

Comparação de algoritmos de aprendizado profundo na detecção de comentários contendo discurso de ódio na Internet

Proposta de Trabalho Final de Graduação

Aluno: Rafael Greca Vieira **Matrícula:** 2018000434

Orientadora: Isabela Neves Drummond

Universidade Federal de Itajubá - UNIFEI
Instituto de Matemática e Computação - IMC
Curso de Ciência da Computação

Início: 26/04/2021 Término: 18/12/2021

Definição do Problema

A popularização da internet e, principalmente, das redes sociais, fez com que o volume de dados disponível aumentasse exponencialmente. Esse volume é composto por diversos tipos de dados, como: imagens, textos, áudios, vídeos, dados geoespaciais, dentre outros. Esse aumento trouxe a necessidade de uma evolução computacional para que fosse possível armazenar e posteriormente extrair informações valiosas desses dados. Uma área que teve uma grande evolução nesse período, mesmo tendo resquícios do seu surgimento antes mesmo da criação da internet, foi o aprendizado de máquina (*Machine Learning*), mais especificamente a área de aprendizado profundo (*Deep Learning*).

Aprendizado de máquina é o nome dado a capacidade de um sistema de inteligência artificial adquirir conhecimento através da extração de padrões em dados brutos (Goodfellow et al., 2016). O aprendizado profundo é uma das subáreas do aprendizado de máquina, que utiliza técnicas que implementam redes neurais profundas, permitindo que esses modelos computacionais consigam extrair informações dos dados com um maior nível de abstração (LeCun et al., 2015).

A análise de sentimentos é um dos campos de estudo da área de Processamento de Linguagem Natural (PLN) dentro do aprendizado profundo. O principal objetivo dessa área de estudo é analisar a opinião

humana e seus sentimentos e emoções sobre um determinado produto ou problema (Torfi et al., 2020). Com o crescimento das redes sociais, fóruns de discussão e sites de venda, onde os clientes podem deixar suas opiniões a respeito de um produto de uma determinada marca, esse campo de estudo vem ganhando ainda mais visibilidade e estudos. A análise de sentimentos possui diversas aplicações, podemos citar algumas delas: analisar o humor do público sobre um determinado movimento político, previsão de vendas de filmes, inteligência de mercado, medição da satisfação do cliente (Ravi et al., 2015), análise de publicações em perfis pessoais, análise sobre investimentos no setor financeiro, dentre outras.

A proposta deste trabalho envolve a utilização de uma base de dados com textos escritos em inglês e que foram extraídos da rede social Twitter. Esse conjunto de dados está disponível no Kaggle e Analytics Vidhya. Nesses textos serão aplicadas técnicas de pré-processamento com o objetivo de possibilitar que informações mais importantes sejam extraídas mais facilmente desses dados. Por fim, modelos de *deep learning* serão projetados e desenvolvidos para que possam ser utilizados na classificação do sentimento contido nos dados. Inicialmente será utilizado o framework Keras (Chollet et al., 2015) para o desenvolvimento desses algoritmos, porém o estudo e a possibilidade da utilização de outros frameworks não serão descartados. Por fim os modelos serão analisados e comparados entre si para que o objetivo definido neste trabalho seja cumprido.

Objetivo

O objetivo deste trabalho final de graduação é realizar uma comparação entre os modelos de aprendizado profundo aplicados na área da análise de sentimentos, mais especificamente na detecção de comentários na internet que podem ou não conter discurso de ódio.

Os modelos escolhidos para serem utilizados foram: *Convolutional Neural Network* (CNN), *Recurrent Neural Network* (RNN) e *Long-Short Term Memory* (LSTM). Eles foram escolhidos por se tratarem de algoritmos bem conhecidos, com uma grande variação de aplicações e por possuírem um acervo de materiais e artigos disponíveis. O projeto se concentra no estudo desses algoritmos de *deep learning*, utilização e aprendizado das técnicas de análise de sentimentos e processamento de linguagem natural. Além disso, os algoritmos serão executados e analisados em relação ao seu tempo de execução e acurácia dos resultados obtidos.

Cronograma

A tabela I apresenta todas as atividades que serão realizadas durante o desenvolvimento do TFG.

Tabela I - Cronograma das atividades do TFG

Atividade	Período
Elaboração da Proposta	Abril/2021
Levantamento Bibliográfico	Maio a Julho/2021
Apresentação Parcial do TFG	Julho/2021
Implementação e Parte Prática	Agosto a Setembro/2021
Obtenção dos Resultados	Setembro/2021
Monografia	Setembro a Novembro/2021
Entrega da Monografia	Novembro/2021
Apresentação Final do TFG	Dezembro/2021

Referências

Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning. MIT Press, 2016. <http://www.deeplearningbook.org>

Yann LeCun, Yoshua Bengio and Geoffrey Hinton. Deep Learning. Nature, 521 (7553):436-444, 2015.

Amirsina Torfi, Rouzbeh A. Shirvani, Yaser Keneshloo, Nader Tavaf and Edward A. Fox. Natural Language Processing Advancements By Deep Learning: A Survey. arXiv eprint arXiv:2003.01200, 2020.

Kumar Ravi, Vadlamani Ravi. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowl-Based Syst. 89. 14-46, 2015.

François Chollet and others. Keras. 2015. <https://keras.io>

Assinaturas

Aluno:

Orientadora: