**Ferramenta Perspective: um estudo de caso para análise de toxicidade em tweets sobre *SARS-Cov-2***

**Rodrigo Bertozzi De Castro¹ e Luciana De Nardin¹**

¹Departamento de Ciência da Computação

Pontifícia Universidade Católica de Minas Gerais (PUC Minas)

37.701-355 – Poços de Caldas – MG – Brasil

{rbcastro, luciana}@pucpcaldas.br

***Abstract.*** The Twitter platform provides the user of the social network with a real-time update of what is happening in the world, such as political news, music, football, the opinion of others on Twitter, among other things. Therefore, we have a large volume of information on each subject to study and draw our own conclusions. In this context, this work consists of using the Perspective API tool to show the toxicity of texts on the social network Twitter, known as tweets, to view the opinion of Twitter users about COVID-19 and its consequences on health, economic aspects and social.

***Resumo.*** A plataforma Twitter proporciona ao usuário da rede social atualização em tempo real do que está acontecendo no mundo, como, notícias de política, músicas, futebol, opinião de outras pessoas no Twitter, entre outras coisas. Sendo assim, temos um grande volume de informações sobre cada assunto para estudar e tirar as próprias conclusões. Neste contexto, este trabalho consiste em usar a ferramenta Perspective API para mostrar a toxicidade dos textos na rede social Twitter, conhecido como tweets, para visualizar a opinião dos usuários do Twitter sobre o COVID-19 e suas consequências sobre os aspectos de saúde, econômicos e sociais.

***Palavras-chave:*** Twitter, perspective, rede social, coronavírus, toxicidade, tweets

**1. Introdução e contextualização**

A rede social Twitter foi criada em 2006 por Jack Dorsey, Evan Williams, Biz Stone e Noah Glass. Desde então, o Twitter é uma das redes sociais mais usadas no mundo com cerca de 290 milhões de usuários¹. No final de 2020, o Twitter contabilizou 192 milhões de usuários ativos diários médios monetizáveis², 40 milhões de contas a mais que o mesmo período em 2019 com 152 milhões de usuários ativos diários médios monetizáveis³.

De acordo com o Ministério da Saúde do Brasil⁴, “os coronavírus são uma grande família de vírus comuns em muitas espécies diferentes de animais, incluindo camelos, gado, gatos e morcegos. Raramente, os coronavírus que infectam animais podem infectar pessoas, como exemplo do MERS-CoV e SARS-CoV. Recentemente, em dezembro de 2019, houve a transmissão de um novo coronavírus (SARS-CoV-2), o qual foi identificado em Wuhan na China e causou a COVID-19, sendo em seguida disseminada e transmitida pessoa a pessoa. A COVID-19 é uma doença causada pelo coronavírus, denominado SARS-CoV-2, que apresenta um espectro clínico variando de infecções assintomáticas a quadros graves. De acordo com a Organização Mundial de Saúde, a maioria (cerca de 80%) dos pacientes com COVID-19 podem ser assintomáticos ou oligossintomáticos (poucos sintomas), e aproximadamente 20% dos casos detectados requer atendimento hospitalar por apresentarem dificuldade respiratória, dos quais aproximadamente 5% podem necessitar de suporte ventilatório.”

Este trabalho, tem como base o uso da Ferramenta Perspective para análise de toxicidade em conversas sobre o COVID-19 na rede social Twitter, para isso, foram coletados tweets entre os dias 04 de março e 05 de março de 2021 para serem analisados.

**2. Perspective API**

A ferramenta Perspective foi criada pela JigSaw e pela equipe de tecnologia Counter Abuse do Google com o objetivo de identificar e diminuir a toxicidade em conversas on-line usando machine learning, no qual pontua o nível de toxicidade em uma frase ou texto, ajudando os desenvolvedores e editores a revisar os comentários com mais facilidade. A ferramenta Perspective fornece modelos de pontuações para vários atributos diferentes, como: toxicidade (toxicity), toxicidade severa (severe toxicity), insulto (insult), profanidade (profanity), ameaça (threat), ataque de identidade (identity attack) e conteúdo sexualmente explícito (sexually explict). A ferramenta está disponível para uso em inglês, espanhol, francês, alemão, português, italiano e russo⁵.

A pontuação é retornada pela resposta da API, indicando a probabilidade de o comentário solicitado ter o atributo fornecido. As pontuações representam uma probabilidade, com um valor entre 0 e 1, uma pontuação mais alta representa uma maior probabilidade que o comentário possui o atributo fornecido⁶.

O uso da ferramenta Perspective é comum na moderação de conteúdo, aumentando seus sistemas de moderação com a ferramenta, criando combinações personalizadas de atributos para atender às necessidades. O objetivo da Perspective API é “tornar a moderação de conteúdo menos um fardo para indivíduos e organizações, mas não substitui completamente o trabalho dos tomadores de decisão humanos. A API de perspectiva ajuda os desenvolvedores e moderadores a tomar melhores decisões em escala, permitindo que o diálogo saudável floresça.” ⁵

Perspective API é usado em grandes jornais no mundo como: The New York Times, Southeast Missourian e El País. Plataformas como Coral, OpenWeb e Disquis usam Perspective para criação de ferramentas e estratégias de moderação de conversas⁷.

**3. Mineração de textos**

Mineração de textos é a descoberta por computador de novas informações, anteriormente desconhecidos, extraindo automaticamente informações de diferentes recursos. Um elemento-chave é a vinculação de informações extraído em conjunto para formar novos fatos ou novas hipóteses a serem exploradas, além disso, por meios mais convencionais de experimentação (Hearst, 2003, tradução nossa).

Na primeira etapa de mineração de textos, ocorre a coleta de dados, de acordo com MARTINS (2003) “o objetivo é recuperar documentos relevantes ao domínio de aplicação do conhecimento a ser extraído”. Após a coleta, é feito o pré-processamento dos dados coletados. Conforme BATISTA (2003), no pré-processamento, é onde os dados são coletados e organizados para formar um conjunto de dados, e assim, podendo identificar dados corrompidos, atributos irrelevantes e valores desconhecidos. Esta fase visa preparar os dados para que a fase de extração de conhecimento seja mais efetiva.

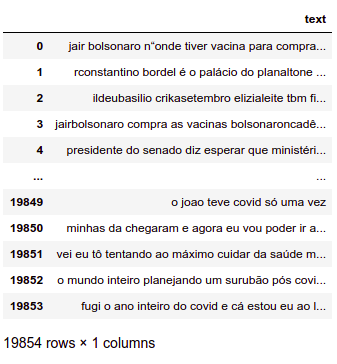
Depois de feito o pré-processamento, na mineração de dados, ocorre a extração de conhecimento, através dos dados pré-processados. De acordo com BATISTA (2003), nesta etapa, são utilizados algoritmos para receber dados e extrair conhecimento desses dados. Após extraído o conhecimento, é feito a análise e interpretação desses dados para que se possa obter análises concretas sobre a pesquisa.

**4. Estudo de caso**

Este estudo de caso tem como objetivo usar ferramenta Perspective para identificar o uso tóxico dos usuários da rede social Twitter sobre SARS-CoV-2 nos tweets com as palavras chaves: COVID-19, Covid, coronavirus, vacinas e vacina que foram coletados em língua portuguesa através da plataforma Twitter Developer e Postman durante os dias 04 e 05 de março de 2021.

No total, foram coletados 19854 tweets com identificação do tweet e o texto escrito, como o necessário para usar a ferramenta Perspective é apenas o texto, a tabela de dados foi modificada.

**Tabela 1 – Tabela de dados contendo apenas o texto dos tweets coletados**



Após a modificação da tabela, também foi retirado caracteres especiais e links que redirecionam para outros sites, deixando apenas as palavras para que a ferramenta Perspective tenha um desempenho melhor. Também foi feito uma WordCloud para identificar as palavras mais usadas, mostrada na figura 1.



**Figura 1 – WordCloud com as palavras mais usadas nos tweets coletados**

**5. Considerações finais**

**Referências**

1. Statistica. Twitter users worldwide. Disponível em: <https://www.statista.com/statistics/303681/twitter-users-worldwide/>. Acesso em: 26 mar. 2021.Na terceira etapa

2. G1. Economia. Tecnologia. Twitter supera projeções de lucro no fim do ano, mas crescimento da base de usuários fica abaixo das expectativas. Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2021/02/09/twitter-supera-projecoes-de-lucro-no-fim-do-ano-mas-crescimento-da-base-de-usuarios-fica-abaixo-das-expectativas.ghtml>. Acesso em: 26 mar. 2021.

3. Meio e Mensagem. Disponível em: <https://www.meioemensagem.com.br/home/midia/2020/05/04/facebook-e-twitter-crescimento-de-usuarios-e-incertezas-nos-negocios.html>. Acesso em: 26 mar. 2021.

4. Ministério Da Saúde. Sobre a COVID-19. Disponível em: <https://coronavirus.saude.gov.br/sobre-a-doenca#o-que-e-covid>. Acesso em: 26 mar. 2021.

5. Perspective API. How it works. Disponível em: <https://www.perspectiveapi.com/how-it-works>. Acesso em: 29 mar. 2021.

6. Developers Perspective API. About the api key concepts. Disponível em: <https://developers.perspectiveapi.com/s/about-the-api-key-concepts>. Acesso em: 29 mar. 2021.

7. Perspective API. Case studies. Disponível em: <https://www.perspectiveapi.com/case-studies>. Acesso em: 29 mar. 2021.

8. Hearst, Marti. "What is text mining." *SIMS, UC Berkeley* 5 (2003). Disponível em: <https://www.jaist.ac.jp/~bao/MOT-Ishikawa/FurtherReadingNo1.pdf>. Acesso em: 29 mar. 2021

9. Martins, Claudia Aparecida, et al. "Uma experiência em mineração de textos utilizando clustering probabilístico clustering hierárquico." *Instituto de Ciências Matemáticas e de Computação. São Carlos: Universidade de São Paulo* (2003).

10. Batista, Gustavo Enrique de Almeida Prado. *Pré-processamento de dados em aprendizado de máquina supervisionado*. Diss. Universidade de São Paulo, 2003.