

Depression Symptoms Identification through Social Media Data

Applying Design Science Research to Develop a Classification Model

Silas Lima Filho
silasfilho@ppgi.ufrj.br
Instituto de Computação, UFRJ
Cidade Universitária, RJ, Brasil

Monica Ferreira da Silva
mfsilvavmail@gmail.com
P. Pós-Graduação Informática, UFRJ
Cidade Universitária, RJ, Brasil

Jonice Oliveira
jonice@ic.ufrj.br
Instituto de Computação, UFRJ
Cidade Universitária, RJ, Brasil

ABSTRACT

The thesis addresses the prevalence of depression as a debilitating condition and underscores the importance of early identification of symptoms for timely interventions. Exploring user-generated content on social networks, the study proposes the use of machine learning models in detecting depressive symptoms. Following the DSR methodology, the research validates the effectiveness of these models compared to existing approaches, involving healthcare professionals and domain experts. The thesis introduces an innovative stacking model using LIWC metrics from social media posts, contributing to the understanding of machine learning-based solutions in identifying symptoms of depressive disorder.

KEYWORDS

transtorno depressivo, redes sociais, análise textual

1 INTRODUÇÃO

Mídias digitais são essenciais para a comunicação e conscientização sobre transtornos como a depressão, que segundo a OMS, atinge cerca de 300 milhões de pessoas [3]. Plataformas online permitem que usuários busquem ajuda e compartilhem suas experiências. Este artigo resume a pesquisa de doutorado sobre a identificação de sintomas de depressão em mídias sociais, abordando a compreensão do transtorno, o mapeamento do estado da arte, a colaboração com profissionais de saúde, a captura e análise de dados e a criação de um modelo de classificação. A pesquisa utilizou abordagens técnicas atuais e incluiu interações com profissionais da saúde, que foi crucial, para enriquecer a análise dos dados sobre discurso depressivo, que destaca o aspecto sociotécnico da pesquisa.

Esta pesquisa investigou como a computação pode contribuir no diagnóstico, identificação, tratamento e relato do transtorno depressivo usando dados de mídias sociais. O objetivo principal foi estudar características da depressão por meio de dados de mídias sociais, visando apoiar a identificação de sintomas.

Este trabalho aborda desafios sociotécnicos destacados por autores como Boscarioli [1] e Cafezeiro [2], que ressaltam a necessidade dessa integração para beneficiar a sociedade. Topi e Tucker [7] destacam que aspectos tecnológicos e sociais devem ser analisados juntos como um fenômeno emergente da interação entre ambos. Essa abordagem foi evidenciada em diálogos com profissionais de saúde, contribuindo para uma melhor compreensão da

depressão e das abordagens terapêuticas. Este trabalho desenvolveu uma pesquisa interdisciplinar, combinando dados de mídias sociais com análises quantitativas e interações com profissionais de saúde mental, resultando em um artefato útil para a comunidade de saúde. Destacamos que nossa pesquisa tem respaldo no comitê de ética de pesquisa da UFRJ. Dado que para construir um artefato final, seria necessário interagir com dados e participantes humanos. O CAAE da submissão na Plataforma Brasil é 54865821.5.0000.5263, e o número do parecer é 5.225.791.

A seguir relatamos parte do estado da arte mapeado na Seção 2, seguida pela Seção 3 que descreve a abordagem metodológica aplicada. Em sequência, a Seção 4 dividida em subseções que descrevem os resultados dos esforços da pesquisa bem como discussões sobre tais resultados, seguida da Seção 5 que conclui o artigo.

2 ESTADO DA ARTE

Para mapear o estado da arte, realizamos uma revisão sistemática da literatura (RSL) sobre a detecção de sintomas de depressão nas mídias sociais. A RSL foi realizada em duas etapas: primeiro, nos repositórios de computação (ACM e IEEE) de 2013 a 2018, totalizando 47 trabalhos; segundo, nos repositórios de saúde (Pubmed, APA e Scopus) de 2019 a 2021. O protocolo de RSL foi atualizado entre as duas etapas, mantendo algumas características em comum.

O aumento das publicações sobre a identificação de usuários depressivos nas mídias sociais, especialmente em 2021, reflete a crescente atenção no campo. Alguns estudos usaram modelos de aprendizado profundo, como *transformers*, enquanto outros utilizaram abordagens de aprendizado de máquina. Alguns também empregaram questionários psicológicos para testar ou filtrar usuários e validar dados com profissionais de saúde. No entanto, nem todos os trabalhos abordaram aspectos éticos ou incluíram discussões correlatas. Identificamos uma lacuna na utilização de técnicas de aprendizado para alavancar métricas do sistema LIWC. Nossa hipótese é a de melhor classificação pela combinação de modelos individuais e suas melhores características em um meta-modelo. Uma diferença ao comparar as duas RSLs é em relação ao envolvimento de profissionais de saúde na construção do modelo. Embora o sistema LIWC seja amplamente utilizado, apenas um estudo usa métricas gerais que abstraem características textuais como tom emocional e autenticidade.

3 METODOLOGIA DE PESQUISA

Utilizando a Design Science Research (DSR), obtivemos uma contribuição significativa para a área de pesquisa. A DSR é focada na criação de artefatos para resolver problemas específicos, além de preservar a relevância na união entre teoria e prática. Nesta tese, as *conjecturas teóricas* são fundamentadas em teorias computacionais

In: VI Concurso de Teses e Dissertações (CTD 2024). Anais Estendidos do XXX Simpósio Brasileiro de Sistemas Multimídia e Web (CTD'2024). Juiz de Fora/MG, Brazil. Porto Alegre: Brazilian Computer Society, 2024.
© 2024 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

e psicológicas. A teoria da computação cobre redes sociais, mídias sociais e algoritmos de aprendizado de máquina, enquanto a psicologia fornece definições de saúde mental, transtornos mentais e distúrbios depressivos. O *artefato* e as *conjecturas teóricas* da DSR devem interagir para validar e orientar sua construção. O *estado da técnica* envolve o processo de design do artefato, que, após ser construído, deve ser validado em um contexto específico.

4 MÉTODO DE IDENTIFICAÇÃO DE SINTOMAS DE DEPRESSÃO

4.1 Interação com Comunidade de Saúde

Um estudo com profissionais da saúde foi realizado após a RSL para entender os métodos de tratamento e identificação do transtorno depressivo utilizados por psicólogos, além de validar as features identificadas. Os desafios incluíram a contatação de profissionais e a explicação do vocabulário de computação (análises, tipos de dados, features) nas entrevistas e questionários.

O estudo ocorreu em duas etapas. Primeiro, entrevistamos três psicólogos para entender suas abordagens na identificação de sintomas e determinação do tratamento, além de validar as features levantadas na RSL. Em seguida, distribuímos um questionário em várias instituições de saúde para coletar dados quantitativos sobre a relevância das diferentes features obtidas em mídias sociais. Quarenta e nove participantes classificaram as features, que foram organizadas em quatro grupos principais: discurso, perfil do usuário, interação e grupos.

Nesse estudo, o grupo de features discursivas foi o mais relevante para identificar a depressão. As features mais importantes de tal grupo foram: *Autorreferência ao estado físico ou mental* e *Referência explícita a alguma doença ou assunto correlato* [5].

4.2 Análise de Discurso de Comunidades

Como as features discursivas foram as mais relevantes para os profissionais e exploradas na RSL, analisamos textos de mídias sociais. Utilizamos a ferramenta Crowdtangle para coletar dados de grupos público¹. A análise linguística foi realizada com o léxico LIWC. O módulo principal de análise do LIWC quantifica quatro aspectos: Pensamento Analítico, que mede o uso de palavras indicando padrões formais e lógicos; Influência, que avalia o status social e confiança na escrita; Autenticidade, que reflete a espontaneidade do discurso; e Tom Emocional, onde scores altos indicam tom positivo e scores baixos, tom negativo.

Analisamos a frequência das palavras nas postagens de 4 comunidades em inglês (835 publicações) e 7 comunidades em português (1945 publicações), coletadas em setembro de 2022. As palavras *anxiety*, *depression*, *people* e *life* foram as mais usadas em ambas as comunidades, variando apenas na ordem.

