

Análise de Sentimentos com VADER: Identificação de Sinais de Depressão e Ansiedade em postagens em Português

Rayane Oliveira do Nascimento¹, Odara Sena dos Santos Feitosa²

¹Instituto Federal de Educação, Ciência e Tecnologia do Ceará - Aracati - CE - Brasil

rayane.oliveira.nascimento08@aluno.ifce.edu.br, odara@ifce.edu.br

Abstract. *The intensive use of social networks, associated with factors such as isolation and internalization of feelings, worsens mental health problems. This study aims to evaluate the performance of the VADER tool in identifying signs of anxiety and depression in short microblog texts written in Portuguese. To this end, a search string was created with terms related to these disorders, from which a set of 10.020 sentences was collected. These sentences underwent a manual annotation process, with the support of a psychologist, and were subsequently classified by VADER. The results indicated that the tool had significant limitations, with an overall accuracy of only 48,8%.*

Resumo. *O uso intensivo das redes sociais, associado a fatores como isolamento e internalização de sentimentos, agrava problemas de saúde mental. Este estudo tem como objetivo avaliar o desempenho da ferramenta VADER na identificação de sinais de ansiedade e depressão em textos curtos de microblog, escritos em português. Para isso, foi criada uma string de busca com termos relacionados a esses transtornos, a partir da qual foi coletado um conjunto de 10.020 frases. Essas frases passaram por um processo de anotação manual, com o apoio de uma psicóloga, e foram posteriormente classificadas pelo VADER. Os resultados indicaram que a ferramenta apresentou limitações significativas, com acurácia geral de apenas 48,8%.*

1. Introdução

No fim de 2019, o mundo enfrentou a pandemia de COVID-19, uma doença respiratória que perdurou em estado de urgência até 2023. O [Worldometers 2024] contabilizou até abril de 2024 aproximadamente 7 milhões de mortes no mundo. Além dos impactos na saúde física, a pandemia também representou um marco para a saúde mental global, evidenciando a urgência em compreender e promover o bem-estar emocional das pessoas.

O isolamento social, necessário no período pandêmico, provocou mudanças significativas na rotina das pessoas. No contexto educacional, esse afastamento impactou diretamente o convívio entre alunos, professores e colegas, deslocando parte do processo de ensino-aprendizagem para o ambiente virtual. Nesse momento, as redes sociais destacam-se como meio de interação, informação e até apoio emocional entre estudantes. No entanto, o uso intenso da tecnologia também provoca efeitos negativos. [Harari 2018], observa que, com o tempo, a tecnologia nos afastou do corpo e da atenção plena, deslocando nosso foco para smartphones e computadores.

O cenário pós-pandemia resultou em um aumento significativo nos casos de instabilidade emocional. Segundo a [WHO et al. 2022], o primeiro ano de pandemia causou

um aumento de 27,6% nos casos de transtorno depressivo maior e 25,6% nos casos de ansiedade. Em outro estudo [WHO 2025], é evidenciado uma relação entre suicídio e transtornos mentais, sendo o suicídio a terceira principal causa de morte entre jovens de 15 a 29 anos — faixa etária que corresponde, em grande parte, ao público mais ativo nas redes sociais.

Esse aumento nos quadros de sofrimento psíquico também tem impactos diretos no contexto educacional. Estudantes, especialmente do ensino médio e superior, enfrentaram desafios emocionais agravados pela pandemia, como solidão, ansiedade de desempenho e dificuldades de adaptação ao ensino remoto. Esses fatores contribuem para quadros de evasão, queda de rendimento e desmotivação. Estudos recentes destacam que o fechamento prolongado das instituições de ensino e a transição abrupta para o ensino remoto afetaram não apenas o aprendizado, mas também a saúde mental dos estudantes, gerando sentimentos de insegurança, estresse e isolamento [Sousa et al. 2024, Rosa et al. 2021].

Nesse cenário, torna-se fundamental compreender como tecnologias podem ser utilizadas não apenas como ferramentas de ensino, mas também como instrumentos de apoio à saúde mental de discentes. A análise de sentimentos, como define [Liu 2012], estuda como as pessoas expressam sentimentos em relação a diversos tipos de entidades. Essa abordagem aplicada a conteúdos de redes sociais se torna útil para identificar sinais precoces de sofrimento emocional, permitindo que gestores e profissionais da educação desenvolvam políticas de acolhimento mais eficazes.

