

作业答案

1.使用思维导图整理朴素贝叶斯的内容

2.动手实现商品评论情感分析案例

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import jieba
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.naive_bayes import MultinomialNB # 多项分布朴素贝叶斯

data = pd.read_csv('data/书籍评价.csv', encoding='gbk')
# print(data)
data['labels'] = np.where(data['评价'] == '好评', 1, 0)
# print(data)

y = data['labels']

stop_words = []
with open('data/stopwords.txt', encoding='utf8') as f:
    lines = f.readlines()
    stop_words = [line.strip() for line in lines]

stop_words = list(set(stop_words))
# print(stop_words)

comment_list = [','.join(jieba.lcut(line)) for line in data['内容']]
# print(comment_list)

transform = CountVectorizer(stop_words=stop_words)
x = transform.fit_transform(comment_list)
names = transform.get_feature_names_out()
# print(names)
x = x.toarray()
# print(x)

x_train = x[0:10, :]
x_test = x[10:, :]
# print(y)
# print("-----")
y = y.values
y_train = y[0:10]
y_test = y[10:]
# print(x_train.shape)
# print(y_train.shape)

model = MultinomialNB(alpha=1.0)
model.fit(x_train, y_train)
```

```
print(model.predict(x_test))  
  
print(model.score(x_test, y_test))
```