama da prestação de cuidados em saúde mental.

### **Integração da fenotipagem digital com intervenções adaptativas**

Um dos progressos mais promissores consiste na integração da fenotipagem digital com JITAIs, que utilizam dados em tempo real para fornecer apoio sensível ao contexto e personalizado às necessidades dos utilizadores. Estas intervenções são particularmente adequadas para lidar com a natureza dinâmica das condições de saúde mental, adaptando-se a flutuações comportamentais e emocionais. Torous et al. (2025) analisaram o papel das JITAIs na saúde mental digital, destacando a sua capacidade de fornecer intervenções oportunas baseadas em sinais comportamentais e ambientais. O estudo sublinhou a importância de envolver os utilizadores finais no co-desenho destas ferramentas, de forma a aumentar a eficácia e a aceitação [103].

Da mesma forma, Deshwal et al. (2025) investigaram a aplicação de JITAIs no tratamento de ansiedade e depressão, utilizando dados como tempo de ecrã, mobilidade e padrões de comunicação derivados de *smartphones*. O sistema conseguiu detetar sinais precoces de deterioração da saúde mental e fornecer intervenções personalizadas, incluindo exercícios de *mindfulness* e lembretes para procurar apoio social. Os resultados mostraram uma melhoria significativa no envolvimento dos utilizadores e nos resultados de saúde mental [106].

### **Avanços na Inteligência Artificial e na Aprendizagem Automática**

A IA desempenha um papel central nos avanços das tecnologias digitais de saúde mental, sobretudo através do desenvolvimento de modelos preditivos mais robustos e de intervenções personalizadas. Modelos de linguagem de grande porte (*Large Language Models* – LLMs), como os avaliados por Yuan et al. (2025), demonstraram elevada eficácia na integração de dados comportamentais heterogéneos e superaram abordagens tradicionais de *machine learning*, oferecendo previsões mais precisas sobre a gravidade da depressão e outras condições. Estes modelos também demonstraram capacidade de analisar dados não estruturados, como mensagens de texto e publicações em redes sociais, fornecendo insights profundos sobre estados emocionais e cognitivos [83].

Outro desenvolvimento relevante está associado ao uso de *explainable AI* (XAI), que aumenta a interpretabilidade dos modelos. Kasaudhan et al. (2025) demonstraram a utilidade da IA explicável na identificação de características comportamentais-chave associadas a depressão e ansiedade. Ao tornar as decisões dos modelos mais transparentes e compreensíveis, a XAI aumenta a confiança de clínicos e pacientes, facilitando a integração destas tecnologias em ambientes de prática clínica [107].

### **Fusão e personalização de dados multimodais**

A fusão de dados multimodais constitui outro avanço significativo. Ao integrar fluxos de dados provenientes de diferentes fontes — como sensores de *smartphones*, dispositivos vestíveis e registos clínicos —, é possível criar perfis detalhados de saúde mental que refletem

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

melhor as experiências individuais. Beames et al. (2025) investigaram a combinação de dados de fenotipagem digital com dados genéticos e clínicos, demonstrando o potencial das abordagens multimodais para aumentar a precisão das previsões e apoiar intervenções altamente personalizadas [84].

Paralelamente, os avanços em *reinforcement learning* têm permitido o desenvolvimento de sistemas adaptativos que se ajustam com base nas interações com os utilizadores. Halkiopoulos et al. (2025) aplicaram esta técnica a uma plataforma de saúde mental digital, permitindo que o sistema aprendesse preferências individuais e otimizasse recomendações ao longo do tempo. Este tipo de adaptação dinâmica mostrou melhorias significativas na adesão dos utilizadores e na eficácia do tratamento, reforçando o potencial das tecnologias digitais para fornecer cuidados responsivos e personalizados [108].

### Direções futuras

O futuro das tecnologias digitais de saúde mental depende de avanços técnicos aliados a um tratamento responsável das dimensões éticas e socioculturais. Entre as prioridades destacam-se: o desenvolvimento de protocolos padronizados para recolha e análise de dados; a incorporação de abordagens de IA explicável para aumentar a transparência; e a criação de soluções inclusivas que abordem desigualdades no acesso e na usabilidade. O co-desenho de ferramentas digitais com diversas partes interessadas — incluindo pacientes, profissionais de saúde e especialistas em ética — será essencial para garantir eficácia, equidade e aceitação. Em suma, os avanços nas tecnologias digitais de saúde mental fortaleceram substancialmente a fenotipagem digital, permitindo intervenções adaptativas, personalizadas e escaláveis. Ao combinar inovação tecnológica com responsabilidade ética, estas ferramentas têm o potencial de transformar os cuidados em saúde mental, melhorando resultados tanto ao nível individual como populacional.

### 3.11 Conclusões

Neste capítulo foram discutidos os principais trabalhos relacionados com o uso de técnicas de *NLP*, *embeddings* e modelos preditivos aplicados ao estudo da saúde mental. Verificou-se que, embora os avanços recentes em *transformers* e LLMs tenham melhorado significativamente a capacidade de analisar textos e identificar padrões associados a transtornos como depressão e ansiedade, permanecem limitações consideráveis. Entre estas destacam-se a escassez de *datasets* públicos adequados, os desafios éticos relacionados com privacidade e representatividade, e as dificuldades em integrar múltiplas modalidades de dados.

Adicionalmente, observou-se que a investigação em fenotipagem digital tem procurado ligar indicadores comportamentais a estados mentais, explorando sinais tanto fisiológicos como linguísticos. Apesar do seu potencial, muitas abordagens permanecem experimentais e carecem de validação clínica robusta.

Assim, esta revisão demonstra a necessidade de novas metodologias que combinem robustez técnica, acessibilidade de dados e foco em dimensões comportamentais. É nesse espaço

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

de inovação que se enquadra o trabalho desenvolvido nesta dissertação, cuja metodologia é detalhada no capítulo seguinte.

## Capítulo 4

### Abordagem / Metodologia

#### 4.1 Introdução

Este capítulo descreve a metodologia seguida para o desenvolvimento da investigação, detalhando as diferentes etapas que permitiram a construção e validação de um modelo *multitask* aplicado à análise de saúde mental em dados textuais. A abordagem adotada teve como ponto de partida a reformulação do objetivo inicial, passando de uma perspectiva multimodal para uma abordagem exclusivamente textual, devido à indisponibilidade de conjuntos de dados públicos multimodais adequados.

São apresentadas as fases de recolha de dados, o processo de *fine-tuning* de modelos de classificação supervisionada no Vertex AI, a iteração cruzada para a criação de um conjunto de dados unificado de transtornos, a associação de comportamentos por meio de técnicas de similaridade semântica e, finalmente, o treino de um modelo *multitask*. A aplicação desta metodologia ao conjunto de dados SNCrawler permitiu a criação de um novo recurso anotado, alinhado com os objetivos do trabalho.

A estrutura do capítulo reflete a sequência lógica do processo desenvolvido, evidenciando as decisões metodológicas, os desafios enfrentados e as soluções adotadas, preparando o terreno para a análise dos resultados que será apresentada no capítulo seguinte.

#### 4.2 Objetivo inicial e reformulação da abordagem

O objetivo inicial deste estudo era desenvolver um modelo multimodal para analisar comportamentos associados a condições de saúde mental, particularmente depressão e ansiedade. Essa abordagem permitiria integrar diferentes tipos de dados, como texto, imagens e áudio, a fim de obter uma representação mais rica e abrangente dos comportamentos relacionados à saúde mental. Essa abordagem segue a tendência atual em inteligência artificial, onde modelos multimodais que combinam informações de diferentes naturezas têm mostrado resultados promissores na melhoria das capacidades de classificação e interpretação.

No entanto, durante a fase de recolha de dados e análise preliminar, verificou-se que não existiam conjuntos de dados multimodais disponíveis publicamente adequados ao âmbito do projeto. Os principais obstáculos encontrados foram:

- **Escassez de conjuntos de dados multimodais públicos:** a maioria dos conjuntos de dados existentes estava restrita a contextos clínicos e não estava disponível para pesquisa aberta devido a questões de confidencialidade e ética.
- **Formatação e compatibilidade:** os poucos conjuntos de dados acessíveis tinham

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

formatos inadequados ou inconsistentes, incompatíveis com os requisitos de integração no pipeline metodológico.

- **Limitações de tamanho:** mesmo quando disponíveis, muitos conjuntos de dados continham um número pequeno de instâncias, tornando impossível a aplicação de técnicas de aprendizagem profunda em escala.

À luz dessas limitações, foi necessário rever a metodologia. Em vez de adotar uma abordagem multimodal, decidimos restringir o estudo a dados textuais por três razões principais:

1. A existência de um maior número de conjuntos de dados públicos contendo textos relacionados à saúde mental.
2. A relevância do texto como meio privilegiado de expressar estados emocionais, particularmente no contexto das redes sociais.
3. A possibilidade de construir um pipeline consistente que ainda produziria resultados relevantes dentro do escopo do estudo proposto.

Durante o processo de reformulação, o conjunto de dados SNCrawler — composto por publicações nas redes sociais recolhidas e armazenadas pelo grupo de investigação ao longo dos anos — assumiu particular relevância. Apesar da sua dimensão e potencial, este conjunto de dados tinha uma limitação crítica: faltavam-lhe etiquetas relacionadas com depressão, ansiedade e comportamentos associados. Por conseguinte, um aspeto fundamental deste trabalho foi o desenvolvimento de um pipeline de anotação automática capaz de transformar o conjunto de dados brutos num recurso estruturado, anotado e útil para a investigação em saúde mental.

Consequentemente, o foco da investigação mudou da análise multimodal para a construção de um conjunto de dados textuais envolvendo um percurso metodológico sintetizado no fluxograma da Figura 4.1, que ilustra as principais etapas desde a recolha de dados até à criação do conjunto de dados final anotado.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

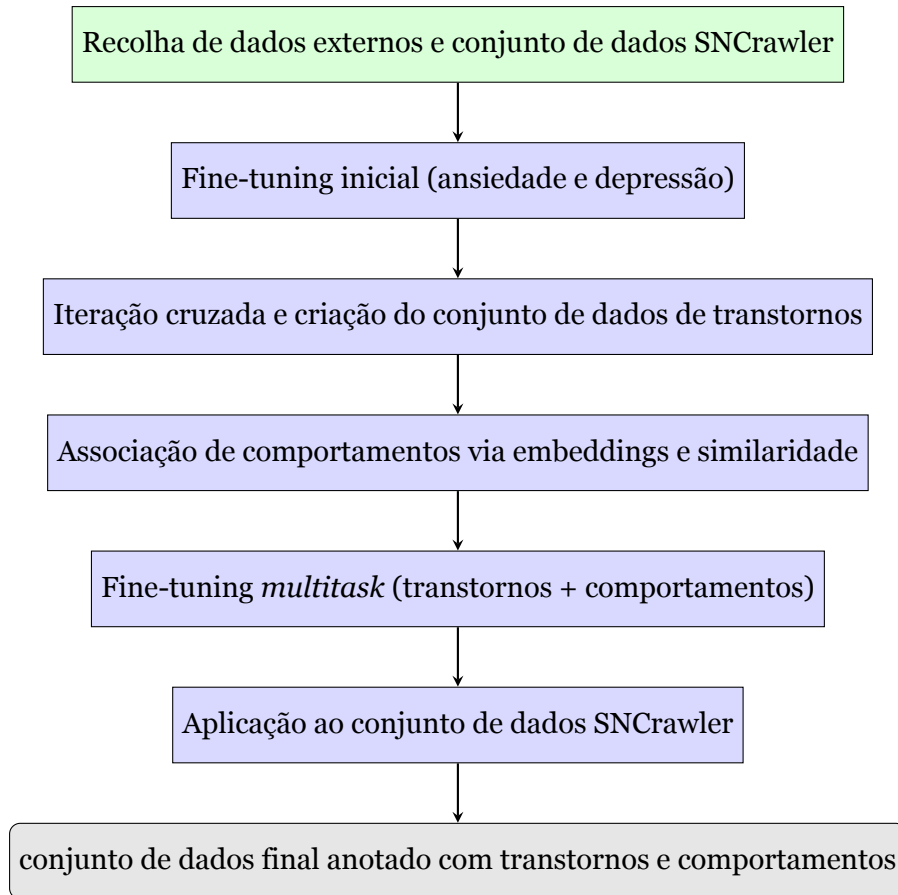


Figura 4.1: Fluxograma da metodologia adotada neste estudo.

### 4.3 Recolha de dados

O primeiro passo metodológico foi recolher dados para permitir o treino e a validação de modelos de classificação de ansiedade e depressão. Como mencionado anteriormente, não foi possível trabalhar com dados multimodais, pelo que a recolha se centrou exclusivamente em dados textuais, dados esses em inglês, pois sendo a língua mais falada no mundo e, consequentemente, nas redes sociais a disponibilidade de dados é muito maior.

#### Critérios de seleção

Os principais critérios para a seleção dos conjuntos de dados foram os seguintes:

1. **Relevância temática:** o conjunto de dados deve conter textos produzidos por pessoas em contextos direta ou indiretamente relacionados à saúde mental.
2. **Anotação prévia ou possibilidade de anotação:** idealmente, os conjuntos de dados devem conter etiquetas associadas a transtornos (por exemplo, ansiedade ou depressão). Quando isso não era possível, era necessário recorrer a processos de anotação automática.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

3. **Origem diversificada:** procurámos recolher dados de diferentes contextos (por exemplo, redes sociais, consultas e fóruns online) para aumentar a diversidade linguística e semântica do material.