Entre as ferramentas disponíveis para esse tipo de análise, destaca-se o léxico VADER (*Valence Aware Dictionary and Sentiment Reasoner*) projetado especificamente para medir sentimentos expressos em textos curtos e informais, como os encontrados em redes sociais do tipo microblog [Hutto and Gilbert 2021].

Dessa forma, este estudo investiga a eficácia do VADER na identificação de sinais de depressão e ansiedade em publicações de redes sociais escritas em português, com foco na plataforma Bluesky. O objetivo principal é avaliar o desempenho da ferramenta na detecção de sentimentos associados a essas condições emocionais, utilizando as métricas de precisão, *recall* e *F1-score*. Considerando o impacto direto desses estados emocionais no desempenho e permanência de estudantes, especialmente no contexto pós-pandêmico, este trabalho também busca discutir o potencial de modelos computacionais no apoio à identificação precoce de sinais de sofrimento psíquico em contextos educacionais mediados por tecnologias digitais.

2. Referencial Teórico

2.1. Ansiedade

O conceito de ansiedade nem sempre é claro, muitas vezes confundida com medo. Enquanto o medo é uma reação imediata a um perigo real, como forma de nos proteger fisicamente, a ansiedade é um estado de humor pessimista em relação ao futuro. De acordo com [Beck 2013], os sintomas se manifestam de diversas formas: Os sintomas fisiológicos incluem, por exemplo, aumento da frequência cardíaca, falta de ar e tensão muscular. Já os cognitivos abrangem medos irracionais e dificuldade de concentração. Os comportamentais dizem respeito à evitação de situações ameaçadoras, enquanto os afetivos envolvem sentimentos como medo, irritabilidade e impaciência.

2.2. Depressão

A origem da depressão pode estar relacionada à fatores biológicos, genéticos e psicológicos, e se apresenta de diferentes maneiras. De acordo com o DSM-5, esses transtornos compartilham sintomas como humor triste, vazio ou irritável, além de alterações somáticas e cognitivas que afetam a vida do indivíduo. O que os diferencia são fatores como duração, momento de ocorrência e causas presumidas [Association. and Association. 2013].

Segundo [Beck and Alford 2016], a depressão pode ser caracterizada por uma alteração específica no humor, como tristeza, solidão e apatia; um autoconceito negativo associado a autorrecriminações e autoacusações; desejos regressivos e autopunitivos, como a vontade de se afastar, esconder-se ou morrer; alterações vegetativas, incluindo anorexia, insônia e perda da libido; e alterações no nível de atividade, como retardo psicomotor ou agitação.

2.3. Análise de Sentimentos

A análise de sentimentos (AS) é uma subárea do Processamento de Linguagem Natural (PLN). De forma geral, a PLN possibilita que os computadores entendam e interpretem a linguagem humana. Já a análise de sentimentos analisa textos a fim de identificar a sua subjetividade. Essa análise pode ser feita em diferentes níveis de detalhamento. No nível de documento, avalia-se a opinião geral expressa em todo o texto; no nível de sentença, analisa-se o sentimento expresso em frases isoladas; e no nível de aspecto, examina-se a opinião relacionados a atributos específicos de uma entidade [Liu 2012].

Esse processo inclui uma série de etapas, ilustradas por [Medhat et al. 2014], que buscam transformar dados textuais em informações interpretáveis sobre sentimentos. Inicialmente identificam-se trechos nos quais os autores expressam julgamentos ou avaliações. Em seguida, entem-se quais sentimentos estão sendo expressos — se positivos, negativos ou neutros. Por fim, os dados são classificados quanto à sua polaridade, atribuindo-se uma valência emocional que reflete o tom geral da opinião.

2.4. VADER

O VADER (Valence Aware Dictionary for Sentiment Reasoning) é uma ferramenta de análise de sentimentos baseada em léxico e regras, com foco em mídias sociais, e sensível à polaridade e intensidade dos textos. [Hutto and Gilbert 2014] documentaram que a ferramenta superou os avaliadores humanos individuais na classificação correta de sentimentos de tuítes (Precisão de Classificação F1 = 0,96 e 0,84, respectivamente).

