

Análise de Sentimentos no *Twitter* com *Tweet Utils UI*: Um Estudo de Caso

Anderson R. P. Sprenger, Gabriel Z. de Souza, Isabel H. Manssour, Vinícius P. Dias

Escola Politécnica – Pontifícia Universidade Católica do Rio Grande do Sul
Av. Ipiranga, 6681 – 90.619-900 – Porto Alegre – RS – Brasil

anderson.sprenger@edu.pucrs.br, zurawski.gabriel@edu.pucrs.br

isabel.manssour@pucrs.br, v.dias005@edu.pucrs.br

Abstract. *This article aims to present a tool for collect, process, and visualize Twitter data: Tweet Utils UI. In addition, it also introduces a new feature to the methods of analysis and visualization by implementing sentiment analysis of the extracted text from Twitter; classifying them as Positive, Neutral, or Negative. To validate this tool and the novel sentiment analysis feature, this paper presents a case study featuring the online reactions of the final episode of Brazilian reality show Big Brother Brasil 23.*

Resumo. *Este artigo tem como intenção apresentar uma ferramenta de coleta, tratamento e visualização de dados extraídos da rede social Twitter: o Tweet Utils UI. Além de também propor uma nova função ao métodos de análise e visualização com a implementação da análise de sentimentos dos textos extraídos do Twitter; os classificando entre Positivo, Neutro ou Negativo.¹ Para validar a ferramenta e a análise de sentimentos adicionada, é apresentado um estudo de caso sobre as reações online ao episódio final do reality show Big Brother Brasil 23.*

Palavras-chave: *Mineração de Dados, Visualização de Dados, Analise de Sentimentos, Twitter*

1. Introdução

Nas últimas décadas, foi observado um aumento considerável na quantidade de dados gerados no mundo, principalmente devido a popularização da internet, a acessibilidade dos *smartphones* e computadores e ao surgimento das redes sociais, as quais se tornaram um fonte constante de dados. Isso proporciona o que conhecemos como *Big Data*, uma grande quantidade e variedade de dados sendo gerada tão rapidamente que nem mesmo damos conta de extrair todo valor contido neles.

Um grande exemplo desse fenômeno é o *Twitter*, onde são publicados cerca de 500 milhões de tuítes em formato de texto por dia. Esses tuítes podem expressar desde opiniões e sentimentos até informações sobre fatos, produtos e eventos. Esses textos refletem a opinião pública e fornecem *insights* sobre a reação a determinados assuntos. Em outras palavras, são dados valiosos para pesquisas, estudos de mercado e tomada

¹O código-fonte com as contribuições feitas ao objeto de estudo está disponível em: <https://github.com/andersprenger/Tweet-Utils-UI>

de decisões empresariais. No entanto, além dos dados em si, também é necessário ter formas objetivas e acessíveis de coletá-los, processá-los e visualizá-los, especialmente para pessoas leigas no campo da análise de dados.

No entanto, análises estritamente quantitativas podem não capturar toda a complexidade dos dados, especialmente quando se trata de compreender as opiniões e emoções humanas. A análise de sentimentos, que envolve a identificação e classificação de emoções expressas em textos, torna-se um componente importante para uma análise de dados abrangente. Para esse propósito, ferramentas relacionadas à manipulação de informações são muito eficientes e podem auxiliar de diversas formas na interpretação desses dados.

O *Tweet Utils UI* é uma ferramenta que permite coletar, processar e visualizar tuítes, sendo útil para qualquer pessoa interessada em compreender determinados públicos na rede social *Twitter*. Neste contexto, a análise de sentimentos é um recurso que agrega muito valor na extração de informações do público em geral, presente nesta rede social. Portanto, nosso objetivo neste trabalho é adicionar essa funcionalidade à ferramenta *Tweet Utils UI*, permitindo a interpretação e classificação de sentimentos, tornando-a mais útil para mineração e visualização de dados sobre opinião pública.

2. Fundamentação Teórica

A Cultura Orientada por Dados, em inglês *Data-Driven Culture*, é uma filosofia empresarial em que as decisões são tomadas diariamente com base em dados objetivos e análises quantitativas. Isso envolve o uso de tecnologias de coleta, análise e interpretação de dados para embasar o processo de tomada de decisões, em vez de depender apenas de palpites, intuição, experiência e julgamentos subjetivos. A abordagem visa utilizar fatos obtidos por meio dos dados para embasar as decisões empresariais [Yu et al. 2021].

Nesse mesmo contexto, a Análise Exploratória de Dados é uma abordagem inicial na busca pela interpretação de um conjunto de dados. Nessa etapa, realiza-se uma investigação sobre o tipo de dados a ser trabalhado, compreendendo suas características, padrões e relacionamentos. Isso permite uma compreensão dos dados para orientar a escolha de técnicas analíticas mais adequadas. Através da análise exploratória, é possível identificar inconsistências, distribuições, valores atípicos e relações inesperadas. Existem diversas técnicas que podem ser utilizadas nessa análise, como estatísticas descritivas, gráficos de dispersão, histogramas, entre outros [Data et al. 2016].

Um método de análise exploratória muito eficiente para textos, como os tuítes na rede social *Twitter*, é a Análise de Sentimentos. Utilizando técnicas de Processamento de Linguagem Natural (PLN), a análise de sentimentos determina se o sentimento expresso é positivo, negativo ou neutro. Ao combinar uma cultura orientada por dados com a análise de sentimentos, as organizações podem obter *insights* sobre a percepção dos clientes, identificar tendências emergentes, monitorar a reputação da marca e tomar medidas proativas para melhorar a experiência do cliente. Essa abordagem permite uma compreensão mais profunda dos sentimentos e opiniões dos clientes em relação à empresa, produtos ou serviços, contribuindo para melhorias contínuas e decisões mais informadas [Agarwal et al. 2011].

A Análise de Sentimentos envolve a utilização de técnicas de aprendizado de máquina para identificar e extrair informações subjetivas de dados textuais. Existem duas

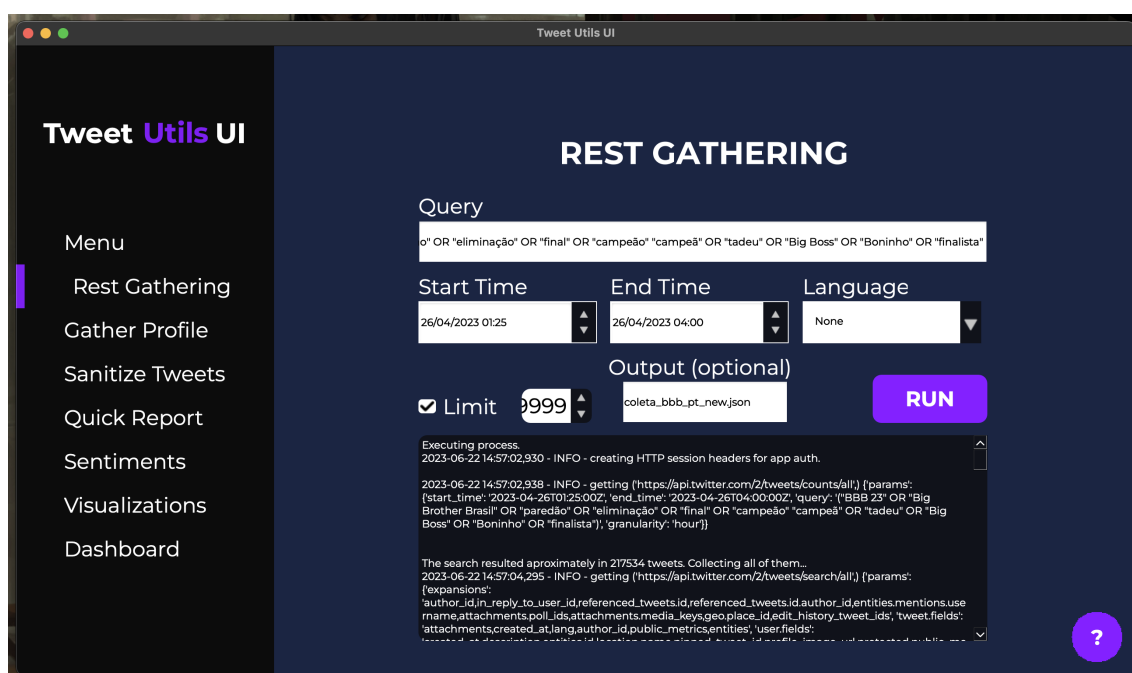


Figura 1. Interface da ferramenta para coleta de tuítes através de busca.

abordagens distintas para a análise de sentimentos: abordagem léxica e aprendizado de máquina. A abordagem léxica baseia-se no uso de recursos léxicos, como listas de palavras positivas e negativas, para determinar o sentimento expresso em um texto. Já a abordagem de aprendizado de máquina envolve o uso de algoritmos e técnicas de aprendizado de máquina para treinar um modelo capaz de identificar automaticamente o sentimento expresso em um texto. Para isso, é necessário coletar conjuntos de dados rotulados, que possuem um conjunto de sentimentos já categorizados. Em seguida, são extraídas determinadas características dos textos pré-processados, como palavras individuais, frequência de palavras e características sintáticas. Os conjuntos de dados são divididos em treinamento e teste para a análise de sentimentos. O conjunto de treinamento é usado para treinar o modelo, enquanto o conjunto de teste é utilizado para avaliar o desempenho do modelo.

3. *Tweet Utils UI*

A ferramenta *Tweet Utils UI*, desenvolvida por [Wagner and Manssour 2022], permite coletar e processar mensagens do Twitter. Seu objetivo principal é realizar uma análise abrangente dos dados e fornecer uma compreensão aprofundada da opinião pública em relação a eventos, produtos e pessoas. Utilizando a *API* do *Twitter*, essa ferramenta permite a coleta, o processamento e a visualização dos tuítes por meio de uma interface de usuário. Isso elimina a necessidade de conhecimento em programação ou o uso de programas complexos de manipulação de dados, tornando a tarefa acessível a usuários sem experiência técnica avançada ou para aqueles que desejam uma abordagem mais direta para coleta de dados.

A ferramenta oferece várias funcionalidades. A função “*Rest Gathering*” realiza a coleta de tuítes com base em uma *string* de busca e um intervalo de datas, sendo também possível definir um limite para o número de tuítes a serem coletados. Nesse caso, a fer-

```

{
  'id': number,
  'text': string,
  'created_at': string,
  'lang': string,
  'author_id': number,
  'retweet_count': number,
  'urls': array,
  'people_cited': array,
  'has_rich_media': boolean
}

```

Figura 2. Formato de um tuíte coletado quando salvo em um arquivo JSON, mostrando as chaves e tipos dos respectivos valores.

ramenta filtra os tuítes mais recentes até atingir o limite determinado. A função “*Gather Profile*” coleta os tuítes de um único perfil de usuário.

Os tuítes coletados pela ferramenta são salvos em um arquivo *JSON*, no formato mostrado na Figura 2. Ele armazena para cada tuíte: o *ID*, texto, a data de criação, o idioma, o *ID* do autor, o número de *retweets*, *URLs* mencionadas, pessoas citadas e se o tuíte possui mídia rica (por exemplo, imagens ou vídeos anexados).

A função “*Sanitize Tweets*” remove as “*stop words*”, que são palavras ignoradas durante a análise do texto, como artigos e preposições, dos tuítes contidos no arquivo *JSON* de entrada. A função “*Quick Report*” extrai estatísticas mais simples com base em um arquivo *JSON* de entrada. A função “*Visualizations*” permite que o usuário selecione as visualizações a serem geradas no formato de um arquivo *HTML*, incluindo gráficos de linha com a quantidade de tuítes ao longo do tempo, mapas de calor para a quantidade de tuítes por minutos e por horas, uma nuvem de palavras das palavras mais mencionadas nos tuítes e um campo para exibir os tuítes mais retuitados. Por fim, a função “*Dashboard*” deveria exibir em um *web browser* as visualizações contidas no arquivo *HTML*, porem a ferramenta não consegue carregar os elementos *JavaScript* usados nas visualizações. Os autores acessaram as visualizações geradas através de *browsers* como *Chrome*, *Firefox* e *Safari*.