### Fontes de dados

Foram utilizados dois tipos de conjuntos de dados:

- **Conjuntos de dados externos:** recolhidos de fontes públicas e contendo texto que foi anotado quanto à presença de depressão e/ou ansiedade. Esses conjuntos de dados foram utilizados para treinar os modelos de classificação iniciais.
- **conjunto de dados SNCrawler:** composto por publicações nas redes sociais recolhidas e armazenadas pelo grupo de investigação HULTIG ao longo dos anos. Apesar da sua dimensão e potencial para análise, este conjunto de dados não foi anotado com informações sobre saúde mental, o que constituiu um dos principais desafios do projeto.

### Pré-processamento

Não foi realizado nenhum pré-processamento extensivo dos dados recolhidos. Presumiu-se que os conjuntos de dados externos estavam prontos para uso. Da mesma forma, os dados do SNCrawler foram usados na sua forma original. Embora essa abordagem possa introduzir ruído nos textos, o objetivo principal era fazer uso total do conteúdo recolhido, mantendo a diversidade expressiva típica das publicações informais nas redes sociais.

### Dimensão dos dados

Depois de integrados os vários conjuntos de dados externos, foi possível criar um conjunto de dados unificado com 30 305 instâncias textuais. Este conjunto de dados serviu de base para as etapas subsequentes de anotação e enriquecimento com comportamentos.

## 4.4 Fine-tuning inicial no Vertex AI

Uma das etapas fundamentais do projeto envolveu o desenvolvimento de dois modelos de classificação, um para ansiedade e outro para depressão. O objetivo era avaliar se os exemplos associados a cada transtorno também apresentavam sinais do outro, separadamente. Essa abordagem permitiu detetar sobreposições entre os transtornos e, subsequentemente, construir um conjunto de dados conjunto e coerente, que serviria de base para as fases subsequentes da metodologia.

O *fine-tuning* foi realizado na plataforma Vertex AI no Google Cloud, usando o modelo base Gemini Flash 2.5. O Vertex AI foi escolhido porque oferece um ambiente de treinamento supervisionado robusto, no qual grande parte da complexidade técnica é abstraída, permitindo que os investigadores se concentrem na preparação de dados e na definição de objetivos experimentais.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

Entre as funcionalidades mais relevantes do Vertex AI destacam-se:

- A gestão automática do pipeline de treino elimina a necessidade de configuração manual do processo de treino, reduzindo assim a margem de erro.
- A monitorização em tempo real das métricas de avaliação permite acompanhar a evolução do modelo ao longo do tempo.
- A interrupção precoce evita o *overfitting*, interrompendo o treino quando não há melhorias significativas no desempenho.

Para permitir o treino, foram construídos dois conjuntos de dados independentes, um relacionado com a ansiedade e outro com a depressão. Cada conjunto de dados consistia em exemplos textuais anotados para indicar a presença ou ausência do transtorno relevante. Seguindo as melhores práticas, **80% dos dados foram utilizados para treino e 20% para validação**, garantindo que uma proporção substancial dos dados fosse utilizada para ajustar os parâmetros do modelo, mantendo um conjunto independente para avaliar o desempenho.

O treino foi realizado no modelo base Gemini Flash 2.5, que foi configurado para um máximo de dez épocas. Este número foi escolhido para alcançar um equilíbrio entre permitir que o modelo aprendesse padrões linguísticos relevantes e evitar sessões de treino excessivamente longas que poderiam resultar em *overfitting*. O Vertex AI aplicou automaticamente o mecanismo de interrupção antecipada, interrompendo o treino na oitava época para ambos os modelos quando foi detetada a estabilização das métricas de validação. A precisão, uma métrica apropriada para tarefas de classificação binária, foi a principal métrica usada para avaliar o desempenho, indicando a eficácia com que o modelo poderia prever a presença ou ausência do distúrbio nos textos.

Os resultados obtidos foram altamente satisfatórios, com os modelos de ansiedade e depressão alcançando 98% de precisão nos seus respetivos conjuntos de validação. Isso confirmou que os modelos podiam aprender padrões linguísticos distintos com um alto grau de precisão e que essa abordagem tinha potencial para apoiar as etapas subsequentes do projeto.

Este ajuste inicial foi fundamental para o progresso do trabalho, pois permitiu-nos criar dois modelos especializados, cada um focado num transtorno específico. Isso garantiu uma separação inicial entre ansiedade e depressão. Essa separação foi crucial para testar posteriormente se os exemplos inicialmente classificados como pertencentes a uma categoria também poderiam apresentar características da outra. Além disso, esta etapa forneceu uma base sólida para a **criação de um conjunto de dados unificado** de transtornos, composto por etiquetas que representam várias combinações das duas condições. Os resultados também validaram a viabilidade de uma abordagem exclusivamente textual, demonstrando que modelos robustos e confiáveis poderiam ser construídos sem recorrer a dados multimodais.

## 4.5 Iteração cruzada e construção do conjunto de dados de transtornos

Após obter os dois modelos especializados — um treinado para ansiedade e outro para depressão —, foi realizada uma fase de iteração cruzada. Essa fase teve como objetivo garantir uma maior consistência nas etiquetas e, acima de tudo, evitar situações em que instâncias anotadas apenas com um dos transtornos pudessem, na realidade, apresentar sinais do outro.

O procedimento envolveu a aplicação do modelo de ansiedade ao conjunto de dados de depressão e do modelo de depressão ao conjunto de dados de ansiedade em paralelo. Isso permitiu-nos verificar se os exemplos inicialmente classificados em apenas uma categoria também continham indicadores que justificassem a atribuição do outro transtorno. Esse processo foi realizado caso a caso, com cada exemplo textual sendo passado pelos dois modelos.

Para implementar esta etapa, foi desenvolvido um script Python para interagir diretamente com os *endpoints* do Vertex AI publicados para os dois modelos. Cada texto foi analisado de acordo com instruções específicas: no caso da ansiedade, o sistema deveria responder com “anxiety” ou “no\_anxiety”; no caso da depressão, com “depression” ou “no\_depression”. Essas respostas foram convertidas em valores booleanos e posteriormente integradas no conjunto de dados final. A biblioteca *tqdm* foi usada para otimizar o processo, permitindo acompanhar em tempo real o progresso da previsão em grandes volumes de dados.

Em seguida, foi construído um novo conjunto de dados unificado com base nas previsões obtidas, com cada entrada associada a uma etiqueta numérica representando a sua condição:

- 0 – nenhuma disorder identificada;
- 1 – apenas ansiedade;
- 2 – apenas depressão;
- 3 – coexistência de ansiedade e depressão.

Este esquema de classificação permitiu-nos resumir as informações de forma clara e sistemática, traduzindo as combinações possíveis entre as duas condições. O resultado foi um conjunto de dados composto por 30 305 instâncias, que reuniu exemplos recolhidos de várias fontes externas e reorganizados de acordo com uma lógica coerente e unificada.

Esta fase de iteração cruzada foi crucial, pois garantiu que o conjunto de dados final refletisse tanto as anotações originais dos dados recolhidos quanto as inferências feitas pelos modelos ajustados. Cada instância passou, portanto, por um processo de validação adicional, aumentando assim a fiabilidade e a robustez do conjunto de dados unificado.

Além disso, a criação desse conjunto de dados conjunto foi um passo decisivo para a metodologia, pois permitiu a integração posterior da dimensão comportamental e a aplicação de técnicas de similaridade semântica. Em termos práticos, essa fase marcou uma transição: deixamos de trabalhar com dados fragmentados em categorias isoladas (como ansiedade ou

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

depressão) e passamos a ter uma base consolidada capaz de representar a complexidade das condições psicológicas em estudo.

### 4.6 Associação de comportamentos

Desde o início, este trabalho tem-se centrado nos comportamentos associados à ansiedade e à depressão. Em vez de se limitar a identificar a presença ou ausência destes transtornos, o objetivo tem sido compreender como é que padrões comportamentais específicos, frequentemente descritos na literatura clínica como manifestações destes transtornos, podem ser discernidos a partir de textos.

Esses comportamentos são de natureza física ou psicológica (por exemplo, falar pouco, falar rapidamente, falta de expressividade, sentir-se triste ou agitado), mas foram adaptados aqui para o domínio textual. A lógica era que as pessoas podem expressar o seu estado em palavras; por exemplo, frases como "estou a falar pouco", "não tenho reação" ou "sinto-me triste" podem ser entendidas como representações textuais de comportamentos observáveis. Esta adaptação tornou assim possível aproximar o fenómeno comportamental da análise de texto.

#### Construção e expansão do vocabulário

A primeira fase deste processo envolveu a definição de um vocabulário inicial de comportamentos. Este vocabulário era composto por termos e expressões linguísticas relacionadas com fenómenos comportamentais frequentemente observados em casos de ansiedade e depressão. Exemplos incluem redução da velocidade da fala, pausas longas, redução da variação de tom e inquietação física, bem como uso excessivo de dispositivos digitais.

Este vocabulário foi então expandido usando o modelo Ollama (especificamente Dolphin-Mistral) através do *script* em *python*. O modelo foi instruído a gerar frases naturais, curtas e informais em inglês que representassem realisticamente cada comportamento. Assim, para cada termo base, foram produzidas várias expressões alternativas, resultando num vocabulário expandido que refletia uma maior diversidade linguística. Esta expansão foi guardada num ficheiro pickle e foi posteriormente utilizada para a associação de comportamentos.

#### Associação por similaridade semântica

Na segunda fase, o *script* em *python* baseado no modelo SentenceTransformer (Paraphrase-MiniLM-L12-V2), foi utilizado para calcular as representações vetoriais, ou *embeddings*, de ambos os textos no conjunto de dados unificado de transtornos e nos comportamentos expandidos. Para cada texto, a similaridade cosseno foi calculada para cada comportamento e o texto com a pontuação de proximidade semântica mais alta foi atribuído.

A similaridade cosseno foi escolhida em detrimento de métricas alternativas, como a distância euclidiana ou a medida de Jaccard, após uma comparação. A similaridade cosseno revelou-se mais adequada para este tipo de tarefa, uma vez que mede a proximidade direcional entre vetores, ignorando a sua magnitude absoluta. Esta característica é especialmente

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

útil em embeddings de linguagem, onde o fator mais importante é a orientação semântica do vetor no espaço de representação.

Um limiar de similaridade de 0.5 foi definido neste processo. Embora este valor seja relativamente baixo, foi adotado como uma decisão pragmática que permitiu que o projeto continuasse. Na verdade, um limiar de 0,5 pode ser considerado semelhante a uma situação de "lançamento de moeda", na medida em que reflete um nível de confiança muito baixo. Idealmente, seria definido um limiar mais exigente, por exemplo, 0.75, para aumentar a precisão da associação de comportamentos e reduzir o risco de atribuições espúrias. No entanto, devido ao tamanho e à natureza dos dados disponíveis, a utilização de um limiar mais elevado resultaria na perda de uma proporção significativa de instâncias, como se pode observar na Figura 4.2, o que comprometeria a construção do conjunto de dados final e tornaria inviáveis as fases subsequentes. A Tabela 4.1 resume os valores numéricos correspondentes.

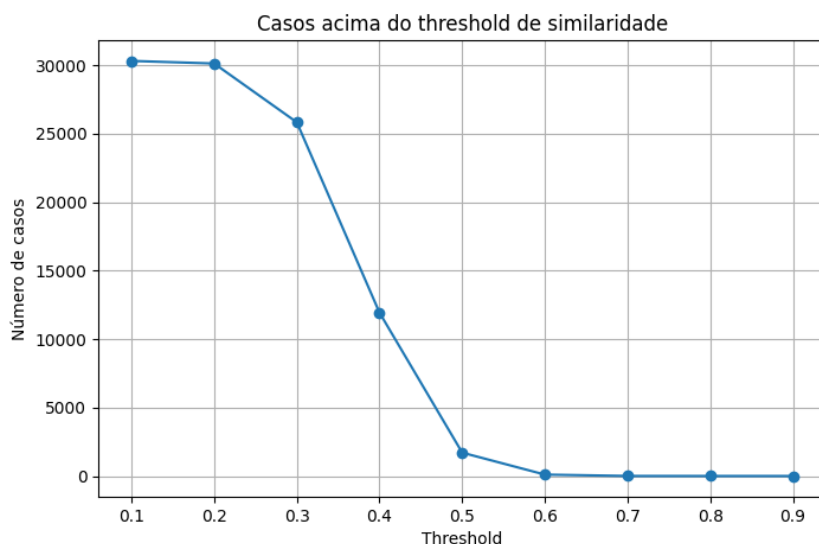


Figura 4.2: Número de casos acima do *threshold* de similaridade.

Threshold	Número de casos
0.1	30 306
0.2	30 117
0.3	25 826
0.4	11 926
0.5	1 705
0.6	103
0.7	4
0.8	1
0.9	0

Tabela 4.1: Número de casos retidos acima de cada *threshold* de similaridade.

Assim, a escolha de 0.5 deve ser entendida como um compromisso metodológico; ela garantiu a viabilidade e a obtenção de resultados. Esta escolha será revisitada em capítulos posteriores, tanto na discussão das limitações da abordagem (4.8) como na proposta de trabalhos futuros (4.9). Aqui, será considerada a possibilidade de aumentar o limiar para  $\geq 0,75$

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

ou adotar estratégias alternativas para melhorar a fiabilidade.

Cada instância do conjunto de dados continha agora o texto original, a etiqueta do transtorno (0 – nenhuma; 1 – ansiedade; 2 – depressão; 3 – ambas) e a etiqueta comportamental adicional.

A introdução desta dimensão comportamental foi um passo essencial para alcançar o objetivo principal do projeto. Embora a perspectiva multimodal tivesse de ser abandonada, a associação de comportamentos específicos aos textos permitiu uma análise mais rica do fenómeno em estudo, indo além da simples distinção binária entre a presença ou ausência de ansiedade e depressão.

Assim, o conjunto de dados final obtido nesta fase serviu de base para a construção de um modelo *multitask* capaz de classificar conjuntamente transtornos e comportamentos, que seria posteriormente aplicado ao conjunto de dados SNCrawler.