A ferramenta também considera elementos que tendem a amenizar ou intensificar os sentimentos: como a palavra *muito*, textos em caixa alta, exclamações e uso de emojis. Essa abordagem é especialmente adequada para a análise de textos curtos, típicos de redes sociais. Como resultado são retornados quatro valores que variam entre -1 (muito negativo) e 1 (muito positivo), sendo eles: *neg*, que representa a proporção de palavras negativas no texto; *neu*, a proporção de palavras neutras; *pos*, a proporção de palavras positivas; e *compound*, que é um resumo da polaridade geral do texto.

O valor *compound* é calculado como a soma ponderada das valências atribuídas a cada palavra do texto, ajustada por intensificadores ou atenuadores, e normalizada em um intervalo entre -1 e 1.

2.5. Bluesky

Bluesky é uma rede social descentralizada. Baseada no Protocolo AT (*Authenticated Transfer Protocol*), que permite que diversos servidores diferentes compartilhem informações entre si, oferecendo ao usuário maior liberdade e autonomia. Conforme [Kleppmann et al. 2024], a rede cresceu para mais de 10 milhões de usuários em outubro de 2024.

Idealizada por Jack Dorsey, ex-CEO do Twitter, o *Bluesky* compartilha semelhanças com a rede que o inspirou, mas apresenta funcionalidades distintas. A plataforma permite que os usuários personalizem seus feeds, escolham ou criem algoritmos de curadoria, e migrem entre servidores, mantendo o controle sobre suas identidades digitais.

Segundo [Duarte 2025], atualmente a rede conta com mais de 33 bilhões de usuários, cerca de um terço deles tem entre 18 e 24 anos. O Brasil ocupa o segundo lugar no ranking de visitantes do Bluesky.

3. Trabalhos relacionados

Para selecionar os trabalhos, foi realizada uma busca nas bases: Portal de Periódicos da CAPES e Google Acadêmico, utilizando os termos "*sentiment analysis*" e "*VADER*", com filtro para publicações dos últimos cinco anos (2020 - 2024).

Em [Teixeira and Chen 2021] foram selecionadas frases do Twitter com base em uma string contendo termos positivos e negativos. Após limpeza dos dados e anotação manual das frases, emojis foram inseridos aleatoriamente no final das frases. Foi realizada uma classificação com o objetivo de avaliar se os emojis influenciam o resultado da análise. Os resultados revelam que a utilização de emojis impacta negativamente a classificação feita pelo VADER, devido à forma como esses símbolos são traduzidos e interpretados pela ferramenta.

No estudo de [Nurcahyawati and Mustaffa 2023], o objetivo foi detectar a orientação do cliente em relação à qualidade do produto ou serviço. Na anotação dos dados é feito um comparativo entre a anotação manual e a realizada pelo VADER, e para classificação foi utilizado o algoritmo *Support Vector Machine*. Nos resultados, fica evidente que a orientação do cliente é em sua maioria positiva e que o modelo aumentou o valor de precisão em relação a anotação manual — de 86% para 88,57%.

Em [Badriya Murdhi Alenzi 2022], foi avaliado o desempenho dos léxicos TextBlob e VADER na classificação de tuítes em árabe sobre vacinas contra a COVID-19. Após a exclusão de conteúdos irrelevantes e tradução dos textos para o inglês, os resultados mostraram variações na precisão de acordo com o contexto. No cenário positivo, o TextBlob atingiu 75% de precisão e o VADER, 70%.

[Ammar Oad 2021] analisou o impacto da adição de 14 sinais de pontuação na classificação do VADER. O conjunto de dados foi selecionado do site Kaggle, tokenizado e submetido à classificação pelo léxico, primeiro sem e depois com cada sinal de pontuação adicionado individualmente ao final da frase. Por fim, observou-se que o uso das pontuações provoca variação nas polaridades. Por exemplo, os sinais de exclamação e interrogação aumentam as polaridades positiva e negativa enquanto diminuem a neutra.

