

Mais do que palavras: uma análise das emoções brasileiras durante a COVID-19

Filipe Andrade Bernardi¹, Vinicius Costa Lima², Rui Pedro Charters Lopes Rijo³, Domingos Alves⁴

filipepaulista12@usp.br; viniciuslima@usp.br; rui.rijo@ipleiria.pt; quiron@fmrp.usp.br

¹ Bioengineering Postgraduate Program, University of São Paulo, São Carlos, Brazil

² Bioengineering Postgraduate Program, University of São Paulo, São Carlos, Brazil

³ Polytechnic Institute of Leiria, Leiria, Portugal

⁴ Department of Social Medicine, Ribeirao Preto Medical School, University of São Paulo, São Paulo, Brazil

Pages: 526-541

Resumo: Este estudo teve como objetivo investigar e analisar as emoções do público frente aos eventos relacionados ao coronavírus no Brasil. A ciência de dados pode ajudar a explicar os fatos e reduzir o ruído que interfere na compreensão do real impacto de surtos virais em nossas vidas. A rede social Twitter foi analisada para verificar a reação da população ao isolamento social e à quarentena. Para explorar as preocupações do público em relação à evolução da COVID-19 e as posturas governamentais adotadas, uma amostra diária de tweets em língua portuguesa foi coletada para realização de análises de sentimentos e emoções. Ressalta-se a importância da disponibilidade e confiabilidade da informação, da decorrente necessidade da transparência de sistemas de informação e do empoderamento de profissionais de saúde e gestores locais para a tomada das melhores decisões de vigilância e/ou clínicas de forma transversal aos diferentes setores da saúde.

Palavras-chave: COVID-19; Processamento de Linguagem Natural; Análise de Sentimentos; Saúde Pública; Sistemas de Informação em Saúde

More than words: an analysis of Brazilian emotions during COVID-19

Abstract: This study aimed to investigate and analyze the emotions of the public regarding events related to the coronavirus in Brazil. Data science can help explain the facts and reduce the noise that interferes with the understanding of the real impact of viral outbreaks on our lives. The social network Twitter has been analyzed to check the population's reaction to social isolation and quarantine. Thus, to explore public concerns about the evolution of COVID-19 and how governmental positions were adopted, a sample of daily tweets in Portuguese was collected to perform analysis of sentiments and emotions. The importance of the availability and reliability of information, the resulting need for transparency of information systems and the empowerment of health professionals and local managers to make

the best surveillance and / or clinical decisions across the different sectors of the sector are emphasized.

Keywords: COVID-19; Natural Language Processing; Sentiment Analysis; Public health; Health Information Systems

1. Introdução

Em 31 de dezembro de 2019, a Organização Mundial de Saúde (OMS) foi notificada sobre uma série de episódios de pneumonia de origem desconhecida na cidade chinesa de Wuhan, Província de Hubei. Dez dias mais tarde, houve a divulgação da detecção de um novo coronavírus (COVID-19). Desde então, casos da doença têm sido registrados no mundo todo, preocupando autoridades de saúde pública de todos os países. (Who, 2020).

Diferentemente da “gripe espanhola” de 1918, a qual se tornou uma epidemia mundial ao longo de um ano, o COVID-19 se espalhou por todos os continentes em poucas semanas, superando a capacidade dos sistemas de saúde de testar, rastrear e conter pessoas com suspeita de infecção (Shanks, 2020). Além de uma alta taxa de transmissão, a heterogeneidade de sintomas dificultou a compreensão da dinâmica de propagação do vírus, acelerando a busca por respostas por parte de pesquisadores de diversas áreas mundialmente (Hao et al., 2020).

Todos os países devem estar preparados para conter a transmissão do vírus e prevenir a sua disseminação, por meio de vigilância ativa com detecção precoce, isolamento e manejo adequado dos casos, monitoramento dos contatos e notificação oportuna (Gates, 2020). As medidas de vigilância e controle de infecção são críticas para uma resposta eficiente de combate a disseminação da doença. Estudos sugerem que, para controlar a propagação acelerada dessa doença e o consequente esgotamento da capacidade dos recursos hospitalares, algumas medidas de prevenção são essenciais (Tambo et al., 2014). Muitos países tem investido em medidas para ajudar a frear e diminuir a propagação da COVID-19.

Ações de higiene pessoal, o uso de máscaras e o distanciamento social estão sendo amplamente divulgadas nos mais diversos veículos de comunicação. Destacam-se ainda, o estabelecimento de estratégias unificadas entre a liderança nacional e autoridades locais, a fim de permitir o direcionamento eficiente de recursos hospitalares (e.g o fornecimento de equipamentos de proteção individual aos profissionais de saúde), a testagem ampla da população, o monitoramento e identificação de grupos de casos (infectados, suspeitos, recuperados); o incentivo a pesquisa fundamental em tempo real (e.g *clinical trials*); e a mobilização popular (Fineberg, 2020).

De acordo com um estudo realizado pelo *Imperial College Covid-19 Response Team*, o isolamento social é a medida mais eficaz para acharar a curva de demanda por atendimento hospitalar crítico na epidemia (Ferguson et al., 2020). Inevitavelmente, diversas mudanças nos hábitos diários de organizações e pessoas ocorreram. Para manter seu funcionamento, muitas empresas e instituições fizeram uma transição repentina, adotando ferramentas para videoconferência e outras soluções digitais.

A pandemia de COVID-19 revelou muitas áreas de preparação de saúde pública que possuem carência de recursos, especialmente em países de renda média e baixa. As intervenções digitais oferecem muitas oportunidades para fortalecer os sistemas de saúde e podem ser recursos vitais na atual emergência de saúde pública (Mahmood et al.,2020). São conhecidas as aplicações da tecnologia digital para o tratamento, diagnóstico, suporte à autogestão e vigilância durante emergências de saúde pública. Muitos países têm sistemas em vigor para atender a uma variedade de funções de saúde (Gu et al., 2014).

