

MMDRS - Metodologia para Monitoramento de Depressão em Redes Sociais

Alexandre M. da Cunha¹, Rodolpho da S. Nascimento¹, Flavio Carvalho¹,
Gustavo Guedes¹

¹Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET/RJ)
R. Gen. Canabarro, 485 - Maracanã, Rio de Janeiro - RJ – Brasil, 20271-204

{alexandre.cunha, rodolpho.nascimento}@eic.cefet-rj.br,
flavio.carvalho@eic.cefet-rj.br, gustavo.guedes@cefet-rj.br

Abstract. *This work proposes a methodology for detecting and intervening in cases of depressive manifestations on social networks, using Affective Computing and Machine Learning techniques. Based on this methodology, the SaveMeNow prototype was developed, which can monitor platforms such as Facebook and Twitter, classify potentially depressive messages, and direct them for professional intervention. The system is divided into two main modules: Monitoring and Assistance, detailed with their respective technologies and processes implemented. The proposal aims to speed up the identification of signs of depression on social networks and provide quick and effective interventions, contributing to the reduction of suicide rates.*

Resumo. *Este trabalho propõe uma metodologia para detecção e intervenção nos casos de manifestações depressivas em redes sociais, utilizando Computação Afetiva e técnicas de Aprendizagem de Máquina. Baseado nessa metodologia, desenvolveu-se o protótipo SaveMeNow, que pode monitorar plataformas como Facebook e Twitter, classificando mensagens potencialmente depressivas e direcionando-as para intervenção profissional. O sistema é dividido em dois módulos principais: Monitoramento e Atendimento, detalhados com suas respectivas tecnologias e processos implementados. A proposta visa agilizar a identificação de sinais de depressão em redes sociais e proporcionar intervenções rápidas e eficazes, contribuindo para a redução das taxas de suicídio.*

1. Introdução

O Transtorno Depressivo Recorrente (CID 10 – F33) é uma doença psiquiátrica crônica [CID10 1992] que tem recebido crescente atenção tanto nos meios de comunicação quanto em grupos de apoio [Glenn et al. 2020]. De acordo com a Organização Mundial de Saúde (OMS), atualmente mais de 300 milhões de pessoas sofrem dessa condição, e entre 2005 e 2015 houve um aumento de 18% no número de diagnósticos de depressão [WHO 2017a]. Esses dados têm incentivado a OMS a promover campanhas globais voltadas ao diagnóstico e tratamento da depressão [WHO 2017b].

Uma grande proporção de indivíduos que tentaram o suicídio apresenta condições psiquiátricas, como depressão ou outros transtornos do humor [Barros et al. 2017, Glenn et al. 2020]. Indivíduos com depressão crônica têm um risco elevado de suicídio e,

em muitos casos, relatam já terem tentado suicídio [Ernst et al. 2019]. Órgãos e entidades públicas, bem como grupos de apoio, têm reagido continuamente para diminuir as taxas de suicídio, buscando estratégias e adotando medidas que ofereçam ajuda a indivíduos depressivos. Entre essas medidas, destacam-se a distribuição automatizada de questionários de registro eletrônico de saúde e a emissão de alertas automatizados aos médicos sobre indicadores de alto risco [Simon et al. 2021].

O suicídio ocasionalmente é anunciado previamente em comentários expostos em plataformas de redes sociais *online*. Alguns usuários utilizam essas redes como método de desabafo e busca por ajuda, tentando aliviar os sintomas depressivos [Ivanich et al. 2021]. Esse cenário não é incomum, dado que diversos estudos destacam que pessoas com depressão tendem a compartilhar mais informações pessoais *online* do que pessoas saudáveis [Park et al. 2012, Caplan 2003]. Usuários com manifestações depressivas podem sofrer tanto que, em alguns casos, expõem seus contatos pessoais (e.g., número de telefone) juntamente com súplicas por ajuda [Ivanich et al. 2021].

