

ML LAB ASSIGNMENT 1

NAME - KUNTAL DAS

ROLL - 002011001032

IRIS_PLANTS : →

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[1] import pandas as pd
import numpy as np
from google.colab import files
import io
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.naive_bayes import GaussianNB, MultinomialNB, BernoulliNB
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier, export_text, export_graphviz
import pydotplus
from IPython.display import Image

[22] # from google.colab import drive
# drive.mount('/content/drive')

[23] # uploaded = files.upload()

[24] # csv_file = pd.read_csv(io.BytesIO(uploaded['iris.csv']))
iris_data = pd.read_csv('/content/drive/MyDrive/ML lab/iris.csv', names=['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'species'])
print(iris_data)

sepal_length sepal_width petal_length petal_width species
completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[25] x = iris_data.drop('species', axis=1)
y = iris_data['species']
print(x)
print(y)

[26] x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.2, random_state=42)
print(x_train)
print(x_test)
print(y_train)
print(y_test)

[150 row x 5 columns]
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[27] gnb = GaussianNB()
gnb.fit(x_train, y_train)
y_pred_gnb = gnb.predict(x_test)
print(y_pred_gnb)

[28] mb = MultinomialNB()
mb.fit(x_train, y_train)
y_pred_mb = mb.predict(x_test)
print(y_pred_mb)

[29] bnb = BernoulliNB()
bnb.fit(x_train, y_train)
y_pred_bnb = bnb.predict(x_test)
print(y_pred_bnb)

[30] accuracy_gnb = accuracy_score(y_test, y_pred_gnb)
print("Gaussian Naive Bayes:")
print("Accuracy: (accuracy_gnb:.2f)")
print(classification_report(y_test, y_pred_gnb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_gnb))

completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[31] print(confusion_matrix(y_test, y_pred_gnb))

Gaussian Naive Bayes:
Accuracy: 1.00

precision recall f1-score support
Iris-setosa 1.00 1.00 1.00 10
Iris-versicolor 1.00 1.00 1.00 9
Iris-virginica 1.00 1.00 1.00 11

accuracy 1.00 1.00 1.00 30
macro avg 1.00 1.00 1.00 30
weighted avg 1.00 1.00 1.00 30

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

[32] accuracy_mb = accuracy_score(y_test, y_pred_mb)
print("Multinomial Naive Bayes:")
print("Accuracy: (accuracy_mb:.2f)")
print(classification_report(y_test, y_pred_mb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_mb))

Multinomial Naive Bayes:
Accuracy: 0.90

precision recall f1-score support
Iris-setosa 1.00 1.00 1.00 10
Iris-versicolor 0.75 1.00 0.86 9
Iris-virginica 1.00 0.73 0.84 11

accuracy 0.92 0.91 0.90 30
macro avg 0.93 0.90 0.90 30
weighted avg 0.93 0.90 0.90 30

Confusion Matrix:
[[10 0 0]
 [0 9 0]
 [0 3 0]]

completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[33] bnb = BernoulliNB()
bnb.fit(x_train, y_train)
y_pred_bnb = bnb.predict(x_test)
print(y_pred_bnb)

[34] accuracy_bnb = accuracy_score(y_test, y_pred_bnb)
print("Bernoulli Naive Bayes:")
print("Accuracy: (accuracy_bnb:.2f)")
print(classification_report(y_test, y_pred_bnb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_bnb))

Bernoulli Naive Bayes:
Accuracy: 0.30

precision recall f1-score support
Iris-setosa 0.00 0.00 0.00 10
Iris-versicolor 0.30 1.00 0.46 9
Iris-virginica 0.00 0.00 0.00 11

accuracy 0.30 0.30 0.30 30
macro avg 0.30 0.33 0.35 30
weighted avg 0.30 0.30 0.34 30

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

[35] accuracy_bnb = accuracy_score(y_test, y_pred_bnb)
print("Bernoulli Naive Bayes:")
print("Accuracy: (accuracy_bnb:.2f)")
print(classification_report(y_test, y_pred_bnb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_bnb))

Bernoulli Naive Bayes:
Accuracy: 0.30

precision recall f1-score support
Iris-setosa 0.00 0.00 0.00 10
Iris-versicolor 0.30 1.00 0.46 9
Iris-virginica 0.00 0.00 0.00 11

accuracy 0.30 0.30 0.30 30
macro avg 0.30 0.33 0.35 30
weighted avg 0.30 0.30 0.34 30

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

completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[36] dtc = DecisionTreeClassifier(criterion='gini')
dtc.fit(x_train, y_train)
y_pred_dtc = dtc.predict(x_test)
print(y_pred_dtc)

['Iris-versicolor' 'Iris-setosa' 'Iris-virginica' 'Iris-versicolor']

[37] accuracy_dtc = accuracy_score(y_test, y_pred_dtc)
print("Decision Tree Classifier:")
print("Accuracy: (accuracy_dtc:.2f)")
print(classification_report(y_test, y_pred_dtc))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_dtc))

Decision Tree Classifier:
Accuracy: 1.00

precision recall f1-score support
Iris-setosa 1.00 1.00 1.00 10
Iris-versicolor 1.00 1.00 1.00 9
Iris-virginica 1.00 1.00 1.00 11

accuracy 1.00 1.00 1.00 30
macro avg 1.00 1.00 1.00 30
weighted avg 1.00 1.00 1.00 30

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

completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[38] dtc = DecisionTreeClassifier(criterion='gini')
dtc.fit(x_train, y_train)
y_pred_dtc = dtc.predict(x_test)
print(y_pred_dtc)

['Iris-versicolor' 'Iris-setosa' 'Iris-virginica' 'Iris-versicolor']

[39] accuracy_dtc = accuracy_score(y_test, y_pred_dtc)
print("Decision Tree Classifier:")
print("Accuracy: (accuracy_dtc:.2f)")
print(classification_report(y_test, y_pred_dtc))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_dtc))

Decision Tree Classifier:
Accuracy: 1.00

precision recall f1-score support
Iris-setosa 1.00 1.00 1.00 10
Iris-versicolor 1.00 1.00 1.00 9
Iris-virginica 1.00 1.00 1.00 11

accuracy 1.00 1.00 1.00 30
macro avg 1.00 1.00 1.00 30
weighted avg 1.00 1.00 1.00 30

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

completed at 10:48 AM
```

```
iris.ipynb
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text
RAM Disk

[40] dtc = DecisionTreeClassifier(criterion='gini')
dtc.fit(x_train, y_train)
y_pred_dtc = dtc.predict(x_test)
print(y_pred_dtc)

['Iris-versicolor' 'Iris-setosa' 'Iris-virginica' 'Iris-versicolor']

[41] accuracy_dtc = accuracy_score(y_test, y_pred_dtc)
print("Decision Tree Classifier:")
print("Accuracy: (accuracy_dtc:.2f)")
print(classification_report(y_test, y_pred_dtc))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_dtc))