4. Contribuições

4.1. Análise de Sentimentos

A ferramenta escolhida para realizar a análise de sentimentos em português é um modelo de aprendizado de máquina proposto por [Duarte and Manssour 2021]. Segundo os autores, o modelo possui uma acurácia de 71%, sendo utilizado um conjunto de dados com 1321 tuítes já classificados previamente.

Antes do treinamento do modelo, técnicas de pré-processamento de texto foram aplicadas a este conjunto de dados, de modo a melhorar a qualidade dos dados e facilitar a análise de sentimentos. Essas técnicas incluíram a remoção de *stop words*, a *stemização*, a *tokenização*, o tratamento de caracteres e a vetorização. O processo de pré-processamento aplicado ao conjunto de dados de treinamento também deve ser aplicado ao texto dos

tuítes antes de serem classificados pela ferramenta. Cada etapa do pré-processamento será descrita em maiores detalhes a seguir:

1. **Remoção das *stop words*:** feito através da função “*sanitize tweets*”, envolve a exclusão de palavras que têm pouco significado para a análise de sentimentos, como artigos, preposições e pronomes. Essas palavras são consideradas irrelevantes para a classificação de sentimentos e são removidas dos tuítes antes do processamento.
2. **Tratamento de caracteres:** refere-se à remoção de caracteres especiais, pontuações e símbolos dos tuítes. Essa etapa é realizada para limpar o texto e garantir que apenas os elementos relevantes para a análise de sentimentos sejam considerados.
3. **Stemização:** processo que reduz as palavras a suas raízes. Por exemplo, as palavras “correr”, “corre” e “correndo” seriam reduzidas ao radical “corr”. Isso é feito para tratar diferentes formas flexionadas de uma mesma palavra como sendo iguais.
4. **Tokenização:** processo de dividir o texto em partes menores, chamadas *tokens*. Os *tokens* podem ser palavras individuais ou até mesmo partes de palavras, dependendo do método utilizado. Isso ajuda na análise de sentimentos, permitindo que os algoritmos considerem cada parte do texto separadamente.
5. **Vetorização:** processo de converter os tuítes em representações numéricas, que são mais adequadas para o processamento pelos algoritmos de aprendizado de máquina. Existem várias técnicas de vetorização, como o CountVectorizer escolhido por [Duarte and Manssour 2021] e implementado neste trabalho. O CountVectorizer é uma técnica de vetorização que converte documentos de texto em vetores numéricos, representando a contagem de ocorrências de palavras ou termos.

Após o pré-processamento do conjunto de dados, são aplicados três algoritmos para a criação do modelo de classificação: Regressão Logística, *Multinomial Naive Bayes* e *Support Vector Machine*. Esses algoritmos, implementados por [Pedregosa et al. 2011], serão descritos em maiores detalhes a seguir:

A Regressão Logística é um modelo estatístico amplamente utilizado para problemas de classificação binária. Ele estima a probabilidade de um objeto, neste caso um texto, pertencer a uma determinada categoria com base em um conjunto de variáveis independentes. No caso da análise de sentimentos, a Regressão Logística é aplicada para determinar a probabilidade de um texto expressar um sentimento positivo ou negativo. [LaValley 2008]

O *Multinomial Naive Bayes* é um algoritmo de classificação probabilístico baseado no *Teorema de Bayes*. Ele assume a independência entre as categorias não binárias para realizar a predição dos sentimentos em um texto. No processo de treinamento, o algoritmo categoriza palavras individuais e realiza a contagem de suas frequências no texto. Com base nessa categorização e contagem, é determinada a qual categoria o texto pertence em termos de sentimentos expressos. [Abbas et al. 2019]

