BRIMO: uma Ferramenta para Análise de Sentimentos

Otávio Alves[1], **Taciana Pontual Falcão**[1], **George Valença**[1], **Ermeson Andrade**[1]

[1] Departamento de Computação – Universidade Federal Rural de Pernambuco (UFRPE)
Recife, PE – Brasil

{otavio.sales,taciana.pontual,george.valenca,ermeson.andrade}@ufrpe.br

Abstract. *Twitter is one of the fastest growing social networks in recent years, being one of the main platforms for people to share their opinions through short posts called tweets. Every trending topic generates rich discussions, and an analysis of people's sentiments about a topic provides a more comprehensive understanding of the subject. However, to our knowledge, there are no tools that make this type of analysis accessible to the general public. Thus, we propose BRIMO, a free web tool with simple and intuitive interface, which allows the analysis of sentiments quickly and objectively, through graphical representations of data. We evaluated BRIMO's usability and usefulness with target users. The results indicate that the tool has great usability according to the criteria used, in addition to being useful for several purposes, in the perception of participants.*

Resumo. *O Twitter é uma das redes sociais com maior crescimento nos últimos anos, sendo uma das principais plataformas para as pessoas compartilharem suas opiniões através de postagens curtas chamadas de tweets. Cada tópico do momento gera discussões, e ter uma análise do sentimento das pessoas sobre o tópico em questão provê um entendimento mais abrangente do assunto. No entanto, não há ferramentas de nosso conhecimento que tornem esse tipo de análise acessível ao público geral. Assim, propomos BRIMO, uma ferramenta web gratuita com interface simples e intuitiva, que permite a análise de sentimentos de forma rápida e direta através de representações gráficas. A usabilidade e utilidade da BRIMO foram avaliadas com representantes do público-alvo. Os resultados indicam que a ferramenta possui uma ótima usabilidade de acordo com os critérios utilizados, além de ser útil para diversos propósitos, na perspectiva dos participantes.*

1. Introdução

Atualmente, as redes sociais são uma parte fundamental no dia a dia das pessoas. Através delas as pessoas se informam, entram em contato com amigos e parentes, procuram emprego, encontram parceiros, e também expressam suas opiniões e sentimentos, sejam eles positivos ou negativos, acerca dos mais variados tópicos, produtos e serviços. O Twitter é uma das redes sociais mais utilizadas no mundo e grande parte de seu conteúdo é composto por experiências pessoais e opiniões, de forma majoritariamente textual e concisa devido a sua limitação de 280 caracteres por *tweet*[Statista 2022]. Isso torna esse ambiente ideal para se ter uma noção das emoções predominantes das pessoas em relação a determinados assuntos e se a percepção geral delas sobre esses assuntos é positiva ou negativa.

Nos dias atuais, as opiniões das pessoas podem mudar em curtos espaços de tempo devido, principalmente, à quantidade de informações que chega até elas pelos mais diversos meios [Burbach et al. 2020]. Por isso, para se fazer uma análise de sentimentos sobre um determinado assunto, torna-se necessário usar informações tão atualizadas quanto possível. A análise de sentimentos a partir de *tweets* em tempo real atende a essa necessidade, e pode auxiliar as pessoas (inclusive entidades públicas e privadas) em suas tomadas de decisão. Por exemplo, através dos dados do Twitter, é possível identificar surtos de dengue em determinadas áreas, e tomar ações a fim de resolver o problema o mais rápido possível e evitar a proliferação dos casos de dengue para outras áreas[Euzebio et al. 2020].

Há muitas pesquisas na área de análise de sentimentos e uma parte relevante delas utiliza o Twitter como base de dados. De um modo geral, essa análise permite extrair emoções e sentimentos de forma automatizada em qualquer documento quer seja formal ou informal. No entanto, não há ferramentas de nosso conhecimento que tornem esse tipo de análise acessível ao público geral. Este trabalho propõe BRIMO, uma ferramenta web para a análise de sentimentos a partir de textos em português com recursos direcionados tanto ao público em geral, quanto a empresas. A ferramenta permite análises a partir de palavras-chave e dois tipos de conjuntos de dados: (i) dados do Twitter; e (ii) qualquer texto fornecido pelo usuário. Na análise via Twitter, há dois módulos: o básico e o avançado. No modo básico, BRIMO requer apenas o uso de palavras-chaves, utilizando como base de dados os 100 *tweets* mais recentes, enquanto no modo avançado BRIMO permite a definição de parâmetros para análises mais avançadas como o intervalo para coleta dos dados e o tamanho da base. Na análise via texto fornecido pelo usuário, a ferramenta só requer que o usuário insira o texto. Como resultado, BRIMO retorna gráficos (análises de sentimentos e polarização), nuvens de palavras e uma lista completa com todos os *tweets* da base (quando aplicável). Todos esses resultados podem ser salvos diretamente pela interface. A ferramenta foi usada por 10 voluntários com o objetivo de avaliar a sua usabilidade e utilidade para análise de sentimentos.

O trabalho está organizado da seguinte maneira: A Seção 2 descreve brevemente algumas informações relacionadas a análise de sentimentos. A Seção 3 detalha os trabalhos relacionados. A Seção 4 apresenta a ferramenta proposta. A Seção 5 expõe os resultados alcançados relacionados a avaliação da ferramenta. Por fim, a Seção 6 descreve o fechamento dos resultados alcançados e brevemente detalha os trabalhos futuros.

2. Análise de Sentimentos

O processo básico da mineração de texto consiste em analisar grandes quantidades de texto em linguagem natural e detectar padrões léxicos ou linguísticos. Como consequência, é possível extrair informações úteis a partir dos padrões detectados [Lokesh 2013]. Nesse contexto, a análise de sentimentos é uma das técnicas frequentemente utilizadas para extrair o conteúdo emocional das palavras[Fang 2015]. Uma aplicação comum desta técnica é avaliar a opinião pública em relação a produtos, marcas e serviços visando gerar *insights* para a tomada de decisão.

A análise de sentimentos requer a execução de um conjunto de etapas para a extração de sentimentos a partir de palavras, sendo o pré-processamento a primeira delas. Nessa etapa, são removidos do texto todos os caracteres especiais e as *stop words*,

que são palavras que não agregam informação para o texto. Em seguida, todos os símbolos de pontuação são removidos e o texto é convertido para letras minúsculas. Posteriormente, técnicas de *stemming* e *lemmatization* podem ser aplicadas, com o objetivo de transformar inflexões ou derivações de uma palavra para uma forma base da palavra que contenha o mesmo significado[Cardozo and Freitas 2021]. Por fim, as sentenças são convertidas em palavras através do processo de tokenização. Após o pré-processamento, é aplicado o método escolhido de análise de sentimentos. Vale destacar que existem vários métodos de análise de sentimentos disponíveis na literatura. Por exemplo, os métodos baseados em dicionários léxicos, e os métodos que utilizam aprendizagem de máquina[Ribeiro and da Silva 2018]. Neste trabalho, especificamente, para a classificação dos sentimentos dos *tweets*, escolhemos utilizar a biblioteca *Syuzhet* do R que é baseada em um dicionário léxico. De acordo com [Misuraca et al. 2020], *Syuzhet* possui uma das melhores acurácias em termos de predição dos sentimentos.

3. Trabalhos Relacionados

Na literatura, existem diversos trabalhos que usam o Twitter para diferentes fins relacionados a análise de sentimentos. Em [França de Souza et al. 2017], é abordada a criação de um método computacional para análise de sentimentos baseado em um dicionário léxico, onde os autores utilizaram o Twitter como fonte de dados para o experimento. Os resultados são comparados com 16 outros métodos de análise de sentimentos textuais. O diferencial dessa abordagem é ser capaz de analisar textos em português brasileiro de forma nativa sem precisar fazer a tradução para o inglês, ganhando uma maior precisão na polarização de palavras de acordo com os sentimentos por elas transmitidos. Em [Wunderlich and Memmert 2020], os autores testam como ferramentas de análise de sentimentos baseadas em dicionário léxico performam em relação a análise manual, tendo como objetivo mostrar o quão úteis elas podem ser para estimular o aumento da adoção das mesmas.

Outros trabalhos propõem ferramentas para as análises de sentimentos usando os *tweets*. Em [Lauand and Oliveira 2014], os autores apresentam um aplicativo para dispositivos móveis que utiliza o Twitter como fonte de dados para fazer as análises das condições do trânsito do estado do Rio de Janeiro. O aplicativo usa os *tweets* de fontes oficiais e não-oficiais de controle do trânsito para inferir as condições de ruas ou avenidas, onde os *tweets* considerados positivos pela análise de sentimentos indicam um trânsito provavelmente livre e os *tweets* considerados negativos indicam um trânsito provavelmente engarrafado. Já em [Murakami et al. 2020], foi desenvolvida uma ferramenta que exibe gráficos de barras, listas de *tweets*, gráficos de pizza e listas de palavras que indicam as emoções positivas/negativas de usuários do Twitter. No entanto, essa ferramenta tem como foco o apoio ao auto-conhecimento através da extração da polarização das emoções dos *tweets* do próprio usuário.

Também existem trabalhos que focam puramente na análise de sentimentos através de dados obtidos do Twitter. Em [Marcec and Likic 2021], os autores analisam os sentimentos do mundo em relação às vacinas da Covid-19, através de *tweets*. Esse trabalho também faz o uso de um dicionário léxico para as análises e traz considerações importantes sobre os resultados da polarização das pessoas em relação às vacinas. Já em [Silva et al. 2021], os autores analisam os sentimentos do povo brasileiro em relação ao SUS (Sistema Único de Saúde) através de *tweets* feitos antes e durante a pandemia do

Covid-19 no período entre Dezembro de 2019 a Outubro de 2020. Em tal trabalho, é utilizado o pacote Syuzhet para as análises de sentimentos, de modo que as mesmas foram divididas de acordo com os principais eventos que aconteceram durante o período da análise.

