

Características para detecção de depressão com dados do Twitter

Characteristics for depression detection using Twitter data

Características para detección de depresión con datos de Twitter

Ataíde Gualberto¹, Jugurta Montalvão²

1 Mestrando, Departamento de Engenharia Elétrica, UFS, São Cristóvão (SE), Brasil.

2 Professor Doutor, Departamento de Engenharia Elétrica, UFS, São Cristóvão (SE), Brasil.

Autor correspondente: (Mestrando) Ataíde Mateus Gualberto dos Santos

E-mail: mateusbirth@academico.ufs.br

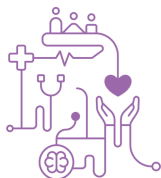
Resumo

Objetivo: Identificar características relevantes na detecção de pessoas com depressão a partir de postagens no Twitter. **Método:** Criação da base de dados, aplicação de técnicas de pré-processamento, seleção de características utilizando testes de hipóteses e o classificador AdaBoost, e verificação do tamanho do vocabulário. **Resultados:** O AdaBoost utilizou 40 classificadores, dos quais 38 verificavam a presença de palavras específicas no texto, atingindo uma acurácia de 73%. Verificou-se que o vocabulário de pessoas com depressão é menor que o de pessoas sem depressão. **Conclusão:** Verificar a presença de certas palavras nos *tweets* de pessoas depressivas é suficiente para obter resultados semelhantes a técnicas mais complexas. Além disso, o vocabulário de pessoas com depressão é menor, conforme demonstrado pela abordagem baseada na entropia de Shannon.

Descritores: Depressão; Reconhecimento Automatizado de padrão; Mineração de Dados.

Abstract

Objective: To identify the most relevant characteristics in detecting people with depression based on their Twitter posts. **Method:** Database creation, data preprocessing techniques, feature selection using hypothesis testing and AdaBoost classifier, and vocabulary size verification. **Results:** AdaBoost used 40 classifiers, 38 of which checked for the presence of specific words in the text, achieving an accuracy of



73%. It was found that the vocabulary of people with depression is smaller than that of people without depression. **Conclusion:** Checking for the presence of certain words in the tweets of depressed people is sufficient to achieve results close to those of more complex techniques. Additionally, the vocabulary of people with depression was shown to be smaller using a Shannon entropy-based approach.

Keywords: Depression; Pattern Recognition, Automated; Data Mining.

Resumen

Objetivo: Identificar características relevantes en la detección de personas con depresión a partir de sus publicaciones en Twitter. **Método:** Creación de la base de datos, aplicación de técnicas de preprocesamiento, selección de características utilizando pruebas de hipótesis y el clasificador AdaBoost, y verificación del tamaño del vocabulario. **Resultados:** AdaBoost utilizó 40 clasificadores, 38 de los cuales verificaban la presencia de palabras específicas en el texto, alcanzando una precisión del 73%. Se encontró que el vocabulario de personas con depresión es menor que el de personas sin depresión. **Conclusión:** Verificar la presencia de ciertas palabras en los tweets de personas depresivas es suficiente para obtener resultados cercanos a técnicas más complejas. Además, se demostró que el vocabulario de personas con depresión es menor utilizando un enfoque basado en la entropía de Shannon.

Descriptor: Depresión; Reconocimiento de Normas Patrones Automatizadas; Minería de Datos.

Introdução

Nos últimos anos, transtornos mentais têm se tornado uma preocupação significativa na saúde pública. Transtornos como a depressão possuem consequências devastadoras tanto para os pacientes quanto para suas famílias e, segundo a OMS¹, os transtornos depressivos são os mais comuns entre as doenças mentais, caracterizados por tristeza, perda de interesse e prazer, sentimentos de culpa ou baixa autoestima, distúrbios do sono ou apetite, cansaço e falta de concentração.

Nesse contexto, as redes sociais tornaram-se uma área relevante na pesquisa de saúde pública, permitindo a observação do bem-estar emocional das pessoas.

¹ <https://www.who.int/news-room/fact-sheets/detail/depression>



Indivíduos com depressão usam essas plataformas para discutir suas condições, compartilhar informações, buscar apoio social e reduzir o isolamento, especialmente após a pandemia de COVID-19.

Nesse cenário, estudos de processamento de linguagem natural, como o de Trifu *et al.*⁽¹⁾, mostram que os padrões de linguagem de pessoas depressivas mudam, com tendência a usar frases mais simples, curtas e frequentemente repetitivas. Esses achados foram observados tanto na linguagem escrita de indivíduos com depressão maior.

Similarmente, em Smirnova *et al.*⁽²⁾, pessoas com depressão leve mostraram um uso maior de coloquialismos, repetições, e sentenças com uma única oração em linguagem escrita. Os autores também notaram que, apesar dos pacientes proverem respostas mais longas, elas continham expressões incompletas, com omissão de palavras (elipses) e enunciados reduzidos.

Tanto nos dois estudos anteriores como em Rude, Gortner e Pennebaker⁽³⁾, o uso de pronomes em primeira pessoa demonstra ser um indicador determinante na classificação de pessoas com depressão. Os autores em Smirnova *et al.*⁽²⁾ sugerem que o alto uso de pronomes nessa categoria está relacionado a perda de significados específicos na fala e também pode ser interpretado como uma manifestação de empobrecimento semântico.

Outro fator que se mostra importante é o sexo do paciente. Na revisão conduzida por Liu *et al.*⁽⁴⁾, foi verificado que, entre alunos do ensino superior, as mulheres estavam mais propensas a ter sintomas depressivos. Na mesma pesquisa ainda encontraram correlação entre sintomas de depressão e níveis padrões de depressão, traços de personalidade (neuroticismo e psicoticismo), eventos estressantes de vida, abuso na infância (incluindo abuso sexual) e pensamentos negativos ou ruminação.

No entanto, a maioria desses estudos se concentra na língua inglesa. Até o momento da escrita deste trabalho, os autores têm conhecimento de apenas dois estudos que utilizaram dados de redes sociais em português brasileiro^(5, 6). Isso revela uma escassez de pesquisas que abordem a língua portuguesa e que incluam outras características das mídias sociais, como interações entre usuários e horários de publicação.



Diante disso, o objetivo deste trabalho é identificar as características mais relevantes na classificação de pessoas com depressão a partir de suas postagens da rede social *Twitter* (atual X). Especificamente, este estudo propõe-se a responder às seguintes perguntas:

- É possível desenvolver um classificador capaz de imitar a tarefa de rotulação manual de mensagens como (controle) escritas por pessoas sem depressão ou (positivos) escritas por pessoas sofrendo de depressão, a partir de características extraídas dos *tweets*?
- Quais são as características mais importantes utilizadas para realizar a classificação acima?
- O vocabulário de uma pessoa com depressão é menor do que uma pessoa sem depressão?

Métodos

Essa seção descreve como foi realizada a coleta e construção da base de dados, seu pré-processamento e sua forma de representação, bem como as características foram selecionadas, como o modelo de classificação foi construído e avaliado.

Coleta e Pré-Processamento de Dados

Inicialmente, foi realizado um procedimento de coleta de dados no *Twitter*, onde foi feita uma busca por postagens de usuários brasileiros que declararam terem sido diagnosticados com depressão (especificamente os termos “diagnosticado”, em ambos os gêneros, e “depressão”). Esta consulta foi submetida ao raspador *snsrape*², que retornou todas as ocorrências dos termos desde o início da pandemia de COVID-19 (03/11/20) até o dia 16/10/22. Além dos *tweets*, foram coletados alguns de seus metadados, como horário de publicação, número de *likes*, número de compartilhamentos, número de menções e o tipo de dispositivo pelo qual a postagem foi realizada.

² <https://github.com/JustAnotherArchivist/snsrape>



Em seguida, as declarações foram verificadas manualmente por uma equipe de 4 colaboradores a fim de selecionar apenas os usuários que de fato disseram terem sido diagnosticados por um especialista da área da saúde mental, como psiquiatras ou psicólogos. Ao fazer isso, frases que estivessem falando sobre outra pessoa, pedindo dinheiro a famosos ou que possuíssem tom cômico foram descartadas. A Tabela 1 apresenta alguns exemplos de *tweets* removidos bem como o motivo da remoção.