Embora diversos trabalhos tenham sido desenvolvidos para detectar depressão em português do Brasil ([da Silva Nascimento et al. 2018, Nascimento et al. 2019, Cortes and de Oliveira Melo 2021, Filho et al. 2024, Santos et al. 2024]), não encontramos propostas de intervenção para o problema na literatura revisada. Neste contexto, o presente trabalho tem como objetivo utilizar fundamentos da Computação Afetiva (CA) [Picard 2000] e técnicas de *Machine Learning* (ML) para propor uma metodologia que possa contribuir para o desenvolvimento de sistemas capazes de detectar potenciais manifestações depressivas em textos em português do Brasil e agilizar a intervenção. Em seguida, apresenta a construção de um protótipo de sistema com base nessa metodologia, o SaveMeNow, um sistema projetado para monitorar redes sociais, analisando entradas em tempo real e classificando mensagens com manifestações depressivas. Após essa triagem automatizada, as mensagens são colocadas em uma fila para que especialistas possam dar o encaminhamento mais adequado. Se constatada a necessidade de apoio emocional e prevenção do suicídio, o especialista pode acessar a entrada original para interagir diretamente com o usuário.

O restante deste trabalho encontra-se dividido da seguinte maneira: a Seção 2 discute alguns trabalhos relacionados à detecção em textos de manifestações depressivas e de conteúdo relacionado a suicídio. A Seção 3, apresenta a metodologia proposta. A Seção 4 descreve uma implementação dessa metodologia. Por fim, a Seção 5 apresenta as considerações sobre o trabalho, limitações e possibilidades de trabalhos futuros.

2. Trabalhos Relacionados

Sistemas desenvolvidos para ajudar a prevenir o suicídio, monitorando textos publicados em redes sociais, foram propostos por Varathan e Talib [Varathan and Talib 2014] e Desmet e Hoste [Desmet and Hoste 2014]. Esses trabalhos utilizam a detecção automática de mensagens com conteúdo relacionado ao suicídio, incluindo o registro de ameaças. Nota-se uma convergência nesses estudos no uso do Twitter como a rede social principal para detectar mais rapidamente indícios de suicídio.

Com uma proposta de identificar e prevenir comportamentos depressivos e suicidas, Calderon-Vilca *et al.* propuseram um programa para avaliar o nível de tendência suicida de usuários adolescentes [Calderon-Vilca et al. 2017]. Utilizaram dados sintéticos

gerados por simulação do comportamento da população adolescente com tendência ao suicídio para testar o sistema. Desta forma, o trabalho não apresenta resultados da análise de publicações reais.

Parraga-Alava et al. propuseram uma abordagem para categorizar potenciais mensagens de suicídio em redes sociais com base em aprendizagem não supervisionada [Parraga-Alava et al. 2019]. Os autores calcularam a similaridade das publicações em relação aos textos presentes em um corpus voltado para a detecção de suicídio utilizando medidas de similaridade semântica. O estudo em questão demonstrou a eficácia de uma abordagem desenvolvida para o agrupamento de mensagens nas línguas espanhola e inglesa. Contudo, diferentemente da pesquisa aqui apresentada, não foram explorados ou discutidos resultados aplicados à língua portuguesa

Sharma et al. desenvolveram um sistema para a detecção precoce de tendências suicidas entre jovens [Sharma et al. 2021]. No estudo apresentado, os autores detalham o treinamento de um modelo utilizando cartas de suicídio autênticas e publicações de diversas fontes, com o objetivo de gerar uma pontuação associada ao risco de suicídio. Entretanto, após a geração dessa pontuação, o sistema não implementa nenhuma ação subsequente, distinguindo-se da abordagem proposta no presente trabalho.

Nos estudos mencionados, além das diferenças previamente destacadas, observa-se a ausência de propostas voltadas ao tratamento de publicações em língua portuguesa, como o que foi desenvolvido neste trabalho. Destaca-se também que, embora os sistemas abordados sejam capazes de processar dados de redes sociais, eles não são projetados para acompanhar múltiplas redes de forma simultânea.