Os trabalhos relacionados apresentados utilizam o Twitter como base de dados para diversos fins como a análise de sentimentos, a execução de estudos de caso, o aprimoramento de métodos e até o desenvolvimento de ferramentas para resolver problemas reais. Porém, nenhuma dessas abordagens/ferramentas permite que uma pessoa possa fazer sua própria análise de sentimentos sobre qualquer tema de forma rápida e intuitiva no idioma português. A ferramenta proposta neste trabalho pode ser aplicada com o objetivo de beneficiar tanto o público em geral, quanto as empresas.

4. BRIMO

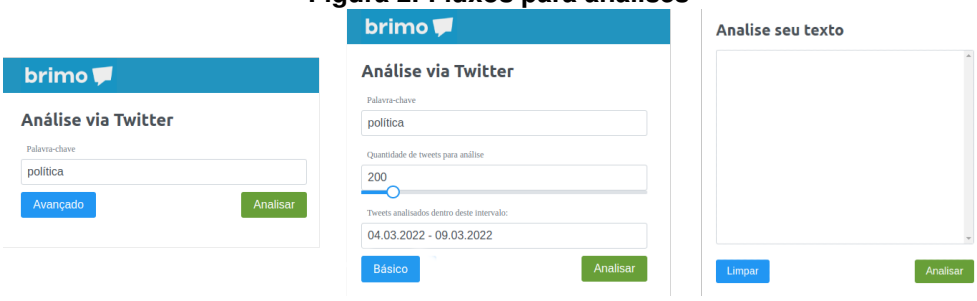
Brimo[1] possui sua estrutura totalmente hospedada no Heroku, que é um PaaS (plataforma como serviço), que possibilita operar todo o ciclo de vida da aplicação inteiramente na nuvem. Nessa estrutura, a ferramenta é dividida entre *front-end* e *back-end*, sendo o *back-end* composto por dois micro-serviços que fornecem APIs para comunicação entre si. Já o *front-end* é responsável pela interface gráfica com a qual o usuário irá interagir e inserir as entradas necessárias para a análise de sentimentos. No desenvolvimento do *front-end* foi utilizado o *framework* Angular que é bastante popular no desenvolvimento de aplicações de página única por proporcionar a arquitetura e as ferramentas necessárias para que seja possível focar apenas no desenvolvimento do código da aplicação. A Figura 1 mostra uma visão geral da interface da ferramenta com uma busca padrão da palavra-chave “política“.

Figura 1. Visão geral da interface da ferramenta

Existem 3 fluxos possíveis para análises na ferramenta, dois para análise via Twitter e um para análise via texto. O primeiro fluxo é a análise via Twitter no modo padrão que tem como entrada apenas a palavra-chave que será usada para busca no Twitter, como mostra a Figura 2, à esquerda. O segundo fluxo é a análise via Twitter no modo avançado que tem como entrada 3 parâmetros: a palavra-chave, a quantidade de *tweets* que serão analisados e um filtro para selecionar o intervalo de dias em que esses *tweets* foram postados, como mostra a Figura 2, ao centro. Por último, a análise via texto, em que o usuário pode inserir um texto de sua escolha (ver Figura 2, à direita) e a ferramenta retorna a análise de sentimentos do texto em questão.

[1]https://brimo-front.herokuapp.com/

Figura 2. Fluxos para análises



O *back-end* foi dividido em duas APIs REST, uma utilizando o *microframework* Flask para Python, e a outra utilizando o pacote Plumber para o R. O R foi escolhido para utilizar o pacote Syuzhet para análise de sentimentos. Mais especificamente, esse pacote implementa o dicionário léxico de emoções NRC (*National Research Council Canada*), que apesar de ter como idioma principal o inglês, ele permite a avaliação de sentimentos em vários idiomas através da tradução automática. Note que o pacote associa palavras com as emoções e sentimentos, de modo a compará-las para identificar as emoções e sentimentos presentes no *tweet* analisado. Adicionalmente, o dicionário tem 14182 palavras com as categorias de sentimentos "positivos" e "negativos" e as emoções de "raiva", "antecipação", "nojo", "medo", "alegria", "tristeza", "surpresa" e "confiança".

O fluxo de processamento da análise via Twitter começa com uma requisição para o servidor através do *endpoint* "/classify/tweet" da API em Python, passando os parâmetros inseridos pelo usuário. Esses parâmetros são utilizados para obter os *tweets* através da API V2 do Twitter. Contudo, foi necessário implementar o processo de obtenção dos *tweets* de forma assíncrona e rodando em plano de fundo, pois o HTTP (protocolo padrão de transporte da Internet utilizado pelas APIs REST para comunicação) é um protocolo melhor utilizado em situações de solicitação ou resposta síncrona, o que não seria possível com a tarefa de obter os *tweets* que pode chegar a ser bastante demorada dependendo da quantidade de *tweets* a serem coletados e analisados.

Consequentemente, a tarefa de obter os *tweets* é colocada em uma fila de trabalho assíncrona utilizando o Celery, que necessita de um servidor de mensagens para armazenar o progresso da tarefa, e o Redis foi adotado para armazenar os dados obtidos da busca por ser extremamente rápido para tarefas de vida curta em que não é necessária a persistência. Assim que essa tarefa é concluída com sucesso, os *tweets* passam por uma etapa de pré-processamento para limpeza dos dados, onde todos os caracteres são convertidos para minúsculos e então são removidos os caracteres especiais, os símbolos de pontuação, as *stop words*, os espaços em branco desnecessários, além da etapa de *stemming* e *lemmatization*. Vale salientar que também foi utilizada uma funcionalidade que a API do *Twitter* fornece para que sejam passados parâmetros diretamente na consulta que removem alguns tipos de *tweets* que não são interessantes para a análise, como conteúdo de mídia e *links*. Em seguida, os dados tratados são enviados para os *endpoints* da API em R, que é a API responsável por devolver os resultados da análise de sentimentos e o *corpus* utilizado para criar as nuvens de palavras.

Ao término desse processo, através de um processo de *polling* do *front-end*, a aplicação sabe que a tarefa foi concluída e recupera o resultado, o qual será utilizado pra

montar os gráficos. Os gráficos de barras, de radar e de rosca foram implementados utilizando a biblioteca ChartJS através da biblioteca de componentes PrimeNG para Angular. A nuvem de palavras, por outro lado, foi implementada com a biblioteca Wordcloud2. O gráfico de barras (Figura 3, à esquerda) mostra o percentual da presença de cada emoção (tristeza, alegria, medo, nojo, raiva e surpresa) nos *tweets* analisados a partir da palavra-chave "política".As emoções antecipação e confiança foram omitidas por serem emoções consideradas secundárias, com o objetivo de simplificar a interpretação dos usuários. O gráfico de radar (Figura 3, à direita) possui a mesma análise do gráfico em barras, porém é uma forma de apresentar vários dados que possuem mais de 1 eixo em um único lugar. O gráfico de rosca (Figura 4, à esquerda) mostra a polarização dos sentimentos com o percentual de sentimentos positivos e negativos. Por fim, a nuvem de palavras (Figura 4, à direita) mostra as palavras que mais aparecem dentro dos *tweets* analisados, dentre elas as mais citadas ficam em destaque.

Figura 3. Gráfico de barras e de radar

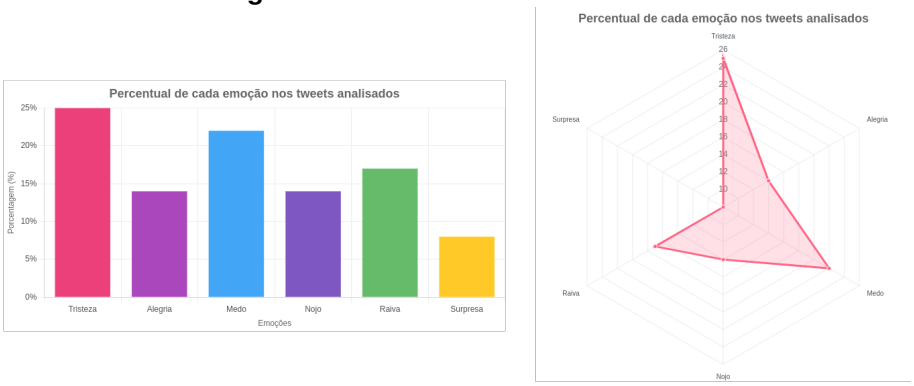
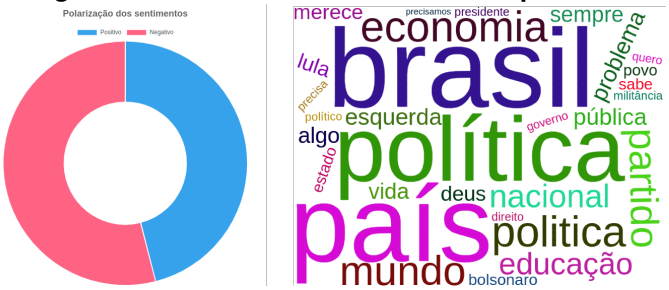


Figura 4. Gráfico de rosca e nuvem de palavras



5. Avaliação e Resultados

A seguir, apresentamos os resultados das avaliações de BRIMO com representantes do público-alvo.

5.1. Avaliação

A avaliação da ferramenta foi realizada com 10 voluntários, dentre eles 5 profissionais de TI e 5 alunos de ensino superior, sendo 3 de Ciência da Computação, 1 de Pedagogia e 1 de Ciências Biológicas. Os profissionais de TI foram recrutados através de um convite em um canal geral interno de uma empresa. Já os alunos de Ciência da Computação se

voluntariaram através de um convite feito em grupo de pesquisa. Por fim, os estudantes de Pedagogia e Ciências Biológicas foram recrutados através de um convite aberto em uma postagem no Twitter. O objetivo da avaliação foi coletar opiniões de representantes do público-alvo, assim como avaliar a usabilidade da interface de BRIMO. Para isso, foram coletadas métricas de uso, assim como respostas dos participantes através de um questionário.

O procedimento consistiu em um teste de usabilidade [Barbosa and Silva 2010], cujo roteiro é apresentado na Tabela 1. Estudos mostram que a partir de 5 voluntários, os resultados do teste de usabilidade tendem a se repetir, conseguindo identificar aproximadamente 80% dos problemas [Nielsen 2000]. O teste foi realizado individualmente e de forma remota através da plataforma Google Meet com duração aproximada de 10 minutos. Com a autorização do participante, a tela foi gravada durante a realização das tarefas, e foi informado que eles poderiam desistir de qualquer tarefa a qualquer momento. Foi esclarecido aos participantes que não seria avaliado o desempenho deles e sim a usabilidade da ferramenta.

Tabela 1. Roteiro de tarefas do teste de usabilidade

Tarefas do teste de usabilidade	
Tarefa 1:	Faça uma análise de sentimentos via Twitter utilizando a palavra-chave "covid".
Tarefa 2:	Navegue até o gráfico de barra e explique seu entendimento sobre o que você está vendo.
Tarefa 3:	Navegue até o gráfico de radar e explique seu entendimento sobre o que você está vendo.
Tarefa 4:	Navegue até o gráfico de polarização e explique seu entendimento sobre o que você está vendo.
Tarefa 5:	Navegue até a nuvem de palavras e explique seu entendimento sobre o que você está vendo.
Tarefa 6:	Navegue até a lista de *tweets* e explique seu entendimento sobre o que você está vendo.
Tarefa 7:	Faça uma análise de sentimentos via Twitter no modo avançado utilizando a palavra-chave "covid" com a quantidade de 200 *tweets* analisados no intervalo dos últimos 3 dias.
Tarefa 8:	Faça a análise de sentimentos de um texto de sua escolha. Caso não possua nenhum em mente, como sugestão, o texto pode ser retirado de portais de notícia, letras de música, enciclopédias, etc.

Seguindo o protocolo do *thinking aloud* ("pensando em voz alta") [Barbosa and Silva 2010], foi solicitado que durante as tarefas, o participante falasse em voz alta seus pensamentos, dificuldades e entendimento sobre o que estava fazendo e vendo. Ao final da chamada, foi enviado o link de um formulário no Google Forms contendo um teste do SUS (Escala de Usabilidade do Sistema) [Brooke 1995] que consiste de 10 itens (Tabela 2). Para cada uma delas, o participante responde em uma escala Likert de 1 a 5, onde 1 significa Discordo totalmente e 5 significa Concordo totalmente. Esse teste foi escolhido por envolver um estudo qualitativo, por identificar possíveis problemas de usabilidade e de funcionalidade de BRIMO.

5.2. Resultados

As Tabelas 3, 4, e 5 mostram, respectivamente, os resultados das tarefas 1, 7 e 8 para as quais foram coletadas um conjunto de métricas: tempo de execução da tarefa; quantidade de erros (cliques durante a interação que não correspondem ao fluxo esperado para a

Tabela 2. Itens da Escala de Usabilidade

10 itens do questionário do SUS (Escala de Usabilidade do Sistema)	
Item 1:	Eu usaria esse sistema com frequência.
Item 2:	Esse sistema é desnecessariamente complexo.
Item 3:	O sistema é fácil de usar.
Item 4:	Eu precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar o sistema.
Item 5:	As funções do sistema estão muito bem integradas.
Item 6:	O sistema apresenta muita inconsistência.
Item 7:	As pessoas aprenderão como usar esse sistema rapidamente.
Item 8:	O sistema é atrapalhado de usar.
Item 9:	Eu me senti confiante ao usar o sistema.
Item 10:	Eu precisei aprender várias coisas novas antes de conseguir usar o sistema.

realização da tarefa); e taxa de sucesso (percentual de voluntários que conseguiram concluir a tarefa). O tempo de referência é o tempo médio de execução sem erros, medido pelos pesquisadores por execução própria. Além disso, são também descritas as dificuldades percebidas nos testes a partir do protocolo *thinking aloud* e da análise da gravação das interações e, por fim, futuras melhorias a partir dessas dificuldades são apresentadas. Vale destacar que Tarefa 1 (ver Tabela 3) não possui os campos de dificuldades e melhorias, pois nenhum dos voluntários teve dificuldade com o fluxo. Já a Tabela 6 mostra os resultados das tarefas 2, 3, 4, 5 e 6, que se referem à percepção geral dos participantes sobre as cinco visualizações exibidas pela ferramenta.

