

# Análise da comunicação das prefeituras do Brasil e a relação com o aumento da COVID-19

Jackson Cardoso, Osmay Camila Bortoncello Globber

Universidade Tecnológica Federal do Paraná

Departamento Acadêmico de Informática

Curitiba, Paraná

jacksonc@alunos.utfpr.edu.br, marycamilainfo@gmail.com

**Resumo** - O presente artigo discute o impacto da comunicação das prefeituras dos municípios brasileiros realizada por meio da rede social Twitter durante o período da pandemia do novo Coronavírus. Adotando técnicas de ciência de dados, fundamentadas em modelos de análises de dados textuais a fim de compreender se prefeituras que se comunicam com maior frequência na rede social Twitter acabam contribuindo de alguma forma para que ocorra a diminuição de casos confirmados e de mortalidade consequentes da pandemia COVID-19.

**Palavras-Chave:** Tweet, Prefeituras, Data Science, Pandemia

## I. INTRODUÇÃO

A pandemia decorrente do espalhamento do novo Coronavírus (COVID-19) acabou afetando cidadãos de todo o planeta, provocando impactos na sociedade, alterando a economia, o trabalho, a educação e a medicina de uma forma imprevista. As prefeituras do Brasil passaram a adotar protocolos rígidos na área da saúde e medidas de segurança a fim de que o contágio da doença fosse impedido e cidadãos infectados pudessem recuperar a saúde o mais breve possível. Sabemos que a tecnologia vem demonstrando cada vez mais, ser uma ferramenta poderosa para a comunicação e trabalhos colaborativos, principalmente com o advento da web 2.0, evidenciada pelo surgimento de redes sociais eletrônicas, onde existe a possibilidade de os próprios usuários elaborarem e organizarem os conteúdos, dentre elas, a Twitter, que chegou e se consolidou como um canal de comunicação favorável para o compartilhamento de notícias, marketing e propaganda para a sociedade. Diante deste cenário, surge a ideia deste trabalho, realizar a análise da comunicação realizada no Twitter por prefeituras brasileiras,

buscando entender se as prefeituras mais engajadas no Twitter tem alguma relação com o comportamento da epidemia do coronavírus nos municípios brasileiros.

## II. TRABALHOS CORRELATOS

[3] menciona que o Twitter tem desempenhado um papel importante na vigilância da saúde por órgãos nacionais e internacionais na ocorrência de epidemias. [2] coleta notícias e tweets sobre COVID-19, em português, de janeiro a maio de 2020, usando a biblioteca Python Tweepy e considerando um conjunto de palavras-chave em português relacionadas ao COVID-19 para filtrar o fluxo do Twitter e obter tweets relevantes sobre a pandemia. [1] apresenta a evolução do novo coronavírus no Brasil, a partir da análise das postagens dos usuários brasileiros no Twitter, utilizando o software de análise qualitativa, Atlas.ti, para extração dos dados do Twitter e para produzir nuvens de palavras presentes nos tweets. Já [4] apresenta exemplos de como coletar informações do Twitter utilizando a API pública e a biblioteca Python Twint. Os trabalhos mostraram ser relevantes visto que enfatizaram analisar o comportamento da pandemia do novo coronavírus, utilizando dados do Twitter essencialmente nos estudos e pela aplicação da ciência de dados nas análises.

## III. PROCESSAMENTO DE DADOS

Foram obtidos dados dos cem primeiros municípios com maior número de óbitos acumulados decorrentes da COVID-19 por meio do portal Brasil.io. Depois desta obtenção, foi realizada a coleta de dados no Twitter, obtendo dados dos tweets realizados pelos perfis das prefeituras

elencadas na fase anterior. Dados de população estimada residente foram obtidos por meio do portal do IBGE de cada município brasileiro, capital ou não. Após a obtenção de dados destas três fontes, implementou-se a junção delas, resultando em uma base de dados para o estudo. A fonte de dados principal deste trabalho é a rede social Twitter, após a coleta de dados, foram realizadas etapas de tratamento para se obter um dataset útil para as análises futuras.

#### A. Modelos

A indexação semântica latente (LSI) foi utilizada essencialmente neste trabalho para a recuperação das informações e como técnica de processamento de linguagem natural. Através deste método é possível identificar padrões nos termos contidos na coleção de dados não estruturados, oriundos dos textos dos tweets dos perfis das prefeituras com maior número de óbitos de COVID-19. A LSI foi selecionada pois é uma técnica capaz de aprender as correlações conceituais em uma coleção de texto visto que os dados do Twitter são metadados organizados de forma complexa.

#### B. Dados

### METODOLOGIA

Os dados utilizados para responder às perguntas de pesquisa, conforme Tabela I, encontram-se disponíveis em três conjuntos de dados. As variáveis da pesquisa são: data da criação do tweet, texto do tweet, quantidade de like do tweet, quantidade de retweet do tweet, quantidade de reply do tweet, número de casos confirmados de COVID-19, quantidade de óbitos por COVID-19, percentual de letalidade de COVID-19, número de casos confirmados de COVID-19 por 100 mil habitantes, quantidade de óbitos de COVID-19 por 100 mil habitantes, população estimada do município, município, UF, região brasileira e o tipo do município, capital ou cidade. Os dados principais são oriundos da rede social Twitter, com limite de 3200 tweets por perfil da prefeitura, do tweet mais novo para o antigo, resultando em 72 mil tweets após a aplicação de limpeza de dados. A pesquisa foi embasada nos dados dos boletins de COVID-19, oriundos do portal Brasil.io, focando nos 100 municípios com maior número acumulado de óbitos, sendo que apenas 90

prefeituras possuem perfil na rede social Twitter. Para a pesquisa também foram obtidos dados auxiliares por meio do portal do IBGE (Instituto Brasileiro de Geografia e Estatística) sobre a população residente estimada de cada município brasileiro.

### IV. RESULTADOS

Para a obtenção dos resultados foram elencadas as seguintes perguntas de pesquisa conforme especificado na Tabela I.

TABELA I – PERGUNTAS DA PESQUISA

ID	Pergunta
1	As prefeituras que fazem mais tweets sobre a COVID-19 têm um menor número de casos confirmados/óbitos?
2	Existe uma diferença de mortalidade/casos confirmados de COVID-19 em cidades que se comunicam mais no Twitter do que aquelas que não se comunicam?
3	Quais os tópicos mais utilizados pelas prefeituras nos tweets quando a pandemia teve o maior número de casos confirmados/óbitos?
4	Quais os tópicos mais utilizados pelas prefeituras nos tweets quando houve a diminuição dos casos confirmados/óbitos?
5	Qual a evolução dos tópicos dos tweets das prefeituras mensalmente e estes tópicos se relacionam com uma alta/baixa de número de casos confirmados de COVID-19?
6	Prefeituras que têm um maior engajamento no Twitter teriam menor número de casos confirmados/óbitos?

Em resposta à pergunta 1: “As prefeituras que fazem mais tweets sobre a COVID-19 têm um menor número de casos confirmados/óbitos?”, a Figura 1 demonstra que

Fig. 1 - Título

Fig. 2 - Título

Em resposta à pergunta 2: “Existe uma diferença de mortalidade/casos confirmados de COVID-19 em cidades que se comunicam mais no Twitter do que aquelas que não se comunicam?”, a Figura 2 apresenta.....

Em resposta à pergunta 3: “Quais os tópicos mais utilizados pelas prefeituras nos tweets quando a pandemia teve o maior número de casos confirmados/óbitos?”, foi possível observar conforme Figura 3 que .....

Em resposta à pergunta 4: “Quais as palavras mais utilizadas pelas prefeituras nos tweets quando houve a diminuição dos casos confirmados/óbitos?”, foi possível observar conforme Figura 4 que .....