Tabela 1 – Exemplos de *tweets* removidos

Tweets	Motivo de retirada
Mas ele não só foi diagnosticado com depressão?	Falando de outra pessoa
Antes que você seja diagnosticado com depressão ou baixa auto estima, primeiro certifique-se que não esteja só cercado por idiotas.	Ironia
Já estava com vários sintomas, mas agora realmente fui diagnosticado com DPR: depressão pós reunião.	Brincadeira
@pessoa_famosa fui diagnosticado com depressão, mas não tenho dinheiro para comprar os remédios. Pode me ajudar?	Possível golpe

As postagens removidas foram colocadas numa lista de falsos positivos, para serem utilizados como exemplos negativos na base de dados. Vale ressaltar que os *user_ids* foram utilizados apenas para buscar as postagens públicas dos usuários, sendo, após a coleta, substituídos por um simples valor inteiro, sem relação com o *id* fornecido pelo *Twitter*.

Dessa forma, por estar utilizando informações de acesso público, de acordo com artigo primeiro da Resolução nº 510, de 7 de abril de 2016, esta pesquisa não precisou ser registrada nem avaliada pelo sistema CEP/CONEP.

Logo após a coleta e armazenagem dos *tweets*, todos dados textuais foram convertidos para letras minúsculas. Além disso, acentos, pontuações, caracteres especiais, *emojis*, menções a outros usuários e *links* foram removidos. No fim, foi empregada a técnica de radicalização nas palavras da base de dados.

Representação Numérica dos Dados

A base de dados é multimodal, contendo dados textuais, numéricos e temporais, exigindo uma representação única para capturar essas características. Inspirado no trabalho de Alsagri⁽⁷⁾, que usa uma matriz termo-documento para indicar a



frequência de palavras em cada *tweet*, este trabalho adota uma abordagem similar, onde cada elemento da matriz indica a presença da k -ésima palavra no n -ésimo *tweet*. A mesma abordagem foi utilizada para representar o tipo de aparelho utilizado para enviar as postagens. Já o horário de publicação foi convertido para um número inteiro no intervalo de 0 a 47, onde cada valor representa um intervalo de 30 minutos desde as 0h até às 23h59min59s. Por fim, os números de *likes*, compartilhamentos e menções já estão em formato numérico.

Seleção de Características

Para selecionar as características determinantes na detecção de depressão, foram utilizados testes de hipóteses. O propósito é o de selecionar apenas as características cujas distribuições de probabilidade sejam estatisticamente diferentes entre os conjuntos de pessoas que disseram ter depressão e os que foram marcados como falsos positivos.

Após esta etapa, ao estimar a distribuição de probabilidade (ou frequência) de cada indicador, é possível considerar cada característica como um classificador fraco, ou seja, um classificador levemente superior ao aleatório. Sendo assim, uma forma de combinar classificadores fracos em um classificador forte é utilizando a técnica do AdaBoost⁽⁸⁾. O AdaBoost cria e ajusta uma combinação linear das respostas dos classificadores fracos.

De fato, o AdaBoost pode combinar um número arbitrariamente alto de classificadores, até mesmo repetindo alguns, de forma a memorizar todos os exemplos de treinamento, mas tendo um péssimo desempenho nos exemplos de teste. Esse fenômeno é comumente conhecido como *overfitting*⁽⁹⁾ e uma forma de evitá-lo é utilizando exemplos de validação. Em especial, neste trabalho, foi utilizada a validação cruzada de k partições (*k-folds*)^(9, 10) para encontrar o número de classificadores fracos a ser utilizado.

Métricas de Avaliação

Para avaliar o desempenho dos modelos criados é necessário utilizar métricas objetivas. Existem quatro números que podem ser observados diretamente da saída do modelo e que representam de forma completa seu desempenho⁽¹¹⁾. São eles os



positivos verdadeiros (PV), positivos falsos (PF), negativos verdadeiros (NV) e negativos falsos (NF).