### 4.7 Fine-tuning *multitask* e aplicação ao conjunto de dados SNCrawler

Uma vez concluída a fase de associação de comportamentos, obteve-se um conjunto de dados enriquecido contendo três dimensões principais: (i) o texto original; (ii) a etiqueta do transtorno (0 – nenhum; 1 – ansiedade; 2 – depressão; 3 – ambos); e (iii) a etiqueta do comportamento. Esse conjunto de dados serviu de base para a fase seguinte, que envolveu o treino de um modelo *multitask* capaz de aprender ambas as tarefas de classificação simultaneamente.

#### Estrutura do conjunto de dados final

O conjunto de dados consolidado reuniu exemplos recolhidos de fontes externas (os conjuntos de dados de ansiedade e depressão usados para o ajuste inicial) e etiquetas de comportamento atribuídas por meio de similaridade semântica. No entanto, devido ao limiar de similaridade estabelecido de 0,5, apenas 1.705 instâncias atingiram esse valor mínimo e puderam ser incluídas no conjunto de dados *multitask*.

Assim, do universo inicial de mais de 30.000 instâncias, apenas uma proporção relativamente pequena pôde ser usada nesta fase. Isso reflete tanto a rigidez do processo de filtragem quanto a escassez de dados textuais explícitos que expressam comportamentos concretos. Apesar dessa limitação, o conjunto final de 1.705 exemplos foi suficientemente representativo para permitir o treinamento supervisionado e serviu como prova de conceito para a metodologia.

#### Fine-tuning *multitask* no Vertex AI

O *fine-tuning* foi realizado na plataforma Vertex AI usando o modelo Gemini-flash-2.0 do Google Cloud. Este modelo foi selecionado devido ao seu suporte direto para tarefas de classificação supervisionada e integração nativa de recursos essenciais para este processo, in-

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

cluindo otimização automática de hiperparâmetros, monitoramento contínuo de métricas de treino e interrupção antecipada para evitar *overfitting*.

Além destas vantagens técnicas, o Gemini-flash-2.0 é atualmente um dos modelos mais populares e competitivos neste campo. O seu desempenho em tarefas de compreensão e classificação textual é comparável ao de alternativas de referência, como o ChatGPT. Outro ponto essencial para a escolha foi os custos ligados a estas tarefas.

A lógica adotada foi a de um modelo *multitask*, em que o mesmo input textual servia de base para duas predições distintas:

1. A etiqueta de disorder, com quatro classes possíveis (0–3);
2. A etiqueta de comportamento.

O conjunto de dados foi dividido de acordo com a estratégia padrão, com 80% das instâncias alocadas para treino e 20% para validação. Durante o treino, o número máximo de épocas foi definido para 10, mas o processo foi interrompido por uma paragem antecipada na oitava época em ambos os casos, evitando assim o *overfitting*. Apesar do volume limitado de dados, foram alcançadas métricas consistentes, com precisões acima de 95% registadas em ambos os domínios de classificação (transtornos e comportamentos).

### Aplicação ao conjunto de dados SNCrawler

O modelo multitask foi treinado e, em seguida, aplicado ao conjunto de dados SNCrawler, que foi desenvolvido ao longo de vários anos e consiste em textos de redes sociais. Embora esse conjunto de dados já tivesse sido parcialmente anotado para outras dimensões linguísticas e sociais, essas anotações não foram utilizadas no presente estudo. Para este estudo, apenas o conteúdo textual foi considerado como ponto de partida.

O objetivo consistiu em transformar esse material em bruto num novo conjunto de dados derivado, orientado para a análise de saúde mental. Assim, cada instância textual do SNCrawler foi processada pelo modelo multitask, que atribuiu simultaneamente as duas etiquetas relativas a *disorder* e ao comportamento.

Este processo resultou num conjunto de dados, no qual os textos do SNCrawler foram caracterizados em termos de condições psicológicas e comportamentos associados. Na prática, porém, o SNCrawler foi utilizado apenas como fonte de texto, fornecendo a base para uma nova camada de anotações.

Este novo conjunto de dados representa o culminar de todo o processo metodológico. A partir de dados textuais que não estavam anotados para ansiedade ou depressão, foi gerado um recurso com etiquetas clínicas e comportamentais. Esta transformação é significativa porque conjuntos de dados públicos que combinam textos de redes sociais, transtornos e comportamentos são extremamente raros.

Portanto, aplicá-lo ao SNCrawler não se tratava apenas de testar o modelo desenvolvido, mas também de criar valor científico. O resultado é um recurso que pode servir de base para novas análises no campo emergente da fenotipagem digital.

### 4.8 Limitações da abordagem

Apesar dos resultados encorajadores, a metodologia desenvolvida apresenta uma série de limitações que devem ser reconhecidas. Essas limitações dizem respeito não apenas a aspectos técnicos, mas também a restrições metodológicas e de dados.

Uma das questões mais relevantes diz respeito à utilização de um limiar de 0,5 na associação de comportamentos a textos. Este nível aproxima-se de uma escolha aleatória, comparável a "atirar uma moeda ao ar", o que significa que algumas das associações podem não refletir suficientemente a relação entre o texto e o comportamento. A opção inicial era adotar um limiar mínimo de 0,75, o que teria garantido uma maior fiabilidade; no entanto, esta configuração teria reduzido drasticamente o número de instâncias utilizáveis. Portanto, foi necessário um compromisso pragmático, mantendo o limiar em 0,5 para permitir que o trabalho continuasse, embora ao custo de introduzir incerteza adicional.

A consequência mais imediata dessa decisão foi uma redução significativa no número de exemplos disponíveis para treinar o modelo *multitask*. Inicialmente, havia mais de 30.000 instâncias, mas apenas 1.705 textos atendiam ao critério de similaridade e podiam ser usados na fase final. Embora esse número fosse suficiente para um exercício de demonstração, era claramente insuficiente para um treinamento robusto, restringindo assim o potencial de generalização do modelo. Portanto, a escassez de dados é uma limitação estrutural que afeta tanto a representatividade quanto a validade estatística dos resultados.

Não foram aplicadas etapas adicionais aos dados no que diz respeito ao pré-processamento textual. Isso ocorreu porque o conjunto de dados SNCrawler já incluía um processo de normalização adequado, garantindo a consistência mínima necessária para a análise. Além disso, uma limpeza mais agressiva poderia comprometer a preservação de sinais linguísticos típicos das redes sociais, como abreviações, erros ortográficos, repetições de caracteres e emojis. Esses sinais muitas vezes constituem pistas relevantes para identificar estados emocionais. Portanto, a ausência de pré-processamento específico não deve ser vista como uma falha metodológica, mas sim como uma escolha deliberada para evitar corromper a natureza dos dados.

Metodologicamente, também houve um desvio do objetivo original. O plano inicial era criar um modelo multimodal que pudesse combinar vários tipos de dados, como texto, áudio, vídeo e sinais fisiológicos, para fornecer uma visão mais abrangente dos fenótipos digitais relacionados à saúde mental. No entanto, a escassez de conjuntos de dados públicos compatíveis com esse objetivo, aliada às dificuldades de acesso a dados multimodais de qualidade, tornou essa abordagem inviável. Consequentemente, o trabalho teve de se restringir apenas ao aspecto textual, o que limita o âmbito da ambição original do projeto.

Outra limitação significativa é a dependência de plataformas externas e conjuntos de dados públicos. Grande parte da metodologia baseou-se no Vertex AI (Google Cloud) para o ajuste fino do modelo Gemini-flash-2.0. No entanto, este modelo é proprietário, o que significa que o acesso ao código e aos pesos não é fornecido. Além disso, foram utilizados vários conjuntos de dados externos relacionados à ansiedade e à depressão, mas esses conjuntos de dados eram frequentemente limitados em tamanho, qualidade ou granularidade. Essa dependência reduziu a flexibilidade da pesquisa, limitou a possibilidade de ajuste fino de parâmetros

críticos e dificultou a reprodutibilidade, pois o acesso a essas ferramentas pode depender de custos ou políticas comerciais.

Por fim, é importante enfatizar que este trabalho é essencialmente exploratório e serve como prova de conceito. Embora a metodologia desenvolvida tenha demonstrado a viabilidade de associar transtornos e comportamentos com base em dados textuais, ela não deve ser interpretada como um sistema pronto para aplicação clínica. Limitações na quantidade de dados, o uso de limites baixos e a ausência de validação clínica rigorosa significam que os resultados não podem ser generalizados sem cautela.

### 4.9 Conclusão

O presente capítulo descreveu em detalhe a metodologia seguida para a construção de um modelo multitask aplicado à análise de fenótipos digitais relacionados com ansiedade e depressão. O percurso metodológico iniciou-se com a recolha de conjuntos de dados de saúde mental já existentes, a partir dos quais foram realizados os primeiros fine-tunings com o objetivo de treinar classificadores específicos para ansiedade e para depressão. Esta etapa permitiu assegurar que as instâncias utilizadas em fases posteriores estavam corretamente categorizadas em termos de transtornos.

Posteriormente, foi criado um conjunto de dados unificado, no qual se introduziu a dimensão dos comportamentos. Para tal, recorreu-se a um vocabulário expandido com auxílio de modelos de linguagem, e a associação entre textos e comportamentos foi realizada através de métricas de similaridade, destacando-se a eficácia da similaridade do cosseno. Esta etapa constituiu um marco central da metodologia, por traduzir comportamentos físicos e psicológicos em expressões textuais, tornando-os analisáveis em contexto computacional. Apesar da necessidade de adotar um threshold relativamente baixo (0.5), que reduziu a robustez das associações, este passo foi fundamental para a viabilização da abordagem.

Com base nos dados assim enriquecidos, procedeu-se ao fine-tuning multitask no Vertex AI, utilizando o modelo Gemini-flash-2.0. Esta escolha assegurou suporte direto a tarefas supervisionadas de classificação, bem como a gestão automática de aspetos críticos como a otimização de hiperparâmetros e o controlo do overfitting. O conjunto de dados final para treino multitask, composto por 1705 instâncias, revelou-se limitado em dimensão, mas suficiente para uma prova de conceito sólida.

Na etapa final, o modelo multitask foi aplicado ao conjunto de dados SNCrawler, que, embora já contivesse algumas anotações linguísticas e sociais, foi aqui utilizado apenas como fonte de texto. O resultado foi a criação de um novo conjunto de dados derivado, no qual cada instância passou a estar caracterizada por duas dimensões em simultâneo: a presença ou ausência de transtornos e os comportamentos associados. Esta transformação representou a concretização metodológica do trabalho, dando origem a um recurso que poderá ser explorado em futuros estudos de digital phenotyping.

Embora o processo tenha enfrentado limitações — nomeadamente a escassez de dados, a dependência de plataformas externas e a impossibilidade de explorar modalidades além do texto —, foi possível demonstrar a viabilidade da abordagem proposta. O capítulo evidenciou

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

assim tanto as decisões metodológicas adotadas como os compromissos necessários para garantir a continuidade do projeto, preparando o terreno para a análise dos resultados e para a discussão crítica que se seguem no capítulo seguinte.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## Capítulo 5

### Resultados e Discussão

O presente capítulo apresenta e discute os resultados obtidos com a aplicação da metodologia descrita no capítulo anterior. São analisados os conjuntos de dados derivados do processo de anotação automática, destacando a distribuição de casos por transtornos e comportamentos. Em seguida, são avaliados os desempenhos dos modelos desenvolvidos, tanto na fase inicial de *fine-tuning* individual para ansiedade e depressão como no modelo final *multitask*.

Além da apresentação das métricas de desempenho, o capítulo explora as implicações dos resultados obtidos, procurando identificar tendências relevantes, limitações observadas e potenciais explicações para os padrões encontrados. Esta análise crítica visa não apenas descrever os resultados, mas também contextualizá-los no âmbito da literatura científica e dos objetivos delineados para a dissertação.

Por fim, é apresentada uma síntese que resume os principais achados, servindo de ponte para a discussão final e para as conclusões do trabalho.

#### 5.1 Fine-tuning para ansiedade e depressão

Dois modelos independentes de classificação binária foram desenvolvidos para a identificação de ansiedade e depressão em textos, respectivamente. Ambos foram obtidos através do *fine-tuning* do modelo base Gemini-2.5-flash utilizando Vertex AI, com os dados divididos em 80% para treino e 20% para validação.

##### Modelo de Ansiedade

O modelo de ansiedade passou por um treino de 17 períodos, com pontos de verificação ativados para monitorar o progresso. As métricas demonstram uma aprendizagem rápida e eficiente. A precisão atingiu valores de 1,0 durante o treino e 0,999 durante a validação nas épocas iniciais, permanecendo estável até o final. Da mesma forma, a perda sofreu uma queda acentuada, aproximando-se rapidamente de zero, estabilizando em torno de 0,007–0,01 para validação. O gráfico de inferência revelou valores consistentes de cerca de 400–450 por etapa, refletindo a distribuição estável dos lotes de dados.

A análise dos pontos de verificação confirmou essa tendência, com a precisão permanecendo constante em 1,0 no treino e 0,999–1,0 na validação, enquanto a perda permaneceu em valores residuais. Esses resultados sugerem que o modelo aprendeu os padrões quase perfeitamente e generalizou de forma consistente para o conjunto de validação.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

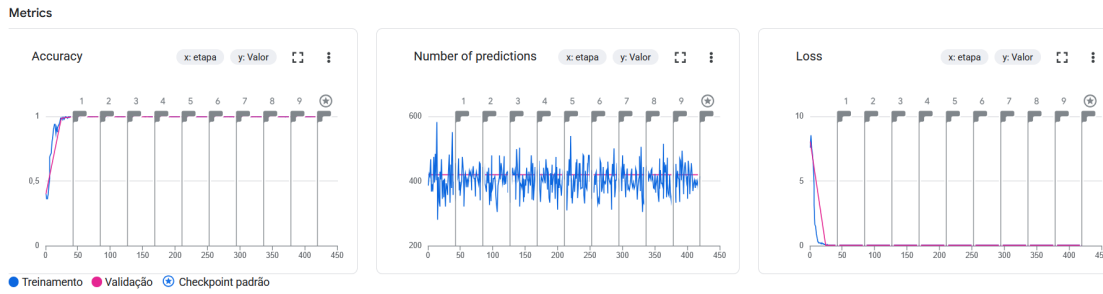


Figura 5.1: Desempenho do modelo de ansiedade: evolução da precisão (*accuracy*), da função de perda (*loss*) e do número de predições durante o treino.

