

FIQUEEMCASA: ANÁLISE DE SENTIMENTO DOS USUÁRIOS DO TWITTER EM RELAÇÃO AO COVID19

G. R. G. Pessanha¹, T. O. Fidelis², C. D. Freire³, E. A. Soares⁴

Universidade Federal de Alfenas^{1,2}, Universidade Federal de São Carlos³, Lancaster University⁴

ORCID ID: <https://orcid.org/0000-0003-0601-0085>³

clarissadourado21@hotmail.com³

Submetido 10/09/2020 - Aceito 14/12/2020

DOI: 10.15628/holos.2020.11147

RESUMO

O avanço da pandemia da COVID-19 pelo mundo fez com que houvesse um aumento exponencial nas buscas sobre o tema na internet. Neste contexto, as mídias sociais, especialmente o *Twitter*, são uma importante fonte de informação, possibilitando o acesso e compartilhamento instantâneo de conteúdos relativos à pandemia, mas também são meio de propagação de notícias falsas. Sendo assim, o objetivo geral deste trabalho é analisar o sentimento dos usuários do *Twitter* em relação a pandemia da Covid-19. A Análise de Sentimentos foi utilizada por meio do processamento de linguagem natural. Os dados do *Twitter* foram coletados no período de janeiro a julho de 2020. As *hashtags* Covid-19, FiqueEmCasa e suas variações foram utilizadas como recurso de busca e seleção dos *tweets*. Os resultados demonstram que o *Twitter* pode ser uma potencial

ferramenta a ser utilizada pela vigilância de surtos e epidemias, pois o conteúdo das mídias sociais pode ser usado para apoiar e aprimorar os sistemas de alerta. O estudo revelou uma relação entre as publicações do *Twitter* e o movimento de disseminação do novo coronavírus no Brasil. O sentimento negativo foi dominante denotando a preocupação da população em relação ao avanço da doença e as consequências por ele geradas. Ressalta-se que os resultados desta pesquisa são importantes para monitorar tendências e identificar os movimentos das postagens em torno de tópicos específicos e contribuem para melhorar a compreensão do conteúdo das mídias sociais durante uma emergência de saúde pública de interesse internacional, como é o caso da COVID19.

PALAVRAS-CHAVE: COVID19.Twitter. Análise de Sentimento.

FIQUEEMCASA: SENTIMENT ANALYSIS OF TWITTER USERS IN RELATION TO COVID19

ABSTRACT

The advancement of the COVID-19 pandemic around the world has led to an exponential increase in searches on the topic on the Internet. In this context, social media, especially Twitter, are an important source of information, allowing instant access and sharing of content related to the pandemic, but they are also a means of spreading false news. Therefore, the general objective of this work is to analyze the sentiment of Twitter users in relation to the Covid-19 pandemic. Sentiment Analysis was used through natural language processing. Twitter data was collected from January to July 2020. The hashtags Covid-19, StayEmCasa and their variations were used as a resource for searching and selecting tweets. The results demonstrate that Twitter can

be a potential tool to be used for surveillance of outbreaks and epidemics, since the content of social media can be used to support and improve alert systems. The study revealed a relationship between Twitter publications and the movement to disseminate the new coronavirus in Brazil. The negative feeling was dominant, denoting the population's concern about the progress of the disease and the consequences generated by it. It is noteworthy that the results of this research are important for monitoring trends and identifying the movements of posts around specific topics and contribute to improving the understanding of the content of social media during a public health emergency of international interest, as is the case with COVID-19.

KEYWORDS: COVID19.Twitter. Sentiment Analysis.

1. INTRODUÇÃO

As mídias sociais apresentaram uma nova forma dinâmica de comunicação que cria enormes quantidades de dados diariamente. Uma das consequências, é a facilidade e o maior acesso a informações que, podem gerar incrementos na educação e conhecimento da população. Neste cenário, o cidadão deixa de ser um mero receptor de conteúdo da internet e das mídias sociais e passa a ser também produtor, uma vez que, a partir da sua interação em mídias sociais, ele cria e compartilha conteúdo com grande alcance (SOUSA JÚNIOR, PETROLL, ROCHA, 2019).

O *Twitter* é uma das mídias sociais mais populares do mundo com, atualmente, 386 milhões de usuários ativos e o Brasil ocupa a 6ª posição na lista de países que lideram o ranking em número de usuários (STATICA, 2020). A referida mídia social tem sido utilizada para a identificação de tendências e formulação proativa de estratégias por meio da interlocução entre os usuários e suas reações em relação as instituições públicas, privadas e seus respectivos representantes. Segundo Odlum & Yoon (2015) o *Twitter* é visto como um meio de transmissão emergente de informações e notícias sobre eventos de saúde pública, evidenciada por sua utilidade durante as atividades de planejamento da pandemia de H1N1. Assim, o amplo alcance do *Twitter* tem o potencial de capturar tendências epidêmicas, coletar informações e disseminar conhecimento (ODLUM & YOON, 2015).

A situação crítica instaurada a partir da pandemia da COVID-19 afetou a população mundial trazendo desafios para todos os países afetados que buscam formas de retardar e controlar a disseminação do vírus (ALHAJJI et al., 2020; PASTOR, 2020; HEYMANN & SHINDO, 2020). No Brasil a situação não foi diferente, o país declarou Emergência em Saúde Pública de Importância Nacional (ESPIN) em decorrência da COVID-19 em 3 de fevereiro de 2020 (BRASIL, 2020) e, a partir de então, muito se tem discutido sobre estratégias de enfrentamento e contenção do vírus.

A pandemia da COVID-19, teve seu epicentro em uma pequena cidade na China se alastrando por todo o mundo. Em meados do mês de março, a Itália se tornou um dos países mais afetados e vários países do continente europeu anunciaram o fechamento das fronteiras, paralelamente a isso, os casos de COVID-19 no Brasil passam por curva crescente em diversos estados, fazendo com que o país também optasse por restringir as fronteiras, incluindo a União Europeia, China e Japão. A quarentena e o isolamento social foram adotados em diversos estados brasileiros com a suspensão de atividades presenciais. Nos meses subsequentes, o país continuou a sofrer severamente com os impactos da doença, principalmente nas áreas da saúde, educação e economia. Foi observado um desencontro entre orientações de ações de combate dos estados e o governo federal. De acordo com os dados do Ministério da Saúde (2020), o número de casos continua a crescer de forma alarmante, atingindo a marca de cerca de 1 milhão de pessoas infectadas em junho de 2020 saltando para mais de 2 milhões de casos confirmados no dia 23 de julho de 2020.

A partir da crescente exposição midiática nacional e internacional, observou-se um aumento exponencial do número de buscas sobre a COVID19 na internet (GARCIA FILHO et al., 2020). Neste contexto, as mídias sociais se configuram como uma importante fonte de informação, uma vez que possibilitam acesso e compartilhamento instantâneo de conteúdos relativos à pandemia. Liu et al.



(2020) afirmam que os meios tradicionais de comunicação podem demorar em fornecer os sentimentos dos usuários de *Internet* ao passo que as mídias sociais podem fornecer informações em escala e em tempo real.

Apesar de ser uma importante ferramenta de disseminação de informações para a conscientização da população, a facilidade de acesso, criação, divulgação e compartilhamento de conteúdo fez com que as mídias sociais se tornassem ambiente propício para a veiculação de notícias imprecisas e falsas, as chamadas *Fake News* (SOUSA JÚNIOR; PETROLL; ROCHA, 2019).

Considerando o contexto de incerteza e preocupação instaurado pela pandemia da COVID-19, as *Fake News* contribuem para a disseminação de informações falsas ou imprecisas, que são ineficientes na conscientização e no combate à doença, além de promover desinformação e gerar medo e caos. Esta combinação de fatores aumenta os desafios enfrentados pelos governos, especialmente, os órgãos de saúde e os órgãos envolvidos na contenção do novo vírus e representa um sério risco à saúde pública.

Sendo assim, monitorar o conteúdo gerado e compartilhado nas mídias sociais se torna importante não somente para identificar falsas informações veiculadas (LANA et al., 2020), mas também para compreender a visão, perspectiva e expectativas da população em relação ao momento crítico atual. Este entendimento possibilita a formulação de políticas públicas de saúde e educação que atendam demandas atuais e combatam a desinformação e o descrédito de informações repassadas pelos meios de comunicação oficiais e/ou tradicionais.

Neste contexto, este trabalho busca responder a seguinte questão: qual é o sentimento dos usuários do *Twitter* com relação a pandemia da COVID19 no Brasil?

Para isso propõe uma análise de sentimento em relação ao conteúdo das postagens dos usuários no *Twitter* no Brasil no período de janeiro a julho de 2020. A Análise de Sentimentos foi utilizada por meio do processamento de linguagem natural. Os dados do *Twitter* foram coletados no período de janeiro a julho de 2020. As *hashtags* Covid-19, FiqueEmCasa e suas variações foram utilizadas como recurso de busca e seleção dos *tweets*. Esse estudo revelou uma relação entre as publicações do *Twitter* e o movimento de disseminação do novo coronavírus no Brasil. Foi possível observar que o sentimento negativo foi dominante em todo o período analisado denotando a preocupação da população em relação ao avanço da doença e as consequências por ele geradas.

Neste contexto, acredita-se que a análise de conteúdo e de sentimentos expressos pelos usuários do *Twitter* desde os estágios iniciais da pandemia do Coronavírus 2019 pode se configurar como uma importante ferramenta no entendimento dos impactos do atual cenário nas emoções, impressões, crenças e pensamentos do público em geral. Estudos desta natureza permitem o mapeamento de opiniões e sentimentos da população que se tornam importantes na formulação de políticas pública para a educação e disseminação, em larga escala, de informações precisas sobre prevenção e recomendações oficiais em relação a saúde pública de modo geral

Por fim, além da seção introdutória, o artigo está estruturado da seguinte forma: na seção 2 apresenta-se o contexto teórico no qual a pesquisa foi desenvolvida; na seção 3 descreve-se a



metodologia empregada nas análises; na seção 4, as análises e discussão dos resultados são apresentadas e, por fim, apresenta-se as considerações finais e direções futuras de pesquisa na seção 5.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Conteúdo e sentimento em Mídias Sociais - Estudos relacionados

Nos últimos anos, observou-se um aumento do interesse de profissionais e pesquisadores acadêmicos na utilização de dados de mídias sociais para as mais diversas finalidades. Inúmeros estudos foram desenvolvidos com o intuito de auxiliar a formulação de políticas públicas relacionados a educação e saúde, especialmente, durante surtos epidemiológicos. No Quadro 1 apresenta-se uma relação de alguns estudos desenvolvidos recentemente que utilizaram dados de mídias sociais, especialmente o *Twitter*, para o monitoramento de tendências, sentimentos, conteúdos e informações acessadas e compartilhadas pelos usuários.

Quadro 1: Alguns artigos sobre análise de sentimentos em mídias sociais e doenças infecciosas

Autor(es)/Ano	Fonte de dados	Propósito do trabalho
Scanfeld, Scanfeld, & Larson (2010)	<i>Twitter</i>	Os autores estudaram o conteúdo de <i>tweets</i> com a menção “antibiótico(s)” para determinar categorias e explorar evidências de mal-entendidos ou mau uso de antibióticos durante o surto infeccioso de H1N1.
Signorini, Segre & Polgreen (2011)	<i>Twitter</i>	Os autores estudaram os conteúdos de <i>tweets</i> para rastrear informações e sentimentos da população em relação a rápida evolução do H1N1 e rastrear e medir a atividade real da doença. O estudo evidenciou que as informações extraídas do <i>Twitter</i> se configuram como uma importante base de dados de interesse público.
Odlum & Yoon (2015)	<i>Twitter</i>	O estudo examinou o conteúdo do <i>Twitter</i> durante um surto de Ebola com o objetivo de fornecer uma imagem instantânea de <i>tweets</i> relacionados ao vírus por meio do monitoramento de tendências na disseminação de informações, estudo da possibilidade de detecção precoce de epidemias e compreensão do conhecimento e das atitudes da população.
Towers et al. (2015)	<i>Twitter</i> e ferramentas de busca	A partir da análise de dados diários de pesquisa na Internet e no <i>Twitter</i> relacionados ao Ebola nos EUA os autores ajustaram um modelo matemático de contágio para verificar se a cobertura noticiosa foi um fator significativo nos padrões temporais dos dados da Internet e do <i>Twitter</i> relacionados ao Ebola. Foram encontradas evidências significativas de contágio, ou se já, vídeos de notícias relacionadas ao Ebola geraram dezenas de milhares de <i>tweets</i> e buscas na Internet relacionadas ao tema.
Chorianopoulos & Talvis (2016)	<i>Twitter</i>	Os autores construíram e disponibilizaram um banco de dados de código aberto que é capaz de detectar sintomas relacionados à gripe e compartilhar os dados em tempo real com a população.
Househ (2016)	<i>Twitter e Google News Trend</i>	O estudo revelou que existe uma relação entre as publicações nas mídias eletrônicas e a atividade do <i>Twitter</i> em torno de eventos significativos, como foi o caso do Ebola. O estudo sugere que as organizações de assistência à saúde devem tirar proveito do relacionamento entre a mídia eletrônica e os eventos de tendência em sites de mídia social como o <i>Twitter</i> para promover campanhas de conscientização.

Shin, Seo & Na (2016)	Google e Twitter	O estudo avaliou a possibilidade de usar um sistema de vigilância digital com base no <i>Google</i> e no <i>Twitter</i> para monitorar surto de MERS-CoV (<i>Middle East respiratory syndrome coronavirus</i>) na Coreia. Os autores encontraram altas correlações entre a pesquisa do <i>Google</i> e os resultados do <i>Twitter</i> e o número de casos em quarentena.
Stefanidis et al. (2017)	Twitter	O objetivo do estudo foi analisar o conteúdo de <i>tweets</i> em relação ao Zika sob a ótica de três aspectos: localização, atores e conceitos, para entender como uma emergência de saúde pública de interesse internacional ocorre nas mídias sociais.
Wong et al. (2017)	Twitter	O objetivo do estudo foi examinar as características dos <i>tweets</i> postados pelos representantes dos departamentos locais de saúde sobre o Ebola. Por meio de uma análise temporal dos <i>tweets</i> sobre o Ebola, os autores constataram a presença de 5 ondas distintas, cada uma correspondendo aos principais eventos noticiosos do Ebola.
Allen et al. (2018)	Twitter	O estudo teve como objetivo a compreensão do conteúdo discutido entre usuários do <i>Twitter</i> acerca de temas relacionados a genética e genômica.
Daughton & Paul (2019)	Twitter	A partir de postagens do <i>Twitter</i> sobre o surto de zika vírus 2015-16, os autores identificaram e descreveram uma relevante mudança no comportamento dos usuários, especificamente, a partir da disseminação de cancelamento de viagens em razão do surto de Zika vírus na América do Sul.
Sousa Junior et al. (2020)	Facebook, Instagram e Twitter	Os autores conduziram uma revisão bibliográfica sobre o SARS-CoV-2 e também analisaram a utilização do termo "coronavírus" em notícias falsas em <i>tweets</i> de usuários brasileiros do <i>Twitter</i> .

Atualmente, diversas pesquisas estão sendo desenvolvidas em relação ao SARS-CoV-2, causador da COVID-19. Os temas são variados, entretanto pode-se destacar o interesse por parte dos pesquisadores acadêmicos em compreender a situação vivenciada sob a ótica da população. Pastor (2020), por exemplo estudou os sentimentos dos usuários do *Twitter* das Filipinas para analisar o efeito da quarentena no estilo de vida da população, o autor concluiu que os conteúdos dos *tweets* denotaram que as necessidades básicas foram afetadas pela falta de apoio do governo e que a maioria dos usuários do *Twitter* tem sentimentos negativos em relação ao COVID-19. Já Xue et al. (2020) buscaram entender o discurso dos usuários chineses do *Twitter* e suas reações psicológicas ao COVID-19, a análise de sentimentos conduzida pelos autores mostrou que o medo pela natureza desconhecida do coronavírus é dominante em todos os tópicos.

Medford et al. (2020) analisaram cerca de 130 mil *tweets* em inglês e concluíram que o número de *tweets* relacionados ao COVID-19 aumentou consideravelmente a partir de 21 de janeiro de 2020 e que quase metade das mensagens analisadas expressaram medo e cerca de 30% expressaram surpresa. Em toda amostra analisada os tópicos mais discutidos foram relacionados aos impactos econômicos e políticos da COVID-19. Já Lwin et al. (2020) conduziram um estudo mais abrangente e analisaram cerca de 20 milhões de *tweets* de mais de 170 países, os resultados demonstraram que o sentimento negativo é dominante. Os autores ressaltam a importância de ações públicas para o atendimento e manutenção da saúde mental da população para evitar que a situação se agrave ainda mais.

Com resultados diferentes, Delizo et al. (2020) conduziram um estudo semelhante nas Filipinas e examinaram a polaridade das opiniões relacionadas ao COVID -19 expressas no *Twitter* e

os resultados mostraram que 52% dos sentimentos dos *tweets* são positivos e 48% expressam sentimentos negativos. Os resultados estão em sintonia com os resultados apresentados por Nemes & Kiss (2020).

Semelhantemente, Alhajji et al. (2020) analisaram as atitudes dos usuários do *Twitter* da Arábia Saudita em relação ao COVID-19 e, a partir de análise de 53.127 *tweets*, concluíram que a população demonstra mais sentimentos positivos que negativos em relação ao novo vírus. Os autores argumentam que os usuários sauditas do *Twitter* mostraram apoio e atitudes positivas em relação às medidas de controle de infecção para combater o COVID-19 e acreditam que esta resposta otimista e favorável da população reflita a confiança popular abrangente e duradoura no governo.

Explorando aspectos específicos do cenário de pandemia vivenciado, Dubey (2020) analisou os sentimentos no *Twitter* de usuários americanos e indianos em relação a confiança nos líderes políticos durante a pandemia e concluíram que, comparativamente, o sentimento negativo dos americanos é maior que o dos indianos e que os usuários da Índia demonstram mais emoções relacionadas a confiança. Já Emtiaz Ahmed et al. (2020) analisou especificamente o sentimento dos americanos em relação a reabertura do comércio e término da quarentena concluindo que no início da quarentena o sentimento de medo era dominante, todavia na eminência da reabertura os sentimentos negativos e de medo diminuíram consideravelmente.

Dubey & Tripathi (2020) analisaram o sentimento no *Twitter* em relação ao novo conceito de trabalhar em casa e concluíram que a prática foi acolhida pelos usuários, uma vez que, a partir da análise de 100.000 *tweets* constataram que as pessoas transmitem sentimentos positivos em relação à nova prática e a maioria das mensagens denotam confiança. E, relacionando o conteúdo desinformativo e *Fake News* ao COVID-19, Recuero e Soares (2020) analisaram como informações relacionados a possíveis “curas” para o coronavírus foram disseminadas no *Twitter*. Os autores analisaram 57.295 mil *tweets* e observaram que existiam alinhamentos entre os discursos desinformativos e discursos políticos.

3. METODOLOGIA

Nesta seção, apresenta-se uma descrição dos procedimentos para a coleta e tratamento dos dados utilizados no decorrer deste trabalho bem como a descrição do método de pesquisa empregado.

3.1 Coleta e processamento dos dados

Os processos metodológicos desenvolvidos para atingir os objetivos desta pesquisa foram divididos em quatro fases (Figura 1): (i) coleta dos dados, (ii) pré-processamento, (iii) construção do algoritmo e (iv) análise de sentimentos e validação dos resultados.

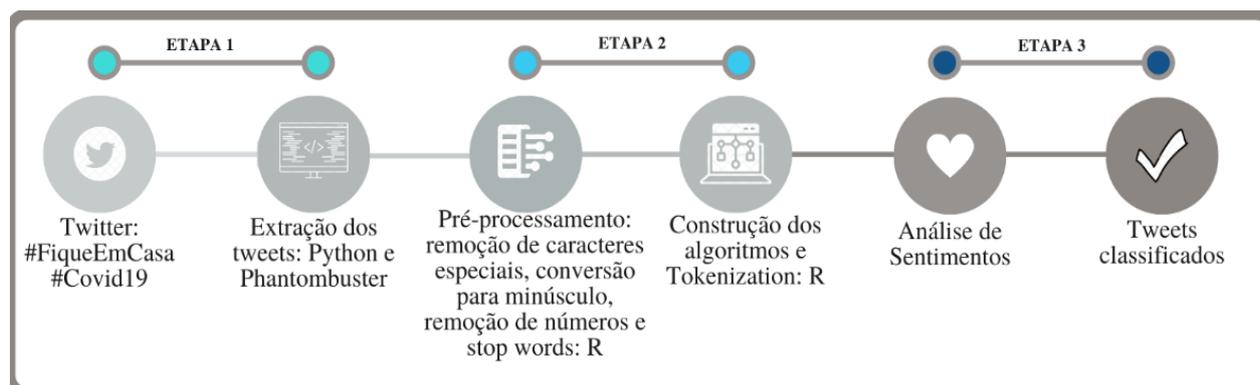


Figura 1: Fases metodológicas da pesquisa

A primeira etapa consistiu na coleta dos dados por meio de uma API (*Application Programming Interface*) disponibilizada pelo próprio *Twitter*. Esta interface permite a coleta e geração de uma base de dados de *tweets*, não incluindo *retweets*, sobre determinado tema. No caso específico deste trabalho, os dados do *Twitter* foram coletados no período de janeiro a julho de 2020. As *hashtags* COVID-19, FiqueEmCasa e suas variações foram utilizadas como recurso de busca e seleção dos *tweets*. A base de dados inicial foi composta por 56.090 *tweets*.