Por fim, o *Support Vector Machine (SVM)* é um algoritmo de aprendizado de máquina supervisionado que pode ser aplicado tanto em tarefas de regressão quanto de classificação. No contexto da análise de sentimentos, o *SVM* é utilizado para reconhecer

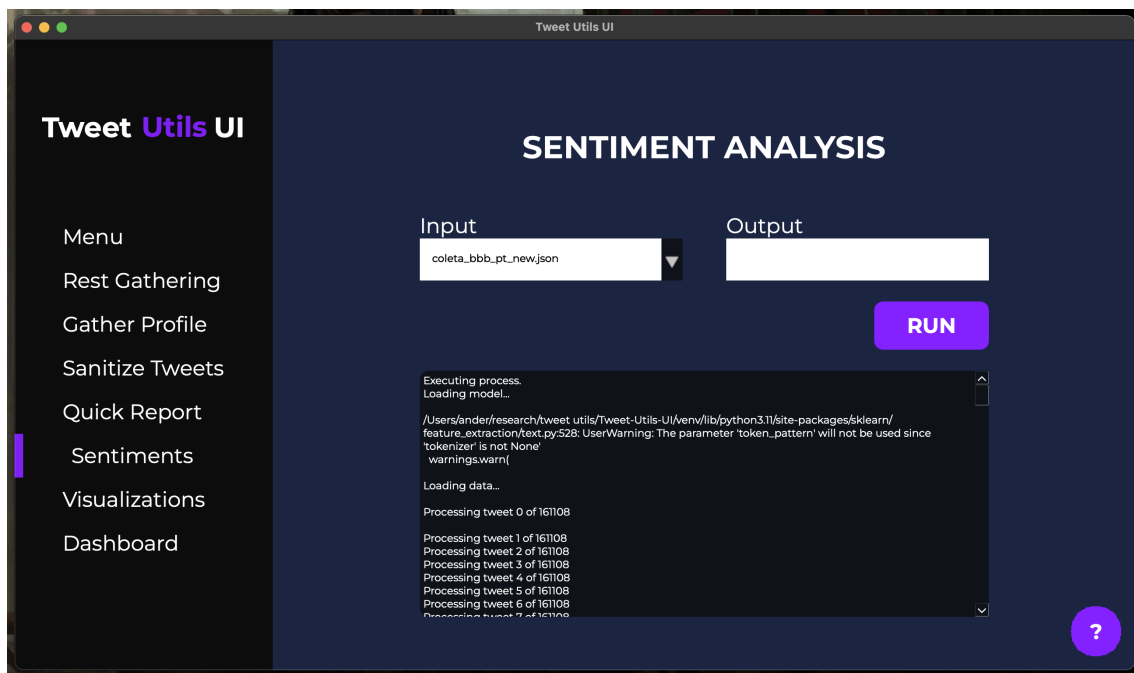


Figura 3. Interface da ferramenta para análise de sentimentos.

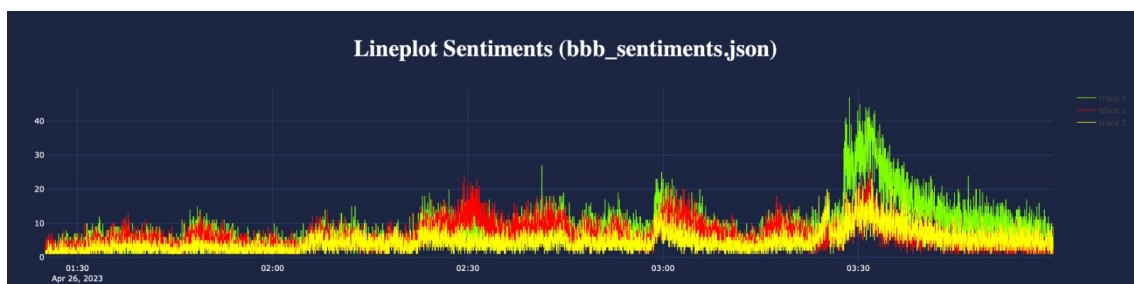


Figura 4. Visualização gerada pela ferramenta para a análise dos sentimentos encontrados nos tuítes.