Resumindo, como se observa na Figura 5.1, o modelo de ansiedade apresentou uma evolução rápida da métrica de precisão, estabilizando em valores elevados já nas primeiras épocas. A função de perda decresceu de forma consistente, convergindo para valores baixos, enquanto o número de predições se manteve estável ao longo do processo. Como demonstrado no gráfico, poderia existir *overfitting*, mas o próprio Vertex AI faz uma avaliação para identificar quando tem de parar antes mesmo de o ocorrer.

### Modelo de Depressão

O modelo de depressão foi treinado ao longo de nove períodos, com o nono ponto de verificação selecionado como padrão. A evolução das métricas demonstrou um desempenho consistente. Ao longo do processo, a precisão do treino aumentou de 0,971 para 0,998, e a precisão da validação subiu de 0,982 para 0,999. A perda diminuiu gradual e consistentemente de 0,053 para 0,006 no conjunto de validação e de 0,078 para 0,006 no conjunto de treino.

Em todas as etapas do treino, houve uma melhoria contínua tanto na precisão quanto na minimização da perda. O número médio de inferências ( 2.000 por etapa) permaneceu estável ao longo de todo o processo, refletindo um processamento de dados consistente.

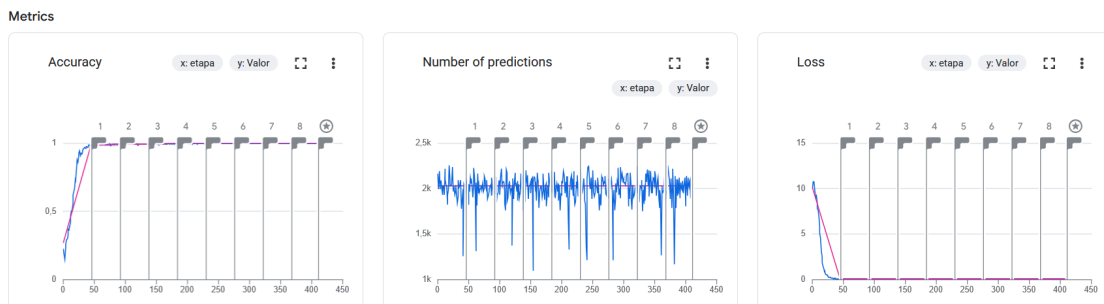


Figura 5.2: Desempenho do modelo de depressão: evolução da precisão (*accuracy*), da função de perda (*loss*) e do número de predições durante o treino.

A Figura 5.2 ilustra o desempenho do modelo de depressão. Tal como no caso anterior, observa-se uma elevada precisão desde as primeiras épocas, acompanhada por uma redução contínua da função de perda. O número de predições manteve-se regular ao longo do treino, evidenciando a consistência do modelo. Em conjunto, estes resultados demonstram a eficácia do ajuste e a fiabilidade do modelo para a identificação de sinais linguísticos de depressão.

### Análise Crítica

Os dois modelos com um tarefa única tiveram um desempenho quase perfeito, alcançando uma precisão próxima a 100% e perda residual mínima na fase de validação. Esses resultados confirmam a eficácia do *Fine-tuning* do modelo base para conjuntos de dados específicos de ansiedade e depressão.

No entanto, é importante observar que valores de precisão tão altos devem ser analisados com cautela. Por um lado, eles refletem a alta qualidade dos conjuntos de dados iniciais e a adequação do modelo. Por outro lado, também podem indicar redundância nos dados ou um nível relativamente baixo de dificuldade em tarefas binárias em comparação com classificações mais complexas. Embora não tenha sido observado nenhum *overfitting* clássico (as curvas de treino e validação permaneceram sobrepostas), existe o risco de os modelos terem sido excessivamente ajustados às características específicas dos dados utilizados, o que poderia afetar a sua capacidade de generalização para outros contextos, mas esse risco é retirado tendo em conta que o próprio Vertex AI não deixa chegar ao *overfitting* fazendo ele uma avaliação para o evitar.

No entanto, os resultados obtidos foram fundamentais para a etapa seguinte. Esses modelos permitiram a identificação consistente de casos de ansiedade e/ou depressão, garantindo a confiabilidade necessária para a construção do conjunto de dados unificado de transtornos. Esse conjunto de dados serviu então como base para as etapas subsequentes da pesquisa. Também podemos observar que dados que inicialmente eram anotados com um transtorno apenas passaram a ter ambos os transtornos, o que permite chegar à conclusão de que dados já anotados com qualquer transtorno não necessariamente não têm os outros transtornos existentes, da mesma forma.

## 5.2 Conjuntos de dados obtidos

O primeiro conjunto de dados obtido é o resultado da iteração cruzada dos modelos da ansiedade e depressão, dando origem a um conjunto de dados do tamanho de 30.305 instâncias. Do total, como representado na Figura 5.3, 28,6% das instâncias, ou 8.671, foram categorizadas como expressando "ambas as condições" (ansiedade e depressão), enquanto 18,1% (5.476 instâncias) referem-se a "apenas à depressão" e 15,8% (4.798 instâncias) "apenas à ansiedade". A classe de "sem transtorno" representou 37,5% do total, com 11.361 instâncias. Os dados obtidos inicialmente estavam equiparados em relação ao número de instâncias com e sem transtorno e também para cada transtorno. A iteração cruzada dos modelos veio acrescentar uma nova classe, a "ambos", o que desequilibrou um pouco o número de instâncias para cada classe, mas não a ponto de se sentir nos próximos passos.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

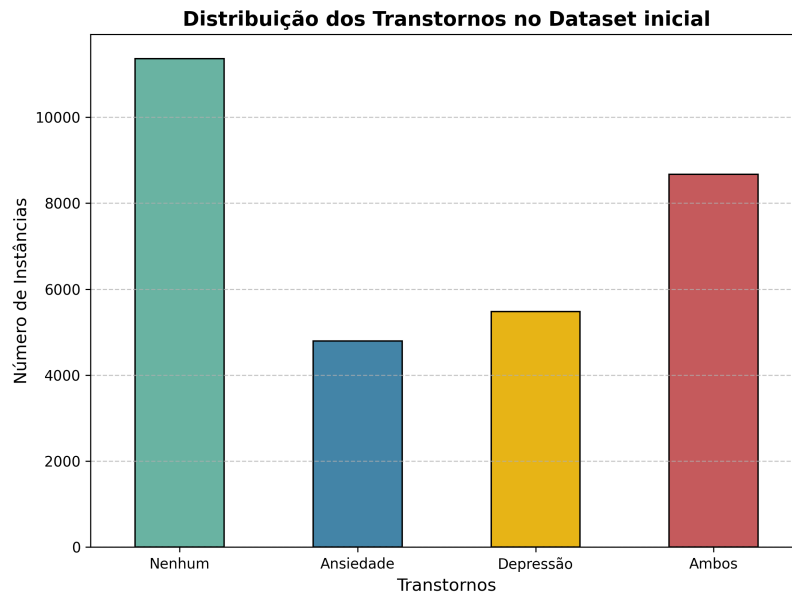


Figura 5.3: Distribuição dos transtornos no conjunto de dados inicial sem comportamentos associados.

Relativamente ao conjunto de dados com os comportamentos, o processo de associação semântica que ligou as instâncias do primeiro conjunto de dados a comportamentos descritos na literatura, revelou-se um pouco condicionante, pois o limiar de similaridade de 0.5 levou a uma diminuição acentuada das instâncias do novo conjunto de dados, ou seja, num conjunto de dados com 1.705 instâncias. Esta redução é particularmente relevante, pois sublinha a dificuldade de encontrar dados textuais que expressem os comportamentos clínicos que foram definidos no enquadramento.

A distribuição dos transtornos neste conjunto de dados mais restrito ficou assim estabelecida: 504 instâncias na classe de "ambos", 466 "sem transtorno", 422 "depressão" e 313 "ansiedade". Podendo verificar que existe um certo equilíbrio em termos dos transtornos, como se pode verificar na Figura 5.4, o que vai permitir resultados melhores do modelo posterior.

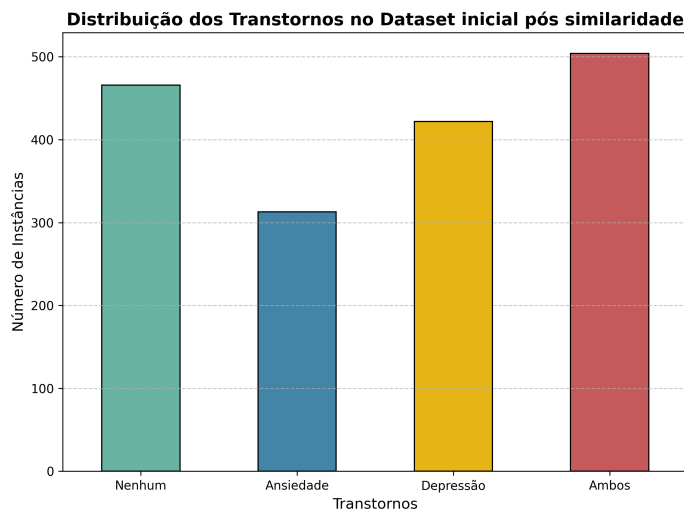


Figura 5.4: Distribuição dos transtornos no conjunto de dados inicial com comportamentos, usado para fine-tuning do modelo multitask.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

Já a distribuição dos 18 comportamentos, representada na Figura 5.5, mostrou que os comportamentos mais comuns, como "postura encurvada" (367 ocorrências) e "expressão facial preocupada" (214), sugerem que certas manifestações comportamentais são mais facilmente detetadas a partir do texto. Por outro lado, o menor número de ocorrências de outros comportamentos, como "caligrafia alterada" (21), "uso excessivo de dispositivos digitais" (28), destaca as limitações inerentes à tentativa de detetar sinais fisiológicos ou motores a partir de dados puramente textuais, um ponto crucial para a discussão e para o contributo deste trabalho.

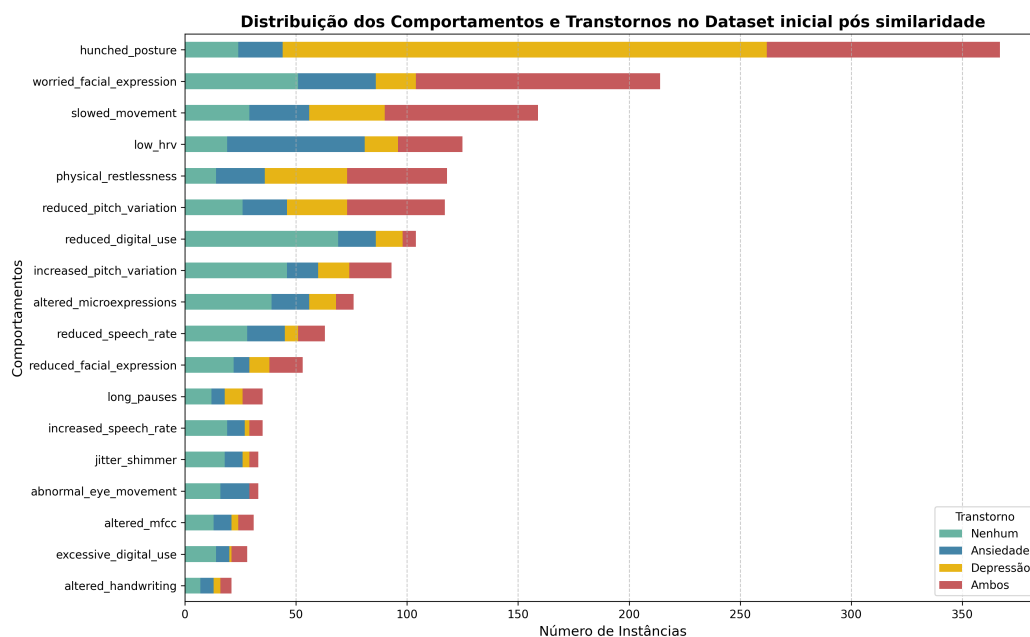


Figura 5.5: Distribuição dos comportamentos no conjunto de dados inicial com comportamentos, usado para fine-tuning do modelo multitask.

Ainda na Figura 5.5 pode-se observar que a análise visual dos dados demonstra uma correspondência entre a frequência de certos comportamentos no texto e os perfis clínicos definidos na literatura. É particularmente evidente que os comportamentos mais frequentes no conjunto de dados, nomeadamente "postura encurvada" e "expressão facial preocupada", estão predominantemente associados às classes de "depressão" e "ambas" as condições (ansiedade e depressão). Esta associação alinha-se diretamente com o enquadramento teórico apresentado, onde a "postura encurvada" é deduzida a partir de linguagem que sugere retração ou isolamento, e a "expressão facial preocupada" é associada à linguagem que expressa inquietação. A prevalência destas ocorrências no texto valida a hipótese de que estes são indicadores linguísticos robustos de estados depressivos ou de ambos os transtornos.

Da mesma forma, comportamentos como "uso excessivo de dispositivos digitais" e "aumento da variação de tom" tem mais incidência nas classes de "ansiedade" e "ambos". O que vai de encontro com o descrito na secção 2.4, que associa a ansiedade à hiperatividade emocional e cognitiva, refletida num maior volume de postagens e no uso de linguagem emocionalmente carregada. Isto reforça a utilidade da abordagem para a fenotipagem digital.

