

Análise Temporal de Sentimentos das Postagens sobre a Dengue em Redes Sociais

Allan Victor Santos de Oliveira¹, Manoel Limeira de Lima Júnior¹

¹Bacharelado em Sistemas de Informação – Universidade Federal do Acre (UFAC)

allan.victor@sou.ufac.br, manael.limeira@ufac.br

Abstract. *This article analyzed the relationship between the volume of dengue-related posts on the social network Bluesky in 2025 and the number of cases of the disease, as well as an investigation into the sentiments, themes, and words present in the posts using Natural Language Processing (NLP) strategies. The analyses revealed a strong correlation between the databases, with the majority of posts being negative and having the highest average intensity. In terms of engagement, positive posts received more likes, while neutral posts received more replies and reposts. Word clouds and topic modeling revealed that the content is mutable over time.*

Resumo. *Este artigo analisou a relação entre o volume de publicações referentes à dengue na rede social Bluesky no ano de 2025 com os casos da doença, além de uma investigação sobre os sentimentos, temas e palavras presentes nas postagens por meio de estratégias de Processamento de Linguagem Natural (PLN). As análises revelaram uma forte correlação entre as bases de dados, no qual a maioria das publicações foram negativas, cuja intensidade média foi a mais elevada. No engajamento, as positivas obtiveram mais curtidas enquanto as neutras mais respostas e republicações, enquanto as nuvens de palavras e modelagem de tópicos revelaram que o conteúdo é mutável com o tempo.*

1. Introdução

A dengue é categorizada como uma doença de importante ameaça à saúde pública em função de seu potencial epidêmico e de seu impacto socioeconômico. No Brasil, essas características provocam sobrecarga nos serviços de saúde voltados para o controle e tratamento dessa doença [Brasil 2024].

A participação da população faz parte das estratégias de prevenção a dengue. Para alcançar o envolvimento da sociedade, são utilizadas campanhas publicitárias e mídias de massa. Essas ações ajudam a comunidade a assimilar práticas preventivas e a auxiliar nas medidas de controle vetorial da dengue [Alves et al. 2022].

Com a popularização das redes sociais, as discussões relacionadas a área da saúde ganharam um novo espaço, permitindo que as pessoas expressem suas percepções e sentimentos [Malakin 2024]. Nesse contexto, surge a oportunidade de utilizar as redes sociais como uma fonte não oficial para detectar precocemente rumores, surtos e emergências em saúde pública, aplicando técnicas de ciência de dados. Termos-chave podem ser capturados em publicações nas redes sociais e indicar viés epidemiológico, baseando-se no cruzamento da frequência desses termos com as curvas epidêmicas [Leal-Neto et al. 2016].

A rede social *Bluesky* se destacou no Brasil após o impasse que levou ao bloqueio do *X* (antigo *Twitter*) em meados de agosto de 2024 [Gerken 2024], estabelecendo-se como uma alternativa a plataforma *X*.

Neste cenário, a pesquisa aplicou técnicas de ciência de dados em textos extraídos das publicações da rede social *Bluesky*, analisando as percepções e sentimentos em relação à epidemiologia da dengue presentes nas postagens no ano de 2025. Para alcançar esse objetivo, foi realizada uma comparação entre o volume de postagens da rede e os números de casos relacionados à epidemiologia da dengue, seguida da aplicação de estratégias de Processamento de Linguagem Natural (PLN), como análise de sentimentos, modelagem de tópicos e nuvem de palavras nas mensagens coletadas.

2. Background

A Dengue é uma doença infecciosa de rápida disseminação, considerada uma ameaça à saúde pública no Brasil [Brasil 2024]. De acordo com [Xavier et al. 2020], uma forma de auxiliar as estratégias de combate a essa doença é por meio do monitoramento das redes sociais, utilizando-as como um indicador em tempo real do cenário epidemiológico.