Os estudos acima evidenciam que o VADER tem sido amplamente utilizado na análise de sentimentos em diferentes domínios, majoritariamente em língua inglesa e em

contextos voltados a produtos ou eventos de interesse geral. No entanto, poucos estudos exploram sua aplicação em contextos sensíveis como saúde mental e em textos em língua portuguesa. Diante disso, esta pesquisa busca contribuir para a literatura ao investigar a eficácia do método na identificação de sinais de ansiedade e depressão em postagens em português.

4. Metodologia

Essa seção descreve de forma estruturada as etapas metodológicas adotadas ao longo do desenvolvimento do estudo, com o objetivo de garantir transparência e a reprodutibilidade da pesquisa.

Seguindo uma ordem cronológica, podemos considerar o passo a passo: (I) planejamento do estudo; (II) definição da string de busca; (III) coleta de dados; (IV) anotação manual das publicações coletadas; (V) classificação dos textos com o léxico VADER;

4.1. Planejamento

Inicialmente, a coleta de dados seria feita no Twitter (atual X), mas a restrição da API em 2023 e a suspensão temporária da plataforma no Brasil em 2024 inviabilizaram essa opção. Por isso, utilizou-se a rede social Bluesky, que oferece acesso público via API e possui formato similar ao Twitter.

Optou-se por analisar textos em português, apesar do VADER ter sido desenvolvido para inglês, visando investigar suas limitações nesse idioma e preencher uma lacuna na literatura.

A anotação dos dados contou com o suporte de uma profissional de psicologia, que ajudou a definir critérios clínicos para selecionar frases relacionadas a sintomas de ansiedade e depressão, garantindo a relevância e ética da análise, sem realizar diagnósticos clínicos.

4.2. String de Busca

Para compor a string de busca voltada à identificação de sentimentos ansiosos e depressivos, foram selecionados termos recorrentes em fontes digitais e no Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM), resultando em uma lista de unigramas, bigramas e trigramas. Esses n-gramas representam palavras isoladas ou combinadas, como “tristeza”, “baixa autoestima” e “não aguento mais”, facilitando a identificação de conteúdos emocionalmente relevantes.

Duas técnicas foram testadas na formação da string: a lematização, que reduz a palavra à sua forma base, conforme encontrada no dicionário, e o stemming, que a reduz à sua raiz, removendo sufixos. Neste trabalho, a lematização foi realizada com a biblioteca **spaCy**, utilizando um modelo pré-treinado para o português (*pt_core_news_sm*), e o stemming com a biblioteca **NLTK**, por meio do algoritmo *RSLPStemmer*.

Entretanto, as técnicas mencionadas só podem ser aplicadas a unigramas. Para identificar o método mais eficiente, foi conduzido um experimento, no qual quatro abordagens foram avaliadas: a aplicação da lematização apenas nos unigramas; a lematização nos unigramas com a adição de bigramas e trigramas; o uso de stemming nos unigramas; e, por fim, a aplicação de stemming nos unigramas combinada com bigramas e trigramas.

Cada método foi testado com cinco coletas pela API do Bluesky, totalizando 500 frases por abordagem. As postagens foram anotadas manualmente pela autora em conjunto com a psicóloga Raynara Oliveira do Nascimento¹, a fim de identificar sinais de ansiedade ou depressão.

A Tabela 1 mostra os resultados. A abordagem com melhor desempenho foi a lematização simples, escolhida como padrão para a coleta final. A string resultante reuniu mais de 50 termos como tristeza, angústia, insegurança, suicídio, insônia, entre outros.

Tabela 1. Comparação entre métodos para construção da string.

Método	Total de frases	Frases ansiosas/depressivas
Lematização	500	64
Lematização + bigramas e trigramas	500	57
Stemming	500	17
Stemming + bigramas e trigramas	500	28

4.3. Coleta de Dados

Com a string de busca definida, iniciou-se a extração dos dados diretamente da API do Bluesky, com a filtragem de postagens em português. Cada requisição retornava, no máximo, 100 resultados e, para evitar duplicação e garantir maior diversidade de textos, as requisições foram realizadas em dias e horários distintos. O processo de coleta ocorreu entre 07/11/2024 e 10/03/2025, resultando em um total de 10.020 frases, que foram armazenadas em uma planilha para posterior classificação.