padrões em um conjunto de dados de treinamento previamente categorizados. O algoritmo mapeia os novos dados de entrada em um espaço junto com os dados de treinamento e busca encontrar similaridades com base na proximidade desses dados. Com base nas correlações encontradas, o *SVM* classifica os novos dados, ou seja, determina a categoria de sentimento associada ao texto. [Noble 2006]

A classificação de sentimentos é feita na função “sentiment analysis”, informando um arquivo *JSON* com uma coleta feita pela ferramenta. Nela, é gerada uma cópia do arquivo contendo uma chave ‘emotion’ com o sentimento encontrado em cada tuíte. Para a visualização da análise de sentimentos, foi criado um novo gráfico de linha (ver Figura 4), com uma linha para cada classe de sentimento: *Positive* (verde), *Negative* (vermelho) e *Neutral* (amarelo). O gráfico é gerado automaticamente pela função “visualizations” quando o arquivo a ser analisado possui classificação de sentimentos.

4.2. Outras Contribuições

Uma das maiores limitações identificadas na ferramenta estava relacionada ao limite de tempo para a extração de tuítes. Antes, os tuítes só podiam ser coletados dentro de um período de uma semana a partir da sua publicação. No entanto, durante a análise da *API* do *Twitter*, foi descoberto que existiam dois *endpoints* para a coleta de tuítes: o “*recent search*”, que se restringe aos últimos 7 dias, e o “*full archive search*”, que não possui limites de datas, mas é restrito ao público acadêmico.

Durante a implementação do consumo da *API* no aplicativo, foi observado que estava sendo utilizado o *endpoint* de “*recent search*”. Levando em consideração o contexto de uso da ferramenta, que envolve coletas de dados de nível acadêmico, foi decidido mudar para o *endpoint* de “*full archive search*”. No entanto, ao utilizar a biblioteca *Tweepy* [Roesslein 2020], foi identificado um problema de limite de taxa de coleta de 300 tuítes a cada 15 minutos, o que dificulta a coleta de grandes quantidades de tuítes, que muitas vezes ultrapassam centenas de milhares.

Durante um teste realizado com a ferramenta *Postman*, foi possível obter 500 tuítes, indicando que essa limitação documentada pelo *Twitter* e possivelmente implementada no código do *Tweepy* não está sendo aplicada efetivamente pelo *Twitter*. Diante dessa situação, foi decidido realizar testes com uma segunda biblioteca de coleta de tuítes chamada *Twarc* [Summers et al. 2023], que possui uma limitação de taxa de coleta que não afeta a viabilidade da coleta. Resultados satisfatórios foram obtidos e serão detalhados na próxima seção. Dessa forma, foi possível implementar uma coleta de dados com mais de uma semana sem grandes limitações de taxa de coleta.

Outra grande limitação encontrada anteriormente estava relacionada à incompatibilidade com diferentes versões de sistemas operacionais. O algoritmo de coleta da ferramenta assumia que a data fosse sempre informada no formato “dd/MM/yyyy HH:mm”. No entanto, não havia um formato padrão de data configurado para o campo de entrada da data, e esse valor era repassado sem nenhum tratamento para a coleta. Isso foi corrigido configurando os objetos “*QDateTimeEdit*”, responsáveis pela entrada das datas, com o formato de data esperado, o que eliminou a incompatibilidade que havia anteriormente em diferentes sistemas operacionais.