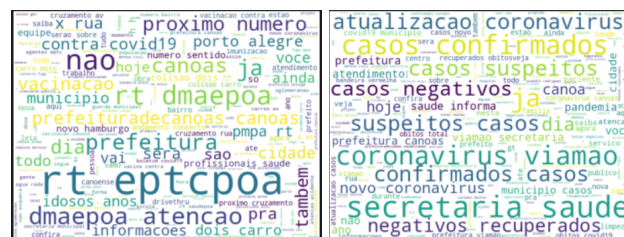


Comparativo entre períodos de pico e diminuicao dos casos

Figura (A) - Wordcloud com palavras mais usadas no aumento de casos - Figura B - Wordcloud com palavras mais usadas na diminuição dos casos



Comparativo do estado da Paraíba entre períodos de pico e diminuição dos casos - Figura A - Wordcloud com palavras mais usadas no aumento de casos - Figura B - Wordcloud com palavras mais usadas na diminuição dos casos



Comparativo do estado da Rio grande do Sul entre períodos de pico e diminuição dos casos - Figura A - Wordcloud com palavras mais usadas no aumento de casos - Figura B - Wordcloud com palavras mais usadas na diminuição dos casos

Em resposta à pergunta 5: “Qual a evolução dos tópicos dos tweets das prefeituras mensalmente e estes tópicos se relacionam com uma alta/baixa de número de casos confirmados de COVID-19?”, foi possível observar conforme Figura 5 que .....

Em resposta à pergunta 6: “Prefeituras que têm um maior engajamento no Twitter teriam menor número de casos confirmados/óbitos?”, a Figura 6 apresenta .....

## V. CONSIDERAÇÕES FINAIS

### A. Conclusões

Com base no estudo realizado, infere-se que os estados que tuitaram menos termos relacionados à COVID-19 quando houve o aumento de casos confirmados da doença; os estados tuitaram mais termos relacionados à COVID-19 quando diminuíram os casos confirmados da doença; um forte engajamento na rede social Twitter não está relacionado ao número de casos confirmados/óbitos; municípios que tuitam mais tendem a ter menos casos confirmados que municípios que não tuitam; o número de óbitos fica igual para municípios que tuitam mais e para os municípios que menos tuitam. Além disso, foi percebido que a distribuição dos tópicos foi se alterando conforme o aumento e distribuição dos casos confirmados, sendo que no começo da pandemia havia bastante tópicos relacionados à COVID-19, e depois, em novembro e dezembro de 2020, os tópicos começaram a mudar de foco, passando de COVID-19 para outros assuntos,

coincidindo com a diminuição dos casos. Consequentemente, no começo de 2021, houve muitos tópicos relacionados a COVID-19 e à vacinação, pois se espera que a massa de pessoas seja imunizada rapidamente.

#### *B. Limitações e Trabalhos Futuros*

As limitações enfrentadas ao decorrer da pesquisa se tratam: constatação de que nem todas as prefeituras possuem perfil no Twitter; e o limite de 3200 tweets por perfil devido às regras do Twitter. Quanto aos trabalhos futuros, pretende-se implementar um classificador de polaridades para análise de sentimentos dos tweets e realizar a análise dos dados de população e verbas fiscais dos municípios.

#### REFERÊNCIAS

- [1] F. R. S. Filho, E. F. Coutinho. “Uma Análise de Tweets sobre o Coronavírus. Revista Sistemas e Mídias Digitais,” 2020. [Em linha]. Disponível em: <https://revistasmd.virtual.ufc.br/arquivos/volume-5/numero-1/rsmd-v5-n1-3.pdf>
- [2] T. de Melo, C. M. S. Figueiredo, “A first public dataset from Brazilian twitter and News on COVID-19 in Portuguese. Elsevier,” 2020. [Em linha]. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352340920310738>
- [3] J. Ortiz et al, “Ciência e COVID-19 no Brasil: a repercussão das decisões da OMS no Twitter,” Revista Latinoamericana de Comunicación, 2020. [Em linha]. Disponível em: <https://revistachasqui.org/index.php/chasqui/issue/download/Chasqui%20145/155#page=476>
- [4] C. C. Xavier et al, “Extração e Classificação de Dados Semânticos do Twitter,” Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web: Minicursos, 2020. [Em Linha]. Disponível em: <https://sol.sbc.org.br/livros/index.php/sbc/catalog/download/8/16/55-1?inline=1>