

DESENVOLVIMENTO DE UM MODELO DE APRENDIZADO DE MÁQUINA, PARA DETECÇÃO DE FAKE NEWS NA LÍNGUA PORTUGUESA

Graduação: Sistemas da Informação

Área temática: Tecnologias da Informação e Comunicação

Resultados: Parcial

Forma de apresentação: Oral presencial

Lucas Schüler Costa¹; Francisco Assis Moreira do Nascimento ²

RESUMO

Notícias falsas têm se tornado um problema cada vez mais grave para a sociedade moderna. Podemos citar como exemplo o caso das eleições americanas de 2016, onde notícias falsas compartilhadas pelo Twitter tiveram importante influência no resultado da eleição. Mitigar os efeitos que as notícias falsas causam na sociedade, depende da capacidade de classificar rapidamente a notícia como falsa. Portanto, o objetivo deste trabalho será desenvolver um modelo utilizando técnicas de aprendizado de máquina. O modelo, deverá ser capaz de classificar uma notícia (escrita em português do Brasil) como verdadeira ou falsa. Será utilizado o dataset FakeRecogna, para realizar as tarefas de treinamento, teste e validação. Este dataset contém uma grande amostragem de notícias, escritas em português do Brasil, e a classificação desta em termos de verdadeira ou falsa. Serão utilizados como algoritmos/técnicas os seguintes: Naive Bayes, Logistic Regression, Support Vector Machines, Random Forest e finalmente o BERT. Após serão comparados os resultados de cada modelo, buscando identificar qual o com melhor performance na tarefa proposta.

Palavras-chave: Machine Learning, Fake News, Ciência de dados.

REFERÊNCIAS

HU, Kai; SLIVA, Amy; WANG, Suhang; et al. Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter, v. 19, n. 1, p. 22–36, 2017. Disponível em: <<https://doi.org/10.1145/3137597.3137600>>.

BOVET, Alexandre; MAKSE, Hernán A. Influence of fake news in Twitter during the 2016 US presidential election. Nature Communications, v. 10, n. 1, p. 7, 2019. Disponível em: <<https://www.nature.com/articles/s41467-018-07761-2>>.

Garcia, Gabriel L.; Afonso, Luis C. S.; Papa, João P.. FakeRecogna: A New Brazilian Corpus for Fake News Detection., v. 13208 LNAI, p. 57-67. Available at: <<http://hdl.handle.net/11449/234317>>.

¹Acadêmico das Faculdades Integradas de Taquara - FACCAT. lucascosta@sou.faccat.br

²Professor Orientador das Faculdades Integradas de Taquara - FACCAT. assis@faccat.br