

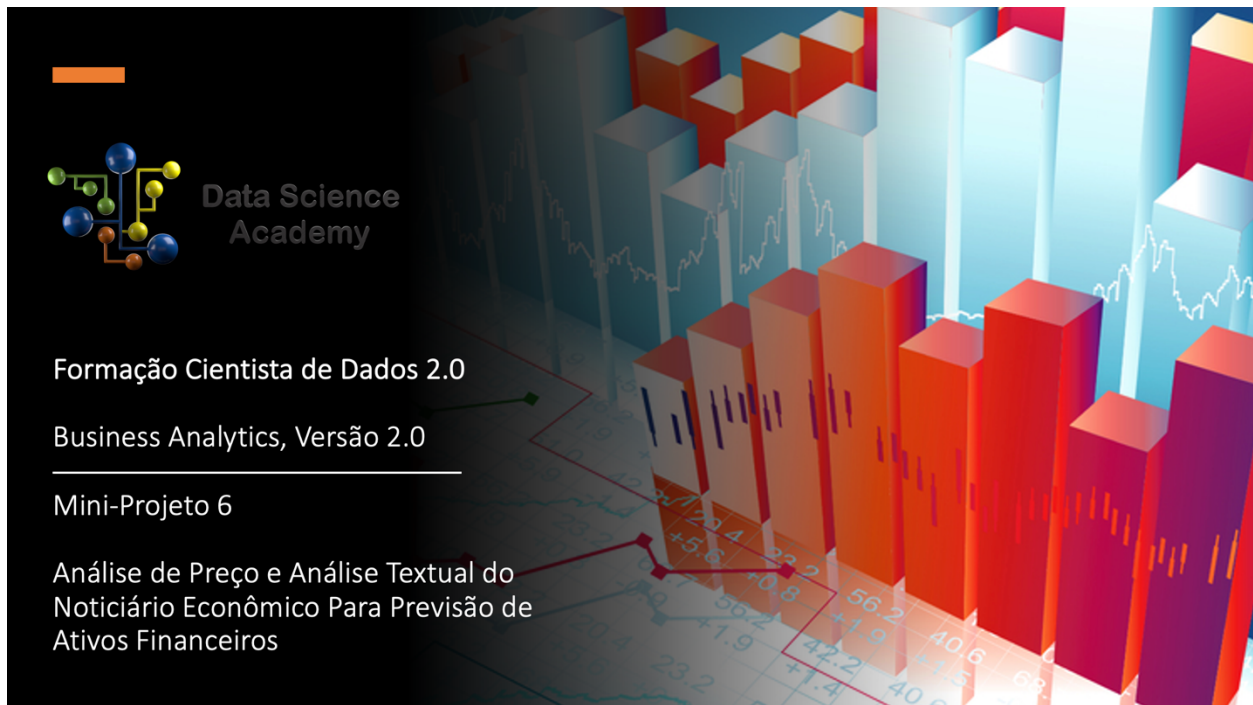
**Data Science
Academy**

www.datascienceacademy.com.br

Business Analytics

Mini-Projeto 6

Análise Textual



A análise de sentimento é um método de análise de texto que detecta a polaridade (por exemplo, uma opinião positiva ou negativa) dentro do texto, seja um documento inteiro, parágrafo ou frase.

A análise de sentimento visa medir a atitude, os sentimentos, as avaliações e as emoções de um falante / escritor com base no tratamento computacional da subjetividade em um texto.

Embora possa parecer fácil no papel, a análise de sentimento é um assunto complicado. Um texto pode conter vários sentimentos ao mesmo tempo. Por exemplo:

“A atuação foi boa, mas o filme poderia ter sido melhor”

A frase acima consiste em duas polaridades!!! Há um sentimento positivo e negativo na mesma sentença.

VADER (Valence Aware Dictionary for Sentiment Reasoning) é um modelo usado para análise de sentimento de texto que é sensível tanto à polaridade (positiva / negativa) quanto à intensidade (força) da emoção. Ele está disponível no pacote NLTK e pode ser aplicado diretamente a dados de texto não rotulados. Usaremos agora em nosso projeto.



A análise do VADER se baseia em um dicionário que mapeia recursos lexicais para intensidades emocionais conhecidas como pontuações de sentimento. A pontuação de sentimento de um texto pode ser obtida somando a intensidade de cada palavra no texto.

Por exemplo: palavras como 'amar', 'desfrutar', 'feliz', 'gostar', todas transmitem um sentimento positivo. Além disso, VADER é inteligente o suficiente para entender o contexto básico dessas palavras, como “não amei” como uma afirmação negativa. Ele também entende a ênfase de letras maiúsculas.

Vamos analisar o sentimento dos textos das notícias e usar essa informação junto com a cotação histórica do ativo financeiro para então construir os modelos preditivos.