Por outro lado, o gráfico também evidencia a raridade de outros comportamentos. Fenó-

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

tipos como "caligrafia alterada", "uso excessivo de dispositivos digitais" (que, embora presente, tem uma contagem de instâncias muito baixa na maioria das classes) ou "tremores" têm uma contagem de instâncias muito baixa em todas as categorias. Este padrão valida a suposição de que certos comportamentos são mais difíceis de encontrar a partir de dados textuais, uma vez que se referem a sinais fisiológicos ou motores que não se traduzem diretamente em palavras. Esta limitação reforça a necessidade de abordagens multimodais para uma fenotipagem digital mais completa, tal como sugerido pela literatura.

### 5.3 Modelo *multitask* (Transtornos + Comportamentos)

O terceiro modelo desenvolvido neste estudo baseou-se numa abordagem *multitask*. Nesta abordagem, cada instância textual foi utilizada para prever dois resultados distintos simultaneamente: a etiqueta do transtorno (0 – nenhuma; 1 – ansiedade; 2 – depressão; 3 – ambas) e a etiqueta do comportamento (18 categorias possíveis). O ajuste fino foi realizado no modelo base Gemini-2.0-flash-001 usando a plataforma Vertex AI e um conjunto de dados enriquecido composto por 1.705 instâncias.

O treino ocorreu ao longo de sete períodos, com o sétimo ponto de verificação selecionado como referência. As métricas demonstram que o modelo melhorou consistentemente ao longo das épocas, estabilizando-se num nível alto, embora não tenha atingido os níveis dos modelos binários.

A precisão da validação atingiu 0,956 e a precisão do treino atingiu 0,961. Estes valores demonstram a robustez do modelo, mesmo num cenário *multitask*. A perda de validação diminuiu progressivamente de 0,319 no ponto de verificação inicial para 0,135 no final, e a perda de treino caiu para 0,107. Estes resultados mostram que o modelo aprendeu de forma estável e generalizou adequadamente para dados não vistos.

O número de inferências por etapa variou de 1.500 a 2.000, refletindo a maior complexidade e variabilidade do processo de treino em comparação com o de modelos individuais.

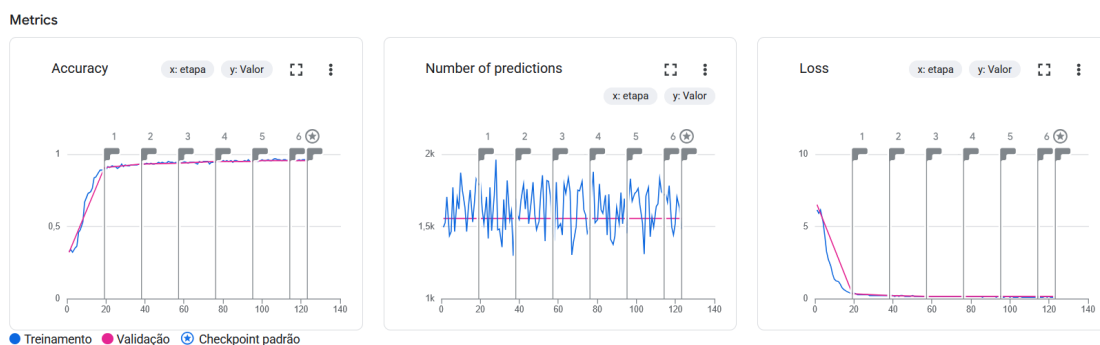


Figura 5.6: Desempenho do modelo *multitask*: evolução da precisão (*accuracy*), da função de perda (*loss*) e do número de predições durante o treino.

Na Figura 5.6 é apresentado o desempenho do modelo *multitask*. Como já foi referido, apesar de treinado com um conjunto de dados mais reduzido, o modelo alcançou valores consistentes de precisão e uma descida clara da função de perda, ainda que com maior oscilação do que nos modelos individuais. O número de predições manteve-se estável, refletindo a adaptação

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

do modelo à classificação simultânea de transtornos e comportamentos. Estes resultados confirmam a viabilidade da abordagem, ainda que limitada pela dimensão dos dados disponíveis.

### **Análise Crítica**

Em comparação com os classificadores binários de ansiedade e depressão, o modelo multitask apresentou valores de precisão ligeiramente inferiores. Esta diferença era expectável, uma vez que a tarefa multitask é intrinsecamente mais complexa, exigindo que o modelo aprenda a otimizar duas saídas distintas em simultâneo.

Por outro lado, os resultados alcançados — com accuracy acima dos 95% — são particularmente relevantes, dado que foram obtidos a partir de um conjunto de dados reduzido (apenas 1 705 instâncias) e com forte desequilíbrio entre as categorias de comportamento. A capacidade de manter um desempenho elevado neste cenário confirma a viabilidade da abordagem multitask, validando a hipótese central do projeto de que é possível associar transtornos e comportamentos em simultâneo a partir de dados textuais.

Contudo, importa reconhecer que a loss residual mais elevada (em comparação com os modelos de tarefa única) indica que o modelo não conseguiu capturar todas as nuances das duas dimensões em jogo. Este aspeto decorre, na maioria, da escassez de dados e da desigualdade na distribuição dos comportamentos, que dificultam a aprendizagem plena de padrões consistentes.

## **5.4 Aplicação ao conjunto de dados SNCrawler**

A aplicação do modelo multitask ao conjunto de dados SNCrawler representou uma etapa fundamental para a validação deste projeto. Mais do que uma simples prova de conceito, este passo permitiu-nos demonstrar a capacidade de generalização da nossa abordagem a um conjunto de dados de redes sociais de grande dimensão, gerando um novo recurso com mais de 14 milhões de instâncias anotadas. Este processo não só validou a metodologia, como também produziu um conjunto de dados robusto, oferecendo perspetivas importantes sobre a manifestação de transtornos e comportamentos associados a plataformas digitais.

A análise da distribuição dos transtornos no conjunto de dados final mostrou uma clara assimetria, tal como ilustrado na Figura 5.7. De um total de 14.213.877 instâncias processadas, a maioria, 10.813.955 (76,1%), foi classificada como "sem transtorno". Este resultado é particularmente relevante, uma vez que se alinha com a natureza de um conjunto de dados generalista, onde os textos que expressam explicitamente sinais de ansiedade ou depressão são, por natureza, minoritários. Contudo, o modelo conseguiu identificar 2.943.264 instâncias (20,7%) com ansiedade, 313.329 (2,2%) com depressão e 526.540 (3,7%) com ambas as condições. Embora a predominância da classe "sem transtorno" evidencie o desafio de encontrar casos relevantes, a capacidade de identificar e classificar milhões de instâncias com transtornos demonstra a viabilidade da abordagem e o potencial do conjunto de dados para análises futuras.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

Numa outra perspectiva estes resultados indicam que num raio de 14.213.877 de postagens, 26.6% podem estar associadas a um transtorno, o que aumenta a importância fenotipagem digital.

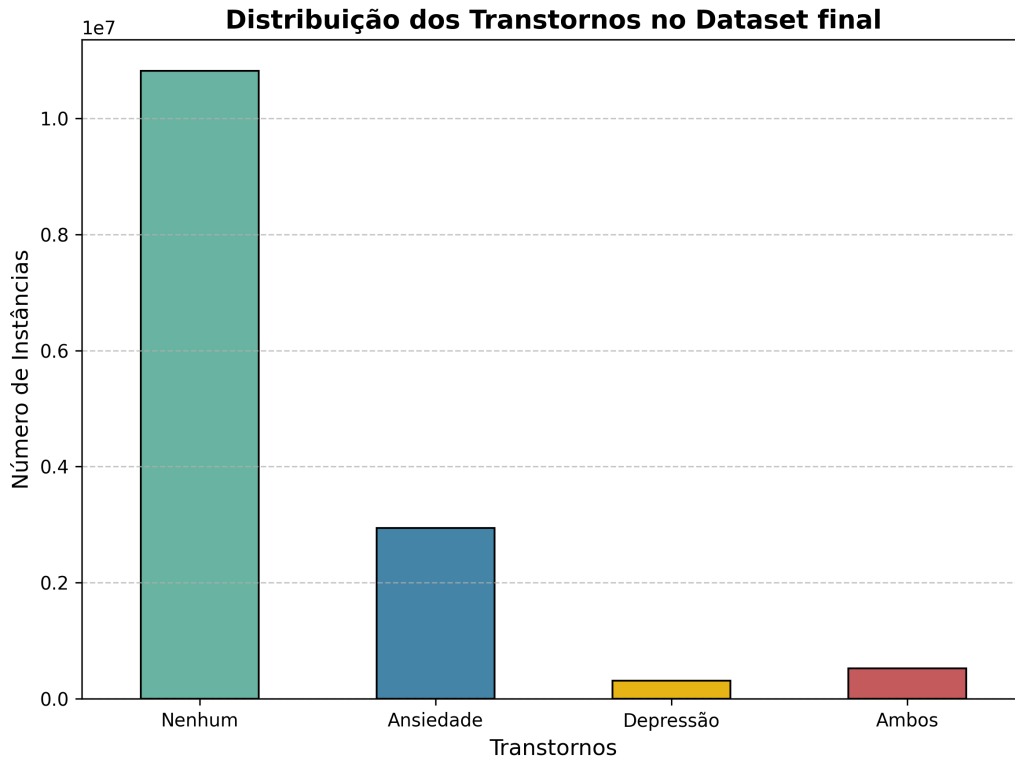


Figura 5.7: Distribuição dos transtornos no conjunto de dados final.

No que respeita aos comportamentos, a distribuição do conjunto de dados final, apresentada na Figura 5.8, mostrou igualmente uma forte assimetria, com alguns comportamentos a ocorrerem com muito maior frequência do que outros. Os mais comuns foram "aumento da variação de tom" (3.546.482 instâncias), "aumento da velocidade da fala" (1.816.448) e "redução da expressão facial" (1.310.109). A alta frequência destes fenótipos sugere que o modelo os capturou como representantes linguísticos de fenómenos emocionais comuns, tal como defendido na literatura que associa a variação do tom e a velocidade da fala a estados de hiperatividade ou lentificação cognitiva. Em contrapartida, a escassez de comportamentos como "caligrafia alterada" (215.477), *jitter\_shimmer* (136.580) ou "Uso excessivo digital" (136.776) reflete a dificuldade inerente à inferência de sinais fisiológicos ou motores a partir de dados puramente textuais.

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

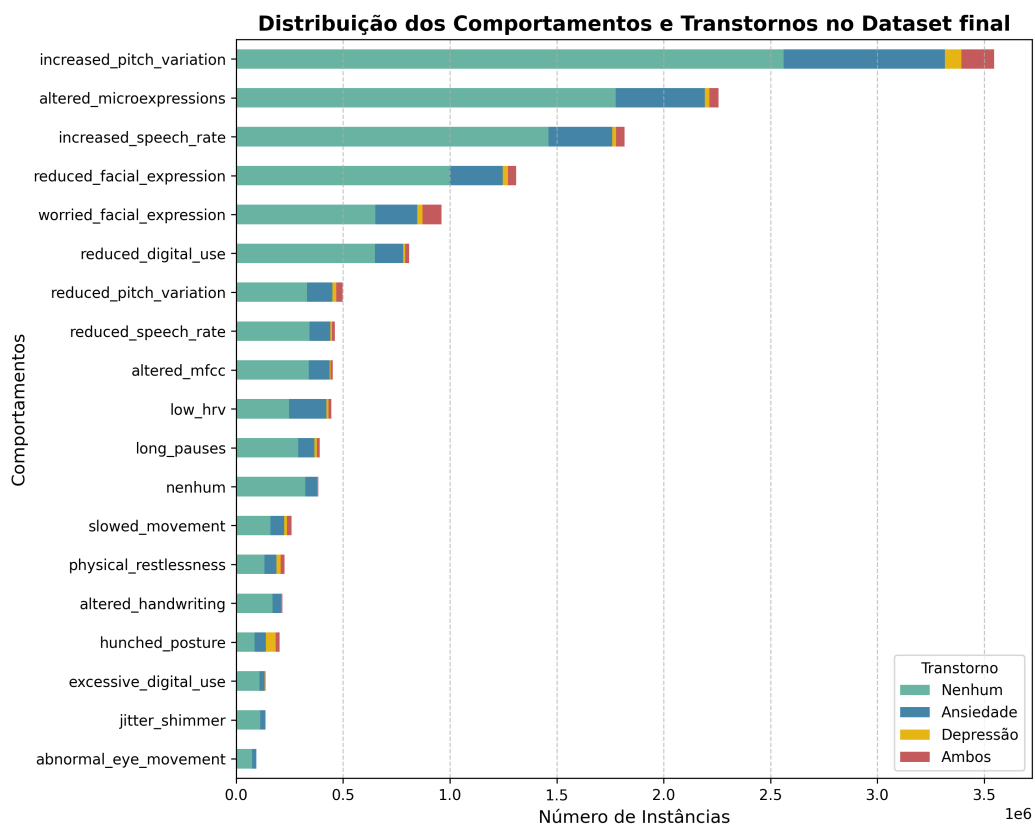


Figura 5.8: Distribuição dos comportamentos no conjunto de dados final.

O gráfico da distribuição da Figura 5.8 oferece uma perspectiva sobre a forma como os comportamentos se traduzem em dados textuais de grande escala. A análise visual desta distribuição, em conjunto com a literatura apresentada nesta tese, revela que a relação entre comportamentos e transtornos é multifacetada e complexa, refletindo as nuances do mundo real.

Uma observação imediata é a predominância esmagadora das instâncias classificadas como "Sem transtorno" em quase todas as categorias de comportamento. Estes dados reforçam a premissa de que os comportamentos não são indicadores exclusivos de patologia, existindo também em contextos normais de comunicação. No entanto, a análise torna-se mais reveladora ao examinar as tendências específicas dentro desta distribuição.