Há um apelo aos governos, agências de saúde e prestadores de cuidados de saúde para alavancar de forma imediata e coerente o poder das ferramentas digitais para fortalecer a capacidade do seu sistema de saúde a fim de responder à pandemia de COVID-19 (Van Bavel et al.,2020). O uso, a viabilidade e a importância dessas e outras ferramentas variam de acordo com as necessidades do país, a infraestrutura existente e outros fatores, como a capacidade de gerenciamento informacional (Mahmood et al.,2020).

Os perigos eminentes do excesso de informação e, conseqüentemente, da desinformação durante o manejo de surtos virais, são chamados de infodemia. A diferença fundamental da pandemia COVID-19 para as demais está na amplificação deste fenômeno, promovido principalmente pelas redes sociais (Zarocostas, 2020). As redes sociais se tornaram uma poderosa ferramenta informacional, além de servirem como um ambiente em que qualquer pessoa pode expressar seus sentimentos e pensamentos abertamente. Ademais, características e escolhas de estilo de vida presentes exclusivamente na atual geração são refletidas pelo seu habitual uso, representando uma drástica mudança nos modos de expressão emocional e respeito pela privacidade (Rudolph & Zacher, 2020).

2. Trabalhos Relacionados

As manifestações públicas frente aos riscos e incertezas geradas por doenças infecciosas desconhecidas, como geralmente ocorrem em pandemias, podem despertar comportamentos e emoções inesperados em uma população. Nesse sentido, a coordenação de estratégias de comunicação efetivas sobre os riscos de contágio são fundamentais para gerar uma resposta comportamental positiva por parte da população (Taylor, 2019). Entender a dinâmica de como as informações são propagadas sobre uma ameaça à saúde favorecem a tomada de decisões assertivas e evitam condutas alarmistas que podem gerar ansiedade e pânico na população.

Durante um surto epidêmico, o monitoramento de redes sociais em tempo real podem ser um valioso instrumento na avaliação das reações populares perante a comunicação de risco, medidas de controle de epidemias e esclarecimentos de rumores (Mcneill et al., 2016). Assim, um importante desafio de pesquisa é determinar como as pessoas buscam ou evitam informações e como essas decisões afetam seu comportamento.

Diversos estudos analisaram respostas comportamentais a outros surtos de doenças infecciosas como os vírus MERS-CoV, dengue, Ebola, Zika e derivados da família de Influenza A. Em geral, as combinações de dados epidemiológicos tradicionais com os fornecidos por sistemas de vigilância digital podem contribuir com o aumento da velocidade e acurácia na detecção e previsão de surtos epidemiológicos (Al-Garadi

et al.,2016). Embora as aplicações dessas novas tecnologias possam exigir recursos humanos e planejamento, o maior desafio apontado ainda é a dificuldade de cooperação coordenada entre os vários níveis governamentais, necessárias para o melhor proveito e uso eficiente dessas ferramentas (Bernardo et al., 2013).

Os primeiros artigos que descrevem o mapeamento de tweets apareceram no Brasil em 2011, durante o monitoramento da dengue. Um estudo destaca que ao coletar dados de quatro dimensões como volume, localização, hora e conteúdo, é possível analisar a forma como as pessoas expressam sua experiência pessoal com a dengue, sendo possível detectar diferentes grupos e, conseqüentemente, auxiliar as autoridades sanitárias na tomada de medidas efetivas de combate aos focos do mosquito (Gomide et al., 2011).

Em 2009, durante o surto de H1N1, um estudo demonstrou, por meio de análise de palavras-chave de tweets, a necessidade de autoridades de saúde pública conhecerem as opiniões, crenças e percepções de sua população a fim de elaborar estratégias de comunicação mais eficazes (Signorini et al., 2011).

Uma análise comparativa da reação da população chinesa entre os surtos de MERS-CoV no Oriente Médio e Europa em 2012 e o surto de H7N9 na China em 2013 expôs que a percepção da comunidade online é mais forte quando o surto da doença ocorre em regiões mais próximas (Fung et al.,2013). Além disso, os primeiros três dias de uma epidemia são apontados como um período crítico para as autoridades tomarem as medidas apropriadas de comunicação, a fim de se esclarecer, prevenir e controlar rumores (Keesara et al., 2020).

Apesar de diversos estudos realizados sobre o surto do vírus Ebola em 2014, nenhum relatou qualquer avaliação útil para as agências de saúde pública durante a resposta emergencial ao vírus (Fung et al., 2016). Da mesma forma, no surto da Síndrome Respiratória do Oriente Médio de 2015 (MERS-CoV), a vigilância digital demonstrou-se um mecanismo acessível capaz de refletir o surto real da doença antes da mesmo da vigilância convencional (Shin et al., 2016).

As variações de locais, atores e conceitos demonstraram ter uma representação fundamental durante a análise de conteúdo do Twitter durante a epidemia de Zika. As influências desses fatores moldaram a formação da opinião pública e a forma como eram absorvidas as notícias sobre os impactos da doença (Stefanidis et al., 2017) e, dentre eles, o principal impacto apontado foi o social (Fu et al., 2016). Foi identificado também uma lacuna de informações e um posicionamento desalinhado entre o público em geral, as autoridades e profissionais de saúde pública (Gui et al., 2017).

Neste sentido, observar mudanças de hábitos e de padrões baseado nas condições individuais de uma população, como diferenças linguísticas e exposições excessivas a informações em situações epidêmicas, são relevantes para a compreensão de modelos analíticos de comportamentos (Daughton & Paul, 2019).

3. Métodos

O Twitter é uma rede social gratuita com mais de 500 milhões de usuários registrados que permite o envio e leitura de tweets, i.e. mensagens curtas limitadas a 280 caracteres,

o que força os usuários a permanecerem focados na mensagem que desejam disseminar. Essa característica torna os tweets propícios à análises baseadas em técnicas de inteligência artificial. Uma dessas análises é o de classificação de sentimentos, um ramo do Processamento de Linguagem Natural (PLN). Neste contexto, “linguagem natural” é qualquer linguagem utilizada para comunicação cotidiana entre humanos, em contraste as linguagens artificiais, tais como as linguagens de programação. A PLN pode ser entendida, em um contexto mais amplo, por qualquer manipulação computacional da linguagem natural (Bird et al., 2009).

Considerando o gigantesco volume de dados produzidos na rede social diariamente, onde milhões de postagens são feitas a cada segundo, foram estabelecidos critérios de seleção e extração da informação. Dentre os critérios adotados, analisamos as palavras mais citadas no Twitter em um período de 17 a 24 de março de 2020, aplicando a uma amostra aleatória de 10 mil tweets com as palavras chave “covid19” e “coronavírus”. Utilizando-se da técnica de alocação latente de Dirichlet (LDA), realizamos a modelagem de tópicos, onde dois foram selecionados para análise expandida. As palavras mencionadas com maior frequência foram selecionadas para compor nosso conjunto final de dados. Os resultados foram representados graficamente por meio da técnica de nuvem de palavras (*wordcloud*).

3.1. Extração de dados

Para explorar as preocupações do público em relação à evolução da COVID-19 e as posturas governamentais adotadas, foram coletadas e armazenadas uma grande amostra de tweets públicos a partir de 24 de março de 2020 que correspondia a um conjunto de termos de pesquisa pré-especificados (e.g., isolamento, quarentena, coronavírus, covid19). Em um segundo momento, foi coletada uma amostra expandida de tweets usando a interface de programação de aplicativos (API) do Twitter, em sua versão gratuita, com a intenção de estimar o sentimento relacionado especificamente ao isolamento e quarentena na rede. Além disso, após discussões com especialistas em saúde pública, novos termos de pesquisa foram adicionados para investigar preocupações sobre os efeitos colaterais provocados pelo surto.

Após o processo de autenticação, extraímos a amostra de tweets por meio de um script desenvolvido na linguagem de programação PHP, posteriormente armazenada em um banco de dados relacional MySQL. Devido às políticas de uso impostas na versão gratuita da plataforma de desenvolvedor do Twitter, a principal limitação deste processo é a extração de tweets no período dos últimos 7 dias, além do limite de envio de até 450 requisições ao Twitter a cada 15 minutos, o que limita o espaço amostral de tweets.

Os valores retornados pela API tratam de informações básicas das postagens (tweets) que incluem, entre outros: o número de identificação do tweet; a data e hora da postagem, o nome do usuário; o tweet com máximo de 280 caracteres; a localização. Cada tweet é marcado com hora e localização geográfica usando o local de residência autodeclarado no perfil do autor.

Todo o fluxo de informações do Twitter é filtrado de acordo com a documentação disponível na API, ou seja, os tweets analisados ainda constituem um subconjunto representativo frente a toda informação disponível. Além disso, como nosso principal interesse era monitorar o tráfego relacionado à atividade durante a pandemia no Brasil,

também excluímos todos os tweets marcados com origem externa ao país e quaisquer tweets não escritos em português. Foram removidos tweets que estejam fora do padrão ASCII e com menos de 5 caracteres

3.2. Pré-processamento dos dados

Os tweets coletados foram usados para produzir um dicionário de palavras em português, do qual foram removidas palavras de parada, as stopwords, consideradas palavras irrelevantes para o desempenho do algoritmo. Também foram removidos hashtags, retweets, links, menções a usuários, etc. Para comprimir o tamanho do dicionário, reduzimos as palavras flexionadas a seu radical, ou seja, ao sentido semântico da palavra, utilizando o processo de stemming. Foi utilizado o algoritmo de Stemming de Porter, disponível no pacote 'SnowballC', na linguagem R (Porter, 2001).

Como o volume de postagens no Twitter varia ao longo do tempo e entre regiões geográficas, as estatísticas de uso foram expressas em termos da fração do total de tweets emitidos no intervalo de tempo e na região geográfica correspondentes. Em seguida, adotamos o modelo de estatística generativa, conhecido também como modelo de tópicos, um método para classificação não supervisionada. Foi utilizado o algoritmo de alocação latente de Dirichlet (LDA) para compilar estatísticas de uso semanal para cada termo do dicionário, ou seja, sumarizar o número de tweets em que cada termo ocorreu.

A categorização de tópicos por LDA é feita através de um modelo estatístico que visa facilitar a identificação de estruturas (ou grupos) dentro de conjuntos textuais. O objetivo é identificar a relevância de um texto frente a um tema, utilizando para tal, a categorização gerada, chamada de tópicos. Na análise por LDA, considera-se que um conjunto textual pode possuir um ou mais tópicos. No caso dos tweets, consideramos que há apenas a presença de um tópico (e.g., tristeza, alegria, medo, raiva) (Girolami & Kaban, 2003).

3.3. Análise de sentimentos e emoções

Uma forma de se realizar a análise de sentimentos e através do uso de dicionários pré-definidos, chamados de corpus ou corpora, no plural. Os dicionários incluem informações relacionadas a diferentes tipos de sentimentos, permitindo a classificação de textos em determinados idiomas (Liu et al., 2010). No sentido lato, refere-se a textos pré-rotulados em sentimentos positivos e negativos, que fornecem informações sobre o sentimento geral de como a resposta foi conduzida.

Basicamente, os algoritmos utilizados para análise de sentimentos baseiam-se na representação de conceitos como polaridade, intensidade, subjetividade e emoções. A classificação de polaridade representa o grau de positividade ou negatividade de um texto e geralmente são expressas como um resultado discreto binário (positivo ou negativo) ou ternário (positivo, negativo ou neutro). Por sua vez, a intensidade refere-se a uma métrica, representada por um intervalo, normalmente um ponto flutuante entre (-1 e 1) ou até entre $-\infty$ e $+\infty$. As análises de subjetividade de sentenças representam o nível de abstração da informação, ou seja, se a informação se aproxima mais de um fato ou de opiniões e sentimentos pessoais, contribuindo assim para a compreensão do contexto,

isto é, textos informais (ex.: coletados de redes sociais) tendem a ser mais subjetivos que textos formais (ex.: coletados de notícias).

