

- XXVII Seminário de Iniciação Científica
- XIV Seminário Integrado de Ensino, Pesquisa e Extensão - SIEPE



## DETECÇÃO DE FAKE NEWS EM UM TWEET UTILIZANDO MACHINE LEARNING E PROCESSAMENTO DE LINGUAGEM NATURAL

Pesquisador(es): SALTON D., Giancarlo (autor principal); LORENCETI D., Ayrton

Instituição de Ensino Superior/Curso: Universidade do Oeste de Santa Catarina (Unoesc),  
Curso de Sistemas de Informação

Área: Ciências Exatas e Tecnológicas

**Introdução:** O número de fake news sendo compartilhadas vem aumentando e os usuários não possuem tempo hábil para verificar todas as notícias por se tratar de um processo demorado. Essa pesquisa teve o intuito de treinar um modelo de Machine Learning que, junto de técnicas de Processamento de Linguagem Natural, detectam se um tweet sobre o coronavírus contém fake news. **Objetivo:** Desenvolver um modelo de Machine Learning, que em conjunto com Processamento de Linguagem Natural, detecta se um tweet sobre o coronavírus em português contém ou não fake news. **Método:** Essa pesquisa é de cunho aplicado, abordagem quantitativa e qualitativa e para atingir os objetivos, foram utilizadas pesquisa exploratória e descritiva. Foram coletados um total de aproximadamente 180 mil registros e destes foram extraídas 344 amostras para treinar o modelo. Para essa pesquisa, não interessa a identidade do autor do tweet, apenas o conteúdo e o id do mesmo. **Resultados:** Durante a fase de treinamento, foram utilizados três modelos pré-treinados, sendo eles: bert-base-multilingual-uncased, bert-base-multilingual-cased e distilbert-basemultilingual-cased. Esses modelos estão todos disponíveis na biblioteca Transformers. Em relação aos algoritmos de classificação, foram utilizados a Árvore de Decisão, o k-Nearest Neighbors, o SGD, o LinearSVC e o SVC. Os modelos treinados obtiveram uma boa taxa de acerto ao identificar tweets que não contém fake news (73% a 93%) ou não se aplicam (69% a 83%). Porém, os modelos não obtiveram sucesso na detecção de fake news no tweets, tendo 0% de taxa de acerto. **Conclusão:** A detecção de fake news em tweets não teve êxito. Isso pode ser resultado de dois fatores: a remoção dos retweets do dataset ou uma amostra pequena de tweets

04 a 08/10/2021 | ISSN

- XXVII Seminário de Iniciação Científica
- XIV Seminário Integrado de Ensino, Pesquisa e Extensão - SIEPE



2237-6593 <https://portalperiodicos.unoesc.edu.br/siepe> classificados. Há alguns estudos na área que comprovam que tweets com fake news são mais disseminados do que tweets que não contêm fake news. Isso significa que esses tweets possuem um alcance maior e atingem mais público, o que se deve principalmente aos retweets. Sem retweets, o dataset de tweets coletados diminui consideravelmente o número de tweets contendo notícias falsas. A outra hipótese é que como o dataset de tweets coletados possui mais de 179 mil tweets, a amostragem selecionada para a classificação não contemplava um número bom de exemplos de tweets contendo fake news. No contexto desse trabalho, devido aos contratempos e a parcela utilizada do dataset, não foi possível detectar fake news em um tweet. Porém, em cenários diferentes, com registros diferentes, pode-se presumir que com a quantidade suficiente de exemplos, é possível sim detectar se um tweet contém fake news. Para trabalhos futuros, sugere-se a utilização de um dataset maior e principalmente com mais exemplos de tweets com fake news para o treinamento do modelo. Por fim, em conjunto com a classificação do conteúdo do tweet em si, poderia ser utilizada uma Rede Neural Convolucional para fazer a extração de textos em imagens dos tweets em busca de fake news, ou a criação de um Web crawler que buscaria o conteúdo de notícias linkadas aos tweets, para posteriormente utilizar

**Palavras-chave:** Machine Learning, Processamento de Linguagem Natural, Twitter, Fake News

**E-mails:** [ayrtonlorenceti@gmail.com](mailto:ayrtonlorenceti@gmail.com), [giancarlo.salton@unoesc.edu.br](mailto:giancarlo.salton@unoesc.edu.br)