Comportamentos como "aumento da variação de tom" (*increased\_pitch\_variation*) e "aumento da velocidade da fala" (*increased\_speech\_rate*), que a literatura da tese associa a estados de hiperatividade e ansiedade, mostram uma frequência notável nas classes de "Ansiedade" e "Depressão", com uma ligeira predominância na primeira. Este resultado valida a capacidade do modelo em capturar os correlatos textuais de estados emocionais característicos, mesmo num ambiente onde os sinais são difusos. Da mesma forma, "redução da expressão facial" (*reduced\_facial\_expression*), associado a uma linguagem mais neutra ou à ausência de expressões emocionais, apresenta uma presença significativa nas classes de "Depressão" e "Ambos", o que se alinha de forma consistente com o enquadramento teórico sobre os comportamentos associados a este transtorno.

Em contrapartida, a análise do gráfico também sublinha as limitações da abordagem. A es-

casas de ocorrências para comportamentos como "caligrafia alterada" (*altered\_handwriting*) ou "jitter/shimmer" em todas as classes de transtornos demonstra a dificuldade de encontrar certas manifestações fisiológicas ou motoras a partir de dados puramente textuais. Esta constatação é fundamental para contextualizar os resultados e reforça a necessidade de abordagens multimodais para uma fenotipagem digital mais abrangente.

### 5.5 Síntese dos resultados

O presente capítulo apresentou e discutiu os resultados obtidos ao longo das diferentes etapas do trabalho. A criação dos conjuntos de dados constituiu um marco fundamental, permitindo primeiro reunir instâncias anotadas com ansiedade e depressão e, posteriormente, expandir essas anotações com a dimensão dos comportamentos. O conjunto de dados unificado de transtornos alcançou mais de 30 mil entradas, enquanto o conjunto de dados multitask, ainda que reduzido a 1 705 instâncias devido ao limiar de similaridade adotado, possibilitou o treino de modelos mais complexos e equilibrados.

Os modelos de tarefa única (ansiedade e depressão) demonstraram desempenhos quase perfeitos, atingindo valores de precisão próximos de 100% e perdas residuais, tanto no treino como na validação. Estes resultados confirmaram a adequação do processo de fine-tuning e a robustez do modelo base utilizado, garantindo a fiabilidade das classificações produzidas. O modelo multitask de transtornos e comportamentos, embora com métricas ligeiramente inferiores, revelou uma precisão de validação superior a 95%, o que é notável para uma tarefa de maior complexidade. Este resultado demonstrou a viabilidade de associar em simultâneo diferentes dimensões clínicas e comportamentais a dados textuais, validando a hipótese central do trabalho.

Finalmente, a aplicação ao conjunto de dados SNCrawler resultou na criação de um recurso de larga escala, com mais de 14 milhões de textos anotados com transtornos e comportamentos. Apesar da predominância da classe "sem transtorno" e da desigualdade na distribuição dos comportamentos, o novo corpus representa um contributo inédito, abrindo caminho para investigações futuras no domínio da fenotipagem digital.

Em síntese, os resultados alcançados confirmam a exequibilidade técnica da abordagem proposta, ao mesmo tempo que evidenciam os seus desafios práticos — desde os desequilíbrios de classes até às limitações na representação de certos comportamentos.

## Capítulo 6

### Conclusões e Trabalho futuro

#### 6.1 Conclusões

O objetivo desta dissertação consistiu em desenvolver um conjunto de dados textual anotado que relacionasse transtornos mentais — em particular, depressão e ansiedade — com comportamentos específicos encontrados a partir de textos provenientes de redes sociais. Para tal, foi necessário reformular a abordagem inicial, que previa uma perspectiva multimodal, em virtude da indisponibilidade de conjuntos de dados públicos adequados a esse propósito. A investigação concentrou-se, assim, numa abordagem exclusivamente textual, tirando partido dos avanços recentes em *Natural Language Processing* (NLP).

O trabalho realizado permitiu criar um processo estruturado que incluiu: a recolha e integração de conjuntos de dados externos, a realização de *fine-tuning* inicial de modelos supervisionados no Vertex AI, a iteração cruzada para a construção de um conjunto de dados unificado de transtornos, a associação de comportamentos por técnicas de similaridade semântica e, finalmente, o treino de um modelo *multitask*. A aplicação do modelo ao conjunto de dados SNCrawler culminou na criação de um novo recurso anotado que, até onde se tem conhecimento, é um dos primeiros a combinar simultaneamente dimensões clínicas (ansiedade e depressão) e comportamentais em dados textuais provenientes de redes sociais.

Os resultados obtidos demonstraram a viabilidade da abordagem, ainda que sujeitos a limitações. O *fine-tuning* inicial dos modelos específicos para ansiedade e depressão alcançou métricas de precisão elevadas, superiores a 95%, confirmando a capacidade dos modelos para identificar padrões linguísticos associados a cada transtorno. Já o modelo *multitask*, embora baseado num conjunto reduzido de 1705 instâncias, mostrou-se capaz de realizar classificações conjuntas de transtornos e comportamentos com resultados consistentes. Estes achados reforçam a relevância da estratégia seguida e evidenciam o potencial do recurso criado como base para futuras investigações em *digital phenotyping*.

Contudo, algumas limitações estruturais devem ser salientadas. A principal prende-se com a necessidade de recorrer a um limiar de similaridade relativamente baixo (0,5) para viabilizar a associação de comportamentos, o que introduz incerteza nas anotações. A escassez de dados após esta filtragem, a dependência de plataformas externas e a impossibilidade de explorar modalidades além do texto constituem igualmente barreiras que limitam a generalização dos resultados. Este trabalho deve, portanto, ser entendido como uma prova de conceito, capaz de abrir caminho, mas não ainda como uma solução pronta para aplicação clínica.

Em síntese, a presente dissertação contribui para o avanço do estudo da saúde mental em dados textuais, apresentando uma metodologia inovadora para a construção de conjuntos de dados enriquecidos com dimensões clínicas e comportamentais. A associação de trans-

tornos e comportamentos num mesmo recurso representa um passo importante no sentido de compreender de que forma padrões linguísticos podem refletir estados psicológicos complexos, respondendo ao apelo crescente por métodos escaláveis e acessíveis na área da saúde mental digital.

### 6.2 Trabalhos futuros

O trabalho realizado demonstrou a viabilidade do uso de textos de redes sociais como fonte de fenótipos digitais associados à ansiedade e à depressão. No entanto, vários aspetos devem ser explorados mais a fundo em pesquisas futuras.

Uma das questões centrais diz respeito ao limiar de similaridade. No presente estudo, foi utilizado um valor de 0,5. Embora isso tenha permitido a construção do conjunto de dados, trata-se essencialmente de uma decisão aleatória, semelhante a jogar uma moeda ao ar. A médio prazo, será essencial aumentar este limiar para valores mais exigentes, como 0,75 ou superior, para garantir maior robustez e fiabilidade nas associações entre textos e comportamentos. Isto só será possível através da recolha ou desenvolvimento de conjuntos de dados mais extensos e diversificados, que compensarão a perda de exemplos associados a limiares mais elevados.

Da mesma forma, a escassez de dados continua a ser um desafio. Embora o conjunto de dados inicial contivesse mais de 30.000 instâncias, apenas 1.705 foram utilizadas para o treino *multitask*. Trabalhos futuros poderiam concentrar-se em estratégias para expandir o corpus, tais como a recolha de novos dados ou a utilização de técnicas de aumento de dados adaptadas ao texto para aumentar a diversidade sem comprometer a integridade semântica. Outra área a ser desenvolvida é a expansão para uma abordagem multimodal, o que traria o projeto de volta à sua ambição inicial. A integração de texto com outros tipos de dados, como áudio, vídeo ou registos fisiológicos, permitiria uma caracterização mais rica e precisa dos fenótipos digitais. No entanto, essa etapa exigirá acesso a conjuntos de dados multimodais, que atualmente são escassos, mas devem se tornar mais prevalentes nos próximos anos.

Também poderiam ser feitas melhorias no modelo *multitask* utilizado. No futuro, o uso de arquiteturas de código aberto mais flexíveis como alternativa aos modelos proprietários, como o Gemini-Flash-1.5, poderia aumentar a reprodutibilidade científica e permitir um maior controlo sobre os parâmetros internos. Além disso, abordagens *multitask* mais sofisticadas, como *hierarchical multitask learning* ou *multi-label classification*, poderiam ser investigadas para capturar com mais precisão a sobreposição entre ansiedade, depressão e diferentes comportamentos.

No que respeita ao conjunto de dados SNCrawler, um caminho promissor será enriquecê-lo com novas camadas de anotação, cruzando as etiquetas de transtornos e comportamentos geradas neste trabalho com outras dimensões já existentes no conjunto de dados original, como características linguísticas e sociais. Esta integração pode oferecer novas perspetivas sobre a interação entre fenómenos linguísticos e estados psicológicos, reforçando a relevância do recurso para a comunidade científica.

Por fim, a validação clínica será essencial. Embora o presente estudo seja uma prova de con-

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

ceito exploratória, a sua aplicação em contextos reais de saúde mental requer uma validação rigorosa por especialistas. Trabalhos futuros poderiam, portanto, envolver a ligação de modelos como o desenvolvido aqui com equipas clínicas para avaliar a sua utilidade prática no monitoramento de sintomas ou no apoio a diagnósticos. Sendo que já existe uma equipa preparada para fazer esta validação clínica dos dados obtidos, estando já a trabalhar nessa validação. No entanto, não foi possível ter essa validação aquando da entrega da dissertação, daí ficar como trabalho futuro a validação clínica dos dados obtidos.

Em resumo, os resultados obtidos abrem caminho para uma ampla gama de pesquisas futuras. Os próximos passos lógicos são melhorias nos limiares, expansão e diversificação de dados, integração multimodal, exploração de modelos de código aberto e validação clínica, o que aproximará o campo da fenotipagem digital de uma aplicação eficaz na compreensão e monitorização da saúde mental.

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

## Bibliografia

- [1] W. H. O. WHO. (2022) Mental disorders. [Online]. Available: <https://www.healthdata.org/research-analysis/gbd> 5
- [2] I. for Health Metrics and Evaluation. (2021) Global burden of disease (gbd). [Online]. Available: <https://www.healthdata.org/research-analysis/gbd> 5
- [3] *Journal of Medical Health Sciences Review*, vol. 2, no. 1, Mar. 2025. [Online]. Available: <https://jmhsr.com/index.php/jmhsr/article/view/199> 6
- [4] H. F. AlQubali, G. A. M. Shibly, A. Alharbi, R. A. H. Arishi, and W. S. Albalawi, “Anxiety and depression among drug addicts: A systematic review,” *International Journal of Innovative Research and Scientific Studies*, vol. 8, no. 3, p. 2980–2986, May 2025. [Online]. Available: <https://www.ijirss.com/index.php/ijirss/article/view/7123> 6
- [5] R. Zughbur, M. Y. Hamam, A. Kagee, M. Hamam, M. Hijazi, Yara M. Hamam, O. Abu-olwan, S. Sayeed, and G. Veronese, “Prevalence and correlates of anxiety, depression, and symptoms of trauma among palestinian adults in gaza after a year of war: a cross-sectional study,” *Conflict and Health*, 2025. 6
- [6] B. Sultana, K. Ahmed, R. M. Pir, M. S. Ahmed, S. R. Nessa, and S. Rahman, Md. “Prevalence of depression and anxiety among chronic illness patients in a tertiary care hospital,” *Saudi Journal of Medicine*, May 2025. 7
- [7] N. Robinson, F. N. Ramos, S. D. Hollon, G. T. Han, and I. D. Ezawa, “Cognitive behavioral immersion for depression: Protocol for a three-arm randomized controlled trial of peer-based coaching in the metaverse,” *JMIR Res Protoc*, vol. 14, p. e65970, May 2025. [Online]. Available: <https://www.researchprotocols.org/2025/1/e65970> 7
- [8] A. S. Awan, F. Brohi, H. Buriro, M. Soomro, and B. Sarki, “Assess the level of depression, anxiety, and stress among undergraduate public nursing students, pakistan,” *Insights-Journal of Health and Rehabilitation*, vol. 3, no. 1 (Health amp;amp; Allied), p. 536–544, Feb. 2025. [Online]. Available: <https://insightsjhr.com/index.php/home/article/view/460> 7
- [9] *Journal of Medical Health Sciences Review*, vol. 2, no. 1, Mar. 2025. [Online]. Available: <https://jmhsr.com/index.php/jmhsr/article/view/199> 7
- [10] M. Alabdulla, S. Reagu, M. Alishaq, N. Al Hammadi, M. H. Elkordy, H. Ghazouani, and A. H. Assar, “The prevalence of depression and anxiety symptoms and their associated factors among patients with cancer in qatar: A cross-sectional study,” *Qatar Medical Journal*, vol. 2025, no. 1, 2025. [Online]. Available: <https://www.qscience.com/content/journals/10.5339/qmj.2025.4> 7