3. Metodologia

A metodologia proposta é desenvolvida para monitorar e agilizar a intervenção de especialistas em casos de conteúdo relacionado a suicídio e manifestações depressivas em textos publicados em plataformas de redes sociais *online*. Esta metodologia foi projetada para permitir a monitoração simultânea de textos em diversas redes sociais, incluindo Facebook, Twitter (X) e Instagram. O objetivo é identificar rapidamente sinais de depressão e comportamento suicida para que intervenções oportunas possam ser realizadas.

Essa seção detalha cada passo do processo de monitoramento e intervenção, conforme ilustrado na Figura 1. O processo começa com a busca automatizada de dados em redes sociais, seguida pela classificação de indícios de depressão utilizando algoritmos de *Machine Learning*. Os dados classificados são então armazenados em um banco de dados acessível para especialistas autorizados, que podem visualizar, assumir casos e descrever soluções apropriadas. Este fluxo de trabalho estruturado assegura que a intervenção seja realizada de maneira eficiente e organizada, proporcionando um suporte robusto e eficaz.

A metodologia é dividida em duas partes principais: o Módulo de Monitoramento (MM) e o Módulo de Atendimento (MA). O Módulo de Monitoramento (MM) abrange os passos 1, 2, 3, 4 e 5, que envolvem a busca de dados em redes sociais, o salvamento dos dados coletados, o pré-processamento, a classificação dos índices de depressão e o armazenamento dos dados classificados em um banco de dados. Já o Módulo de Atendimento (MA) consiste nos passos 6, 7, 8 e 9, que incluem o login do especialista, a visualização das mensagens classificadas, a assunção dos casos e a descrição das soluções propostas para cada caso identificado. A Figura 1 apresenta uma descrição detalhada dos passos.

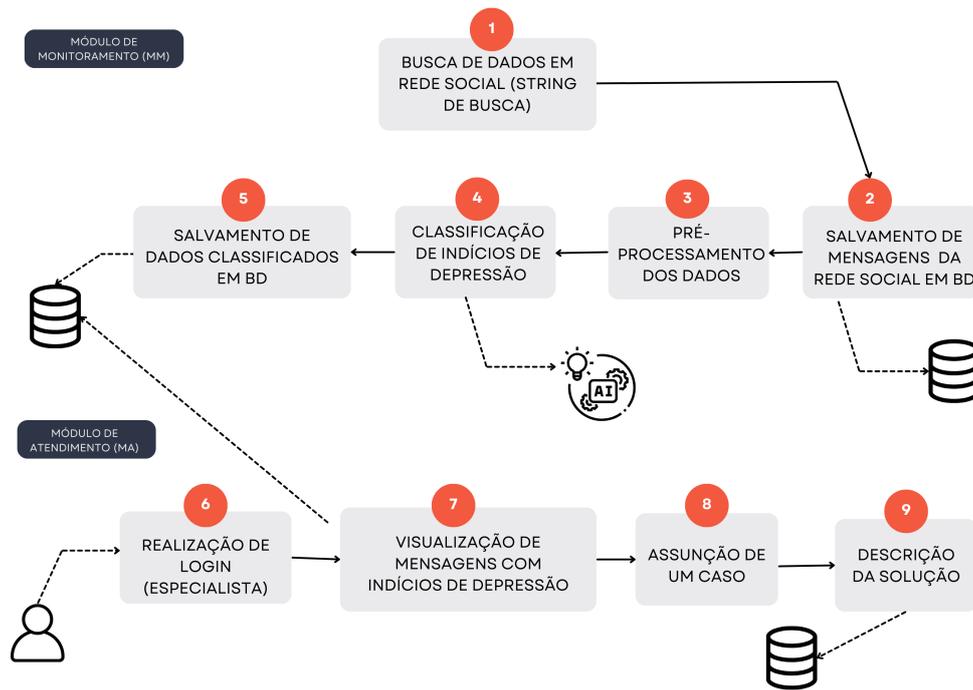


Figura 1. Fluxo de Trabalho da metodologia proposta.

