

 04

Removendo a pontuação

Júlia está com dúvidas sobre a necessidade de remoção da pontuação, pois quando utiliza a função **CountVectorizer** do **SKLearning** a pontuação já é retirada.

Explique para Júlia se há **necessidade de retirar a pontuação? Por quê?**

Seleciona uma alternativa

A Não é necessário retirar a pontuação, pois quando se utiliza a função CountVectorizer do SKLearning a pontuação já é retirada automaticamente, não causando impacto na classificação.

B Sim, é necessário retirar a pontuação, pois mesmo usando o CountVectorizer para realizar a representação em Bag of Words, nós criamos a função pareto, que por sua vez não remove a pontuação automaticamente. Portanto, para que o pareto represente a frequência das palavras, o mais próximo possível da frequência da representação Bag of Words é importante remover as pontuações do nosso corpus.

C Mesmo usando o CountVectorizer para realizar a representação em Bag of Words, nós criamos a função pareto, que remove a pontuação automaticamente. Então para que o pareto represente a frequência das palavras o mais próximo possível da frequência da representação Bag of Words, é importante remover as pontuações do nosso corpus, visto que o CountVectorizer não remove a pontuação por padrão.