5. Estudo de Caso

Para validar as contribuições apresentadas na seção 4, foi feito um estudo de caso, de modo a analisar a opinião pública de telespectadores do episódio final do *reality show Big Brother Brasil 2023 (BBB23)*. O programa foi escolhido por ser uma das atrações mais populares da televisão brasileira, além de ter os telespectadores decidindo através de votações as eliminações e o vencedor da edição. Isso resulta em um grande engajamento nas redes sociais.

```
"BBB 23" OR "Big Brother Brasil" OR "paredão" OR "eliminação"
OR "final" OR "campeão" OR "campeã" OR "tadeu" OR "Big Boss"
OR "Boninho" OR "finalista"
```

A coleta de tuítes envolveu uma busca com as palavras-chaves acima entre aspas e usando o operador lógico “*OR*”, no período entre “22:25 do dia 25 de abril de 2023” até “1:00 do dia 26 de abril de 2023”, horário de Brasília. Esse período foi convertido ao

horário *UTC* antes de ser informado a ferramenta (ver Figura 1). Esta busca encontrou 161109 tuítes², em uma busca realizada no dia 04 de junho de 2023, 40 dias após a final do reality show. O conjunto de dados coletado foi pré-processado pela função “*sanitize tweets*”, para então ter seus sentimentos classificados pela função “*sentiment analysis*”, resultando por fim em um *gráfico* de linhas gerado pela função “*visualizations*”³ (ver Figura 4).

Apesar de a sobreposição das linhas tornar difícil visualizar individualmente os dados, é possível selecionar individualmente os sentimentos a serem exibidos no gráfico. Os pico de tuítes que ocorreu no anúncio do vencedor da edição também dificulta a análise em um primeiro momento, porém pode-se selecionar com o cursor do mouse um período do gráfico eventuais picos que aumentam em demasia a escala do gráfico.

Na amostra coletada, foi possível em um primeiro momento identificar um pico de mensagens com sentimento positivo no anúncio da Amanda como vencedora do BBB23. Ao selecionar a parte do gráfico anterior ao anúncio, é notado um pico de mensagens negativas às 23:30 do dia 25 de abril, sendo uma possível reação negativa dos telespectadores a um fato ocorrido durante a transmissão ao vivo do programa.

6. Trabalhos Futuros

Existem duas sugestões para melhorias futuras no *Tweet Utils UI*. Primeiramente, a interface do usuário (UI) pode ser aprimorada para tornar a ferramenta ainda mais fácil de usar. Isso pode incluir a renderização das visualizações dentro do próprio programa (atualmente existe essa funcionalidade no menu “*dashboard*”, mas não está funcionando corretamente), a opção de selecionar as datas como entrada e uma descrição mais clara e intuitiva durante o uso da ferramenta, tornando-a amigável e orientada ao usuário. Isso permitiria uma navegação simples e compreensão clara das funcionalidades oferecidas.

Além disso, empacotar a aplicação em um formato mais acessível, como um instalador ou um aplicativo executável, pode facilitar a instalação e uso para usuários com pouco conhecimento técnico avançado. Isso eliminaria a necessidade de configurações complexas e permitiria que os usuários iniciassem a utilização da ferramenta de forma mais rápida e simples.

Em uma segunda análise, é interessante buscar melhorar a precisão da análise de sentimentos realizada pelo *Tweet Utils UI*, uma vez que o método atual apresenta uma acurácia de 71%. A análise de sentimentos é fundamental para compreender a opinião pública, e é importante que os resultados sejam o mais precisos possível. Para alcançar isso, é necessário aprimorar os algoritmos com mais treinamentos e técnicas utilizadas na análise de sentimentos.

Isso pode envolver o treinamento do modelo com conjuntos de dados mais abrangentes,⁴ usando algoritmos mais recentes, como por exemplo a melhoria da classificação

²Os arquivos json com os tuítes coletados, com e sem a classificação, estão disponíveis em: <https://github.com/andersprenger/effective-dollop>

³A visualização gerada com a análise de sentimentos está disponível em: <http://sprenger.poa.br/studious-octo-sniffle/insights/bbb23-final/sentiment-analysis.html>

⁴Um exemplo de conjunto de dados relevante para um novo treinamento, com cerca de 800 mil tuítes pré-classificados em português, está disponível em: <https://www.kaggle.com/datasets/augustop/portuguese-tweets-for-sentiment-analysis>

de palavras-chave e a consideração de contextos específicos para interpretar adequadamente o sentimento expresso em cada tweet. Um trabalho mais aprofundado nessa área pode resultar em uma análise de sentimentos mais precisa e confiável, aumentando assim a qualidade das informações fornecidas aos usuários do *Tweet Utils UI*.

7. Conclusão

Em suma, este artigo apresentou uma contribuição dos autores deste artigo na ferramenta *Tweet Utils UI*, desenvolvida por [Wagner and Manssour 2022] para coleta, processamento e visualização de dados do Twitter. A extensão adicionada ao *Tweet Utils UI* foi a análise de sentimentos, que permite a interpretação e classificação dos sentimentos expressos nos tuítes coletados.

A análise de sentimentos é uma técnica importante para compreender as opiniões e emoções presentes nos dados textuais. Neste trabalho, utilizamos um modelo de aprendizado de máquina desenvolvido por [Duarte and Manssour 2021], que segundo seus autores apresenta uma acurácia de 71% na classificação de sentimentos em tuítes em português. Foi também adicionado um gráfico de linha para visualização dos dados classificados. A partir deste fazer um breve estudo de caso sobre a opinião pública da final do reality show *Big Brother Brasil 23*.

Apesar da internet ser um ambiente com grande pluralidade de opiniões, foi possível identificar picos de mensagens onde houveram convergências na opinião pública sobre a transmissão ao vivo do programa. Houve um grande pico de tuítes com sentimento positivo quando a Amanda foi anunciada como vencedora do BBB23, e um pico menor de mensagens negativas no decorrer do programa (25 de abril de 2023, às 23:30, horário de Brasília), provavelmente refletindo uma reação negativa a este momento específico do programa.

Como trabalho futuro, é possível continuar aprimorando a análise de sentimentos, explorando outros modelos de aprendizado de máquina e técnicas de processamento de linguagem natural. Além disso, podem ser adicionadas mais funcionalidades ao *Tweet Utils UI* para oferecer ainda mais opções de visualização e análise dos dados coletados. Com essas melhorias contínuas, o *Tweet Utils UI* continuará a ser uma ferramenta poderosa para extrair *insights* valiosos do Twitter.

Referências

- Abbas, M., Memon, K. A., Jamali, A. A., Memon, S., and Ahmed, A. (2019). Multinomial naive bayes classification model for sentiment analysis. *IJCSNS Int. J. Comput. Sci. Netw. Secur*, 19(3):62.
- Agarwal, A., Xie, B., Vovsha, I., Rambow, O., and Passonneau, R. J. (2011). Sentiment analysis of twitter data. In *Proceedings of the workshop on language in social media (LSM 2011)*, pages 30–38.
- Data, M. C., Komorowski, M., Marshall, D. C., Saliciccioli, J. D., and Crutain, Y. (2016). Exploratory data analysis. *Secondary analysis of electronic health records*, pages 185–203.
- Duarte, G. H. and Manssour, I. H. (2021). Sentimentalpeaks: Um modelo de aprendizado de máquina para classificação e visualização de sentimentos em tweets.

- LaValley, M. P. (2008). Logistic regression. *Circulation*, 117(18):2395–2399.
- Noble, W. S. (2006). What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Roesslein, J. (2020). Tweepy: Twitter for python! *URL: <https://github.com/tweepy/tweepy>*.
- Summers, E., Brigadir, I., Hames, S., van Kemenade, H., Binkley, P., tinafigueroa, Ruest, N., Walmir, Chudnov, D., Thiel, D., Betsy, Chartier, R., celeste, Lin, H., Alice, Chosak, A., Lenz, M., McCain, R. M., Milligan, I., Segerberg, A., Shahrokhian, D., Walsh, M., Lausen, L., Woodward, N., eggplants, Ramaswami, A., Nguyen, B., Hereñú, D., Milajevs, D., and Elwert, F. (2023). Docnow/twarc: v2.14.0.
- Wagner, P. P. and Manssour, I. H. (2022). Tweet utils ui: Uma ferramenta para coleta e visualização de tweets.
- Yu, W., Wong, C. Y., Chavez, R., and Jacobs, M. A. (2021). Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture. *International Journal of Production Economics*, 236:108135.