3.1. Módulo de Monitoramento (MM)

O Módulo de Monitoramento (MM) realiza a busca automática de dados em redes sociais, armazena esses dados, classifica os índices de depressão utilizando algoritmos de *Machine Learning* e salva os dados classificados em um banco de dados para análise posterior. A descrição dos passos consiste em:

- Passo 1 - Busca de Dados em Rede Social: O processo inicia com a busca automática de dados em redes sociais. São utilizadas *strings* de busca específicas para identificar postagens que contenham indícios de depressão ou comportamento suicida.
- Passo 2 - Salvamento de Mensagens da Rede Social em BD: Os dados coletados são armazenados em um banco de dados (BD). Este armazenamento inicial é crucial para assegurar que todas as informações relevantes sejam mantidas para análise posterior.
- Passo 3 - Pré-processamento dos dados: No pré-processamento dos dados, diversas etapas são realizadas para preparar os textos coletados das redes sociais para a análise e classificação. Por exemplo, pode-se destacar a remoção de hiperlinks e caracteres especiais (e.g., *, \$).
- Passo 4 - Classificação de Indícios de Depressão: Utilizando algoritmos de *Machine Learning*, o sistema analisa os dados armazenados e classifica as postagens de acordo com a presença de indícios de depressão. Esta etapa é fundamental para filtrar os dados que realmente necessitam de intervenção. É importante destacar que esse passo necessita de um modelo pré-treinado de *Machine Learning*.
- Passo 5 - Salvamento de Dados Classificados em BD: Os dados que foram classificados como contendo indícios de depressão são salvos em um banco de dados

separado. Este banco de dados é acessível para os especialistas envolvidos no processo de intervenção.

3.2. Módulo de Atendimento (MA)

O Módulo de Atendimento (MA) permite que especialistas autorizados façam login, visualizem as mensagens classificadas, assumam casos e descrevam as soluções propostas, garantindo uma intervenção organizada e eficiente. A descrição dos passos consiste em:

- Passo 6 - Realização de Login (especialista): Especialistas previamente autorizados fazem login no sistema. Este controle de acesso garante que apenas pessoas qualificadas possam visualizar e interagir com os dados sensíveis.
- Passo 7 - Visualização de Mensagens com Índices de Depressão: Após o login, os especialistas têm acesso a uma interface para visualizar todas as mensagens que foram classificadas como contendo indícios de depressão. Esta visualização é essencial para a identificação rápida e eficaz dos casos que necessitam de intervenção imediata.
- Passo 8 - Assunção de um Caso: O especialista pode assumir a responsabilidade por um caso específico, sinalizando no sistema que aquele caso está sendo tratado. Este controle permite um acompanhamento organizado e evita duplicidade de esforços.
- Passo 9 - Descrição da Solução: Finalmente, o especialista fornece uma descrição detalhada da solução ou intervenção proposta. Esta descrição é registrada no sistema para acompanhamento e análise posterior, permitindo a geração de relatórios e a avaliação da eficácia das intervenções realizadas.

4. Implementação de protótipo

Esta seção detalha a implementação do protótipo SaveMeNow, dividindo a descrição nos dois módulos principais da metodologia: o Módulo de Monitoramento (MM) e o Módulo de Atendimento (MA). Cada módulo é descrito em seu atual estado da pesquisa. Vale destacar que esta é uma pesquisa em andamento.

4.1. Módulo de Monitoramento (MM)

Para o desenvolvimento do MM, utilizamos a linguagem de programação **Python**, que escolhemos por sua ampla disponibilidade de bibliotecas para operações de *Machine Learning* (ML) e por ser multiplataforma. Inicialmente, geramos seis modelos de classificação, pois esses modelos são necessários no passo 4 da metodologia proposta. Utilizamos a biblioteca **scikit-learn**¹ para isso. Os classificadores de ML que selecionamos para a versão implementada neste trabalho foram: Multinomial Naive Bayes, SVM, Random Forest, Decision Tree, Ada Boosting e Multilayer Perceptron (MLP).