Por fim, as emoções indicam um sentimento específico presente em uma mensagem. Alguns algoritmos utilizam-se da composição de métricas de intensidade e polarização para traduzir os sentimentos em emoções. Diversas abordagens lexicais estão disponíveis na literatura (Gonçalves et al., 2013), no entanto, para nossa análise adotamos o NRC Emotion Lexicon - Emolex (Mohammad & Turney, 2013), um método léxico que classifica textos em 8 categorias afetivas. Essas emoções foram definidas no modelo conhecido como Roda das Emoções de Plutchik (Chafale & Pimpalkar, 2014) e foram inspiradas na teoria da psicologia evolutiva das emoções, que descrevem emoções como alegria, tristeza, raiva, medo, confiança, desgosto, antecipação e surpresa. No R foi utilizado o pacote Syuzhet.

3.4. Comparação e validação dos sentimentos

A validação dos sentimentos positivos e negativos obtidos por meio do script desenvolvido na linguagem R foram confrontados no serviço de API disponibilizado pela Microsoft, LUIS. Acrônimo para Language Understanding Intelligent Service (LUIS), o serviço baseado em nuvem aplica técnicas inteligência artificial personalizada a textos de linguagem natural para prever o significado geral de sentenças, e extrair informações detalhadas relevantes (Williams et al., 2015). O mesmo conjunto de tweets foi submetida a análise do LUIS e para representar visualmente seus resultados, adotamos a ferramenta mantida pela Google.

4. Resultados e Discussão

A ciência de dados pode ajudar a explicar os fatos e reduzir o ruído que interfere na compreensão do real impacto de surtos virais em nossas vidas. Os resultados deste estudo demonstram que a vigilância das redes sociais deve ser totalmente incorporada ao sistema de preparação e resposta a epidemias de uma população, o que é coerente com estudos realizados em outras situações epidêmicas (Al-Garadi et al., 2016). A principal diferença deste estudo é investigar e analisar a evolução das reações populares frente às medidas adotadas para a contenção do avanço da pandemia do novo coronavírus, associadas a acontecimentos e conceitos comportamentais discutidos na literatura. Com base na democratização e na influência informacional que o Twitter e outras redes sociais exerceram sobre a opinião pública e o comportamento durante pandemias anteriores, analisamos os desafios e oportunidades para a comunidade de saúde.

Nunca na história moderna ocorreram medidas radicais de isolamento e distanciamento social que afetassem o mundo todo. Consequentemente, o crescimento do fenômeno de infodemia foi potencializado através das redes sociais. Para entendermos o contexto e o impacto as políticas de combate adotadas, precisamos entender os fatores associados, ao que alguns autores denominam como a crise tripla do COVID-19, que envolvem assuntos econômicos, sanitários e comportamentais (Conti et al., 2020). Apesar de singulares, todos esses fatores formam uma unidade “sine qua non” na minimização dos impactos gerados, onde ressalta-se a importância da interlocução coesa e efetiva de líderes a fim de manter a confiança social.

A informação é considerada a pedra angular neste processo de minimização de impactos e, quando incorporadas, as redes sociais permeiam o engajamento das pessoas e podem auxiliar na compreensão de como as informações são absorvidas pela população, refletindo assim aspectos positivos e negativos. A figura 1 reforça bem esse cenário, onde o aumento gradativo da média do sentimento negativo ocorre concomitantemente ao avanço da pandemia no país no mês de abril e a maior necessidade do empenho da população em cumprir as medidas de isolamento.

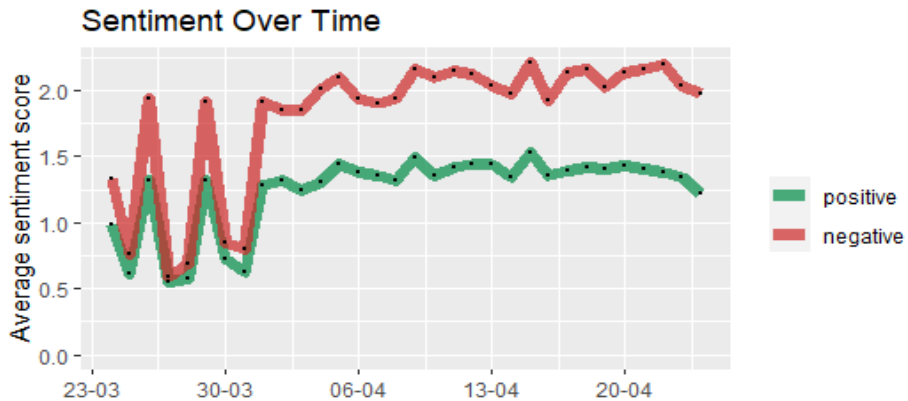


Figura 1 – Evolução dos sentimentos durante o mês de abril

As ações humanas são guiadas por evidências sociais, efeito popularmente conhecido como “efeito manada”, ou seja, reações irracionais semelhantes de um grupo que ocorrem por um ou mais indivíduos. A literatura tem se concentrado em estudos sobre o comportamento de manada principalmente em assuntos relacionados a mercados financeiros, comportamentos de compra compulsiva e decisões discricionárias de grupos isolados, no entanto, são insuficientes as evidências que descrevem como este efeito influencia hábitos e comportamentos em saúde, principalmente em situações de emergência (Chang et al., 2020). Mudanças de comportamento em larga escala, como decorre da necessidade de confinamento social, impõe cargas psicológicas significativas aos indivíduos.

A percepção de um indivíduo sobre determinado assunto em um cenário globalizado como o Twitter é influenciada pelo seu círculo de contatos, isto é, a consolidação de um assunto como um fato é sustentado basicamente pelo número de pessoas ou tweets que esse indivíduo interage e considera suficientemente capaz de concordar entre si de forma coerente (Machanick, 2020). Em situações de incerteza, o sentimento de medo pode ser propagado e virtualmente potencializado (Taylor, 2019). Em sua forma mais branda, a demonstração de medo pode decorrer de uma coleção de sentimentos de preocupação, e, portanto, provocar sintomas de ansiedade (hipocondria) e, conseqüentemente, um aumento na procura por serviços de saúde (Schimmenti et al., 2020). Uma análise cumulativa das emoções de maior predominância também no mês de abril corrobora

com isso e reflete a situação de calamidade que esses eventos podem gerar. A figura 2 elucida esse cenário.

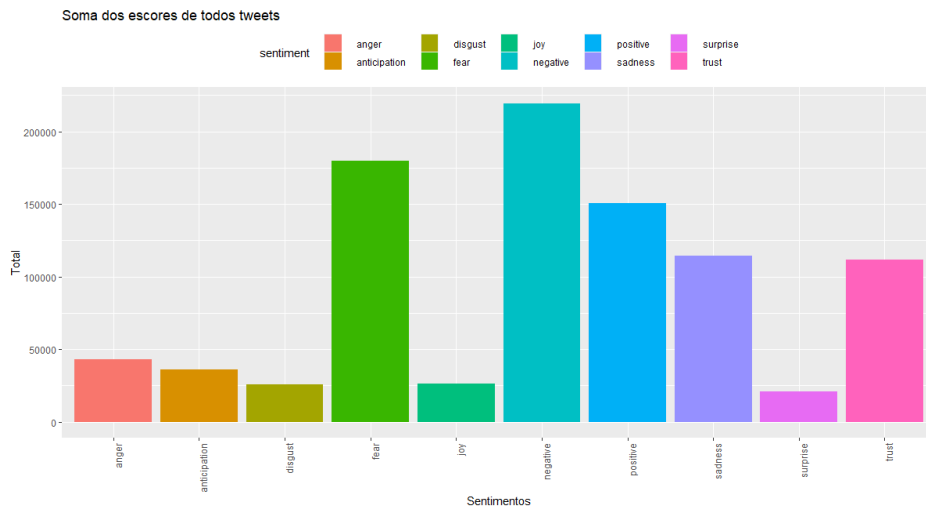


Figura 2 – Somatório das emoções ao longo do mês de abril

Com o surgimento de conglomerados virtuais capazes de inflar discursos que nem sempre corroboram com recomendações de especialistas, a construção de um ambiente de desinformação e de hostilidades é iminente. Além disso, discussões como a restrição a liberdades e o posicionamento de lideranças tem fundamentado a polarização de grupos cada vez mais atuantes durante a pandemia de COVID-19 (Recuero et al., 2020).

Uma característica marcante de discursos públicos sobre política científica é a certeza extrema daqueles que carecem de uma estrutura conceitual para entender um problema, essa manifestação é conhecida como síndrome do impostor. Também descrita como efeito Dunning-Kruger, o desafio de argumentar a política científica contra tal oposição torna-se árduo, pois quanto menos o conhecimento, maior a certeza (Dunning, 2011). Assim indivíduos com posicionamentos semelhantes se isolam do resto da sociedade e possuem acesso somente a opiniões e informações que reforçam o posicionamento do grupo provocando um efeito de verdade ilusória, ou em alguns casos um efeito de câmara de eco ideológico. Como consequência desses efeitos estão polarização das redes sociais, o hiperpartidarismo, a desinformação e teorias da conspirações, que cada vez mais tornam-se fatos mais consolidados do que a ciência cuidadosamente pensada (Recuero et al., 2020).

Em um momento caótico, escolher caminhos baseando-se em evidências não deveria ser uma opção, e sim uma obrigação dos líderes que precisam, ao tomar as decisões, gerar impactos positivos na vida das pessoas (Allcott et al., 2020). Globalmente, muitas pessoas têm sentimentos negativos e falta de confiança em sua liderança nacional, refletidas em atitudes que ignoram recomendações da OMS como recomendações sem

embasamento científico e a promoção de aglomerações (Lancet, 2020). A negligência e minimização dos riscos de uma pandemia tem sido respostas comuns em países de baixa e média renda (LMICS) e, até mesmo em alguns casos, em países desenvolvidos (Mather, 2020). No Brasil, as fronteiras permaneceram abertas sem exigências de que os viajantes cumprissem qualquer tipo de quarentena. Além disso grandes eventos, como o Carnaval, ocorreram sem quaisquer restrições sanitárias, sustentando hipóteses de que a circulação do vírus tenha ocorrido antes mesmo de medidas e confirmações oficiais (Ebrahim & Memish, 2020).

Em alguns países, há a percepção de que o combate a ameaças epidêmicas é irrelevante frente a outras doenças presentes há mais tempo no dia a dia de suas populações, como o HIV, tuberculose, dengue e malária. O empenho em ações de combate a surtos é visto com desinteresse por LMICS quando comparados a países de alta renda. De fato, as ameaças epidêmicas são potencialmente devastadoras para o cenário econômico destes países, comprometendo diversos setores com a diminuição da produtividade, investimentos em saúde elevados, perda de receita com turismo e viagens e incentivos negativos ao investimento (Frieden et al., 2020). Através da modelagem LDA de 8 tópicos, essa preocupação pode ser notada com a presença de termos como “economia”, “trabalhar” e “social”, conforme apresentado na figura 3.