Tabela 3. Tarefa 1: Análise via Twitter com palavra-chave

		Tempo de execução da tarefa	Conseguiu executar com sucesso	Quantidade de erros	Tempo médio de execução	Tempo de referência	Taxa de sucesso
Tarefa 1:	Voluntário 1:	8 segundos	Sim	0	5,8 segundos	6 segundos	100%
	Voluntário 2:	8 segundos	Sim	0			
	Voluntário 3:	3 segundos	Sim	0			
	Voluntário 4:	3 segundos	Sim	0			
	Voluntário 5:	4 segundos	Sim	0			
	Voluntário 6:	4 segundos	Sim	0			
	Voluntário 7:	15 segundos	Sim	0			
	Voluntário 8:	3 segundos	Sim	0			
	Voluntário 9:	4 segundos	Sim	0			
	Voluntário 10:	6 segundos	Sim	0			

A Figura 5 mostra o resultado das avaliações dos voluntários sobre os itens do questionário SUS (Tabela 2). O questionário segue uma ordem específica de itens que não deve ser modificada para evitar que o voluntário de forma enviesada responda todas as alternativas com a mesma resposta. Assim, os itens são intercalados com uma de cunho positivo e uma outra de cunho negativo. A ordem também importa no cálculo da pontuação que veremos a seguir.

Todos os itens ímpares, em que o resultado desejado é que o voluntário selecione as opções mais à direita (Concordo, Concordo totalmente), estão mostrando um resultado bastante positivo, em especial, os itens 3, 7 e 9 que tratam do quão fácil é aprender a utilizar BRIMO. Já nos itens pares, em que o resultado desejado é que sejam selecionadas as opções mais à esquerda (Discordo, Discordo totalmente), não houve nenhuma escolha neutra ou contrária ao ideal. Adicionalmente, a escala SUS possui uma pontuação que pode ir de 0 a 100, sendo que 85 pontos ou mais representa usabilidade excepcional,

Tabela 4. Tarefa 7: Análise no modo avançado

		Tempo de execução da tarefa	Conseguiu executar com sucesso	Quantidade de erros	Tempo médio de execução	Tempo de referência	Taxa de sucesso
Tarefa 7:	Voluntário 1:	31 segundos	Sim	3	32,1 segundos	30 segundos	100%
	Voluntário 2:	21 segundos	Sim	0			
	Voluntário 3:	44 segundos	Sim	1			
	Voluntário 4:	23 segundos	Sim	0			
	Voluntário 5:	23 segundos	Sim	0			
	Voluntário 6:	26 segundos	Sim	0			
	Voluntário 7:	40 segundos	Sim	1			
	Voluntário 8:	32 segundos	Sim	1			
	Voluntário 9:	29 segundos	Sim	0			
	Voluntário 10:	52 segundos	Sim	2			
Dificuldades:	Todos os erros cometidos pelos voluntários foram causados pelo fluxo de selecionar o intervalo das datas no componente do calendário. O componente requer que o usuário selecione duas datas, sendo que a segunda deve ser ao menos 1 dia após a primeira para formar um intervalo válido de datas. Caso a ordem seja invertida, ou seja, a segunda data selecionada seja uma data anterior a primeira, o componente não reconhece como um intervalo válido e mantém apenas a segunda data selecionada aguardando pela seleção da próxima data válida pra completar o intervalo. Os participantes que erraram essa etapa inverteram essa ordem de seleção.						
Melhorias:	Colocar uma mensagem flutuante ao abrir o componente do calendário, indicando que deve ser selecionada primeiro a data de início e depois a data de fim.						

Tabela 5. Tarefa 8: Análise de texto de escolha do participante

		Tempo de execução da tarefa	Conseguiu executar com sucesso	Quantidade de erros	Tempo médio de execução	Tempo de referência	Taxa de sucesso
Tarefa 8:	Voluntário 1:		Não		25,1 segundos	35 segundos	80%
	Voluntário 2:	22 segundos	Sim	0			
	Voluntário 3:	27 segundos	Sim	0			
	Voluntário 4:	21 segundos	Sim	0			
	Voluntário 5:	25 segundos	Sim	0			
	Voluntário 6:		Não				
	Voluntário 7:	35 segundos	Sim	0			
	Voluntário 8:	6 segundos	Sim	0			
	Voluntário 9:	37 segundos	Sim	0			
	Voluntário 10:	28 segundos	Sim	0			
Dificuldades:	Os voluntários que não conseguiram executar com sucesso essa tarefa não compreenderam o objetivo da funcionalidade, que é produzir uma análise de sentimentos utilizando um texto qualquer. Eles entenderam que seria uma análise de sentimentos via Twitter, porém com um grupo de palavras-chave.						
Melhorias:	Colocar uma mensagem no campo de texto, indicando que ali deve ser inserido texto que será analisado.						

e uma pontuação abaixo de 70 representa uma usabilidade inaceitável. Vários estudos mostram que a SUS é extremamente confiável, sendo uma excelente métrica pra validar a usabilidade em vários contextos [Grier et al. 2013]. A Tabela 7 mostra que o resultado final obtido para BRIMO foi de 93,75, estando acima dos 85 pontos necessários para ser considerado uma usabilidade excepcional. Isso mostra que BRIMO é uma ferramenta fácil de usar e que atende as necessidades do público-alvo.