Para implementar o Passo 1 do Módulo de Monitoramento (MM), utilizamos a API do Twitter² (atualmente denominado X). Para implementação do Passo 2 e Passo 5, adotamos um banco de dados relacional, empregando a linguagem SQL, e o sistema gerenciador de banco de dados foi a versão 10 do MariaDB³.

¹<https://scikit-learn.org/>

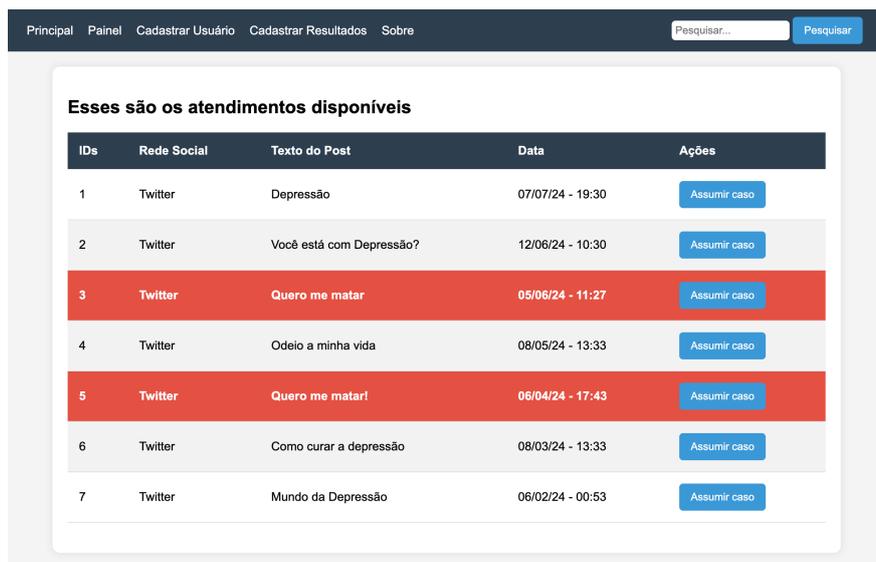
²<https://developer.x.com/docs/twitter-api>

³<https://mariadb.org/>

Durante a etapa de pré-processamento (Passo 3), tratamos os dados para remover *stop words*, hiperlinks e caracteres especiais (e.g., #, %, *, \$) aplicando a biblioteca NLTK⁴. Para lidar com diferentes formas de escrita que se desviam do vocabulário ortográfico padrão (e.g., goxtou, companhia), utilizamos um dicionário de português-BR do software LibreOffice⁵ em conjunto com a medida Jaro-Winkler [Winkler 1999] para encontrar a maior similaridade entre a variação e a forma ortográfica padrão.

4.2. Módulo de Atendimento (MA)

No estágio atual dessa pesquisa, implementamos o Módulo de Atendimento (MA) apenas em sua interface gráfica. A Figura 2 exibe as opções disponíveis no painel do menu localizado na parte superior da tela principal do sistema: “Principal”, “Painel”, “Cadastrar Usuário”, “Cadastrar Resultados” e “Sobre”, além de um campo para pesquisa com seu respectivo botão de acionamento. Os dados foram dispostos no painel em formato de tabela com os seguintes nomes de colunas: “IDs”, “Rede Social”, “Texto do post” e “Data”. Respectivamente, esses campos representam um identificador único dos casos disponíveis para atendimento, o nome da rede social de onde a entrada foi extraída, o texto da entrada e o campo com data e hora da entrada. À direita destes campos, estão posicionados botões de “Ações”.



The screenshot shows the main interface of the SaveMeNow system. At the top, there is a navigation bar with menu items: "Principal", "Painel", "Cadastrar Usuário", "Cadastrar Resultados", and "Sobre". To the right of the navigation bar is a search input field labeled "Pesquisar..." and a "Pesquisar" button. Below the navigation bar, the main content area displays a table titled "Esses são os atendimentos disponíveis". The table has five columns: "IDs", "Rede Social", "Texto do Post", "Data", and "Ações". There are seven rows of data. The third and fifth rows are highlighted in red, indicating urgent cases. Each row has a blue "Assumir caso" button in the "Ações" column.