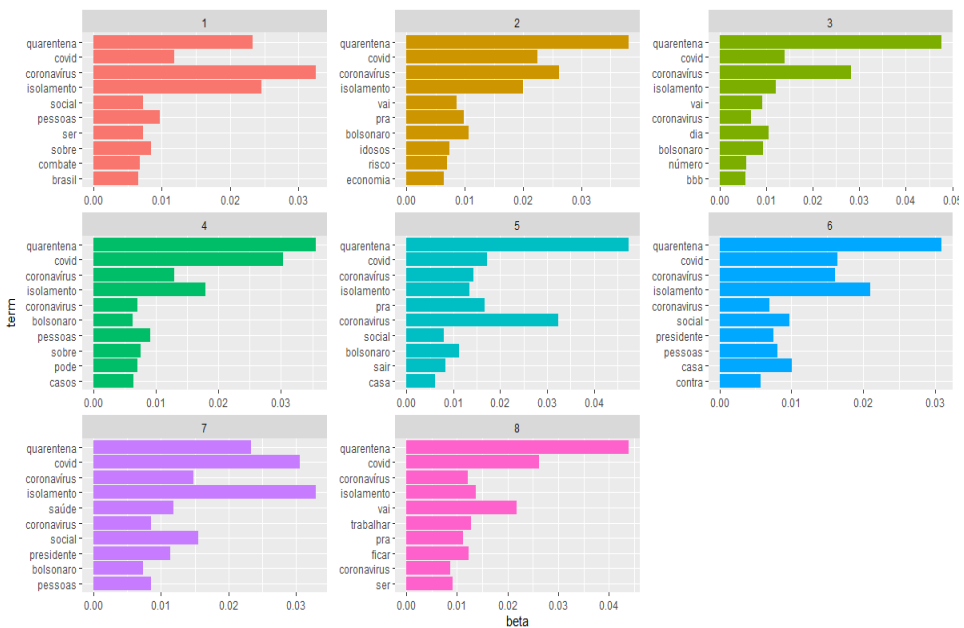


Figura 3 – Tópicos mais relevantes durante a quarentena no mês de abril

Um estudo recente que avaliou a associação entre recessão econômica e mortalidade de adultos calcula que, no Brasil, a cada 100.000 habitantes, 2 mortes adultas ocorrem

para cada 1% de aumento da taxa de desemprego (Hone et al., 2019). Segundo especialistas, em um cenário sem políticas compensatórias que auxiliem empresas e empregados, poderiam ocorrer um aumento da então taxa de desemprego de 11,6%, para 16,1% (Correio Braziliense, 2020), o que, segundo o modelo, causaria ao menos 18,5 mil mortes. Tais estimativas reforçam a necessidade de líderes de LMICS na adoção de políticas preventivas que minimizem os efeitos dos lockdowns e ao mesmo tempo mitigue as reações populares com ações comunicativas efetivas (Who, 2020). Neste sentido, o capital social e sistemas de proteção, principalmente a populações mais vulneráveis, são primordiais. Portanto, a compreensão da real maneira como também se comportam as comunidades podem permitir mensurar nível de resiliência de sua população após um desastre natural, bem como algo como uma pandemia. Nas figuras 4 e 5, reforçamos a “tempestade de emoções”, principalmente negativas, que podem surgir durante o período de confinamento. Ambos algoritmos presentes em cada figura, gerado por nossa aplicação e pelo LUIS, respectivamente, foram confrontados para demonstrar a predominância de fatores negativos.

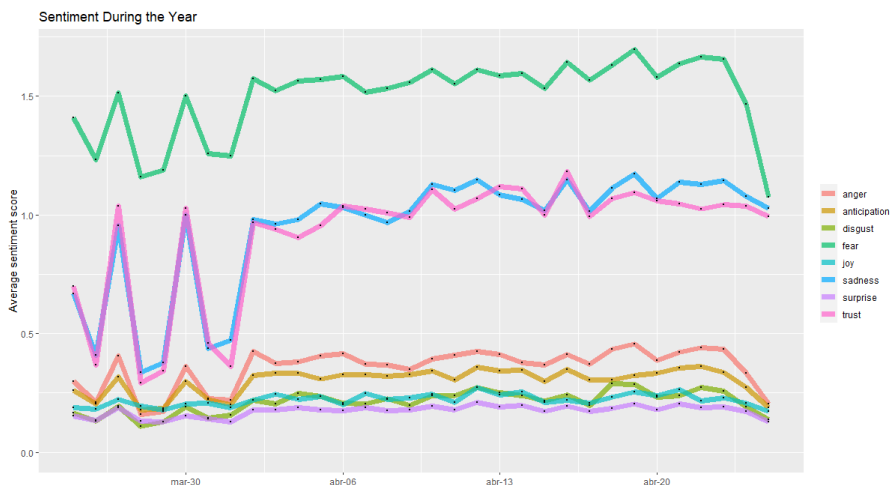


Figura 4 – “Tempestade de emoções durante a quarentena”

Finalmente, outro fator fundamental que está intrinsecamente ligado ao “efeito manada”, empatia e influência sobre pessoas, que pode mudar os rumos de uma situação de emergência de saúde pública, como uma pandemia, é a liderança (Gates, 2020). A liderança é um mecanismo de persuasão capaz de induzir e influenciar o comportamento humano. Em saúde, o campo de atuação de líderes permanece uma incógnita, onde a divergência de informação entre governo e população sobre práticas de controle e prevenção são um desafio (Rhodes & Hart, 2014). O impacto do papel desempenhado por líderes nesse contexto é, portanto, crucial na redução de ruídos e comportamentos indesejáveis, como a notável falta de sinergia e alinhamento na resposta oficial ocasionada entre os diferentes níveis de governo (Allcott et al., 2020).

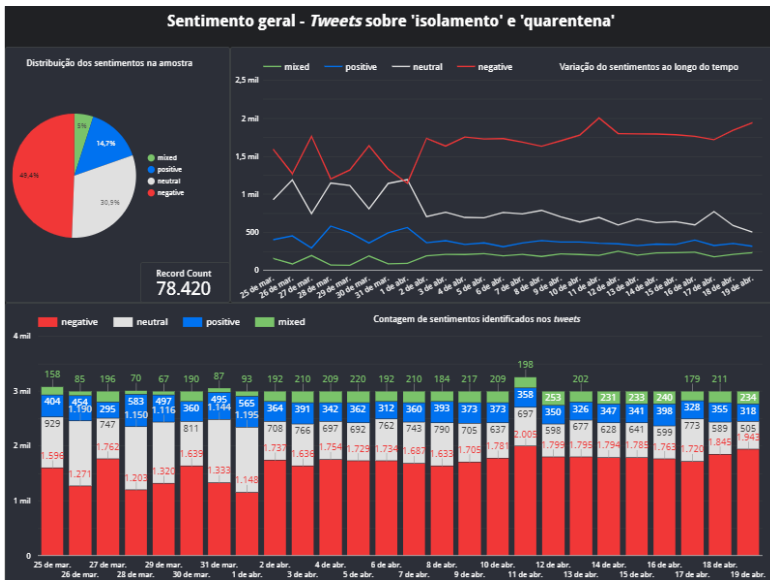


Figura 5 – Análise LUIS de sentimentos durante o isolamento

5. Considerações Finais

O Brasil é um dos poucos países de nível médio de desenvolvimento que detém um Sistema Único de Saúde (SUS) e sistemas eletrônicos de informação, no entanto ainda carece de sistemas de vigilância e estratégias de comunicação em saúde eficazes. O presente trabalho não se refere apenas ao uso de técnicas de computação existentes em problemas relevantes de saúde, como a COVID-19, mas envolve também o desenvolvimento de novas tecnologias e métodos para a definição de paradigmas na área de e-Science no âmbito da e-Health, por meio de mecanismos inovadores para e gerenciamento automatizado de conhecimento.