No entanto, na maioria das tarefas de classificação, são empregados índices mais fáceis de interpretar e que são calculados com os números acima. Estes são a acurácia, a precisão, a revocação e a medida F, cujas definições são apresentadas nas equações 1 a 4^(7, 10, 11).

$$\text{Acurácia: } A = \frac{PV+NV}{PV+NV+PF+NF} \quad (1)$$

$$\text{Precisão: } P = \frac{PV}{PV+PF} \quad (2)$$

$$\text{Revocação: } R = \frac{PV}{PV+NF} \quad (3)$$

$$\text{Medida F: } F = 2 \frac{P \cdot R}{P+R} \quad (4)$$

A acurácia é a razão entre o número de acertos e o número de exemplos, a precisão é a taxa de exemplos positivos verdadeiros em relação ao número de exemplos que o modelo previu que eram verdadeiros, a revocação é a taxa de positivos verdadeiros sobre todos os exemplos que eram verdadeiros e a medida F é a média harmônica entre a precisão e a revocação.

Tamanho do Vocabulário

Existem várias formas diferentes de se verificar o tamanho do vocabulário de um corpus. Neste trabalho foram utilizadas duas, sendo elas, o número de palavras únicas e a quantidade efetiva de palavras. A primeira apenas afere o tamanho do conjunto de palavras utilizadas, enquanto a segunda também leva em consideração a frequência relativa das palavras. Em específico, a equação 5 contém a definição de quantidade efetiva de palavras utilizada nesta pesquisa⁽¹²⁾:

$$C = 2^{-\sum_{x \in V} p_d(x) \cdot \log_2(p_d(x))} \quad (5)$$

Onde V é o vocabulário do corpus e $p_d(x)$ é a frequência relativa da palavra x ocorrer no discurso de uma pessoa com depressão.



Resultados e Discussão

Representação Numérica dos Dados

Na construção da base de dados, foram coletados 5122 *tweets* com 4829 usuários diferentes. Destes, 1879 usuários foram incluídos como pessoas que disseram ter depressão. Foi coletada ainda uma amostra de 1879 pessoas que foram marcadas como falsos positivos na etapa de rotulação manual para que a base de dados ficasse balanceada.

Com a base montada, os dados multimodais vindos do *Twitter* foram, então, transformados para uma forma numérica única. Cada *tweet* foi transformado num vetor $\{0, 1\}^{5902}$, sendo o número de colunas igual ao número de palavras únicas da base. Já os dispositivos foram representados como vetores $\{0, 1\}^{20}$, ou seja, havia 20 tipos de dispositivos diferentes na base.

Seleção de Características

Para a seleção de características foram utilizados testes de hipóteses com um $\alpha = 0,05$, que é um valor padrão na literatura. Foram selecionadas as palavras com os 100 menores p-valores, ou seja, as palavras com as maiores diferenças entre as distribuições de probabilidade (estimadas como frequências relativas) associadas a usuários que relataram ter depressão e aos falsos positivos. A Quadro 1 contém as dez palavras com os menores p-valores.

Quadro 1: As 10 palavras com menor p-valor.

As dez palavras com menor p-valor
● Fui
● Eu
● Ele
● Foi
● Pos
● Me
● Minha
● Anos
● Sou
● Ansiedade

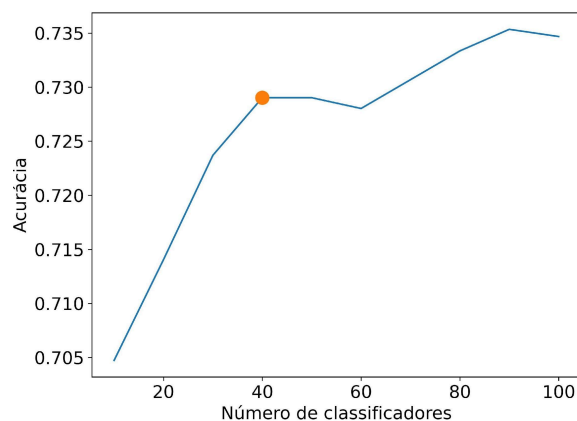
No caso dos dispositivos, apenas os do tipo *Android* passaram no teste de hipóteses com um p-valor $< 2\%$. A variável horário de publicação também apresentou



relevância estatística com um p-valor de 2%. Finalmente, o número de compartilhamentos teve um p-valor $< 0,02\%$. As outras características (número de *likes* e de menções) não apresentaram diferença estatística relevantes, com p-valores $> 5\%$.

Dessa forma, foram criados conjuntos de treinamento e teste com 103 características aprovadas nos testes de hipóteses. O conjunto de treinamento recebeu 80% dos dados (3006 exemplos) e o de teste 20% (752 exemplos). O treinamento foi dividido em 4 partes para validação cruzada, buscando evitar *overfitting*. A Figura 1 apresenta a evolução da acurácia na partição de validação conforme aumenta o número de classificadores.

Figura 1 - Acurácia do AdaBoost no conjunto de validação versus o número de classificadores.



Nota-se que a partir de 40 classificadores a acurácia começa a diminuir no conjunto de validação antes de voltar a subir. Porém, a velocidade de aumento é muito menor, indicando uma diminuição na capacidade de generalização do modelo mesmo com o aumento do número de parâmetros. Com isso, foi treinado outro AdaBoost com exatamente 40 classificadores e seu desempenho foi aferido. Os resultados encontram-se na Tabela 2.

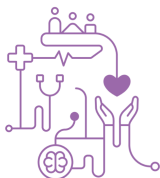


Tabela 2: Métricas de desempenho do AdaBoost no conjunto de teste.

Acurácia	Revocação	Precisão	Medida F
0,731	0,739	0,728	0,733

Pode-se observar que todas as métricas giram em torno de 73%, sugerindo que o modelo comete praticamente a mesma quantidade de erros dos tipos I e II. Estudos correlatos^(7, 10, 11) revelaram resultados similares em relação à medida F, variando de 73% a 79%. Entretanto, as demais métricas empregadas nessas pesquisas apresentaram variações significativas, indicando que os algoritmos utilizados acarretaram em erros de um dos tipos em maior proporção que do outro.

Uma das características do AdaBoost é que ele atribui coeficientes lineares a cada uma das respostas dos classificadores. Olhando a magnitude desses coeficientes, é possível inferir suas importâncias. A Tabela 3 apresenta alguns dos classificadores mais importantes do AdaBoost com 40 classificadores.

Tabela 3 - Alguns dos melhores classificadores utilizados pelo AdaBoost com 40 classificadores.

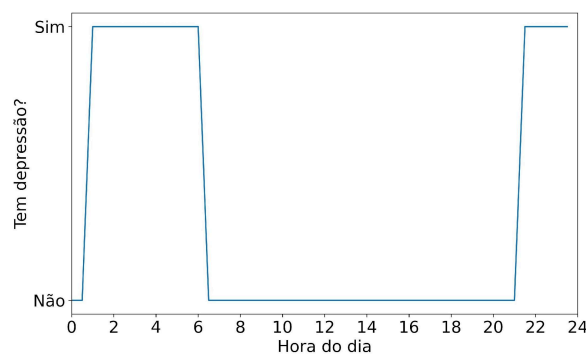
Classificador	Posição	Importância
Fui	1º	10,39%
Eu	2º	6,92%
Pos	3º	5,84%
Ansiedade	4º	4,74%
Sou	5º	4,22%
Ele	6º	3,92%
Foi	7º	3,81%
Anos	8º	3,49%
[horário de publicação]	29º	1,47%
[dispositivo]	30º	1,37%



Vale lembrar que para os classificadores do tipo textual, verificam se a presença da determinada palavra está mais associada a exemplos positivos ou negativos. De forma semelhante, o classificador do tipo dispositivo verifica se a maioria dos *tweets* de treinamento foi enviada por um aparelho *Android*. Já o classificador referente ao horário de publicação da postagem verifica se num horário específico o número de pessoas que disseram ter depressão era maior do que os de falsos positivos. A Figura 2 mostra a saída do classificador do tipo horário para cada um dos horários do dia.

Em paralelo a isso, outro conjunto de experimentos foi realizado na fase de seleção de características. Foi utilizada a correção de Bonferroni⁽¹³⁾ para compensar a quantidade de testes de hipóteses realizados. O α ficou em torno de $8,5 \times 10^{-6}$, e apenas 38 palavras passaram no teste de hipóteses. Outra alteração que ocorreu foi que o número de classificadores utilizados pelo AdaBoost caiu de 40 para 30, utilizando o mesmo procedimento de escolha descrito acima. No mais, as 20 características (todas palavras) mais importantes utilizadas pelo AdaBoost foram iguais tanto antes quanto depois da correção de Bonferroni e a acurácia do AdaBoost com a correção de Bonferroni caiu um ponto percentual, ficando em 72%.

Figura 2 – Evolução da classificação de pessoas depressivas ao longo do dia.



Nota-se que das 21h30min até às 6h30min praticamente todos os intervalos são associados com a classe de pessoas depressivas, exceto das 0h às 0h30min, possivelmente devido a alguma especificidade da amostra coletada para este trabalho. Um intervalo similar, das 23h às 6h, foi encontrado no trabalho de Leis *et al.*⁽¹⁴⁾, que utilizou dados do *Twitter* na língua espanhola. Uma explicação para isso pode ter



relação com o fato da população com depressão sofrer com distúrbios do sono³. Porém, é necessário lembrar que este classificador contribuiu apenas 1,47% nas decisões do AdaBoost. De fato, a acurácia do classificador horário é de apenas 52,9%, sendo levemente superior ao classificador binário aleatório.

Tamanho do Vocabulário

Na seção anterior, foi mostrado que o número de palavras únicas na base de dados é 5902. Ao realizar a contagem de palavras únicas das pessoas com depressão e sem depressão, foram encontradas, respectivamente, 3755 e 4382 palavras. Esses números, apesar de serem simples de obter, assumem que todas as palavras possuem a mesma probabilidade de ocorrer no texto, o que não é verdade.

Para contornar esse problema, foi utilizada uma métrica baseada na entropia de Shannon, descrita na equação 5, para calcular o tamanho efetivo do vocabulário dos usuários com e sem depressão. Esse cálculo retornou 367 palavras para os depressivos e 430 para os não depressivos. Esses valores podem ser entendidos como uma média ponderada da quantidade de palavras, refletindo a probabilidade ajustada de ocorrência no texto. A relevância estatística deste resultado foi aferida por um teste de hipóteses que verificou se as distribuições de frequência de uso das palavras das pessoas com e sem depressão são iguais. Este teste retornou um p-valor $< 10^{-12}$, indicando que, na verdade, as distribuições são diferentes.

Esses resultados reforçam aqueles encontrados em Trifu *et al.*⁽¹⁾, onde pessoas com transtorno depressivo maior demonstraram maior uso de frases curtas, truncadas, áridas, além de tautologias e elipses, em comparação com o grupo controle. Similarmente, o trabalho de Smirnova *et al.*⁽²⁾ encontrou maior uso de repetições léxicas e semânticas, além do uso predominante de frases com apenas uma oração, no grupo de pessoas com depressão leve. Já em Leis *et al.*⁽¹⁴⁾, os autores encontraram que o número de caracteres utilizados nas sentenças do grupo com depressão é menor do que o controle.

Uma explicação para o fato de diversos estudos encontrarem resultados similares, apesar das diferenças metodológicas, se dá por conta de o cérebro de pessoas com depressão funcionar de forma diferente de uma pessoa saudável⁽¹⁵⁾ e

³ <https://www.who.int/news-room/fact-sheets/detail/depression>



isso influencia, entre outras coisas, as formas de comunicação da pessoa depressiva. De fato, o estudo de Trifu *et al.*⁽¹⁾ fala sobre como mudanças linguísticas estão associadas a déficits na memória de trabalho, alternância de tarefas, planejamento estratégico, atenção e velocidade psicomotora, além de mostrar que essas mudanças na linguagem refletem a gravidade e a natureza dos déficits cognitivos presentes na depressão.

Limitações e Trabalhos Futuros

Este estudo apresenta algumas limitações que devem ser claramente apontadas. As limitações incluem: a base de dados composta apenas por pessoas que autodeclararam ter depressão, sem confirmação clínica; a restrição na busca de termos utilizados para coletar *tweets* sobre depressão; a ausência de dados socioeconômicos e demográficos dos usuários; a possibilidade de haver outras doenças além da depressão entre os participantes; e a falta de investigação sobre os tipos específicos de depressão dos usuários.

Trabalhos futuros podem tentar encontrar características específicas para diferentes tipos de depressão. Outra possibilidade é a utilização de representações semânticas e *embeddings*, muito utilizados em ferramentas de geração de texto e imagens como o *ChatGPT*⁴ ou *DALL-E*⁵, de forma a representar os dados numa dimensão que reflita mais fielmente a complexidade real desses dados. Por fim, é possível aumentar a quantidade de dados, utilizando a mesma plataforma (*Twitter*), ou buscando dados de outras redes sociais.