Os resultados mostram comportamentos similares nas comunidades em ambos os idiomas. Há uma concentração de scores baixos em Pensamento Analítico, indicando postagens com baixo teor formal. Na métrica de Influência, os valores também são baixos, indicando influência social moderada. Autenticidade possui scores altos, sugerindo veracidade no vocabulário. A maioria das postagens tem scores baixos em Tom Emocional, indicando discurso negativo.

¹www.crowdtangle.com/resources [acessado em 11-08-2024]

4.3 Modelo de Classificação

Nosso modelo de classificação foi treinado com dados científicos [4], baseado no modelo de [8], que usa sentimentos de postagens em mídias sociais como pesos em aprendizado ativo. Analisamos postagens de três subfóruns: *depression*, *fitness* e *divorce*, rotuladas como não depressivas (0) e depressivas (1). As postagens foram avaliadas em *Emotional Tone*, *Analytical Thinking*, *Clout* e *Authentic*.

Usamos o modelo BERT para vetorizar as postagens com a biblioteca *sentence-transformers* [6]. Realizamos o produto de Hadamard entre esses vetores e os vetores LIWC para capturar a influência combinada do conteúdo das postagens e das métricas LIWC. Após construir os modelos, reparamos o modelo para classificar dados de fóruns sobre depressão do Facebook buscando identificar padrões heurísticos. Observamos que mensagens depressivas geralmente tinham baixos escores de influência social, enquanto previsões negativas apresentavam altos escores de pensamento analítico, como mensagens de boas-vindas e apoio.

De 835 registros, o modelo classificou 23 posts como não depressivos. Em comunidades em português, de 1760 mensagens, 123 foram classificadas como não depressivas, muitas buscando orientação ou discutindo luto e tristeza. Apesar de mais erros devido à tradução, os resultados foram similares, com mensagens frequentemente buscando orientação profissional ou discutindo luto e perda.

5 CONCLUSÃO

Este artigo apresentou as principais contribuições da pesquisa de doutorado, destacando a interação com a comunidade de saúde e as análises de dados de mídias sociais para explorar padrões de saúde mental. Pesquisas futuras buscarão dados de alta qualidade sobre populações com depressão ou usuários diagnosticados clinicamente. A extração de heurísticas de discursos depressivos resultou em um modelo de classificação para identificar indicadores de saúde mental em ambientes digitais.

REFERENCES

- [1] C. Boscaroli, R. Araujo, and R. S. Maciel. 2017. *I GrandSI-BR Grand Research Challenges in Information Systems in Brazil 2016-2026*. SBC, Porto Alegre. <https://doi.org/10.5753/sbc.2884.0>
- [2] Isabel Cafezeiro, Jose Viterbo, Leonardo Costa, Luciana Salgado, Marcelo Rocha, and Rodrigo Monteiro. 2017. *Strengthening of the Sociotechnical Approach in Information Systems Research*. SBC, 133–147. <https://doi.org/10.5753/sbc.2884.0.11>
- [3] Munmun De Choudhury. 2014. Opportunities of Social Media in Health and Well-Being. *XRDS* 21, 2 (dec 2014), 23–27. <https://doi.org/10.1145/2676570>
- [4] Daniel M Low, Laurie Rumker, John Torous, Guillermo Cecchi, Satrajit S Ghosh, and Tanya Talkar. 2020. Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study. *Journal of medical Internet research* 22, 10 (2020), e22635.
- [5] Silas P. Lima Filho, Mônica Ferreira da Silva, Jonice Oliveira, and Livia Ruback. 2022. A Study about Gathering Features in Depression Detection' Problem with Health Professionals Community. *iSys - Brazilian Journal of Information Systems* 15, 1 (Oct. 2022), 10:1–10:26. <https://doi.org/10.5753/isys.2022.2285>
- [6] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. <https://arxiv.org/abs/1908.10084>
- [7] H. Topi and A. Tucker. 2014. *Computing Handbook: Information Systems and Information Technology*. CRC Press. <https://books.google.com.br/books?id=7hzcBQAAQBAJ>
- [8] Shou-Chun Yen, Kuo-Chung Chu, and Po-Yao Tsai. 2021. Prediction model of social network suicide ideation by small sample. In *2021 IEEE 22nd International Conference on Information Reuse and Integration for Data Science (IRI)*. IEEE, 385–389. <https://doi.org/10.1109/IRI51335.2021.00060>