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [11] M. F. Wittenberg, C. E. Milliren, K. Waddicor, and S. L. Fitzgerald, “Anxiety, depression, and mental health service use among pregnant adolescents/young adults at an urban pediatric hospital based clinic,” *Clinical Child Psychology and Psychiatry*, vol. 30, no. 3, pp. 683–698, 2025, PMID: 40375641. [Online]. Available: <https://doi.org/10.1177/13591045251341007> 7
- [12] H. F. AlQubali, G. A. M. Shibly, A. Alharbi, R. A. H. Arishi, and W. S. Albalawi, “Anxiety and depression among drug addicts: A systematic review,” *International Journal of Innovative Research and Scientific Studies*, vol. 8, no. 3, p. 2980–2986, May 2025. [Online]. Available: <https://www.ijirss.com/index.php/ijirss/article/view/7123> 7
- [13] H. Gokhale, S. Singare, K. Chirinda, and K. Pillai, “The impact of social media engagement on mental health,” *International Journal For Multidisciplinary Research*, vol. 7, no. 3, May 2025. 7, 8
- [14] A. Ndindeng, “The impact of social media on mental health,” *Mental Health and Digital Technologies*, 2025, ahead-of-print. 7, 8, 9
- [15] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, “Predicting depression via social media,” *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 7, no. 1, pp. 128–137, Aug. 2021. [Online]. Available: <https://ojs.aaai.org/index.php/ICWSM/article/view/14432> 8
- [16] B. S. Satpute, W. P. Rahane, and R. Bharati, “Examining social media posts for identification of anxiety and depression utilizing machine learning techniques,” in *2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, 2023, pp. 295–300. 8, 9
- [17] D. G, S. K, P. P, and T. K, “Machine learning based mental health analysis using social media data,” in *2025 International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2025, pp. 1445–1450. 8, 9
- [18] L. Tudehope, N. Harris, L. Vorage, and E. Sofija, “What methods are used to examine representation of mental ill-health on social media? a systematic review,” *BMC Psychology*, vol. 12, no. 1, p. 105, 2024. [Online]. Available: <https://doi.org/10.1186/s40359-024-01603-1> 8
- [19] S. Dai, K. Li, Z. Luo, P. Zhao, B. Hong, A. Zhu, and J. Liu, “Ai-based nlp section discusses the application and effect of bag-of-words models and tf-idf in nlp tasks,” *Journal of Artificial Intelligence General science (JAIGS) ISSN:3006-4023*, vol. 5, no. 1, p. 13–21, Jun. 2024. [Online]. Available: <https://ojs.boulibrary.com/index.php/JAIGS/article/view/149> 13
- [20] G. S. S. Chamarthi and J. P. Pramod, “Exploring word-level representations in modern natural language processing,” *International Journal for Research in Applied Science and Engineering Technology*, 2024. 13

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [21] K. Deulkar, M. Narvekar, P. Gandhi, D. Gada, and S. Kamath, “Evaluating the influence of text representation techniques on diverse machine learning algorithms for stress detection in social media users,” in *2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, 2024, pp. 1–7. 14
- [22] M. M. Hossain, Sanjara, M. S. Hossain, S. Chaki, M. S. Rahman, and A. B. M. S. Ali, “Revolutionizing mental health sentiment analysis with bert-fuse: A hybrid deep learning model,” *IEEE Access*, vol. 13, pp. 85 428–85 446, 2025. 14
- [23] A. D. Bhargavi, “Comparative study of static and contextual text vectorization for sentiment analysis,” *International Journal for Research in Applied Science and Engineering Technology*, 2025. 14
- [24] T. Mikolov, K. Chen, G. S. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” in *International Conference on Learning Representations*, 2013. [Online]. Available: <https://api.semanticscholar.org/CorpusID:5959482> 14, 17
- [25] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in *Neural Information Processing Systems*, 2013. [Online]. Available: <https://api.semanticscholar.org/CorpusID:16447573> 14
- [26] J. Pennington, R. Socher, and C. Manning, “GloVe: Global vectors for word representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang, and W. Daelemans, Eds. Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1532–1543. [Online]. Available: <https://aclanthology.org/D14-1162/> 14, 17
- [27] T. Shi and Z. Liu, “Linking glove with word2vec,” *ArXiv*, vol. abs/1411.5595, 2014. [Online]. Available: <https://api.semanticscholar.org/CorpusID:13019011> 14
- [28] K. Ethayarajh, D. Duvenaud, and G. Hirst, “Towards understanding linear word analogies,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, A. Korhonen, D. Traum, and L. Màrquez, Eds. Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 3253–3262. [Online]. Available: <https://aclanthology.org/P19-1315/> 14
- [29] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer, “Deep contextualized word representations,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, M. Walker, H. Ji, and A. Stent, Eds. New Orleans, Louisiana: Association for Computational Linguistics, Jun. 2018, pp. 2227–2237. [Online]. Available: <https://aclanthology.org/N18-1202/> 14

- [30] K. Ethayarajh, “How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 55–65. [Online]. Available: <https://aclanthology.org/D19-1006/> 14
- [31] A. Vaswani, N. M. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Neural Information Processing Systems*, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:13756489> 14
- [32] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186. [Online]. Available: <https://aclanthology.org/N19-1423/> 14, 17
- [33] A. Radford and K. Narasimhan, “Improving language understanding by generative pre-training,” 2018. [Online]. Available: <https://api.semanticscholar.org/CorpusID:49313245> 15
- [34] A. Bandyopadhyay and J. Nair, “Word embedding for bengali language using domain-related corpus,” in *2023 International Conference on Inventive Computation Technologies (ICICT)*, 2023, pp. 896–901. 15
- [35] A. Roy, Y. Park, and S. Pan, “Learning domain-specific word embeddings from sparse cybersecurity texts,” *ArXiv*, vol. abs/1709.07470, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:2776354> 15
- [36] K. Yang, S. Ji, T. Zhang, Q. Xie, Z. Kuang, and S. Ananiadou, “Towards interpretable mental health analysis with large language models,” in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2023. [Online]. Available: <http://dx.doi.org/10.18653/v1/2023.emnlp-main.370> 15
- [37] A. Agrawal, “Illuminate: A novel approach for depression detection with explainable analysis and proactive therapy using prompt engineering,” *International Journal of Psychiatry*, 2024. 15
- [38] S. M. Rezaeinia, A. Ghodsi, and R. Rahmani, “Improving the accuracy of pre-trained word embeddings for sentiment analysis,” *ArXiv*, vol. abs/1711.08609, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:1018179> 15

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [39] M. Kowsher, M. S. I. Sobuj, M. F. Shahriar, N. J. Prottasha, M. S. Arefin, P. K. Dhar, and T. Koshiba, “An enhanced neural word embedding model for transfer learning,” *Applied Sciences*, vol. 12, no. 6, 2022. [Online]. Available: <https://www.mdpi.com/2076-3417/12/6/2848> 15
- [40] K. Yang, S. Ji, T. Zhang, Q. Xie, Z.-Z. Kuang, and S. Ananiadou, “Towards interpretable mental health analysis with chatgpt,” 2023. [Online]. Available: <https://api.semanticscholar.org/CorpusID:258715257> 15
- [41] D. Lozoya, A. Berazaluze, J. Perches, E. Lúa, M. Conway, and S. D’Alfonso, “Generating mental health transcripts with SAPE (Spanish adaptive prompt engineering),” in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, K. Duh, H. Gomez, and S. Bethard, Eds. Mexico City, Mexico: Association for Computational Linguistics, Jun. 2024, pp. 5096–5113. [Online]. Available: <https://aclanthology.org/2024.naacl-long.285/> 15
- [42] T. Liu, D. Jain, S. R. Rapole, B. L. Curtis, J. Eichstaedt, P. V. Kulkarni, and S. C. Guntuku, “Detecting symptoms of depression on reddit,” in *Proceedings of the 15th ACM Web Science Conference 2023*, 2023. 15
- [43] N. Hussain, A. Qasim, G. Mehak, M. Zain, G. Sidorov, A. Gelbukh, and O. Kolesnikova, “Multi-level depression severity detection with deep transformers and enhanced machine learning techniques,” *AI*, vol. 6, no. 7, 2025. [Online]. Available: <https://www.mdpi.com/2673-2688/6/7/157> 15
- [44] A. N. Sutranggono and R. Sarno, “Detection and sentiment analysis based on mental disorders aspects using bidirectional gated recurrent unit and semantic similarity,” *International Journal of Intelligent Engineering and Systems*, 2024. 16
- [45] Y. Zhu and J. Zhang, “Evaluating the relevance of health-related topics using three similarity measures,” *Information Development*, vol. 0, no. 0, p. 02666669251316264, 0. [Online]. Available: <https://doi.org/10.1177/02666669251316264> 16
- [46] S. Kim, J. Cha, D. Kim, and E. Park, “Understanding mental health issues in different subdomains of social networking services: Computational analysis of text-based reddit posts,” *J Med Internet Res*, vol. 25, p. e49074, Nov 2023. [Online]. Available: <https://www.jmir.org/2023/1/e49074> 16
- [47] T. Sheng, W. Cai, J. Huang, and Z. Dong, “Detecting adolescent depression in social media: A hierarchical ensemble learning approach,” *Theoretical and Natural Science*, vol. 43, pp. 312–325, 2024. 16
- [48] S. Rai, K. Shelat, D. Jain, A. Kishen, Y. M. Cho, M. Redkar, S. Hardikar-Sawant, L. Ungar, and S. C. Guntuku, “Cross-cultural differences in mental health expressions on social media,” in *Proceedings of the 3rd Workshop on Cross-Cultural Considerations in NLP (C3NLP 2025)*, V. Prabhakaran, S. Dev, L. Benotti,

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- D. Hershcovich, Y. Cao, L. Zhou, L. Cabello, and I. Adebara, Eds. Albuquerque, New Mexico: Association for Computational Linguistics, May 2025, pp. 132–142. [Online]. Available: <https://aclanthology.org/2025.c3nlp-1.10/> 16
- [49] Y. Cao, J. Dai, Z. Wang, Y. Zhang, X. Shen, Y. Liu, and Y. Tian, “Machine learning approaches for depression detection on social media: A systematic review of biases and methodological challenges.” *Journal of Behavioral Data Science*, vol. 5, no. 1, p. 67–102, Feb. 2025. [Online]. Available: <https://jbds.isdsa.org/jbds/article/view/11016>
- [50] D. Chandrasekaran and V. Mago, “Evolution of semantic similarity—a survey,” *ACM Computing Surveys*, vol. 54, no. 2, p. 1–37, Feb. 2021. [Online]. Available: <http://dx.doi.org/10.1145/3440755> 16
- [51] A. Severyn, M. Nicosia, and A. Moschitti, “Learning semantic textual similarity with structural representations,” in *Annual Meeting of the Association for Computational Linguistics*, 2013. [Online]. Available: <https://api.semanticscholar.org/CorpusID:6431532> 17
- [52] A. Oudin, R. Maatoug, A. Bourla, F. Ferreri, O. Bonnot, B. Millet, F. Schoeller, S. Mouchabac, and V. Adrien, “Digital phenotyping: data-driven psychiatry to redefine mental health,” *Journal of Medical Internet Research*, vol. 25, p. e44502, 2023. 18
- [53] J.-P. Onnela and S. L. Rauch, “Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health,” *Neuropsychopharmacology*, vol. 41, no. 7, pp. 1691–1696, 2016. 18
- [54] S. Saeb, M. Zhang, C. J. Karr, S. M. Schueller, M. E. Corden, K. P. Kording, D. C. Mohr *et al.*, “Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study,” *Journal of medical Internet research*, vol. 17, no. 7, p. e4273, 2015. 18
- [55] R. Wang, M. S. Aung, S. Abdullah, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, M. Merrill, E. A. Scherer *et al.*, “Crosscheck: toward passive sensing and detection of mental health changes in people with schizophrenia,” in *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, 2016, pp. 886–897. 18
- [56] D. J. Stein, J. A. Naslund, and J. Bantjes, “Covid-19 and the global acceleration of digital psychiatry,” *The Lancet Psychiatry*, vol. 9, no. 1, pp. 8–9, 2022. 18
- [57] A. G. Reece and C. M. Danforth, “Instagram photos reveal predictive markers of depression,” *EPJ Data Science*, vol. 6, no. 1, p. 15, 2017. 18
- [58] M. H. Aung, M. Matthews, and T. Choudhury, “Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies,” *Depression and anxiety*, vol. 34, no. 7, pp. 603–609, 2017. 18

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [59] K. O’Loughlin, M. Neary, E. C. Adkins, and S. M. Schueller, “Reviewing the data security and privacy policies of mobile apps for depression,” *Internet interventions*, vol. 15, pp. 110–115, 2019. 18
- [60] M. E. Morris, Q. Kathawala, T. K. Leen, E. E. Gorenstein, F. Guilak, M. Labhard, and W. Deleeuw, “Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness,” *Journal of medical Internet research*, vol. 12, no. 2, p. e10, 2010. 18
- [61] P. Dagum, “Digital biomarkers of cognitive function,” *NPJ digital medicine*, vol. 1, no. 1, p. 10, 2018. 18
- [62] S. K. Ernala, A. F. Rizvi, M. L. Birnbaum, J. M. Kane, and M. De Choudhury, “Linguistic markers indicating therapeutic outcomes of social media disclosures of schizophrenia,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 1, no. CSCW, pp. 1–27, 2017. 18
- [63] R. H. Birk and G. Samuel, “Digital phenotyping for mental health: reviewing the challenges of using data to monitor and predict mental health problems,” *Current psychiatry reports*, vol. 24, no. 10, pp. 523–528, 2022. 18
- [64] L. Kola, “Global mental health and covid-19,” *The Lancet Psychiatry*, vol. 7, no. 8, pp. 655–657, 2020. 18
- [65] H. Onyeaka, C. Muoghalu, M. Malekani, C. O. Azuike, P. O. Eseaton, O. E. Ayisire, S. V. Enemu, T. Basiru, J.-P. Omuojine, C. Ezeaka *et al.*, “Use of wearable devices among individuals with depression and anxiety: a population level study,” *Psychiatry Research Communications*, vol. 2, no. 4, p. 100081, 2022. 19
- [66] M. Faurholt-Jepsen, M. Frost, C. Ritz, E. M. Christensen, A. Jacoby, R. L. Mikkelsen, U. Knorr, J. Bardram, M. Vinberg, and L. V. Kessing, “Daily electronic self-monitoring in bipolar disorder using smartphones—the monarca i trial: a randomized, placebo-controlled, single-blind, parallel group trial,” *Psychological medicine*, vol. 45, no. 13, pp. 2691–2704, 2015. 19
- [67] R. L. Mandryk and M. V. Birk, “The potential of game-based digital biomarkers for modeling mental health,” *JMIR mental health*, vol. 6, no. 4, p. e13485, 2019. 19
- [68] J. Kamath, R. L. Barriera, N. Jain, E. Keisari, and B. Wang, “Digital phenotyping in depression diagnostics: integrating psychiatric and engineering perspectives,” *World Journal of Psychiatry*, vol. 12, no. 3, p. 393, 2022. 19
- [69] B. Nagamatu and I. Paraboni, “Detecção precoce de transtornos de saúde mental em português,” *Linguamática*, vol. 16, no. 2, p. preprint, Dez. 2024. [Online]. Available: <https://www.linguamatica.com/index.php/linguamatica/article/view/440> 19
- [70] S. L. Filho, M. Silva, and J. Oliveira, “Identificação de sintomas de depressão por dados de mídias sociais: Aplicando design science research para desenvolver um