IDs	Rede Social	Texto do Post	Data	Ações
1	Twitter	Depressão	07/07/24 - 19:30	Assumir caso
2	Twitter	Você está com Depressão?	12/06/24 - 10:30	Assumir caso
3	Twitter	Quero me matar	05/06/24 - 11:27	Assumir caso
4	Twitter	Odeio a minha vida	08/05/24 - 13:33	Assumir caso
5	Twitter	Quero me matar!	06/04/24 - 17:43	Assumir caso
6	Twitter	Como curar a depressão	08/03/24 - 13:33	Assumir caso
7	Twitter	Mundo da Depressão	06/02/24 - 00:53	Assumir caso

Figura 2. Tela principal do SaveMeNow, exibindo o código de identificação da publicação, a plataforma de rede social, a data da publicação e o botão para assumir o caso. No exemplo em questão, os casos identificados como urgentes pelo sistema são exibidos em vermelho.

5. Conclusões

Neste trabalho, apresentamos uma metodologia para a detecção e intervenção em casos de depressão, utilizando Computação Afetiva e técnicas de *Machine Learning*. Baseado nessa metodologia, desenvolvemos o protótipo SaveMeNow, implementado para monitorar o Twitter, classificando mensagens potencialmente depressivas e direcionando-as para intervenção profissional.

⁴<http://www.nltk.org/howto/portuguese.html>

⁵<https://cgit.freedesktop.org/libreoffice/dictionaries/plain/ptBR/ptBR.dic>

Uma das limitações no desenvolvimento foi o custo elevado da API de nível básico do Twitter no valor de \$100. Isso torna inviável a utilização contínua sem financiamento adicional. O pacote básico do SaveMeNow precisa ser ajustado para operar dentro dessas restrições financeiras, limitando a abrangência do monitoramento.

Como trabalhos futuros, pretende-se concluir o desenvolvimento do Módulo de Atendimento, atualmente apenas com interface gráfica. Além disso, pretende-se utilizar dicionários afetivos para melhorar a acurácia dos classificadores, como por exemplo o dicionário do LIWC 2015 em português do Brasil [Carvalho et al. 2024]. Também propomos aprimorar as técnicas de classificação textual, que atualmente enfrentam desafios contextuais pelos algoritmos utilizados. Técnicas baseadas em redes neurais e modelos de linguagem, mais robustas para essas tarefas, podem ser implementadas, incluindo o uso de APIs de modelos de linguagem de grande escala (LLMs).

Para superar as limitações, futuros trabalhos incluem a submissão de projetos de fomento para obter recursos financeiros que cubram os custos da API do Twitter e o desenvolvimento completo do Módulo de Atendimento. Essas iniciativas são essenciais para garantir monitoramento contínuo e eficaz. Além disso, proporcionarão intervenções rápidas, contribuindo para a redução das taxas de suicídio.

Referências

- Barros, J., Morales, S., Echávarri, O., García, A., Ortega, J., Asahi, T., Moya, C., Fishman, R., Maino, M. P., and Núñez, C. (2017). Suicide detection in Chile: proposing a predictive model for suicide risk in a clinical sample of patients with mood disorders. *Brazilian Journal of Psychiatry*, 39(1):1–11.
- Calderon-Vilca, H. D., Wun-Rafael, W. I., and Miranda-Loarte, R. (2017). Simulation of suicide tendency by using machine learning. In *2017 36th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–6. IEEE.
- Caplan, S. E. (2003). Preference for Online Social Interaction: A Theory of Problematic Internet Use and Psychosocial Well-Being. *Communication Research*, 30(6):625–648.
- Carvalho, F., Junior, F. P., Ogasawara, E., Ferrari, L., and Guedes, G. (2024). Evaluation of the Brazilian Portuguese version of Linguistic Inquiry and Word Count 2015 (BP-LIWC2015). *Language Resources and Evaluation*, 58(1):203–222.
- CID10 (1992). *Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - CID-10*. Organização Mundial da Saúde. Disponível em: <https://icd.who.int/browse10/2019/en>. Acesso em: 25 jul. 2024.
- Cortes, O. A. C. and de Oliveira Melo, W. E. (2021). Utilizando análise de sentimentos e SVM na classificação de tweets depressivos. *Anais do Computer on the Beach*, 12:102–110.
- da Silva Nascimento, R., Parreira, P., dos Santos, G. N., and Guedes, G. P. (2018). Identificando sinais de comportamento depressivo em redes sociais. In *Anais do VII Brazilian Workshop on Social Network Analysis and Mining*. SBC.
- Desmet, B. and Hoste, V. (2014). Recognising suicidal messages in Dutch social media. In *9th international conference on language resources and evaluation (LREC)*, pages 830–835.