Na fase de pré-processamento (fase 2), optou-se como recorte geográfico o Brasil, logo, foram selecionados apenas os *tweets* em língua portuguesa. Em seguida, foram removidos os caracteres especiais, números, *stopwords* e conteúdos visuais e também se substituiu os caracteres maiúsculos por minúsculo. Esta etapa é importante para otimizar o desempenho da classificação e, conseqüentemente, incrementar a qualidade da análise de sentimentos. Nesta etapa alguns *tweets* foram excluídos e a composição do banco de dados utilizado nas etapas posteriores da pesquisa é apresentada na Tabela 1.

Tabela 1: Base de dados analisada

Mês/Ano	Quantidade de <i>Tweets</i>
Janeiro/2020	119
Fevereiro/2020	339
Março/2020	8.724
Abril/2020	13.823
Maió/2020	8.881
Junho/2020	9.442
Julho/2020*	9.926
Total	51.245

*A etapa de coleta foi encerrada em 22 de julho de 2020

Posteriormente, iniciou-se a etapa 3 do trabalho que consiste, especificamente, em encontrar classes ou polaridades para os dados e classificá-los como positivos, neutros ou negativos. Esta fase será detalhada no tópico a seguir.

3.2 Análise de Sentimentos (AS)

A Análise de Sentimentos tem sido aplicada em diversos estudos na área de Administração Pública (AKAY, DRAGOMIR & ERLANDSSON, 2015; OLIVEIRA & BERMEJO, 2017; MALINI, CIARELLI & MEDEIROS, 2017; OLIVEIRA et al., 2015; OLIVEIRA, BERMEJO & SANTOS, 2017; SILVA NETO, 2017; FONSECA, ALVES & LIMA, 2017). O principal motivo para a utilização de AS está na possibilidade de conversão de um grande volume de dados não estruturados e textuais em modelos capazes de agregar a opinião coletiva e, assim, gerar informações úteis sobre o comportamento da população para a construção de previsões e identificação de tendências (YU, DUAN & CAO, 2013).

A AS se subdivide em duas categorias de técnicas de extração de sentimentos textuais. A primeira categoria, o conjunto de técnicas supervisionadas, se baseia em conceitos de aprendizagem de máquina e faz uso de regras e rótulos previamente definidos para a distinção dos sentimentos. A segunda categoria, técnicas não supervisionadas, não carece do tratamento e dos rótulos pré-estipulados, pois são baseadas em tratamentos léxicos que envolve o cálculo da polaridade de um texto a partir da orientação semântica das palavras nele contidas (ARAÚJO, GONÇALVES & BENEVENUTO, 2013). Esta é uma das suas principais vantagens uma vez que desta forma não mantém a aplicação restrita ao contexto para o qual foram treinados (ARAÚJO, GONÇALVES & BENEVENUTO, 2013).

Neste estudo adotou-se a abordagem léxica, técnica não supervisionada, para a verificação da polaridade (positivo e negativo) de opiniões e pensamentos dos usuários do *Twitter* em relação a pandemia da COVID-19. Por se tratar de uma análise em língua portuguesa, para esta etapa da pesquisa utilizou-se o pacote *LexiconPT* e o dicionário léxico *SentiLex-PT02*.

A base de dados pré-processada foi utilizada para a fase de classificação dos *tweets*. Inicialmente, o conteúdo dos *tweets* foram comparados com as palavras disponíveis no dicionário léxico *SentiLex-PT02* e transformadas em símbolos (*Tokenization*) para a posterior leitura do algoritmo (PERES et al., 2019). O processo de classificação é feito conforme a equação (1).

$$\sum_{k=1}^N t(i, j) \quad \text{Equação (1)}$$

Em que $t(i, j)$ representa a palavra/termo (i) e a sua polaridade (j) de acordo com a sua disponibilidade no léxico. Após o somatório dos termos de uma sentença os textos são classificados.

Para a validação e avaliação do desempenho das classificações para as empresas estudadas adotou-se as diretrizes apresentadas por Araújo, Gonçalves e Benevenuto (2013). Para a avaliação do desempenho dos métodos empregados considerou-se os valores de falsos positivos (FP) e negativos (FN) e verdadeiros positivos (VP) e negativos (VN). A partir destes valores foram calculadas as métricas apresentadas na Tabela 2.

Tabela 2: Métricas de desempenho das classificações



Métrica	Cálculo
Taxa de verdadeiro positivo (<i>Recall</i>)	$R = VP/(VP+FN)$
Taxa de falsos positivos (<i>precision</i>)	$P = VP/(VP+FP)$
Acurácia	$A = (VP+VN)/(VP+VN+FP+FN)$
<i>F-measure</i>	$F = 2(RP)/(R+P)$

Os resultados da classificação e as respectivas análises dos sentimentos dos usuários do *Twitter* em relação a pandemia da COVID-19 são apresentadas do tópico a seguir.

4. ANÁLISE DOS RESULTADOS

Em situações extremas como a causada pela pandemia da COVID-19, o monitoramento de aspectos relacionados à saúde mental e ao bem-estar psicológico da população também se torna importante (HOLMES et al, 2020; BAVEL et al., 2020; SIGNORINI, SEGRE & POLGREEN, 2011). Neste sentido, o *Twitter* é uma ferramenta útil para entender o sentimento da população em tempo real, pois é amplamente utilizada pelos usuários para comunicar suas visões, preocupações, perspectivas, críticas e emoções (ODLUM & YOON 2015; STEFANIDIS et al. 2017; WONG et al. 2017). Assim, este estudo utilizou o conteúdo postado por usuários do *Twitter* em relação a pandemia da COVID-19 para analisar os sentimentos dos indivíduos.

Neste sentido, foram analisados *tweets* relacionados a pandemia da COVID-19 no período de janeiro a julho de 2020. Antes de março de 2020, nossos dados mostram uma pequena porcentagem de mensagens relacionados ao COVID-19. Entretanto nos meses seguintes os números de *tweets* aumentaram à medida que a doença avançava e ganhava mais proporção midiática. Os dados apresentados na Figura 2 demonstram um comportamento ascendente similar entre as curvas acumuladas de casos confirmados de COVID-19 e *tweets* relacionados ao tema.