O conjunto de dados resultante está disponível publicamente para consulta e download [Oliveira 2025].

4.4. Anotação Manual

A planilha gerada na etapa anterior, continha três colunas principais: uma com as frases, outra com a classificação feita pela autora e a terceira com a classificação da psicóloga. As marcações seguiram o seguinte padrão: frases com indícios de ansiedade ou depressão receberam a tag *A/D*; dúvidas foram sinalizadas com *?*; e, frases fora do escopo do estudo permaneceram sem marcação. Além disso, foram incluídas colunas auxiliares para contabilizar o número de frases classificadas por cada avaliadora e os consensos obtidos entre elas. Frases sem consenso foram excluídas do experimento.

A avaliação seguiu alguns critérios para a receberem a TAG ou não:

Critérios para receberem a TAG "A/D"(Frases ansiosas/depressivas)

- **Uso de primeira pessoa:** O autor expressa diretamente seus sentimentos.
- **Afirmção explícita do transtorno:** O autor menciona que tem ansiedade ou depressão.
- **Descrição de sintomas ou sentimentos:** Frases que descrevem sensações típicas de ansiedade ou depressão.
- **Tristeza prolongada:** Relato de tristeza ou sentimentos depressivos recorrentes.

¹Formada pela Unijaguaripe, CRP 11/18932.

Cr terios para frases livre de TAG (N o ansiosas/depressivas)

- **Fala sobre terceiros:** Refer ncia   ansiedade ou depress o de outras pessoas.
- **Uso humor stico ou ir nico:** Tom de brincadeira, ironia ou humor ao abordar o tema.
- **Discuss o cient fica, informativa ou opinativa:** abordam os transtornos de forma te rica ou expressam apenas opini o sobre o tema.
- **M sicas ou trechos de textos citados:** Quando a frase n o reflete o sentimento real do autor, mas sim uma cita  o.
- **Conte do vago:** Frases curtas ou amb guas, sem sinais claros.
- **Fora do contexto analisado:** Aus ncia de rela  o com o tema.

4.5. Classifica  o com VADER

O modelo VADER foi aplicado diretamente sobre as frases originais, sem pr -processamento, a fim de preservar elementos informais como emojis e g rias, que podem sinalizar sintomas de ansiedade e depress o.

O c digo foi escrito em Python e utilizou duas bibliotecas principais: **Pandas**, para leitura e manipula  o dos dados, e **vaderSentiment**, para o c lculo do sentimento de cada frase. Neste trabalho, seguimos o crit rio proposto na documenta  o do modelo: frases com pontua  o *compound* superior a 0,05 foram classificadas como *positivas*, inferiores a -0,05 como *negativas*, e entre esses valores como *neutras*.

Para alinhar com os objetivos deste estudo, frases classificadas como *negativas* foram interpretadas como contendo sintomas emocionais *Com sintoma*, enquanto frases positivas e neutras foram agrupadas como *Sem sintoma*.

Devido   distribui  o desequilibrada entre frases com e sem sintomas, realizamos um balanceamento dos dados, aplicando a t cnica de subamostragem — por meio da fun  o *resample* da biblioteca *sklearn.utils*. A quantidade de frases da classe majorit ria — *Sem sintoma* — foi reduzida para igualar o n mero de inst ncias da classe minorit ria — *Com sintoma*. Essa decis o assegurou uma avalia  o mais equilibrada do desempenho do modelo.

A performance da classifica  o foi avaliada por meio de m tricas como precis o, recall e F1-score, al m da visualiza  o da matriz de confus o.

5. Resultados

Os resultados da classifica  o autom tica indicam limita  es significativas na identifica  o de frases associadas a sintomas de ansiedade ou depress o. A acur cia geral do modelo foi de 0.488 (48,8%), o que revela uma performance abaixo do esperado para um modelo bin rio. A tabela 2 apresenta as m tricas por classe:

Com rela  o   precis o do modelo, ao prever que uma frase n o apresenta sintomas, ele acerta em quase metade dos casos: 0.492. Para a classe *Com sintoma*, a precis o foi ainda menor: 0.466, o que indica que a maioria das frases classificadas como contendo sinais de sofrimento emocional n o correspondia a casos reais.