- Ernst, M., Kallenbach-Kaminski, L., Kaufhold, J., Negele, A., Bahrke, U., Hautzinger, M., Beutel, M. E., and Leuzinger-Bohleber, M. (2019). Suicide attempts in chronically depressed individuals: What are the risk factors? *Psychiatry Research*, page 112481.
- Filho, S. L., Silva, E., Oliveira, J., and Silva, M. (2024). Depressset: Um conjunto de dados de análises textuais sobre postagens depressivas. In *Anais do XIII Brazilian Workshop on Social Network Analysis and Mining*, pages 214–220, Porto Alegre, RS, Brasil. SBC.
- Glenn, C. R., Kleiman, E. M., Kellerman, J., Pollak, O., Cha, C. B., Esposito, E. C., Porter, A. C., Wyman, P. A., and Boatman, A. E. (2020). Annual research review: A meta-analytic review of worldwide suicide rates in adolescents. *Journal of child psychology and psychiatry*, 61(3):294–308.
- Ivanich, J. D., O’Keefe, V., Waugh, E., Tingey, L., Tate, M., Parker, A., Craig, M., and Cwik, M. (2021). Social network differences between American Indian youth who have attempted suicide and have suicide ideation. *Community mental health journal*, pages 1–6.
- Nascimento, R., Carvalho, F., and Guedes, G. (2019). Identificando sintomas depressivos: um estudo de caso no Youtube. In *Anais do VIII Brazilian Workshop on Social Network Analysis and Mining*, pages 119–130, Porto Alegre, RS, Brasil. SBC.
- Park, M., Cha, C., and Cha, M. (2012). Depressive moods of users portrayed in Twitter. In *Proceedings of the ACM SIGKDD Workshop on healthcare informatics (HI-KDD)*, volume 2012, pages 1–8, New York, NY, USA. ACM.
- Parraga-Alava, J., Caicedo, R. A., Gómez, J. M., and Inostroza-Ponta, M. (2019). An unsupervised learning approach for automatically to categorize potential suicide messages in social media. In *2019 38th International Conference of the Chilean Computer Science Society (SCCC)*, pages 1–8. IEEE.
- Picard, R. W. (2000). *Affective computing*. MIT Press, Cambridge, MA.
- Santos, W. R. d., de Oliveira, R. L., and Paraboni, I. (2024). SetembroBR: a social media corpus for depression and anxiety disorder prediction. *Language Resources and Evaluation*, 58(1):273–300.
- Sharma, M., Pant, B., Singh, V., and Kumar, S. (2021). STP: Suicidal Tendency Prediction among the youth using social network data. In *Next Generation Information Processing System*, pages 161–169. Springer.
- Simon, G. E., Stewart, C. C., Gary, M. C., and Richards, J. E. (2021). Detecting and assessing suicide ideation during the COVID-19 pandemic. *The Joint Commission Journal on Quality and Patient Safety*, 47(7):452–457.
- Varathan, K. D. and Talib, N. (2014). Suicide detection system based on Twitter. In *2014 Science and Information Conference*, pages 785–788. IEEE.
- WHO (2017a). Depression and other common mental disorders: global health estimates.
- WHO (2017b). Depression Let’s talk. Campaigning Essential.
- Winkler, W. E. (1999). The state of record linkage and current research problems. Technical report, Statistical Research Division, U.S. Bureau of the Census.