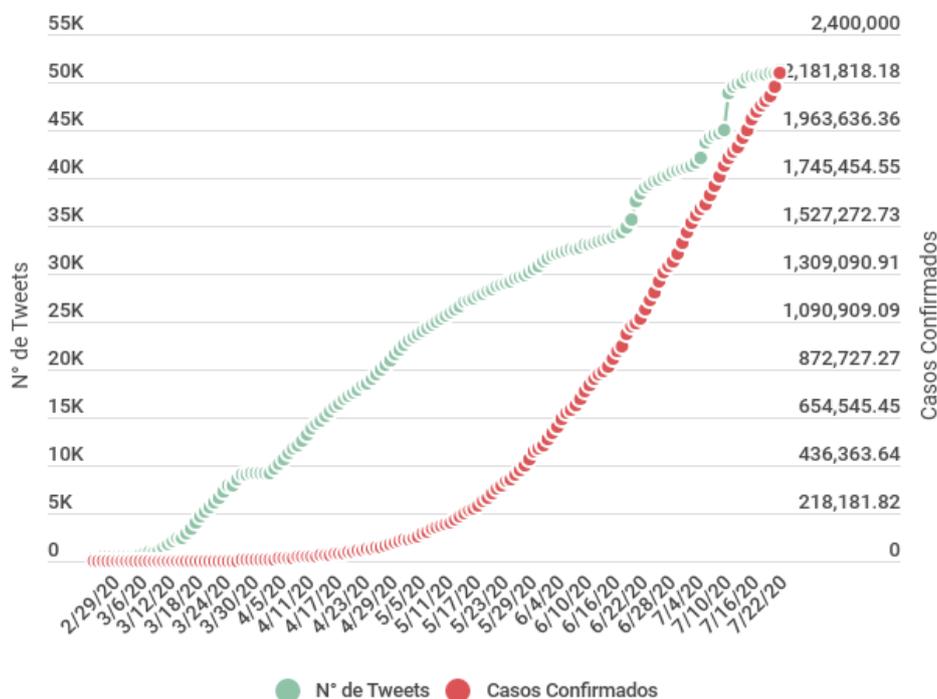


Figura 2: Séries acumuladas de casos confirmados de COVID-19 e tweets no Brasil

É possível notar que a série acumulada de casos confirmados de COVID-19 apresenta uma tendência exponencial de crescimento ao passo que a série acumulada de *tweets* apresenta uma tendência linear de crescimento, porém ambas demonstraram comportamentos ascendentes. As características de tendência (linear e exponencial) das curvas de número de *tweets* e casos de confirmados de COVID-19 no Brasil são semelhantes aos padrões encontrados por Medford et al. (2020) em *tweets* em língua inglesa e os casos confirmados do vírus a nível mundial.

Partindo para a análise de conteúdo e classificação dos sentimentos contidos nas mensagens, inicialmente, adotou-se a técnica *Bag of Words* para a conversão dos textos em vetores para facilitar o estudo da frequência de todas as palavras distintas presentes no texto. A partir desta etapa do trabalho, dividiu-se a base de dados em duas subcategorias. A primeira subcategoria é composta apenas pelo grupo de mensagens classificados como positivas. A segunda subcategoria é formada apenas pelas mensagens classificadas como negativas. Essa subdivisão foi feita com o intuito de compreender as frequências específicas de palavras em cada uma das polaridades. Na Figura 3, apresenta-se apenas a frequência de palavras das mensagens classificadas como positivas.

Por meio da análise da Figura 3 nota-se que a maioria das palavras mais frequentes encontradas nos *tweets* classificados com sentimentos positivos em relação a pandemia da COVID-19 se relaciona a campanha “Fique em casa” que ganhou adesão nacional por meio das mais diferentes mídias a partir de março de 2020. Além disso, palavras como “saúde”, “amigo”, “recuperar”, dentre outras demonstram sinais de otimismo nas mensagens enviadas pelos usuários. Tais movimentos positivos podem estar relacionados às diversas campanhas publicitárias de conscientização e suporte adotadas pelas mais diversas instituições. Além de promover a



conscientização, estes movimentos podem, potencialmente, resgatar o espírito de coletividade e esperança da população.

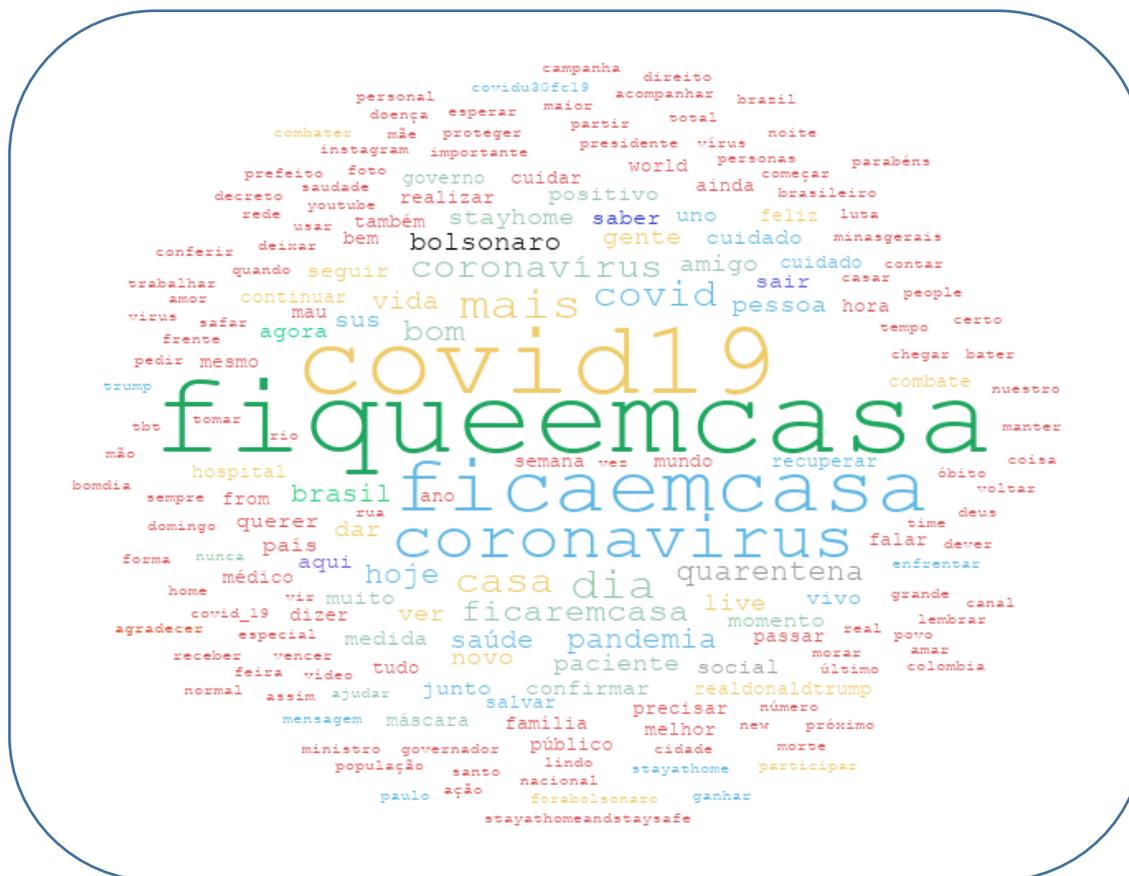


Figura 3: Nuvem de palavras dos *tweets* classificados como positivos.

Na Figura 4, apresenta-se uma nuvem de palavras com a frequência de termos utilizados nos *tweets* classificados como negativos. Por meio do construto, é possível observar que a frequência do movimento “Fique em casa” também é expressiva. Entretanto, pode-se também observar a presença de palavras como: “morte”, “casos”, “vida”, “sus”, “pandemia”, “quarentena”, dentre outras. Tais constatações denotam a preocupação dos usuários do *Twitter* em relação a atual conjuntura instaurada e refletem as consequências que a disseminação do vírus trouxe para a população. Além do movimento #FiqueEmCasa, a *hashtag* “O Brasil pede socorro” também está presente nas mensagens com sentimentos negativos em relação à pandemia.

Ainda, é possível notar nas Figuras 3 e 4 uma forte ligação entre o *tweets* analisados e o cenário político brasileiro, tal relacionamento se evidencia mais fortemente nos comentários classificados como negativos (Figura 4).

estes resultados tornam imperativa a adoção de políticas e estratégias públicas no trato e manutenção da saúde mental e psicológica da população.

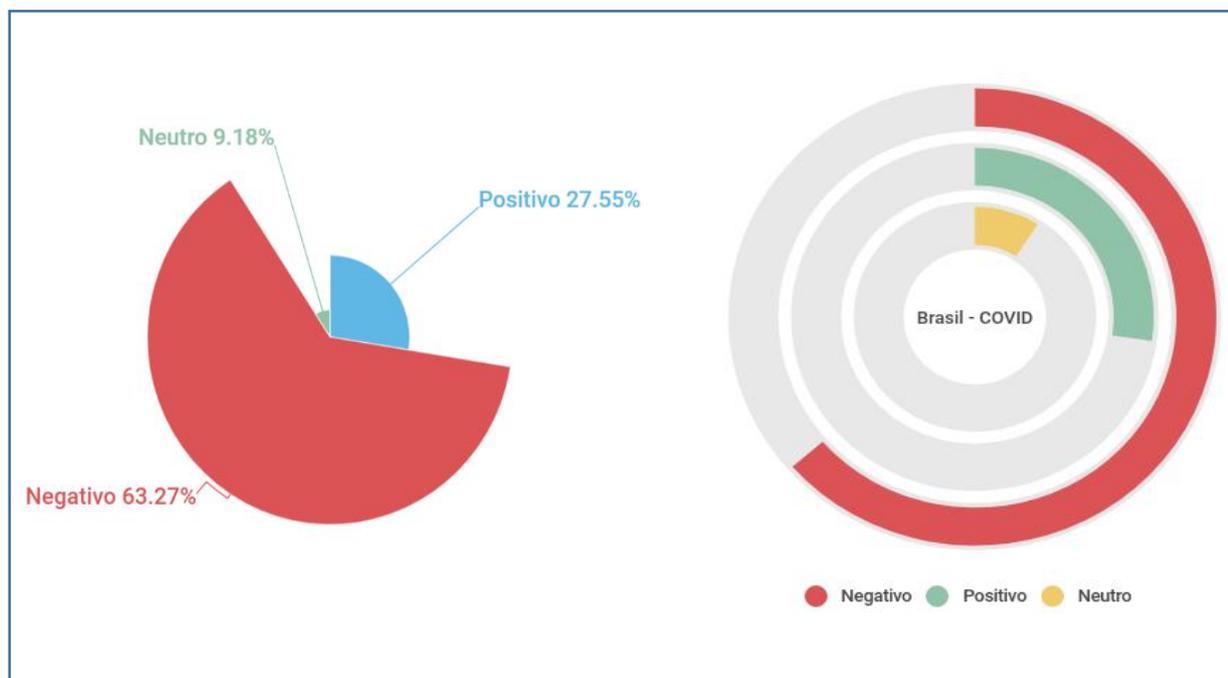


Figura 5: Classificação dos tweets analisados.

Já em comparação aos estudos de Delizo et al. (2020) nas Filipinas e Alhajji et al. (2020) na Arábia Saudita, os resultados desta análise se encontram divergentes. Os autores constataram uma predominância de sentimento positivo na população. Nesse caso, é importante ressaltar que as especificidades dos países estudados são importantes de serem considerados em análises comparativas.

Neste sentido, Alhajji et al. (2020) afirmam que a cultura religiosa do país pode interferir na visão otimista da realidade. Além disso, os autores argumentam que, especificamente, na Arábia Saudita a confiança da população nos representantes públicos e o alinhamento de estratégias de enfrentamento nas mais diferentes esferas podem ter afetado o modo como a população reagiu positivamente a crise vivida. Talvez estas diferenças culturais e políticas expliquem as divergências de sentimentos encontrados entre os estudos. É nítida a influência do cenário político brasileiro da formulação dos comentários compartilhados no *Twitter* (Figuras 3 e 4), especialmente, nos comentários negativos.



Figura 6: Evolução semanal da proporção dos sentimentos.

Na Figura 6 apresenta-se a evolução semanal dos sentimentos no conjunto de dados analisado. Nota-se que o sentimento negativo é dominante em todo o período estudado e, em alguns momentos, atingiu-se o patamar de 70% na classificação ajustada. Estas descobertas revelam que as emoções negativas estão presentes de maneira predominante em todo o período em que o país vivenciou a pandemia da COVID-19, tal evidência apoia a necessidade de ações que mantenham



o bem-estar mental da população neste momento de crise (HOLMES et al, 2020; BAVEL et al., 2020). Bavel et al. (2020) e Duan & Zhu (2020) argumentam que se emoções negativas dominantes não forem abordadas, há um grande potencial de surgimento de resultados indesejados, como desconfiança em relação ao manejo da doença e, conseqüentemente, perda de credibilidade dos órgãos públicos, especialmente, os de trato e contenção do vírus.

5. CONCLUSÃO

O objetivo deste trabalho foi analisar o sentimento dos usuários do *Twitter* em relação à pandemia de COVID-19 no Brasil no período de janeiro a julho de 2020. Para isso foi empregado como técnica a Análise de Sentimentos (AS) por meio do processamento de linguagem natural.

O estudo revelou a existência de uma relação entre as publicações do *Twitter* e o movimento de disseminação do novo coronavírus no Brasil, em que foi possível observar que o sentimento negativo revelado pelas expressões que acompanharam o movimento #fiqueemcasa, como por exemplo “Brasil pede socorro”, “morte” e “risco”, foi dominante em todo o período analisado denotando a preocupação da população em relação ao avanço da doença e as conseqüências por ele geradas. Logo as mídias sociais, como o *Twitter* são espaços de manifestação, mobilização e interação entre pessoas, esse estudo contribui para melhorar a compreensão desse fenômeno diante de um contexto de emergência de saúde pública e de interesse internacional, como é o caso da pandemia da COVID-19 bem como acompanhar a existência e disseminação de *fake news*.

Os resultados são importantes para monitorar tendências e identificar os movimentos das postagens em torno de tópicos específicos e assim buscar compreender como eles se refletem nas ações cotidianas dos indivíduos. Este monitoramento contínuo pode auxiliar as organizações de saúde a gerir proativamente momentos de crise tanto em relação a adoção de medidas para tranquilizar, conscientizar e zelar pelo bem-estar da população em geral quanto para agir tempestivamente para contornar os sentimentos e ações gerados em torno de notícias falsas (*Fake News*) e imprecisas, evitando com que estas sejam amplamente divulgadas já que tem como conseqüências a desinformação, medo e falta de credibilidade nas mídias oficiais.

Além disso, o *Twitter* pode ser uma potencial ferramenta a ser utilizada pela vigilância de surtos e epidemias, pois o conteúdo das mídias sociais pode ser usado para apoiar e aprimorar os sistemas de alerta. A ferramenta pode ser usada amplamente como um meio de informação que tem potencial de alcance global, dado ao seu número de usuários e atuação geográfica, para a disseminação de informações para o maior número de pessoas possível. Esta estratégia permite uma rapidez da disseminação das informações, ampliando o acesso ao conhecimento e amenizando o sentimento de medo, buscando ainda diminuir a propagação de *fake news* já que esse tipo de notícia pode ser um agravante e pode potencializar a evolução de epidemias.