A rede social *Bluesky* é um *microblogging* semelhante ao *X*. Sua popularidade no Brasil cresceu após as restrições enfrentadas pelo *X* em 2024 [Lemos 2024], se destacando por ir na contramão das novas políticas de monetização das redes sociais [Santana and de Freitas 2024], se tornando um ambiente propenso a realização de pesquisas, sendo necessário técnicas automatizadas para a extração destes dados.

A mineração de texto refere-se ao processo de extração de informações úteis a partir de documentos textuais, sendo considerada uma extensão da mineração de dados, é empregada para identificar padrões em dados não estruturados ou semiestruturados, comumente escritos em linguagem natural [Aranha and Passos 2006]. Neste contexto, o uso de ferramentas que auxiliam o tratamento da linguagem humana são fundamentais.

Segundo [Caseli et al. 2023], o Processamento de Linguagem Natural (PLN) é uma área da Inteligência Artificial (IA) que busca resolver problemas que necessitam de um tratamento computacional da linguagem humana. Existem diversos campos para a aplicação destas técnicas, como a extração de sentimentos em textos.

Nesta pesquisa, utilizou-se uma ferramenta chamada LeIA (Léxico para Inferência Adaptada), adaptada para extração de sentimentos em português, com foco nos textos expressos em redes sociais [Almeida 2018]. De forma auxiliar, análise estatísticas são importantes para fornecer uma visão mais ampla do cenário em que se deseja pesquisar.

Na modelagem de tópicos, o LDA (*Latent Dirichlet Allocation*) é um modelo probabilístico criado para identificar grupos temáticos em grandes volumes de dados textuais [Blei et al. 2003]. Apesar disso, o LDA não fornece uma interface visual, apenas agrupa as palavras em suas coocorrências.

Por outro lado, a nuvem de palavras é uma representação visual da frequência com que uma palavra aparece em um texto, sendo ela uma técnica de análise exploratória de dados que fornece uma visão intuitiva e agradável dos principais termos, auxiliando o leitor a avaliar rapidamente se o conteúdo é relevante [Heimerl et al. 2014].

3. Trabalhos Relacionados

Ferramentas de monitoramento das redes sociais possibilitam realizar uma avaliação rápida e de baixo custo das opiniões, sentimentos e atitudes da sociedade, a fim de desenvolver políticas públicas mais eficazes [de Santana and de Souza 2017].

No cenário epidemiológico da dengue, o trabalho de [Antunes et al. 2014] propôs um sistema automatizado para rastrear menções à doença em mídias sociais, principalmente no *Twitter*, entre os meses de abril e setembro de 2012. O estudo demonstrou que o volume de publicações possuía uma correlação direta com o aumento dos casos reais da doença, funcionando como uma espécie de termômetro para os surtos.

O trabalho de [Angeluci and de Farias 2024] investigou a percepção pública sobre a dengue no ano de 2024, utilizando a ferramenta *Stilingue* para coletar informações de publicações nas redes sociais durante o período de 1º de janeiro à 15 de abril de 2024. Além disso, as postagens com maior engajamento possuíam teor político, enquanto que a maioria dos sentimentos foram classificados como negativos (70%).

Na pesquisa realizada por [Melo 2025], cujo objetivo era analisar a percepção e a repercussão sobre o desmatamento e as queimadas a partir de publicações na rede social *Bluesky*. Os resultados mostraram que a reação pública ao desastre ambiental constitui um ecossistema discursivo complexo e multifacetado, no qual boa parte dos debates gira em torno da indignação política, em que, para o público, as queimadas são mais vistas como uma falha de governança do que como um desastre natural.

O presente trabalho traz uma análise semelhante àquelas realizadas nesses estudos. Entretanto, ao adicionar a dimensão temporal, proporcionando uma análise específica dos dados, destacando como os debates sobre esse assunto evoluem ao longo do tempo.

4. Materiais e Métodos

Este trabalho utilizou o PLN para analisar as percepções da sociedade sobre a dengue na rede social *Bluesky*. Os dados coletados provinham de duas fontes: os casos de dengue foram obtidos por meio da plataforma oficial do governo DATASUS¹, enquanto as publicações na rede social foram coletadas por um algoritmo em *Python* e requisições a uma API (*Application Programming Interface*) da própria plataforma social.

Para acessar os *endpoints* da API, é necessário possuir um recurso chamado “Senhas de Aplicativo” para acessar o servidor de dados por meio de *scripts*. A partir desse ponto, o fluxo de *endpoints* utilizados foi estruturado da seguinte maneira: i) obter um *token* de acesso por meio do envio das credenciais; ii) coletar as postagens públicas com base nos critérios de busca: Termos, onde as palavras-chave foram *dengue* e *aedes*, os sintomas não foram incluídos para não capturar postagens referentes a outras doenças; Período, abrangendo integralmente o ano de 2025; e Idioma, incluindo o parâmetro *lang: pt* para filtrar as postagens em língua portuguesa; e iii) obter a árvore de respostas das publicações, possibilitando a análise das discussões.

Vale ressaltar que os dados coletados são públicos, e, com o objetivo de preservar o anonimato dos usuários, seus nomes foram trocados por identificadores genéricos.

¹Link do Site do Governo: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>

As postagens foram submetidas a um pré-processamento específico para cada etapa da pesquisa. Na análise de sentimentos, o LeIA foi utilizado para remover *links* e menções a usuários. Na nuvem de palavras, foram removidas as pontuações, *emojis* e palavras sem valor semântico (*stopwords*). Na modelagem de tópicos, além dos processos anteriores, foi aplicado um método de lematização, que reduz as palavras à sua forma raiz.

Após o pré-processamento, foi realizada a aplicação de técnicas de ciência de dados com o objetivo de extrair informações úteis, sendo elas:

- **Correlação entre as bases:** utilizando o Coeficiente de Correlação de *Spearman* (r) para medir a correlação entre os casos de dengue e as publicações.
- **Análise de Sentimentos:** calculando, por meio da ferramenta LeIA, a polaridade e a intensidade dos sentimentos presentes nas postagens.
- **Engajamento:** extraíndo a média de curtidas, republicações e respostas em relação aos sentimentos das postagens.
- **Nuvens de Palavras:** empregando a biblioteca *WordCloud* para a geração de imagens que reflitam as palavras mais utilizadas.
- **Modelagem de Tópicos:** utilizando o LDA para agrupar palavras que coocorrem nas publicações e rotular um tópico.

Neste trabalho, a ferramenta *Google Gemini Pro* foi utilizada na criação do código para a coleta de dados; *scripts* de aplicação dos métodos de PLN (análise de sentimentos, modelagem de tópicos, nuvem de palavras); e para a geração dos gráficos. A ferramenta *Writefull Premium* foi utilizada para fins de revisão ortográfica e gramatical. Os autores revisaram as correções e assumem integralmente a responsabilidade pela precisão das informações e pela solidez intelectual do conteúdo. O código, as bases de dados, os gráficos e as imagens elaboradas estão disponíveis em um repositório público².

5. Resultados e Discussões

Por meio dos métodos de coleta de dados, foram obtidas cerca de 11.592 publicações referentes a dengue na plataforma *Bluesky* em 2025 no Brasil, das quais 8.125 (70,09%) correspondiam a postagens únicas e 3.467 (29,91%) são respostas a essas postagens. Em relação aos casos de dengue no mesmo ano, obtivemos 1.663.248 notificações da doença, das quais 88,28% (1.468.415) correspondem a primeira metade do ano, enquanto a segunda metade apresenta 11,72% (194.869) do total de registros.

