

# Acurácia no TextBlob

Acurácia é uma métrica fundamental na avaliação da performance de modelos de classificação de texto. Ela mede a proporção de previsões corretas em relação ao total de previsões feitas pelo modelo.

## Fórmula da Acurácia

A fórmula matemática para calcular a acurácia é:  $\text{Acurácia} = (\text{Previsões Corretas}) / (\text{Total de Previsões})$  Por exemplo, se um modelo fez 100 previsões e acertou 80, sua acurácia seria de 80%. Isso significa que o modelo conseguiu classificar corretamente 80% dos textos.



# Importância da Acurácia

A acurácia desempenha um papel crucial na classificação de texto, fornecendo informações valiosas sobre a qualidade e confiabilidade do modelo.

## Avaliação de Modelos

A acurácia permite avaliar a eficácia de diferentes modelos de classificação. Ao comparar a acurácia de vários modelos, podemos identificar qual modelo é o mais preciso para a tarefa em questão.

## Confiabilidade das Previsões

A acurácia indica a confiabilidade das previsões do modelo. Um modelo com alta acurácia tem maior probabilidade de gerar previsões precisas e confiáveis, enquanto um modelo com baixa acurácia pode gerar previsões imprecisas e pouco confiáveis.

## Comparação de Modelos

A acurácia é essencial para comparar diferentes modelos de classificação. Ao comparar a acurácia de vários modelos treinados com os mesmos dados, podemos determinar qual modelo é o mais preciso para a tarefa específica.

# Cálculo da Acurácia

Para calcular a acurácia do modelo, é necessário comparar as classificações preditas pelo modelo com as classificações reais dos textos. O cálculo da acurácia pode ser realizado utilizando a função `accuracy_score` da biblioteca `scikit-learn`.

1

## Comparação

A função `accuracy_score` compara as classificações preditas pelo modelo (`rotulos_preditos`) com as classificações reais dos textos (`rotulos_verdadeiros`).

2

## Cálculo

A função calcula a proporção de previsões corretas em relação ao total de previsões.

3

## Resultado

O resultado é a acurácia do modelo, expressa como um valor entre 0 e 1, onde 1 representa 100% de acurácia.

# Limitações da Acurácia

Embora a acurácia seja uma métrica útil, é importante estar ciente de suas limitações, pois ela pode fornecer uma imagem distorcida da performance do modelo em determinadas situações.



## Conjuntos de Dados Desbalanceados

A acurácia pode ser enganosa em conjuntos de dados desbalanceados, ou seja, conjuntos onde a frequência de uma classe é significativamente maior que a frequência de outras classes. Em tais casos, um modelo pode alcançar alta acurácia simplesmente classificando todos os exemplos como a classe majoritária, mesmo que o modelo não esteja generalizando bem para outras classes.



## Gravidade dos Erros

A acurácia não considera a gravidade dos erros. Um modelo pode ter alta acurácia, mesmo que cometa erros graves em certos casos. Por exemplo, um modelo que classifica um email como spam pode ter um impacto significativamente maior do que um modelo que classifica um email como não spam, mesmo que ambos estejam errados.