Tabela 2. Desempenho do classificador VADER por classe.

Classe	Precisão	Recall	F1-score	Suporte
Sem sintoma	0.492	0.805	0.611	1326
Com sintoma	0.466	0.170	0.249	1326
Acurácia geral	-	-	0.488	2652

A classe *Sem sintoma* obteve recall de 0.805, indicando que o modelo identificou corretamente a maioria das frases neutras ou positivas. Em contrapartida, a classe *Com sintoma* apresentou recall de apenas 0.170, demonstrando grande dificuldade do modelo em reconhecer sinais emocionais negativos — aspecto fundamental para os objetivos deste estudo.

A Figura 1 apresenta a matriz de confusão correspondente, que revela que o modelo classificou erroneamente 1.101 frases com sintoma como *Sem sintoma* (falsos negativos), e 258 frases sem sintoma como *Com sintoma* (falsos positivos). O número de acertos foi limitado a 225 verdadeiros positivos e 1.068 verdadeiros negativos, o que reforça a assimetria de desempenho entre as classes.

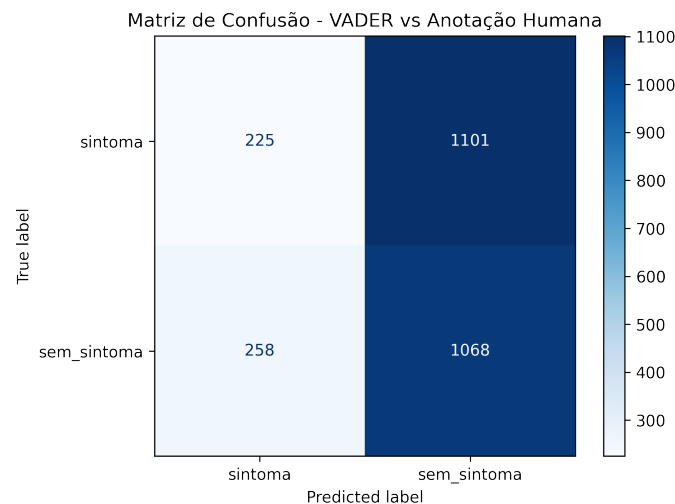


Figura 1. Matriz de Confusão

Para ilustrar melhor essas limitações, a Tabela 3 apresenta uma amostra de frases do conjunto de dados, acompanhadas das anotações manuais e das classificações atribuídas pelo VADER, bem como observações sobre os possíveis motivos de erro:

Tabela 3. Exemplos de limitações na classificação do VADER em frases reais.

Frase	Anotação Manual	Anotação VADER	Observação
Desculpa, me bateu uma crise de futuro.	Sintoma	Neutro	Expressa mal-estar emocional, mas sem vocabulário emocional explícito.

Frase	Anotação Manual	Anotação VADER	Observação
Lá vai eu ter outra crise pensando nisso	Sintoma	Neutro	Apesar do conteúdo emocional, a ausência de termos negativamente marcados impede a correta detecção.
Segunda semana do ano e ontem tive a pior crise de ansiedade da minha vida quase vomitei etc [emoji de rosto maravilhado] vamo q vamo	Sintoma	Neutro	Apesar de relatar um episódio negativo, o uso de humor — marcado pelo emoji e pela expressão "vamo q vamo" — confunde a ferramenta.

Essas limitações explicam, em parte, o baixo desempenho observado nas métricas. A primeira está relacionada ao idioma: como o VADER foi desenvolvido para o inglês, ele não reconhece bem gírias, construções informais e expressões típicas do português. A segunda diz respeito à natureza dos sintomas analisados: muitas frases expressam sofrimento emocional de forma sutil, sem vocabulário claramente negativo, o que dificulta a detecção por modelos baseados em regras simples. Além disso, o VADER foi originalmente aplicado em contextos genéricos, como avaliações de produtos e comentários online, o que limita sua eficácia em cenários mais sensíveis como o da saúde mental.

6. Considerações Finais

Este estudo destacou a importância de ferramentas para identificar sinais de sofrimento emocional, especialmente diante do aumento desses casos no pós-pandemia. Com foco no público jovem, mais ativo nas redes sociais e afetado por questões de saúde mental, que em grande parte compõe o ambiente estudantil. Embora a análise dos dados não se limite ao tema educação, é inegável que os impactos da saúde mental refletem nesse espaço, comprometendo o aprendizado, o engajamento e o bem-estar dos estudantes.

