plement-naive-bayes-classification

October 31, 2023

[21]:
```
Ankush Kumar
AP21110011026
CSE-P
```

[ ]:
```
# Define the data you provided
data = [
    {"age": "<=30", "income": "high", "student": "no", "credit_rating": "fair",␣
 ↪"comp": "no"},
    {"age": "<=30", "income": "high", "student": "excellent", "credit_rating":␣
 ↪"no", "comp": "yes"},
    {"age": "31...40", "income": ">40", "student": "high", "credit_rating":␣
 ↪"no", "comp": "fair"},
    {"age": ">40", "income": ">40", "student": "low", "credit_rating": "yes",␣
 ↪"comp": "no"},
    {"age": "31...40", "income": "low", "student": "yes", "credit_rating":␣
 ↪"excellent", "comp": "no"},
    {"age": "31...40", "income": ">40", "student": "low", "credit_rating":␣
 ↪"excellent", "comp": "yes"},
    {"age": "<=30", "income": "medium", "student": "no", "credit_rating":␣
 ↪"fair", "comp": "no"},
    {"age": "<=30", "income": "low", "student": "yes", "credit_rating": "fair",␣
 ↪"comp": "yes"},
    {"age": ">40", "income": "medium", "student": "yes", "credit_rating":␣
 ↪"fair", "comp": "yes"},
    {"age": "<=30", "income": "medium", "student": "yes", "credit_rating":␣
 ↪"excellent", "comp": "yes"},
    {"age": "31...40", "income": "medium", "student": "no", "credit_rating":␣
 ↪"excellent", "comp": "yes"},
    {"age": "31...40", "income": ">40", "student": "high", "credit_rating":␣
 ↪"fair", "comp": "yes"},
    {"age": "medium", "income": "low", "student": "yes", "credit_rating":␣
 ↪"excellent", "comp": "no"},
]
```

[ ]:
```
# Calculate the total number of instances
total_instances = len(data)
```

[22]:
```
# Separate the data into 'yes' and 'no' classes
yes_data = [d for d in data if d['comp'] == 'yes']
no_data = [d for d in data if d['comp'] == 'no']
```

[23]:
```
# Calculate the prior probabilities
P_yes = len(yes_data) / total_instances
P_no = len(no_data) / total_instances
```

[24]:
```
print(f"P(buys_computer = 'yes') = {P_yes}")
print(f"P(buys_computer = 'no') = {P_no}")
```

```
P(buys_computer = 'yes') = 0.5384615384615384
P(buys_computer = 'no') = 0.38461538461538464
```

[25]:
```
# Define the input features
input_features = {
    "age": "<=30",
    "income": "medium",
    "student": "yes",
    "credit_rating": "fair",
}
```

[26]:
```
# Calculate conditional probabilities
def calculate_conditional_probability(feature, value, class_data):
    count = sum(1 for d in class_data if d[feature] == value)
    total = len(class_data)
    return count / total
```

[27]:
```
P_age_yes = calculate_conditional_probability("age", input_features["age"],␣
 ↪yes_data)
P_income_medium_yes = calculate_conditional_probability("income",␣
 ↪input_features["income"], yes_data)
P_student_yes_yes = calculate_conditional_probability("student",␣
 ↪input_features["student"], yes_data)
P_credit_fair_yes = calculate_conditional_probability("credit_rating",␣
 ↪input_features["credit_rating"], yes_data)
```

[28]:
```
P_age_no = calculate_conditional_probability("age", input_features["age"],␣
 ↪no_data)
P_income_medium_no = calculate_conditional_probability("income",␣
 ↪input_features["income"], no_data)
P_student_yes_no = calculate_conditional_probability("student",␣
 ↪input_features["student"], no_data)
P_credit_fair_no = calculate_conditional_probability("credit_rating",␣
 ↪input_features["credit_rating"], no_data)
```

[29]:
```
print(f"P(age = '<=30' | buys_computer = 'yes') = {P_age_yes}")
print(f"P(income = 'medium' | buys_computer = 'yes') = {P_income_medium_yes}")
print(f"P(student = 'yes' | buys_computer = 'yes') = {P_student_yes_yes}")
print(f"P(credit_rating = 'fair' | buys_computer = 'yes') =␣
 ↪{P_credit_fair_yes}")
```

```
P(age = '<=30' | buys_computer = 'yes') = 0.42857142857142855
P(income = 'medium' | buys_computer = 'yes') = 0.42857142857142855
P(student = 'yes' | buys_computer = 'yes') = 0.42857142857142855
P(credit_rating = 'fair' | buys_computer = 'yes') = 0.42857142857142855
```

[30]:
```
print(f"P(age = '<=30' | buys_computer = 'no') = {P_age_no}")
print(f"P(income = 'medium' | buys_computer = 'no') = {P_income_medium_no}")
print(f"P(student = 'yes' | buys_computer = 'no') = {P_student_yes_no}")
print(f"P(credit_rating = 'fair' | buys_computer = 'no') = {P_credit_fair_no}")
```

```
P(age = '<=30' | buys_computer = 'no') = 0.4
P(income = 'medium' | buys_computer = 'no') = 0.2
P(student = 'yes' | buys_computer = 'no') = 0.4
P(credit_rating = 'fair' | buys_computer = 'no') = 0.4
```

[31]:
```
# Calculate P(X|Ci)
P_X_given_yes = P_age_yes * P_income_medium_yes * P_student_yes_yes *␣
 ↪P_credit_fair_yes
P_X_given_no = P_age_no * P_income_medium_no * P_student_yes_no *␣
 ↪P_credit_fair_no
```

[32]:
```
print(f"P(X|buys_computer = 'yes') = {P_X_given_yes}")
print(f"P(X|buys_computer = 'no') = {P_X_given_no}")
```

```
P(X|buys_computer = 'yes') = 0.033735943356934604
P(X|buys_computer = 'no') = 0.012800000000000004
```

[33]:
```
# Calculate P(X|Ci) * P(Ci)
P_X_given_yes_times_P_yes = P_X_given_yes * P_yes
P_X_given_no_times_P_no = P_X_given_no * P_no
```

[34]:
```
print(f"P(X|buys_computer = 'yes') * P(buys_computer = 'yes') =␣
 ↪{P_X_given_yes_times_P_yes}")
print(f"P(X|buys_computer = 'no') * P(buys_computer = 'no') =␣
 ↪{P_X_given_no_times_P_no}")
```

```
P(X|buys_computer = 'yes') * P(buys_computer = 'yes') = 0.018165507961426325
P(X|buys_computer = 'no') * P(buys_computer = 'no') = 0.004923076923076925
```

[35]:
```
# Determine the predicted class
```

```
prediction = 'yes' if P_X_given_yes_times_P_yes > P_X_given_no_times_P_no else␣
↪'no'
```

[36]:
```
print(f"The predicted class is: {prediction}")
```

```
The predicted class is: yes
```