Para trabalhos futuros sugere-se o desenvolvimento de estudos de sentimentos nos estados brasileiros a fim de se conhecer potenciais influências das diferentes estratégias adotadas pelos governadores no sentimento da população. Além disso, seria interessante expandir o estudo para



outras mídias sociais, *Facebook e Instagram*, por exemplo e adotar para fins de comparação outras bibliotecas e dicionários na classificação dos sentimentos dos *tweets*.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Akay, A., Dragomir, A. & Erlandsson, B. (2015). Network-Based Modeling and Intelligent Data Mining of Social Media for Improving Care. *IEEE Journal of Biomedical and Health Informatics*, 19(1), 210-218. doi: 10.1109/JBHI.2014.2336251.
- Alhajji, M., Al Khalifah, A., Aljubran, M. & Alkhalifah, M. (2020). Sentiment Analysis of Tweets in Saudi Arabia Regarding Governmental Preventive Measures to Contain COVID-19. *Preprints*. doi: 10.20944/preprints202004.0031.v1.
- Allen, C. G., Andersen, B., Houry, M. J., & Roberts, M. C. (2018). Current Social Media Conversations about Genetics and Genomics in Health: A Twitter-Based Analysis. *Public health genomics*, 21(1-2), 93–99. doi: 10.1159/000494381
- Araújo, M., Gonçalves P. & Benevenuto, F. (2013). Métodos para análise de sentimentos no Twitter. *Simpósio Brasileiro de Sistemas Multimídia e Web (WEBMEDIA)*, Salvador.
- Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., ... & Drury, J. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 1-12.
- Brasil (2020). MINISTÉRIO DA SAÚDE. Painel Coronavírus. Disponível em: <https://covid.saude.gov.br/>
- Chorianopoulos, K., & Talvis, K. (2016). Flutrack.org: Open-source and linked data for epidemiology. *Health Informatics Journal*, 22(4), 962–974. doi: 10.1177/1460458215599822.
- Dai, H., & Charnigo, R. (2018). A sentiment analysis of merscov outbreak through twitter social media monitoring. *JP Journal of Biostatistics*, 15(2), 107-125. doi: 10.17654/bs015020107.
- Daughton, A. R., & Paul, M. J. (2019). Identifying Protective Health Behaviors on Twitter: Observational Study of Travel Advisories and Zika Virus. *Journal of medical Internet research*, 21(5). doi: 10.2196/13090
- Delizo, J. D. et al. (2020). Philippine Twitter Sentiments during Covid-19 Pandemic using Multinomial Naïve-Bayes. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(1.3), 408 – 412.
- Duan, L., & Zhu, G. (2020). Psychological interventions for people affected by the COVID-19 epidemic. *Lancet Psychiatry*, 7(4), 300-302.
- Dubey, A.D., & Tripathi, S. (2020). Analysing the Sentiments towards Work-From-Home Experience during COVID-19 Pandemic. *Journal of Innovation Management*, 8, 13-19. doi: 10.24840/2183-0606_008.001_0003



- Dubey, A. D. (2020). Decoding the Twitter Sentiments Towards the Leadership in the Times of COVID-19: A Case of USA and India . *SSRN* 3588623. Recuperado de <https://ssrn.com/abstract=3588623>. doi: 10.2139/ssrn.3588623
- Emtiaz A., M., R. I., Rabin, M. & Naz Chowdhury, F.(2020). COVID-19: Social Media Sentiment Analysis on Reopening. *arXiv e-prints*. 2020arXiv200600804E
- Fonseca, P. A. A., Alves, V. d. L., & Lima, L. M. D. (2017). Cultura do Estupro: uma análise de conteúdo sobre a percepção dos usuários via Twitter. *Idealogando: revista de ciências sociais da UFPE*, 1(1), 75-84.
- Garcia, C. F., Vieira, L. J. E. S., & Silva, R. M. D. (2020). Buscas na internet sobre medidas de enfrentamento à COVID-19 no Brasil: descrição de pesquisas realizadas nos primeiros 100 dias de 2020. *Epidemiologia e Serviços de Saúde*, 29(3), 8e2020191. doi: 10.5123/S1679-49742020000300011.
- Heymann, D. L., & Shindo, N. (2020). COVID-19: what is next for public health? *The Lancet*, 395(10224), 542-545. doi: 10.1016/S0140-6736(20)30374-3
- Holmes, E. A., O'Connor, R. C., Perry V.H., Tracey, I., Wessely, S., Arseneault, L., & et al. (2020). Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. *Lancet Psychiatry*, 7(6), 547-560.
- Househ, M. (2016). Communicating Ebola through social media and electronic news media outlets: A cross-sectional study. *Health informatics journal*, 22(3), 470–478. doi: 10.1177/1460458214568037.
- Lana, R. M., Coelho, F. C., Gomes, M. F. D. C., Cruz, O. G., Bastos, L. S., Villela, D. A. M., & Codeço, C. T. (2020). Emergência do novo coronavírus (SARS-CoV-2) e o papel de uma vigilância nacional em saúde oportuna e efetiva. *Cadernos de Saúde Pública*, 36(3). doi: 10.1590/0102-311X00019620.
- Liu, S., Yang L., Zhang C., Xiang Y.T., Liu Z., Hu S., & Zhang B. (2020). Online mental health services in China during the COVID-19 outbreak. *Lancet Psychiatry*, 7(4), e17–e18.
- Lwin, M.O., Lu, J., Sheldenkar, A., Schulz, P. J., Shin, W., Gupta, R., & Yang. Y. (2020). Global Sentiments Surrounding the COVID-19 Pandemic on Twitter: Analysis of Twitter Trends. *JMIR Public Health Surveill*, 6(2), e19447. doi: 10.2196/19447.
- Malini, F., Ciarelli, P., & Medeiros, J. (2017). O sentimento político em redes sociais: big data, algoritmos e as emoções nos tweets sobre o impeachment de Dilma Rousseff. *Liinc em Revista*, 13(2), 323-342.
- Mathur, A., Kubde, P., & Vaidya, S. (2020). Emotional Analysis using Twitter Data during Pandemic Situation: COVID-19. *5th International Conference on Communication and Electronics Systems (ICCES)*, 845-848. doi: 10.1109/ICCES48766.2020.9138079.