O léxico VADER apresentou desempenho limitado em frases sobre saúde mental em português, evidenciando suas restrições fora do contexto original para o qual foi projetado. Isso reforça a importância de usar ferramentas adaptadas à língua e cultura locais, especialmente para detectar sinais sutis de sofrimento emocional, contribuindo de forma mais eficaz para ações preventivas, inclusive no contexto educacional.

Referências

- Ammar Oad, Imtiaz Hussain Koondhar, P. K. B. L. H. A. O. A. K. S. R. B. (2021). Vader sentiment analysis without and with english punctuation marks.
- Association., A. P. and Association., A. P. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. American Psychiatric Association Arlington, VA, 5th ed. edition.
- Badriya Murdhi Alenzi, Muhammad Badruddin Khan, M. H. A. H. A. K. J. S. M. A. A. A. (2022). Automatic annotation performance of textblob and vader on covid vaccination dataset. *Intelligent Automation & Soft Computing*, 34(2):1311–1331.

- Beck, A. and Alford, B. (2016). *Depressão: Causas e Tratamento*. Artmed Editora.
- Beck, A. T. (2013). *Terapia cognitiva para trastornos de ansiedad*. Desclee de brouwer.
- Duarte, F. (2025). Bluesky user age, gender, & demographics. Exploding Topics. Publicado em 2 meses atrás (em 2025). Disponível em: <https://explodingtopics.com/blog/bluesky-users>. Acesso em: 13 jul. 2025.
- Harari, Y. N. (2018). *21 lições para o século 21*. Editora Companhia das Letras.
- Hutto, C. and Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1):216–225.
- Hutto, C. J. and Gilbert, E. (2021). *VADER: Valence Aware Dictionary and sEntiment Reasoner*. Disponível em: https://vadersentiment.readthedocs.io/_/downloads/en/latest/pdf/. Acesso em: 7 jul. 2025.
- Kleppmann, M., Frazee, P., Gold, J., Graber, J., Holmgren, D., Ivy, D., Johnson, J., Newbold, B., and Volpert, J. (2024). Bluesky and the at protocol: Usable decentralized social media. In *Proceedings of the ACM Conext-2024 Workshop on the Decentralization of the Internet (DIN '24)*, page 9, New York, NY, USA. ACM.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan amp; Claypool Publishers.
- Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093–1113.
- Nurcahyawati, V. and Mustaffa, Z. (2023). Vader lexicon and support vector machine algorithm to detect customer sentiment orientation. *Journal of Information Systems Engineering and Business Intelligence*, 9(1):108–118.
- Oliveira, R. (2025). raynasc/anxiety-depression-dataset-pt: Versão inicial (v1.0).
- Rosa, C. d. M., Santos, F. F. T. d., and Gonçalves, A. M. (2021). Los efectos de la pandemia de la covid-19 en la permanencia en la educación superior. el escenario de una universidad federal brasileña. *Revista Iberoamericana de Educación*, 86(2):61–76.
- Sousa, C., Fierro, M., Veraksa, A., Sá, S., Oliveira, L., Gomes, C., Pimentel, G., de Araújo, G. C. C., and Holanda, P. (2024). JOVENS UNIVERSITÁRIOS, PANDEMIA e EDUCAÇÃO REMOTA: CONSTATAÇÕES e DESAFIOS. *Integración y Conocimiento*, 2(13):1093–1113.
- Teixeira, G. O. and Chen, Z. (2021). Avaliação do impacto de emojis utilizando a ferramenta de análise de sentimentos vader.
- WHO (2025). Suicídio. <https://www.who.int/news-room/fact-sheets/detail/suicide>. Acesso em: 7 jul. 2025.
- WHO et al. (2022). Mental health and covid-19: early evidence of the pandemic’s impact: scientific brief, 2 march 2022. Technical report, World Health Organization.
- Worldometers (2024). Coronavirus update. <https://www.worldometers.info/coronavirus/>. Acesso em: 7 jul. 2025.