

Assignment No. 6

Contents for Theory:

1. Concepts used in Naïve Bayes classifier
2. Naive Bayes Example
3. Confusion Matrix Evaluation Metrics

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.naive_bayes import GaussianNB
from sklearn.metrics import accuracy_score, precision_score,
recall_score, confusion_matrix
```

```
url = 'C:/Users/Ashwini/Desktop/Iris.csv'
df = pd.read_csv(url)
```

```
print("First 5 rows of the dataset:")
print(df.head())
```

First 5 rows of the dataset:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa

```
print("\n Checking for null values:")
print(df.isnull().sum())
```

Checking for null values:

```
Id          0
SepalLengthCm  0
SepalWidthCm  0
PetalLengthCm  0
PetalWidthCm  0
Species      0
dtype: int64
```

```
df['Species'] = df['Species'].astype('category').cat.codes
```

```
X = df.drop(columns=['Species'])
```

```
Y = df['Species']
```

```

X_train, X_test, y_train, y_test = train_test_split(X, Y,
test_size=0.2, random_state=42)
gaussian = GaussianNB()
gaussian.fit(X_train, y_train)
y_pred = gaussian.predict(X_test)

accuracy = accuracy_score(y_test, y_pred)
precision = precision_score(y_test, y_pred, average='micro')
recall = recall_score(y_test, y_pred, average='micro')
cm = confusion_matrix(y_test, y_pred)

print("\n Model Performance:")
print(f" Accuracy: {accuracy:.4f}")
print(f" Precision: {precision:.4f}")
print(f" Recall: {recall:.4f}")

```

```

Model Performance:
Accuracy: 1.0000
Precision: 1.0000
Recall: 1.0000

```

```

print("\n Confusion Matrix:")
print(cm)

```

```

Confusion Matrix:
[[10  0  0]
 [ 0  9  0]
 [ 0  0 11]]

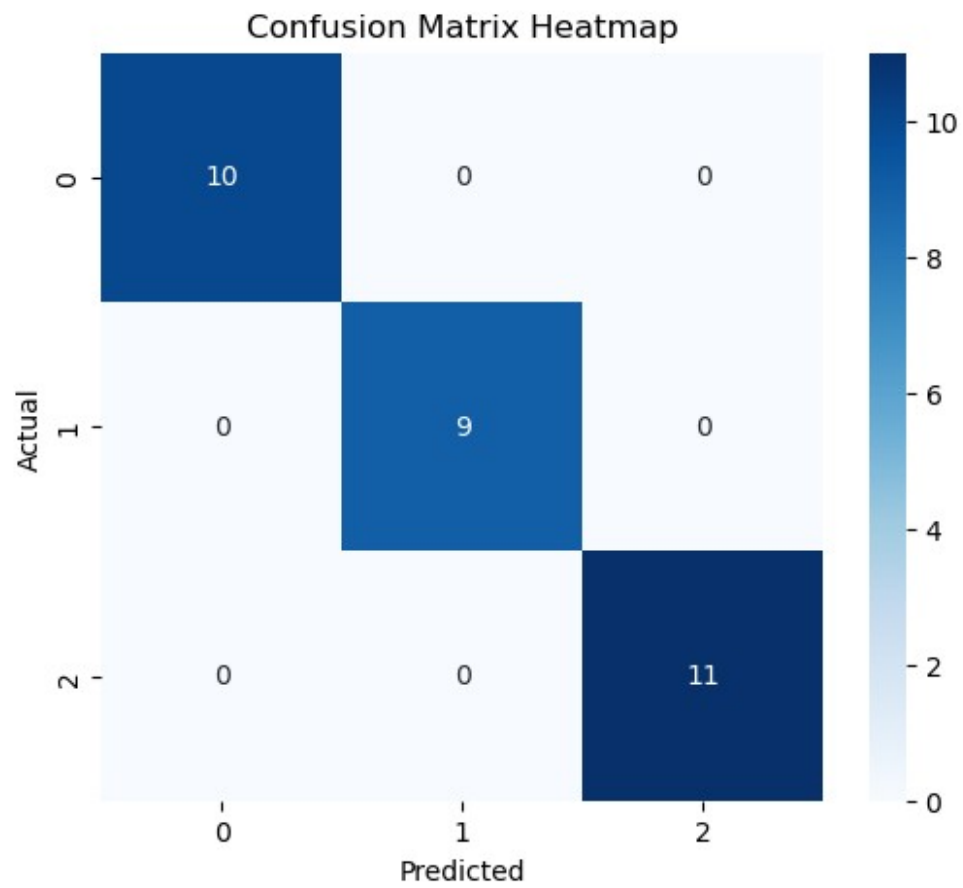
```

```

import seaborn as sns

plt.figure(figsize=(6, 5))
sns.heatmap(cm, annot=True, fmt="d", cmap="Blues",
xticklabels=np.unique(df['Species']),
yticklabels=np.unique(df['Species']))
plt.xlabel("Predicted")
plt.ylabel("Actual")
plt.title("Confusion Matrix Heatmap")
plt.show()

```



Name: Habibsaeed Mukebil
R_no. : 13235