Tabela 6. Tarefas 2 a 6: Compreensão das visualizações

Tarefa 2 - Gráfico em barras:	Percepções dos voluntários:	Todos os voluntários entenderam que esse gráfico mostra a divisão das emoções relacionadas à palavra-chave "covid" no Twitter.
	Considerações / Melhorias futuras:	O gráfico de barra se mostrou bastante eficaz, visto que todos os voluntários conseguiram compreender sem dificuldades a informação que ele apresenta.
Tarefa 3 - Gráfico de radar:	Percepções dos voluntários:	Os voluntários perceberam que esse gráfico apresenta a mesma análise do gráfico anterior, porém com uma visualização diferente. Mas alguns mencionaram preferir o gráfico anterior.
	Considerações / Melhorias futuras:	O gráfico de radar não trouxe valor adicional à análise, pois o gráfico em barra mostra as mesmas informações de forma melhor, segundo os participantes. Porém, é uma visualização extra que tem o potencial de agregar e não prejudica a usabilidade da ferramenta.
Tarefa 4 - Gráfico de rosca:	Percepções dos voluntários:	Os voluntários perceberam que esse gráfico mostra o percentual de sentimentos positivos e negativos relacionados à palavra-chave "covid".
	Considerações / Melhorias futuras:	Alguns tiveram dúvidas sobre o resultado das porcentagens, pois acharam um pouco inconsistente em relação à divisão das emoções nos gráficos anteriores. Isso se deve ao fato que duas emoções ("antecipação" e "confiança") foram omitidas da análise na tentativa de simplificar a visualização, por serem consideradas emoções secundárias. Porém, essas considerações mostraram que é necessário retornar tais emoções ao gráfico.
Tarefa 5 - Nuvem de palavras:	Percepções dos voluntários:	Os voluntários compreenderam que a nuvem de palavras mostra as palavras que mais apareceram nos *tweets* com a palavra-chave "covid" e que as palavras maiores são as que aparecem com mais frequência. Um dos voluntários disse que não achou muito agradável a disposição das palavras na nuvem, devido ao fato de haver palavras menores dentro dos espaços das palavras maiores.
	Considerações / Melhorias futuras:	Os voluntários compreenderam bem a informação passada, porém a disposição das palavras na nuvem pode ser melhorada para ser mais agradável visualmente.
Tarefa 6 - Lista de tweets:	Percepções dos voluntários:	Os voluntários tiveram algumas percepções distintas sobre a lista de *tweets*. Primeiro, de que eram os *tweets* utilizados como base para as análises. Segundo, era a lista dos últimos *tweets* sobre a palavra-chave "covid". Por fim, era a lista dos *tweets* mais relevantes sobre a palavra-chave "covid". Também foi levantado por um dos voluntários a possibilidade de ter em cada *tweet* as informações de data e hora.
	Considerações / Melhorias futuras:	O entendimento de que os *tweets* da lista foram os utilizados para a análise foi percebido pela maioria, porém seria interessante acrescentar essa informação no cabeçalho do componente, de forma que fique mais claro.

6. Conclusões

A ferramenta BRIMO visa facilitar e popularizar o uso da análise de sentimentos, de forma que não seja necessário conhecimento prévio para que as pessoas possam utilizar essa tecnologia em seu dia a dia para os mais diversos fins. Foram feitos testes com usuários para avaliar a usabilidade e utilidade da ferramenta com objetivo de saber se ela funciona como deveria e se ela atende aos objetivos propostos nesse trabalho. Os resultados foram bastante positivos, desde as métricas coletadas no roteiro de tarefas até as pontuações obtidas na escala SUS. Como trabalho futuro pretendemos acrescentar mais palavras ao dicionário léxico e uma etapa extra de pré-processamento dos *tweets*, visto que ao utilizar o Twitter como base de dados temos um vocabulário muito informal e cheio de ruídos, que pode atrapalhar a precisão da análise. Além disso, almejamos superar as limitações de coleta do Twitter, como o limite de *tweets* e o período de análise.

