

Processamento de Linguagem Natural: 5A – Classifying emotion intensity on tweets

José Miguel da Costa Simões - up201704317

José Rafael Fidalgo Fonseca Matias – up201706413

Paulo André Areias Gomes Leal Pinto - up201706415

Especificação do Trabalho

O tema do nosso trabalho é o “Affect in tweets – El-oc (an emotion intensity ordinal classification task)”

Os problemas de aprendizagem a serem enfrentados serão

- Geração de um *dataset* útil e adequado a partir de milhões de *tweets* em Inglês

- Exploração de diferentes técnicas de geração do *dataset*

- Implementação de diferentes algoritmos de aprendizagem

- Comparação da eficiência de cada algoritmo e tempo despendido pelo mesmo

Propomo-nos então a criar uma inteligência artificial, treinada
com datasets por nós gerados

Trabalhos encontrados na pesquisa

[CodaLab Datasets](#)

[The Best Way to Prepare a Dataset Easily](#)


[Machine Learning for Natural Language Processing](#)



Algoritmos e ferramentas

Algoritmos: A implementação de diferentes algoritmos para efeitos de comparação e análise de eficiência é uma parte chave do nosso trabalho. Assim, propomo-nos a explorar os algoritmos de aprendizagem supervisionada Naïve Bayes, Árvores de Decisão e Redes Neurais. Havendo a possibilidade de explorar mais algoritmos gostaríamos ainda de incluir neste conjunto Support Vector Machines.

Ferramentas: Uma vez que o segundo projeto será desenvolvido em Python, iremos utilizar o IDE Pycharm para o desenvolvimento do mesmo. Iremos também efetuar a exploração de outras bibliotecas/ferramentas como IPython, Jupyter Notebook, Sickit-learn, NLTK e StanfordNLP.



Trabalho já realizado

Foram criadas funções de extração de tweets e a respetiva classificação de intensidade emocional

Definimos funções para dar *parse* dos respetivos tweets e classificações, agrupando-os em estruturas de dados

Foi ainda desenvolvido um sistema rudimentar de IA baseado em Naïve Bayes

O sistema ainda tem um desempenho relativamente fraco, tendo sucesso em tweets claros e explícitos

Todo este desenvolvimento foi feito e testado para a emoção de fúria