5.1. Correlação Entre Postagens e Casos de Dengue

O coeficiente de correlação de *Spearman* (r) entre o volume de postagens e os casos de dengue obteve um valor r igual a 0,8289, que, segundo as diretrizes estabelecidas por [Schober et al. 2018], valores entre de 0,7 e 0,89 indicam uma forte correlação, além de ser altamente significativa ($p < 0,001$). Ao analisar a defasagem temporal (lag), observa-se que ocorre uma maior reação das postagens 3 semanas após o aumento de casos da doença, com um valor r de 0.8602. Dessa forma, o volume de postagens e os casos de dengue estão estatisticamente interligados durante o período analisado, como pode ser visto na Figura 1.

²Link do repositório: <https://anonymous.4open.science/r/scripts-base-dados-e-imagens-4F38/>

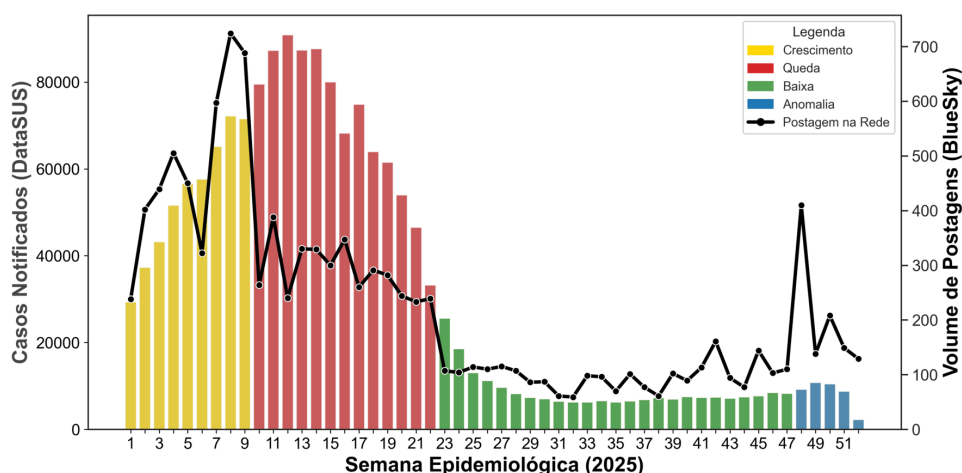


Figura 1. Correlação entre Postagens e Números de casos

Observa-se que a organização do gráfico em cores simboliza as partições da base de dados, definido após uma inspeção visual da Figura 1 levando em consideração as quantidades de casos e de publicações. Essa segmentação visa observar a evolução das discussões presentes nas postagens em quatro blocos temporais:

- **Crescimento:** cor amarela, concentra grande parte das postagens (37,66%) e o crescimento dos casos, com o período abrangendo o intervalo da semana 0 à 9;
- **Pico e Queda:** cor vermelha, quando houve uma diminuição nas publicações, enquanto os casos atingiram o auge e, em seguida, começaram a decair. Seu período temporal é da semana 10 até a 22;
- **Baixa:** cor verde, simboliza a diminuição na quantidade de postagens e de casos da doença; sua faixa de tempo abrange da semana 23 até a semana 47;
- **Anomalia:** cor azul, reflete as publicações no período de estabilidade dos casos de dengue, que vai da semana 48 até a 52.

5.2. Análise de Sentimentos

A análise de sentimentos LeIA foi aplicada nas postagens, classificando-as como positivas, negativas e neutras. A Figura 2 ilustra o resultado das classificações em cada bloco, no qual observa-se a similaridade entre os três primeiros, nos quais a maioria dos textos foi classificado com sentimento negativo, seguida pelos neutros e, por fim, positivos, implicando que os sentimentos expressos permaneceram constantes ao longo do ano.

Entretanto, no bloco “Anomalia”, a média de sentimentos positivos aumentou, enquanto as médias de polaridade negativa regrediram em relação aos demais blocos temporais. Isso implica que, durante este período, houve uma discussão sobre um determinado tópico que quebrou o padrão observado nos três primeiros blocos.

Em conjunto com a análise de sentimentos, investigou-se a intensidade dos mesmos, verificando qual polo era mais intenso nas mensagens. Isso foi realizado por meio da métrica de pontuação *compound* do modelo LeIA. O bloco “Pico e Queda” obteve a menor intensidade média das mensagens classificadas como positivas (0,4102) e a maior das classificadas como negativas (0,5085), podendo ser interpretado como um período mais intenso nas críticas e reclamações. Por outro lado, o bloco “Anomalia” obteve a

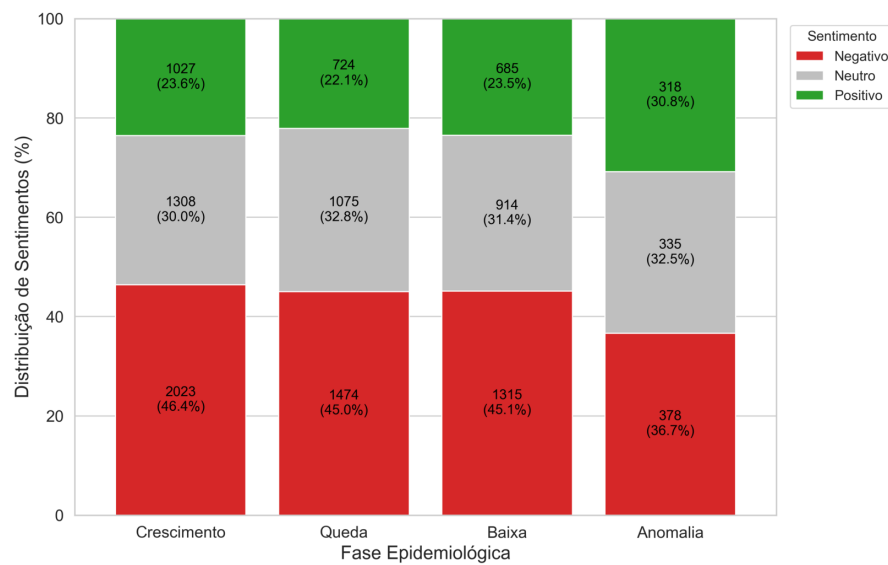


Figura 2. Sentimentos das postagens sobre a dengue nos blocos temporais.

maior média positiva (0,4343) e a menor intensidade média negativa (0,4655), indicando uma faixa temporal mais suave nas discussões, ainda que de forma sutil.

5.3. Engajamento por Sentimento

Para entender a relação entre os sentimentos e o engajamento das postagens, a Tabela 1 apresenta as métricas de *Likes* (curtidas), *Replies* (respostas) e *Reposts* (republicações).

Bloco	Sentimento	Likes	Replies	Reposts
Crescimento	Negativo	3,31	0,95	2,92
	Neutro	4,31	1,51	6,82
	Positivo	4,56	1,20	5,24
Pico e Queda	Negativo	2,92	1,04	1,24
	Neutro	1,55	1,90	1,79
	Positivo	3,13	1,63	1,78
Baixa	Negativo	2,84	0,91	1,38
	Neutro	2,97	1,76	4,05
	Positivo	3,60	1,20	2,72
Anomalia	Negativo	10,75	1,26	10,31
	Neutro	9,83	1,55	20,75
	Positivo	15,72	1,14	17,00

Tabela 1. Engajamento dos sentimentos sobre a dengue por bloco.

As publicações de sentimento negativo, apesar de serem a maioria, não receberam médias maiores de curtidas, respostas e republicações em nenhum dos blocos. Outro fato observado é que as publicações com polaridade positiva obtiveram a maior média de curtidas, enquanto as neutras obtiveram médias superiores de republicações e respostas em todos os blocos temporais. Entretanto, no bloco “Anomalia”, o desvio padrão médio dos valores de engajamento foi cerca de 288% maior que os dos outros blocos, sinalizando a

ocorrência de publicações virais neste período. Isso pode ser interpretado como a sociedade se engajando em mensagens positivas, como alguém se recuperando de uma doença, e em notícias sobre a dengue, que seguem uma linguagem neutra, e em publicações viralmente pontuais, evidenciado no bloco “Anomalia”.

5.4. Nuvens de Palavras

Para compreender o vocabulário e visualizar a frequência de ocorrência das palavras, foram geradas nuvens de palavras para cada bloco temporal, conforme pode ser visto na Figura 3. É possível observar a evolução dos principais termos utilizados nas publicações, como no bloco “Pico e queda”, onde as palavras em destaque estão mais voltadas para a prevenção, com os termos *vacina* e *contra*, a provável utilização do termo *não* como um elemento de reforço (não deixe água parada) e a palavra *casos* relacionada ao avanço da doença. Esses termos demonstram uma preocupação inicial com o crescimento do número de casos e com medidas de prevenção, como a vacinação e os cuidados em casa.



Figura 3. Nuvens de palavras dos blocos de estudo.

No bloco “Pico e Queda”, as palavras relatam o aumento dos índices de casos de dengue, juntamente com relatos sobre a propagação da doença, como o aumento do tamanho das palavras ‘casos’, ‘morte’, ‘doença’ e ‘melhora’. No bloco “Baixa”, é possível observar uma mudança expressiva nas palavras em destaque. Devido à extensão desse intervalo temporal, somada às baixas postagens e aos casos da doença, os debates acabaram fluindo por vários temas distintos. Entre eles, ainda é possível identificar temáticas sobre a vacina, o volume de casos, as medidas de prevenção e os relatos de recuperação.

Por fim, na nuvem de palavras do bloco “Anomalia”, o centro dos debates nesta faixa temporal foca principalmente no tema da vacinação; isso se dá ao fato de que, neste período, foi anunciada pelo governo uma nova vacina nacional contra a dengue, evidenciado pelo grande contraste das palavras que remetem a essa informação (vacina, unica, dose, butantan). Isso nos mostra que a sociedade está dando atenção e importância a novos inventos científicos relacionados à dengue.

5.5. Modelagem de Tópicos

A modelagem de tópicos refere-se à identificação dos principais temas abordados nas discussões das postagens e permite agrupar palavras que apresentam correlação com base em seus padrões de distribuição. A escolha do número de temas por bloco foi realizada por meio de imagens interativas gerada pelo *pyLDAvis*, devido a um teste de varredura que falhou em encontrar o número ideal de tópicos, no qual os valores não possuíam discrepância que justifica-se sua escolha. A Tabela 2 mostra os principais temas encontrados, onde é possível acompanhar a evolução das discussões sobre a dengue nas postagens.

No bloco “Crescimento”, observa-se temas sobre o controle e prevenção da doença, que se encaixam com o crescente número de registros neste bloco, junto com um tema que possivelmente reflete o desabafo das pessoas sobre ter contraído a doença.

No bloco “Pico e Queda”, os temas passam por uma evolução, como o “Diagnóstico”, que se dividiu nos temas “Sintomas” e “Comparação”, nos quais, além de relatar as dores que estavam sentindo, realizaram comparativos com outras doenças ou até mesmo não conseguindo diferencia-los. Pode-se observar que as temáticas de prevenção foram substituídas por discussões sobre relatos de fatalidades e pelas atualizações dos boletins.

Crescimento		Pico e Queda	
Tema	Palavras-chave	Tema	Palavras-chave
Diagnostico	passar, dor, febre, dengar	Sintomas	dor, corpo, cabeça, tomar
Prevenção	vacina, armadilha, repelente	Comparação	virose, aede, covid, gripe
Gestão	saúde, Brasil, notícia, morte	Fatalidade	confirmar, morte, óbito
		Boletim	registrar, confirmar, morte
Baixa		Anomalia	
Tema	Palavras-chave	Tema	Palavras-chave
Prevenção	combate, aede, ação, doença	Nova Vacina	vacina, dose, butantan
Sintomas	dor, passar, achar, melhorar	Debate	ciência, nature, brasileiro
Vacinação	vacina, tomar, esperar, sus	Combate	ação, reforçar, risco, aede
Experiência	dengar, pegar, passar	Sintomas	achar, febre, mau, morrer

Tabela 2. Tópicos e palavras-chave identificados pelo LDA nos blocos temporais.

No bloco “Baixa”, os temas mostram-se mais dispersos, abrangendo desde ações preventivas até a experiência de ter contraído a doença. Um dos temas anteriores que persistiu foi “Sintomas”, mas agora as palavras insinuam uma melhora em relação à doença.

Por fim, no bloco “Anomalia”, os temas giram em torno de uma nova vacina anunciada, juntamente com discussões científicas, menções aos combates vetoriais e a persistência do tema “Sintomas” referente aos relatos de pessoas que contraíram a doença.

Os agrupamentos temáticos presentes nos blocos temporais evidenciam que as discussões não ficam centradas em uma única temática; elas possuem vários polos que evoluem ao longo do tempo, de acordo com a situação em que se encontram.

6. Considerações Finais

Este trabalho investigou os sentimentos expressos em publicações sobre a dengue na rede social *Bluesky* no ano de 2025. Para isso, técnicas de ciência de dados foram utilizadas para coletar, processar e visualizar informações úteis.

Esses *insights* mostram que, a integração de sistemas de monitoramento epidemiológico com técnicas de mineração de dados e PLN possibilita obter uma visão mais detalhada do cenário da doença, como: o grau de percepção da população em relação a dengue, com a correlação; seu nível de satisfação em relação as campanhas, por meio da análise de sentimentos; e o foco das discussões, com as nuvens e a modelagem de tópicos.

A análise entre os casos e as postagens mostrou uma forte correlação, indicando que o fluxo de postagens acompanha a curva epidemiológica dos casos da doença. Já na análise de sentimentos, observou-se que a maioria das publicações apresentava polaridade negativa e exibiam uma intensidade maior do que aquelas classificadas como positivas. Isso pode ser interpretado como a utilização deste espaço para desabafar, expressar suas dores, preocupações e indignações a respeito da doença.

Quanto ao engajamento, os sentimentos positivos tiveram uma média maior de curtidas em todos os blocos (40,4% maior), enquanto os neutros possuem médias superiores nas respostas (45,75% maior) e nas republicações (67,85% maior). Pode-se concluir que a sociedade dá mais atenção a mensagens de polaridade positiva e neutra.

Com as nuvens de palavras e a modelagem de tópicos, podemos observar que o conteúdo das publicações não é estático; ele está em constante mudança, tanto nas palavras e suas frequências quanto nos temas associados a elas. Essa mudança vai de acordo com o cenário epidemiológico em que se encontra e é influenciada por notícias de grande importância, como visto no bloco “Anomalia”.

Entre as limitações da pesquisa estão: o viés demográfico da plataforma, que não possui um geolocalizador do usuário, e a filtragem das postagens, que se limitou a capturar publicações classificadas como língua portuguesa, as quais não são necessariamente brasileiras. Nas correlações, não foram investigadas as relações de causalidade, pois não estavam no escopo desta pesquisa. Além disso, o algoritmo LeIA apresenta dificuldade para interpretar orações complexas, o que dificulta a classificação correta das postagens. O LDA é um modelo puramente probabilístico que permite abordar tópicos com grande variedade de termos, focando apenas em palavras que aparecem com alta frequência e que ocorrem em grupos. Por fim, a coleta de dados foi determinada por um conjunto restrito de termos de busca (aedes e dengue), de modo que as publicações que utilizaram outros termos relacionados à dengue não foram coletadas pelo mecanismo de captura.

