

Processamento de Linguagem Natural

Yuri Malheiros

Exercício - Análise de Sentimentos

Descrição

O objetivo deste exercício é treinar e avaliar um classificador Naive Bayes para uma tarefa de análise de sentimentos.

O texto utilizado no treinamento e no teste deve ser pré-processado para incluir o prefixo “NAO_” de acordo com a seguinte regra:

O prefixo “NAO_” deve ser adicionado a cada palavra após um token de negação (não, nem, nunca, jamais, tampouco) até a próxima pontuação. São consideradas pontuações: . , ? ! ;

O treino deve ser realizado com 80% dos dados e o teste com 20%. Você deve calcular a acurácia do classificador aplicado aos dados de teste e comparar com o resultado do Naive Bayes sem a utilização do prefixo “NAO_”.

Dados

Os dados utilizados são os dados do corpus de reviews de produtos da B2W disponibilizado em: <https://github.com/yurimalheiros/pln/tree/main/datasets>

Entrega

Todo o código e resultados devem ser reportados usando um Jupyter Notebook. Crie uma célula no início com seu nome e matrícula. Coloque o seguinte nome seu notebook: `pln-01-seunome-seusobrenome`.

Esse exercício é individual, não faça em grupo.