naive_bayes_test

July 3, 2024

[1]:
```
#import necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
```

[2]:
```
#Create dataframe and split into training and test
df= pd.read_excel('naive_bayes_algorithm/test-data.xlsx')
df=df.iloc[: , [7,8,9,10,11,12]]
print(df.head(2).to_string())
target= df.GA
inputs=df.drop('GA', axis="columns")
x_train, x_test, y_train, y_test = train_test_split(inputs, target, test_size=0.
↪2)
```

```
   Saves  Save%   CS  PSxG  Opposition XG   GA
0    3.0  100.0  1.0   0.3            1.1  0.0
1    3.0   42.9  0.0   3.4            2.5  4.0
```

[3]:
```
#Create and test model for accuracy
model=GaussianNB()
model.fit(x_train,y_train)
print(model.score(x_test,y_test))
print(y_test)
print(model.predict(x_test))
```

```
0.7283950617283951
174    1.0
343    1.0
317    3.0
106    0.0
347    0.0
      …
175    1.0
146    0.0
196    7.0
50     0.0
292    2.0
Name: GA, Length: 81, dtype: float64
```

```
[1. 1. 2. 0. 0. 1. 0. 0. 1. 1. 3. 3. 3. 3. 0. 0. 3. 3. 1. 2. 3. 3. 0. 1.
 1. 1. 2. 1. 0. 0. 0. 0. 1. 1. 0. 2. 2. 3. 1. 1. 2. 1. 1. 1. 0. 1. 3. 0.
 0. 0. 1. 0. 1. 0. 1. 1. 2. 2. 1. 0. 1. 0. 1. 0. 0. 1. 1. 1. 3. 0. 1. 0.
 1. 0. 0. 4. 1. 0. 4. 0. 1.]
```

[ ]: