

16

Utilizando os mesmos dados nos 2 algoritmos

Vimos nesse capítulo a importância da utilização dos mesmo dados para compararmos dois algoritmos de comparação, vejamos se estavamos fazendo isso atualmente:

- Dados do nosso algoritmo:

```
from sklearn.naive_bayes import MultinomialNB
modelo = MultinomialNB()
modelo.fit(treino_dados, treino_marcacoes)

resultado = modelo.predict(teste_dados)

acertos = resultado == teste_marcacoes

total_de_acertos = sum(acertos)
total_de_elementos = len(teste_dados)
```

- Dados do algoritmo base:

```
acerto_base = max(Counter(Y).itervalues())
taxa_de_acerto_base = 100.0 * acerto_base / len(Y)
```

Aparentemente são os mesmos dados, porém, na prática são diferentes! Pois o `teste_marcacoes` faz parte de 90% do `Y`, ou seja, precisamos fazer com que o algoritmo base utilize também esses mesmo 90%. Antes de alterar os valores, move o código do algoritmo base para o final do arquivo `classifica_buscas.py`:

```
# restante do código

print("Taxa de acerto do algoritmo: %f" % taxa_de_acerto)
print(total_de_elementos)

acerto_base = max(Counter(Y).itervalues())
taxa_de_acerto_base = 100.0 * acerto_base / len(Y)
print("Taxa de acerto base: %f" % taxa_de_acerto_base)
```

Por fim, altere os valores do algoritmo base de `Y` para `teste_marcacoes`. Rode o código e cole o resultado abaixo:

Responda

INserir código

Formatação