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- modelo de classificação,” in *Anais Estendidos do XX Simpósio Brasileiro de Sistemas de Informação*. Porto Alegre, RS, Brasil: SBC, 2024, pp. 1–15. [Online]. Available: [https://sol.sbc.org.br/index.php/sbsi\\_estendido/article/view/28591](https://sol.sbc.org.br/index.php/sbsi_estendido/article/view/28591) 19
- [71] J. Melcher, R. Hays, and J. Torous, “Digital phenotyping for mental health of college students: a clinical review,” *BMJ Ment Health*, vol. 23, no. 4, pp. 161–166, 2020. 20
- [72] T. R. Insel, “Digital phenotyping: technology for a new science of behavior,” *Jama*, vol. 318, no. 13, pp. 1215–1216, 2017. 20
- [73] S. H. Jain, B. W. Powers, J. B. Hawkins, and J. S. Brownstein, “The digital phenotype,” *Nature Biotechnology*, vol. 33, no. 5, pp. 462–463, 2015. [Online]. Available: <https://doi.org/10.1038/nbt.3223> 20
- [74] T. R. Insel, “Digital phenotyping: Technology for a new science of behavior,” *JAMA*, vol. 318, no. 13, pp. 1215–1216, 10 2017. [Online]. Available: <https://doi.org/10.1001/jama.2017.11295> 20
- [75] J. Torous, M. V. Kiang, J. Lorme, and J.-P. Onnela, “New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research,” *JMIR Mental Health*, vol. 3, no. 2, p. e16, May 2016. [Online]. Available: <http://mental.jmir.org/2016/2/e16/> 20
- [76] T. R. Insel, “Digital phenotyping: A global tool for psychiatry,” *World Psychiatry*, vol. 17, no. 3, pp. 276–277, 2018. 21
- [77] S. D. Dlima, S. Shevade, S. R. Menezes, and A. Ganju, “Digital phenotyping in health using machine learning approaches: Scoping review,” *JMIR Bioinform Biotech*, vol. 3, no. 1, p. e39618, Jul 2022. [Online]. Available: <https://bioinform.jmir.org/2022/1/e39618> 21
- [78] A. M. Bernardos, M. Pires, D. Ollé, and J. R. Casar, “Digital phenotyping as a tool for personalized mental healthcare,” in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, ser. PervasiveHealth’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 403–408. [Online]. Available: <https://doi.org/10.1145/3329189.3329240> 21
- [79] N. Martinez-Martin, T. R. Insel, P. Dagum, H. T. Greely, and M. K. Cho, “Data mining for health: staking out the ethical territory of digital phenotyping,” *npj Digital Medicine*, vol. 1, no. 1, p. 68, 2018. [Online]. Available: <https://doi.org/10.1038/s41746-018-0075-8> 21
- [80] P. Bufano, M. Laurino, S. Said, A. Tognetti, and D. Menicucci, “Digital phenotyping for monitoring mental disorders: Systematic review,” *J Med Internet Res*, vol. 25, p. e46778, Dec 2023. [Online]. Available: <https://www.jmir.org/2023/1/e46778> 22, 23

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [81] Y. Zhang, C. Stewart, Y. Ranjan, P. Conde, H. Sankesara, Z. Rashid, S. Sun, R. J. B. Dobson, and A. A. Folarin, “Large-scale digital phenotyping: identifying depression and anxiety indicators in a general uk population with over 10,000 participants,” 2024, preprint on arXiv, submitted 24 September 2024. 22, 28
- [82] H. Wisniewski, P. Henson, and J. Torous, “Using a smartphone app to identify clinically relevant behavior trends via symptom report, cognition scores, and exercise levels: A case series,” *Frontiers in Psychiatry*, vol. Volume 10 - 2019, 2019. [Online]. Available: <https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2019.00652> 22
- [83] Y. Yuan, Y. Gao, H. Moen, E. Isometsä, P. Marttinen, and T. Aledavood, “Leveraging large language models for digital phenotyping: Detecting depressive state changes for patients with depressive episodes,” *medRxiv*, 2025. [Online]. Available: <https://www.medrxiv.org/content/early/2025/05/13/2025.05.10.25327354> 22, 25, 30
- [84] J. R. Beames, O. Dabash, M. J. Spoelma, A. Shvetcov, W. Y. Zheng, A. Slade, J. Han, L. Hoon, J. F. Kupper, R. Parker, B. Mitchell, N. G. Martin, J. M. Newby, A. E. Whitton, and H. Christensen, “Feasibility of collecting and linking digital phenotyping, clinical, and genetics data for mental health research: Pilot observational study,” *JMIR Form Res*, vol. 9, p. e71377, Jun 2025. [Online]. Available: <https://doi.org/10.2196/71377> 22, 25, 31
- [85] B. Kadirvelu, T. B. Bel, A. Freccero, M. D. Simplicio, D. Nicholls, and A. A. Faisal, “Digital phenotyping for adolescent mental health: A feasibility study employing machine learning to predict mental health risk from active and passive smartphone data,” 2025. [Online]. Available: <https://arxiv.org/abs/2501.08851> 23
- [86] J. Melcher, R. Hays, and J. Torous, “Digital phenotyping for mental health of college students: a clinical review,” *Evidence Based Mental Health*, vol. 23, no. 4, 2020. [Online]. Available: <https://doi.org/10.1136/ebmental-2020-300180> 23, 24
- [87] N. Martinez-Martin, H. T. Greely, and M. K. Cho, “Ethical development of digital phenotyping tools for mental health applications: Delphi study,” *JMIR Mhealth Uhealth*, vol. 9, no. 7, p. e27343, Jul 2021. [Online]. Available: <https://doi.org/10.2196/27343> 23, 24, 26, 27, 29
- [88] J. R. Beames, J. Han, A. Shvetcov, W. Zheng, A. Slade, O. Ibrahim, J. Rosenberg, B. O’Dea, S. Kasturi, L. Hoon, A. Whitton, H. Christensen, and J. M. Newby, “Use of smartphone sensor data in detecting and predicting depression and anxiety in young people (12–25 years): A scoping review,” *SSRN Electronic Journal*, 2024, preprint, posted 18 April 2024. [Online]. Available: <https://ssrn.com/abstract=4798261> 24
- [89] L. Marciano, E. Vocaj, M. A. Bekalu, A. La Tona, G. Rocchi, and K. Viswanath, “The use of mobile assessments for monitoring mental health in youth: Umbrella

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- review,” *J Med Internet Res*, vol. 25, p. e45540, Sep 2023. [Online]. Available: <https://www.jmir.org/2023/1/e45540> 24
- [90] E. Perlmutter, B. Dwyer, and J. Torous, “Social media and youth mental health: Assessing the impact through current and novel digital phenotyping methods,” *Current Treatment Options in Psychiatry*, vol. 11, no. 2, pp. 34–51, 2024. [Online]. Available: <https://doi.org/10.1007/s40501-024-00312-1> 24, 28
- [91] F. Annisa, “Understanding youth mental health: A narrative review,” *Nurse and Health: Jurnal Keperawatan*, vol. 14, no. 1, pp. 78–84, Jun. 2025. [Online]. Available: <https://ejournal-kertacendekia.id/index.php/nhjk/article/view/767> 24
- [92] M. Ali, C. Lucasius, T. P. Patel, M. Aitken, J. Vorstman, P. Szatmari, M. Battaglia, and D. Kundur, “Speech as a multimodal digital phenotype for multi-task llm-based mental health prediction,” 2025. [Online]. Available: <https://arxiv.org/abs/2505.23822> 25
- [93] J. P. M. Mendes, I. R. Moura, P. Van de Ven, D. Viana, F. J. S. Silva, L. R. Coutinho, S. Teixeira, J. J. P. C. Rodrigues, and A. S. Teles, “Sensing apps and public data sets for digital phenotyping of mental health: Systematic review,” *J Med Internet Res*, vol. 24, no. 2, p. e28735, Feb 2022. [Online]. Available: <https://www.jmir.org/2022/2/e28735> 25
- [94] Y. Hu, S. Zhang, T. Dang, H. Jia, F. D. Salim, W. Hu, and A. J. Quigley, “Exploring large-scale language models to evaluate eeg-based multimodal data for mental health,” in *UbiComp '24: Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2024, pp. 412–417, companion Proceedings. 26
- [95] A. Lanata, G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo, “Complexity index from a personalized wearable monitoring system for assessing remission in mental health,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 132–139, 2015. 26
- [96] G. Valenza, M. Nardelli, A. Lanatà, C. Gentili, G. Bertschy, R. Paradiso, and E. P. Scilingo, “Wearable monitoring for mood recognition in bipolar disorder based on history-dependent long-term heart rate variability analysis,” *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 5, pp. 1625–1635, 2014. 27
- [97] R. Wang, H. Aung, Min S. S. Abdullah, R. Brian, A. T. Campbell, T. Choudhury, M. Hauser, J. Kane, M. Merrill, E. A. Scherer, S. Tseng, Vincent W. and D. Ben-Zeev, “Crosscheck: Toward passive sensing and detection of mental health changes in people with schizophrenia,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*, 2016, pp. 886–897. 27

## Fenotipagem digital para a saúde mental: Indicadores comportamentais

- [98] L. Hassan, A. Milton, C. Sawyer, A. J. Casson, J. Torous, A. Davies, B. Ruiz-Yu, and J. Firth, “Utility of consumer-grade wearable devices for inferring physical and mental health outcomes in severe mental illness: Systematic review,” *JMIR Ment Health*, vol. 12, p. e65143, Jan 2025. [Online]. Available: <https://mental.jmir.org/2025/1/e65143> 27
- [99] I. Barnett, J. Torous, P. Staples, L. Sandoval, M. Keshavan, and J.-P. Onnela, “Relapse prediction in schizophrenia through digital phenotyping: a pilot study,” *Neuropsychopharmacology*, vol. 43, no. 8, pp. 1660–1666, 2018. [Online]. Available: <https://doi.org/10.1038/s41386-018-0030-z> 27
- [100] S. De Riba Mayoral and B. Revelles Benavente, “Hacia una pedagogia afectiva del movimiento,” *Tercio Creciente*, vol. 8, no. 2, Jun. 2019. [Online]. Available: <https://revistaselectronicas.ujaen.es/index.php/RTC/article/view/4413> 27, 28
- [101] M. Sheikh, M. Qassem, and P. A. Kyriacou, “Wearable, environmental, and smartphone-based passive sensing for mental health monitoring,” *Frontiers in Digital Health*, vol. Volume 3 - 2021, 2021. [Online]. Available: <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2021.662811> 27
- [102] S. Patel and K. Saunders, “Apps and wearables in the monitoring of mental health disorders,” *British Journal of Hospital Medicine*, vol. 79, no. 12, p. 672, 2018. [Online]. Available: <https://doi.org/10.12968/hmed.2018.79.12.672> 27
- [103] J. Torous, J. Linardon, S. B. Goldberg, S. Sun, I. Bell, J. Nicholas, L. Hassan, Y. Hua, A. Milton, and J. Firth, “The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality,” *World Psychiatry*, vol. 24, no. 2, pp. 156–174, 2025. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/wps.21299> 28, 30
- [104] D. Currey and J. Torous, “Increasing the value of digital phenotyping through reducing missingness: a retrospective review and analysis of prior studies,” *BMJ Mental Health*, vol. 26, no. 1, p. e300718, 2023. [Online]. Available: <https://doi.org/10.1136/bmjment-2023-300718> 29
- [105] E. Mokadam, H. G. Musa, A. Kumalasari, and F. Istighfaroh, “Public policy challenges in the digital age: Balancing innovation, governance, and equity,” *JIAN (Jurnal Ilmiah Administrasi Negara)*, vol. 9, no. 1, pp. 102–109, Jan. 2025. [Online]. Available: <https://ojs.ejournalunigoro.com/index.php/JIAN/article/view/1172> 29
- [106] P. Deshwal, “Data interfaces in mental health: Supporting awareness, not surveillance,” *Journal of Computer Science and Technology Studies*, vol. 7, no. 4, p. 37, 2025. [Online]. Available: <https://doi.org/10.32996/jcsts.2025.7.4.37> 30
- [107] K. Kasaudhan, “Ai-driven tools for detecting and monitoring mental health conditions through behaviour patterns,” *International Journal of Preventive Medicine and*

## **Fenotipagem digital para a saúde mental: Indicadores comportamentais**

*Health*, 2025. [Online]. Available: <https://doi.org/10.54105/ijpmh.b1054.05030325>  
30

- [108] C. Halkiopoulos and E. Gkintoni, “The role of machine learning in ar/vr-based cognitive therapies: A systematic review for mental health disorders,” *Electronics*, vol. 14, no. 6, 2025. [Online]. Available: <https://www.mdpi.com/2079-9292/14/6/1110> 31