- Medford, R. J., Saleh, S. N., Sumarsono, A., Perl, T. M., & Lehmann, C. U. (2020). An “Infodemic”: Leveraging High-Volume Twitter Data to Understand Early Public Sentiment for the COVID-19 Outbreak. *Open Forum Infectious Diseases*. Recuperado de <https://www.medrxiv.org/content/10.1101/2020.04.03.20052936v1>. doi: 10.1093/ofid/ofaa258.
- Nemes, L. & Kiss, K. (2020). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*. Recuperado de <https://www.tandfonline.com/doi/ref/10.1080/24751839.2020.1790793?scroll=top>. doi: 10.1080/24751839.2020.1790793.
- Odlum, M., & Yoon, S. (2015). What can we learn about the Ebola outbreak from tweets? *American journal of infection control*, 43(6), 563–571. doi: 10.1016/j.ajic.2015.02.023.
- Oliveira, D. J. S. & Bermejo, P. H. S. (2020). Mídias sociais e administração pública: análise do sentimento social perante a atuação do Governo Federal brasileiro. *Organizações & Sociedade*, 24(82), 491-508, 2017.
- Oliveira, D. J. S., Bermejo, & P. H. S., & Santos, P. A. (2017). Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls. *Journal of Information Technology & Politics*, 14(1), 34-45.
- Oliveira, D. J. S., Bermejo, P. H. S., & Santos, P. A. (2015). Sentiment analysis, social media, and public administration. In D. C emal, K. Ejub, R. Dragan, & S. Boban (Ed.), *Handbook of research on democratic strategies and citizen-centered E-government services* (Cap. 13, pp. 231-249). Hershey: IGI Global.
- Pastor, C. K. (2020). Sentiment Analysis of Filipinos and Effects of Extreme Community Quarantine Due to Coronavirus (COVID-19) Pandemic. Recuperado de <https://ssrn.com/abstract=3574385>. doi: 10.2139/ssrn.3574385.
- Peres, V., Vieira, R., & Bordini, R. (2019). An alises de Sentimentos: abordagem lexical de classifica  o de opini  o no contexto mercado financeiro brasileiro. Recuperado em de http://www.comp.ita.br/labsca/waiaf/papers/VitorPeres_paper_6.pdf
- Recuero, R., & Soares, F. (2020). O Discurso Desinformativo sobre a Cura do COVID-19 no Twitter: Estudo de caso. *E-Comp os*, 23. Recuperado de <https://preprints.scielo.org/index.php/scielo/preprint/view/84>. doi: 10.1590/SciELOPreprints.84
- Scanfeld, D., Scanfeld, V., & Larson, E. L. (2010). Dissemination of health information through social networks: twitter and antibiotics. *American journal of infection control*, 38(3), 182–188. doi: 10.1016/j.ajic.2009.11.004



- Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza A H1N1 pandemic. *PLoS One*, 6(5). doi: 10.1371/journal.pone.0019467
- Silva Neto, S. R. da. (2017). Uma abordagem computacional para identificação de indício de preconceito em textos baseada em análise de sentimentos. 2017. 95 f. Dissertação (mestrado em Informática) - Instituto de Computação, Programa de Pós-Graduação em Informática, Universidade Federal de Alagoas, Maceió.
- Shin, S., Seo, D., & An, J. (2016). High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. *Sci Rep* 6. <https://doi.org/10.1038/srep32920>.
- Sousa Júnior, J. H., Petroll, M. D. L. M., & Rocha, R. (2019) A. Fake News e o Comportamento Online dos Eleitores nas Redes Sociais durante a Campanha Presidencial Brasileira de 2018. In: XXII SEMEAD – SEMINÁRIOS EM ADMINISTRAÇÃO, USP, São Paulo, Brasil, 22. Recuperado de <http://login.semead.com.br/22semead/anais/arquivos/501.pdf>.
- Statista (2020) Statista. Recuperado em 20 de junho 2020 de <https://www.statista.com/statistics/1043366>
- Towers, S., Afzal, S., Bernal, G., Bliss, N., Brown, S., Espinoza, B., Jackson, J., Judson-Garcia, J., Khan, M., Lin, M., Mamada, R., Moreno, V. M., Nazari, F., Okuneye, K., Ross, M. L., Rodriguez, C., Medlock, J., Ebert, D., & Castillo-Chavez, C. (2015). Mass Media and the Contagion of Fear: The Case of Ebola in America. *PLoS one*, 10(6), e0129179. doi: 10.1371/journal.pone.0129179.
- Wong, R., Harris, J. K., Staub, M., & Bernhardt, J. M. (2017). Local Health Departments Tweeting About Ebola: Characteristics and Messaging. *Journal of public health management and practice*, 23(2), e16–e24. doi: 10.1097/PHH.0000000000000342.
- Xue, J. et al. (2020). Public discourse and sentiment during the COVID-19 pandemic: using Latent Dirichlet Allocation for topic modeling on Twitter. *Social and Information Networks*. Recuperado de <https://arxiv.org/abs/2005.08817>.
- Yu, Y., Duan, W., & Cao, Q. (2013). The impact of social and conventional media on firm equity value: a sentiment analysis approach. *Decision Support Systems*, 55(4), 919- 926.

COMO CITAR ESTE ARTIGO:

G. R. G. Pessanha, T. O. Fidelis, C. D. Freire, E. A. Soares (2020). #Fiqueemcasa: análise de sentimento dos usuários do Twitter em relação ao covid19. *Holos*. 36(5), 1-20.

SOBRE OS AUTORES

G. R. G. PESSANHA

Docente no curso de Administração Universidade Federal de Alfenas, Doutor em Administração pela Universidade Federal de Lavras, Mestre em Administração pela Universidade Federal de Lavras e bacharel em Administração pela Universidade Federal de Lavras. E-mail: pessanhaunifal@gmail.com



ORCID ID: <https://orcid.org/0000-0002-6480-357X>

T. O. FIDELIS

Bacharel Interdisciplinar em Ciência e Economia pela Universidade Federal de Lavras e Graduando em Ciências Atuariais pela Universidade Federal de Lavras, membro do Insurtech Innovation Program - PUC-RJ.

E-mail: taylorfidelis@hotmail.com

ORCID ID: <https://orcid.org/0000-0001-6071-1478>

C. D. FREIRE

Doutoranda em Engenharia de Produção pela Universidade Federal de São Carlos (UFSCar) com período de estágio doutoral na Lancaster University Management School (LUMS) - Inglaterra. Mestre em Gestão de Organizações e Sistemas Públicos pela Universidade Federal de São Carlos (UFSCar), MBA em Gestão de Projetos Inovadores pela Fundace (FEAR/USP) e Graduada em Administração pela Faculdade de Ciências Agrárias e Veterinárias da Universidade Estadual Paulista Júlio de Mesquita Filho (FCAV/UNESP). Atuou como professora em cursos de graduação e especialização na área de Administração, atua como produtora de material didático e conteudista, e atua como Chefe da Seção de Tesouraria e Compradora da Faculdade de Direito de Ribeirão Preto da Universidade de São Paulo (FDRP/USP). É uma das coordenadoras do Núcleo de Estudos Organizacionais (NEO). Tem como áreas de interesse: Estudos Organizacionais, Empreendedorismo, Tecnologia e Inovação, Sociologia Econômica, Gestão de Pessoas, Estratégia Organizacional. E-mail:

clarissadourado21@hotmail.com

ORCID ID: <https://orcid.org/0000-0003-0601-0085>

E. A. SOARES

Doutor em Engenharia Química pela Universidade Federal do Rio Grande do Norte/UFRN em co-tutela com a Université de Toulon (França); Engenheiro do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte/IFRN (Brasil); Membro do Núcleo de Inovação Tecnológica do IFRN. E-mail:

edu.soares999@gmail.com

ORCID ID: <https://orcid.org/0000-0002-2634-8270>

Editor(a) Responsável: Leandro Costa

Pareceristas *Ad Hoc*: LENINA SILVA E KALYANE OLIVEIRA