Conclusão

Este trabalho propõe o estudo de quais são as características relevantes na classificação de pessoas com depressão, utilizando dados do *Twitter* incluindo suas postagens, número de *likes*, menções e compartilhamentos, além do horário de publicação do *tweet* e do tipo de dispositivo pelo qual a postagem foi realizada. Testes

⁴ <https://chatgpt.com/>

⁵ <https://labs.openai.com/>



de hipóteses e o AdaBoost foram utilizados para selecionar os melhores indicadores presentes nos dados.

Foi mostrado que, verificando apenas a presença de determinadas palavras nos *tweets*, bem como o horário de publicação e se o dispositivo utilizado foi do tipo *Android*, é possível desenvolver um classificador capaz de simular a rotulação manual de exemplos verdadeiros positivos contra falsos positivos com acurácia e medida F de 73%, o que está próximo de resultados encontrados na literatura.

Baseado em trabalhos relacionados, investigou-se o tamanho do vocabulário de pessoas com depressão usando duas abordagens: contagem direta de palavras únicas nos *tweets* e entropia de Shannon, que considera a frequência das palavras. Constatou-se que o vocabulário é menor no grupo com depressão em comparação ao grupo controle em ambos os métodos analisados.

A contribuição principal deste trabalho está nos fatos de que (i) apenas juntando classificadores fracos é possível simular a rotulação manual realizada por humanos e (ii) o tamanho do vocabulário de pessoas com depressão é menor do que o da população geral.

Agradecimentos

A CAPES, pela bolsa de pesquisa de mestrado cujo número do processo é 88887.712271/2022-00.

À Professora Jislane Silva Santos Menezes e aos meus colegas Lucas Macena Vasconcelos e Viviane Andrade de Souza, todos do Instituto Federal de Sergipe, pela ajuda na coleta e rotulação dos dados.

Referências

1. Trifu RN, et al. Linguistic indicators of language in major depressive disorder (MDD). An evidence based research. *J Evid Based Psychother.* 2017;17(1).
2. Smirnova D, et al. Language patterns discriminate mild depression from normal sadness and euthymic state. *Front Psychiatry.* 2018;9:105.
3. Rude S, Gortner EM, Pennebaker J. Language use of depressed and depression-vulnerable college students. *Cogn Emot.* 2004;18(8):1121-1133.



4. Liu Y, et al. Predictors of depressive symptoms in college students: A systematic review and meta-analysis of cohort studies. *J Affect Disord.* 2019;244:196-208.
5. Santos WRD, de Oliveira RL, Paraboni I. SetembroBR: a social media corpus for depression and anxiety disorder prediction. *Lang Res Evaluat.* 2023.
6. Mann P, Paes A, Matsushima EH. See and read: detecting depression symptoms in higher education students using multimodal social media data. In: *Proceedings of the International AAAI Conference on Web and Social Media.* 2020. p. 440-451.
7. Alsagri HS, Ykhlef M. Machine learning-based approach for depression detection in Twitter using content and activity features. *IEICE Trans Inf Syst.* 2020;103(8):1825-1832.
8. Schapire RE. The boosting approach to machine learning: An overview. *Nonlinear Estim Classif.* 2003;149-171.
9. Domingos P. A few useful things to know about machine learning. *Commun ACM.* 2012;55(10):78-87.
10. Moreira LB, Namen AA. Sistema preditivo para a doença de Alzheimer na triagem clínica. *J Health Inform.* 2016;8(3).
11. Islam MR, et al. Depression detection from social network data using machine learning techniques. *Health Inf Sci Syst.* 2018;6:1-12.
12. Montalvao J, et al. On the representation of sparse stochastic matrices with state embedding [Pré-print]. Available at SSRN 4605637.
13. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic Physiol Opt.* 2014;34(5):502-8.
14. Leis A, et al. Detecting signs of depression in tweets in Spanish: behavioral and linguistic analysis. *J Med Internet Res.* 2019;21(6):e14199.
15. Disner SG, Beevers CG, Haigh EA, Beck AT. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci.* 2011;12(8):467-77.