A partir deste trabalho, emerge a possibilidade de futuras pesquisas que abordem uma metodologia semelhante em medidas com maior engajamento, a utilização de técnicas e algoritmos de mineração de dados, além de uma análise de outras enfermidades nas redes.

Referências

- Almeida, R. J. A. (2018). Leia - léxico para inferência adaptada. <https://github.com/rafjaa/LeIA>.
- Alves, J. A., Andrade, N. F. d., Lorenzo, C. F. G., Mendonça, A. V. M., and Sousa, M. F. d. (2022). Percepção da comunidade sobre suas ações preventivas contra dengue, zika e chikungunya nas cinco regiões do Brasil. *Physis: Revista de Saúde Coletiva*, 32(03):e320312.

- Angeluci, A. C. B. and de Farias, M. T. (2024). Percepção pública da dengue em 2024: analisando big data com social listening. *Comunicação & Inovação*, 25:e20249660–e20249660.
- Antunes, M. N., Silva, C. H. d., Guimarães, M. C. S., and Rabaço, M. H. L. (2014). Monitoramento de informação em mídias sociais: o e-monitor dengue. *TransInformação*, 26:9–18.
- Aranha, C. and Passos, E. (2006). A tecnologia de mineração de textos. *Revista Eletrônica de Sistemas de Informação*, 5(2).
- Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.
- Brasil (2024). Volume 55 - nº 11: Monitoramento das arboviroses e balanço de encerramento do comitê de operações de emergência (coe) dengue e outras arboviroses 2024. Disponível em: <https://tinyurl.com/385ph3jt>. Acesso em: 18 set. 2025.
- Caseli, H. d. M., Nunes, M. d. G. V., and Pagano, A. S. (2023). O que é PLN? *Processamento de linguagem natural: conceitos, técnicas e aplicações em português*.
- de Santana, M. B. and de Souza, C. G. B. (2017). Uso das redes sociais por órgãos públicos no brasil e possibilidades de contribuição do monitoramento para gestão. *Gestão. Org*, 15(6):99–107.
- Gerken, T. (2024). O que é o Bluesky, a rede social rival do X para onde estão indo usuários em protesto a Elon Musk. Disponível em: <https://tinyurl.com/ykwbffj6m>. Acesso em: 24 fev. 2026.
- Heimerl, F., Lohmann, S., Lange, S., and Ertl, T. (2014). Word cloud explorer: Text analytics based on word clouds. In *2014 47th Hawaii international conference on system sciences*, pages 1833–1842. IEEE.
- Leal-Neto, O. B., Dimech, G. S., Libel, M., Oliveira, W., and Ferreira, J. P. (2016). Detecção digital de doenças e vigilância participativa: panorama e perspectivas para o Brasil. *Revista de Saúde Pública*, 50:17.
- Lemos, V. (2024). Bluesky: a rede social que conseguiu um milhão de usuários e bateu recordes de interações após bloqueio do X no brasil. Disponível em: <https://tinyurl.com/442xkzw4>. Acesso em: 27 fev. 2026.
- Malakin, L. A. (2024). *Avaliação de Classificadores na Análise de Sentimentos em Redes Sociais Durante a Pandemia da COVID-19*. PhD thesis, Universidade de São Paulo.
- Melo, B. H. F. d. (2025). Percepção e repercussão: uma análise de sentimentos, tópicos e engajamento em publicações brasileiras sobre desmatamento e queimadas no Bluesky em agosto e setembro de 2024.
- Santana, B. S. and de Freitas, L. A. (2024). Capítulo 34 PLN em redes sociais.
- Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, 126(5):1763–1768.
- Xavier, F., Olenski, J. R. W., Acosta, A. L., Sallum, M. A. M., and SARAIVA, A. (2020). Análise de redes sociais como estratégia de apoio à vigilância em saúde durante a covid-19. *Estudos avançados*, 34:261–282.