

UNISINOS – Tópicos Especiais II

Exercício prático GA

Nomes: Franciele Fagundes, Gabriel Castro, Matheus Hahn, Savanna Denega

Relatar os resultados segundo estas etapas previstas e escrever uma conclusão sobre o experimento, comentando sobre:

a. as ferramentas e recursos utilizados;

- Software Jupyter;
- Biblioteca NLTK;
- Biblioteca TextBlob;
- Biblioteca Pandas;
- Biblioteca Matplotlib;

b. as referências (sites, livros, artigos) utilizados;

- <https://www.nltk.org/>
- http://www.nltk.org/howto/portuguese_en.html
- <https://textblob.readthedocs.io/en/dev/>

c. os resultados obtidos com o experimento;

- Realizamos uma análise de polaridade das sentenças do dataset (frases de review dos hotéis da Europa) com a biblioteca TextBlob;
- Para a comparação do resultado da avaliação da ferramenta TextBlob, realizamos a análise manual de um subset do dataset utilizado;
- Apresentamos a comparação entre as frases, representando em uma matriz de confusão as frases selecionadas no experimento.

d. possibilidades de trabalhos futuros com este experimento;

- Possibilidade de continuar com o experimento, implementando um modelo de classificação correta a partir dos dados analisados, com por exemplo, utilização de Redes Neurais.

e. aspectos do aprendizado pessoal durante o experimento.

Aprendizado do uso das ferramentas:

- Jupyter;
- NLTK;
- TextBlob;
- Pandas;
- Matplotlib.

Aprendizado de conceitos:

- Trabalhar com as fases de organização de dados, tratamento de texto, de como utilizar a função de stopwords, e deletar palavras desnecessárias à avaliação;
- Comparação de dados, podendo fazer a real comparação da ferramenta com os dados reais da análise;
- Verificação de que a ferramenta TextBlob em modo puro (sem alinhamento direcionado a esse tipo de modelo de dataset) acaba por não oferecer o resultado real esperado da validação das frases, sendo necessário em um segundo momento fazer uma implementação, com por exemplo Redes Neurais para trabalhar com os dados apresentados.