



**Data Science  
Academy**

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

Processamento de Linguagem Natural

Mini-Projeto1

Detectando Sarcasmo em Passagens de  
Texto com TensorFlow 2.0



Este projeto destina-se a diferenciar se uma passagem de texto é sarcasmo ou não.

Estudos sobre a detecção de sarcasmo usam principalmente conjuntos de dados do Twitter coletados usando a supervisão baseada em hashtag, mas esses conjuntos de dados são "barulhentos" em termos de rótulos e idioma. Além disso, muitos tweets são respostas a outros tweets e a detecção de sarcasmo exige a disponibilidade de tweets contextuais.

Para superar as limitações relacionadas ao ruído nos conjuntos de dados do Twitter, um grupo de pesquisadores da Universidade de Princeton nos EUA, criou um conjunto de dados de notícias para a Detecção de Sarcasmo. Os dados foram coletados em dois sites de notícias:

O site TheOnion tem como objetivo produzir versões sarcásticas dos eventos atuais e foram coletadas todas as manchetes das categorias Notícias em Breve e Notícias em Fotos (que são sarcásticas). Coletaram ainda manchetes de notícias reais (e não sarcásticas) do HuffPost.

Aqui o link do estudo: <https://arxiv.org/pdf/1704.05579.pdf>

Neste Mini-Projeto você encontra toda a construção do modelo com arquitetura de Deep Learning e usando TensorFlow 2. Cada célula está comentada e com os links de referência.

Ao final, o modelo recebe novas frases e é capaz de prever se a frase é sarcasmo ou não.

Este Mini-Projeto é apenas para aquecer os motores e preparar para o que estudaremos em todo treinamento.