Além disso, foi possível usar e identificar a necessidade de desenvolvimento de novos métodos de avaliação de desempenho para modelos preditivos em termos de calibração e utilidade epidemiológica para tomada de decisão baseada em evidências, coletadas durante situações epidêmicas. Isso pode ser medido, mesmo que em um curto espaço de tempo, por exemplo, por meio da publicação de artigos científicos em periódicos de impacto internacional, tanto na área de e-Science, quanto na área de computação aplicada à saúde e, principalmente, em revistas de saúde interessadas, incluindo resultados da pesquisa de implementação, como impacto epidemiológico e econômico da incorporação de novas estratégias/tecnologias no sistema de saúde. Toda a iniciativa da construção e manutenção das análises, sempre foi pensada de forma colaborativa, sendo que com o passar do tempo essa rede de colaboração se expandiu e temos tido a oportunidade de contribuir de maneira direta com o entendimento da complexidade da pandemia e seus reflexos na sociedade.

É importante destacar que em toda a trajetória até aqui das análises tem sido aprimoradas com técnicas de deep learning e novas redes têm sido testadas. Além disso um sistema de monitoramento em tempo real de tweets está em desenvolvimento a fim de servir como “termômetro” frente a forma como vem sendo conduzida situações que emergência. As limitações deste trabalho referem-se as variações de idioma e cenário, sendo necessários mais testes de acurácia e precisão. O fator determinante para esta limitação justifica-se ainda pela escassez de modelos pré-treinados na língua portuguesa para situações de saúde.

Referências

- Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., & Yang, D. Y. (2020). Polarization and public health: Partisan differences in social distancing during the Coronavirus pandemic. NBER Working Paper, (w26946).
- Al-garadi, M. A., Khan, M. S., Varathan, K. D., Mujtaba, G., & Al-Kabsi, A. M. (2016). Using online social networks to track a pandemic: A systematic review. *Journal of biomedical informatics*, 62, 1-11.
- Bernardo, T. M., Rajic, A., Young, I., Robiadek, K., Pham, M. T., & Funk, J. A. (2013). Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *Journal of medical Internet research*, 15(7), e147.
- Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. “O’Reilly Media, Inc.”.
- Bufulin, A. P., da Cunha Braz, M. B., & da Vitória, F. M. (2020). Coronavírus e direito de família: as implicações do enfrentamento da emergência de saúde pública de importância internacional decorrente do surto de Covid-19 no regime de convivência familiar. *civilistica.com: revista eletrônica de direito civil*, 9(1), 1-15.
- Chafale, D., & Pimpalkar, A. (2014). Review on developing corpora for sentiment analysis using plutchik’s wheel of emotions with fuzzy logic. *International Journal of Computer Sciences and Engineering (IJCSE)*, 2(10), 14-18.
- Chang, C. L., McAleer, M., & Wang, Y. A. (2020). Herding behaviour in energy stock markets during the Global Financial Crisis, SARS, and ongoing COVID-19. *Renewable and Sustainable Energy Reviews*, 134, 110349.
- Conti, T. V. (2020). Crise Tripla do Covid-19: um olhar econômico sobre políticas públicas de combate à pandemia.
- BRAZILIENSE, C., 2020. 5 Milhões Podem Entrar Na Fila Do Desemprego Em Apenas Três Meses. [online] Acervo. Available at: <https://www.correiobraziliense.com.br/app/noticia/economia/2020/04/05/internas_economia,842458/5-milhoes-podem-entrar-na-fila-do-desemprego-em-apenas-tres-meses.shtml> [Accessed 29 October 2020].
- 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), e13090.

- Dunker, C. I. L. (2020) A arte da quarentena para principiantes (Pandemia Capital). São Paulo: Boitempo Editorial.
- Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one’s own ignorance. In *Advances in experimental social psychology* (Vol. 44, pp. 247-296). Academic Press.
- Ebrahim, S. H., & Memish, Z. A. (2020). COVID-19—the role of mass gatherings. *Travel medicine and infectious disease*.
- Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., ... & Dighe, A. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.
- Fineberg, H. V. (2020). Ten weeks to crush the curve.
- Frieden, T. R., Tappero, J. W., Dowell, S. F., Hien, N. T., Guillaume, F. D., & Aceng, J. R. (2014). Safer countries through global health security. *The Lancet*, 383(9919), 764-766.
- Fu, K. W., Liang, H., Saroha, N., Tse, Z. T. H., Ip, P., & Fung, I. C. H. (2016). How people react to Zika virus outbreaks on Twitter? A computational content analysis. *American journal of infection control*, 44(12), 1700-1702.
- Fung, I. C. H., Fu, K. W., Ying, Y., Schaible, B., Hao, Y., Chan, C. H., & Tse, Z. T. H. (2013). Chinese social media reaction to the MERS-CoV and avian influenza A (H7N9) outbreaks. *Infectious diseases of poverty*, 2(1), 31.
- Fung, I. C. H., Duke, C. H., Finch, K. C., Snook, K. R., Tseng, P. L., Hernandez, A. C., ... & Tse, Z. T. H. (2016). Ebola virus disease and social media: a systematic review. *American journal of infection control*, 44(12), 1660-1671.
- Gates, B. (2020). Responding to Covid-19—a once-in-a-century pandemic?. *New England Journal of Medicine*, 382(18), 1677-1679.
- Girolami, M., & Kabán, A. (2003, July). On an equivalence between PLSI and LDA. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaiion retrieval* (pp. 433-434).
- Gonçalves, P., Araújo, M., Benevenuto, F., & Cha, M. (2013, October). Comparing and combining sentiment analysis methods. In *Proceedings of the first ACM conference on Online social networks* (pp. 27-38).
- Gomide, J., Veloso, A., Meira Jr, W., Almeida, V., Benevenuto, F., Ferraz, F., & Teixeira, M. (2011, June). Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In *Proceedings of the 3rd international web science conference* (pp. 1-8).
- Gu, H., Chen, B., Zhu, H., Jiang, T., Wang, X., Chen, L., ... & Jiang, J. (2014). Importance of Internet surveillance in public health emergency control and prevention: evidence from a digital epidemiologic study during avian influenza A H7N9 outbreaks. *Journal of medical Internet research*, 16(1), e20.

- Gui, X., Wang, Y., Kou, Y., Reynolds, T. L., Chen, Y., Mei, Q., & Zheng, K. (2017). Understanding the patterns of health information dissemination on social media during the Zika outbreak. In *AMIA Annual Symposium Proceedings* (Vol. 2017, p. 820). American Medical Informatics Association.
- Hao, X., Cheng, S., Wu, D., Wu, T., Lin, X., & Wang, C. (2020). Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature*, 584(7821), 420-424.
- Hone, T., Mirelman, A. J., Rasella, D., Paes-Sousa, R., Barreto, M. L., Rocha, R., & Millett, C. (2019). Effect of economic recession and impact of health and social protection expenditures on adult mortality: a longitudinal analysis of 5565 Brazilian municipalities. *The Lancet Global Health*, 7(11), e1575-e1583.
- Keesara, S., Jonas, A., & Schulman, K. (2020). Covid-19 and health care's digital revolution. *New England Journal of Medicine*, 382(23), e82.
- Lancet, T. (2020). COVID-19: learning from experience. *Lancet* (London, England), 395(10229), 1011.
- Liu, B. (2010). Sentiment analysis and subjectivity. *Handbook of natural language processing*, 2(2010), 627-666.
- Machanick, P. (2020). More Covid. *South African Computer Journal*, 32(2).
- Mahmood, S., Hasan, K., Carras, M. C., & Labrique, A. (2020). Global Preparedness Against COVID-19: We Must Leverage the Power of Digital Health. *JMIR Public Health and Surveillance*, 6(2), e18980.
- Mather, T. P. S., Marin, B. G., Perez, G. M., Christophers, B., Paiva, M. L., Oliva, R., ... & Marqués, C. G. (2020). Love in the time of COVID-19: negligence in the Nicaraguan response. *The Lancet Global Health*, 8(6), e773.
- McNeill, A., Harris, P. R., & Briggs, P. (2016). Twitter influence on UK vaccination and antiviral uptake during the 2009 H1N1 pandemic. *Frontiers in public health*, 4, 26.
- Mohammad, S. M., & Turney, P. D. (2013). *Nrc emotion lexicon*. National Research Council, Canada, 2.
- Porter, M. F. (2001). *Snowball: A language for stemming algorithms*.
- Recuero, R., Soares, F., & Zago, G. (2020). Polarização, Hipertidarismo e Câmaras de Eco: Como circula a Desinformação sobre Covid-19 no Twitter.
- Rhodes, R. A., & Hart, P. T. (Eds.). (2014). *The Oxford handbook of political leadership*. Oxford University Press.
- Rudolph, C. W., & Zacher, H. (2020). "The COVID-19 generation": A cautionary note. *Work, Aging and Retirement*.
- Shanks, G. D. (2020). COVID-19 versus the 1918 influenza pandemic: different virus, different age mortality patterns. *Journal of travel medicine*, 27(5), taaa086.
- Schimmenti, A., Billieux, J., & Starcevic, V. (2020). The four horsemen of fear: An integrated model of understanding fear experiences during the COVID-19 pandemic. *Clinical Neuropsychiatry*, 17(2), 41-45.

- Shin, S. Y., Seo, D. W., An, J., Kwak, H., Kim, S. H., Gwack, J., & Jo, M. W. (2016). High correlation of Middle East respiratory syndrome spread with Google search and Twitter trends in Korea. *Scientific reports*, 6, 32920.
- Signorini, A., Segre, A. M., & Polgreen, P. M. (2011). The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PLoS one*, 6(5), e19467.
- Stefanidis, A., Vraga, E., Lamprianidis, G., Radzikowski, J., Delamater, P. L., Jacobsen, K. H., ... & Crooks, A. (2017). Zika in Twitter: temporal variations of locations, actors, and concepts. *JMIR public health and surveillance*, 3(2), e22.
- Tambo, E., Ugwu, E. C., & Ngogang, J. Y. (2014). Need of surveillance response systems to combat Ebola outbreaks and other emerging infectious diseases in African countries. *Infectious diseases of poverty*, 3(1), 1-8.
- Taylor, S. (2019). *The psychology of pandemics: Preparing for the next global outbreak of infectious disease*. Cambridge Scholars Publishing.
- Van Bavel, J. J., 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.
- Williams, J. D., Kamal, E., Ashour, M., Amr, H., Miller, J., & Zweig, G. (2015, September). Fast and easy language understanding for dialog systems with Microsoft Language Understanding Intelligent Service (LUIS). In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 159-161).
- World Health Organization. (2020). Considerations for public health and social measures in the workplace in the context of COVID-19: annex to considerations in adjusting public health and social measures in the context of COVID-19, 10 May 2020 (No. WHO/2019-nCoV/Adjusting_PH_measures/Workplaces/2020.1). World Health Organization.
- Zarocostas, J. (2020). How to fight an infodemic. *The Lancet*, 395(10225), 676.

© 2021. This work is published under <https://creativecommons.org/licenses/by-nc-nd/4.0/>(the “License”). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.