Figura 5. Resultado do questionário

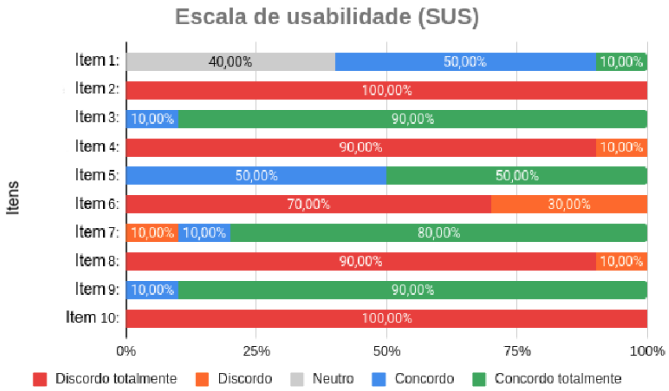


Tabela 7. Pontuação SUS

Itens / SUS / Voluntários	1	2	3	4	5	6	7	8	9	10	Pontuação SUS
1	4	1	4	1	4	2	2	1	4	1	80
2	4	1	5	1	4	1	4	1	5	1	92,5
3	5	1	5	2	5	1	5	2	5	1	95
4	3	1	5	1	5	1	5	1	5	1	95
5	4	1	5	1	5	2	5	1	5	1	95
6	4	1	5	1	4	2	5	1	5	1	92,5
7	3	1	5	1	5	1	5	1	5	1	95
8	3	1	5	1	4	1	5	1	5	1	92,5
9	4	1	5	1	5	1	5	1	5	1	97,5
10	3	1	5	1	4	1	5	1	5	1	92,5
										Pontuação final SUS:	93,75

Referências

[Barbosa and Silva 2010] Barbosa, S. and Silva, B. (2010). *Interação humano-computador*. Elsevier Brasil.

[Brooke 1995] Brooke, J. (1995). Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189.

[Burbach et al. 2020] Burbach, L., Halbach, P., Ziefle, M., and Calero Valdez, A. (2020). Opinion formation on the internet: The influence of personality, network structure, and content on sharing messages online. *Frontiers in Artificial Intelligence*, 3.

[Cardozo and Freitas 2021] Cardozo, L. and Freitas, L. (2021). Análise de sentimentos: Avaliando o desempenho de pré-processamento e de algoritmos de aprendizagem de máquina sobre o dataset tweetsentbr. In *Anais do X Brazilian Workshop on Social Network Analysis and Mining*, pages 169–174, Porto Alegre, RS, Brasil. SBC.

[Euzebio et al. 2020] Euzebio, C., Agy, S., Boldorini Jr, C., Porto, L., Alcarás, J. R., Martinez, A., and Ruiz, E. (2020). Statistical analysis of small twitter data collection to identify dengue outbreaks. In *Anais do VIII Symposium on Knowledge Discovery, Mining and Learning*, pages 17–24. SBC.

[Fang 2015] Fang, X. Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(5):1–14.

[França de Souza et al. 2017] França de Souza, K., Ramos Pereira, M. H., and Hasan Dalip, D. (2017). Unilex: Método léxico para análise de sentimentos textuais sobre conteúdo de tweets em português brasileiro. *Abakós*, 5(2):79–96.

[Grier et al. 2013] Grier, R. A., Bangor, A., Kortum, P., and Peres, S. C. (2013). The system usability scale: Beyond standard usability testing. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1):187–191.

[Lauand and Oliveira 2014] Lauand, B. P. and Oliveira, J. (2014). Inferindo as condições de trânsito através da análise de sentimentos no twitter. *iSys - Brazilian Journal of Information Systems*, 10(1):56–74.

[Lokesh 2013] Lokesh, K. Parul, K. B. (2013). Text mining: Concepts, process and applications. *Journal of Global Research in Computer Science*, 4(3):36–39.

[Marcec and Likic 2021] Marcec, R. and Likic, R. (2021). Using twitter for sentiment analysis towards astrazeneca/oxford, pfizer/biontech and moderna covid-19 vaccines. *Postgraduate medical journal*, pages postgradmedj–2021–140685.

[Misuraca et al. 2020] Misuraca, M., Forciniti, A., Scepi, G., and Spano, M. (2020). Sentiment Analysis for Education with R: packages, methods and practical applications. *arXiv preprint arXiv:2005.12840*.

[Murakami et al. 2020] Murakami, H., Ejima, N., and Kumagai, N. (2020). *Self-understanding Support Tool Using Twitter Sentiment Analysis*, pages 327–332. Springer, New York, NY.

[Nielsen 2000] Nielsen, J. (2000). Why you only need to test with 5 users. ”Disponível em: `https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/`”. Acesso em: 19 mai. 2022.

[Ribeiro and da Silva 2018] Ribeiro, A. P. and da Silva, N. F. F. (2018). Um estudo comparativo sobre métodos de análise de sentimentos em tweets. *Revista de Sistemas de Informação da FSMA*, (22):35–48.

[Silva et al. 2021] Silva, H., Andrade, E., Araújo, D., and Dantas, J. (2021). Sentiment analysis of tweets related to sus before and during covid-19 pandemic. *IEEE Latin America Transactions*, 20(1):6–13.

[Statista 2022] Statista (2022). Most popular social networks worldwide as of january 2022, ranked by number of monthly active users. ”Disponível em: `https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/`”. Acesso em: 19 mai. 2022.

[Wunderlich and Memmert 2020] Wunderlich, F. and Memmert, D. (2020). Innovative approaches in sports science—lexicon-based sentiment analysis as a tool to analyze sports-related twitter communication. *Applied sciences*, 10(2):431.