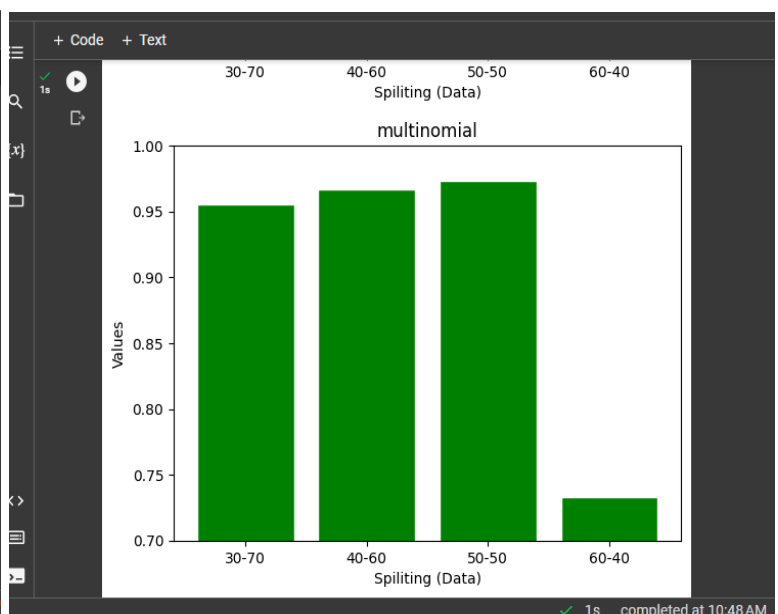
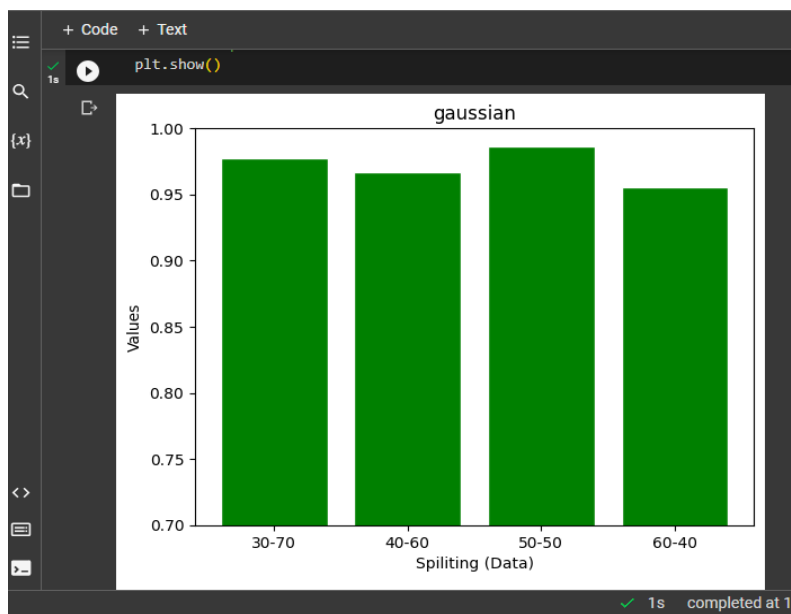
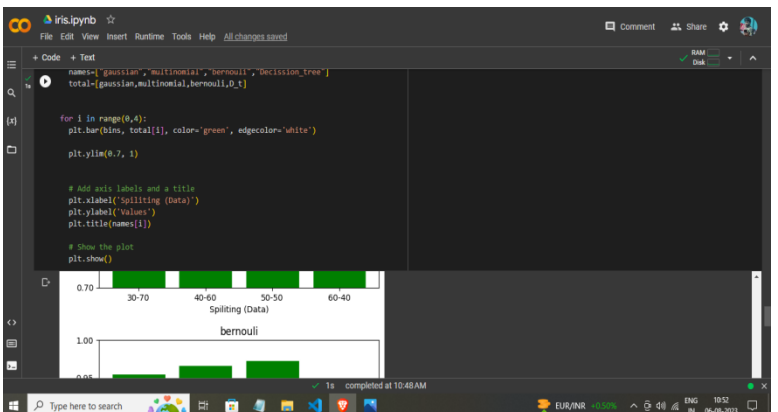
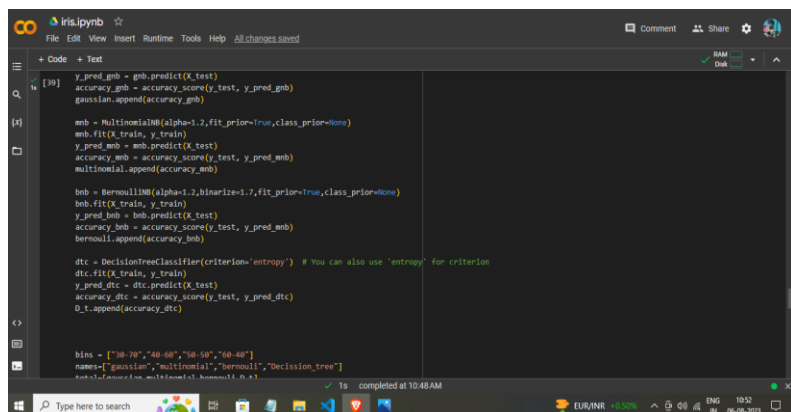
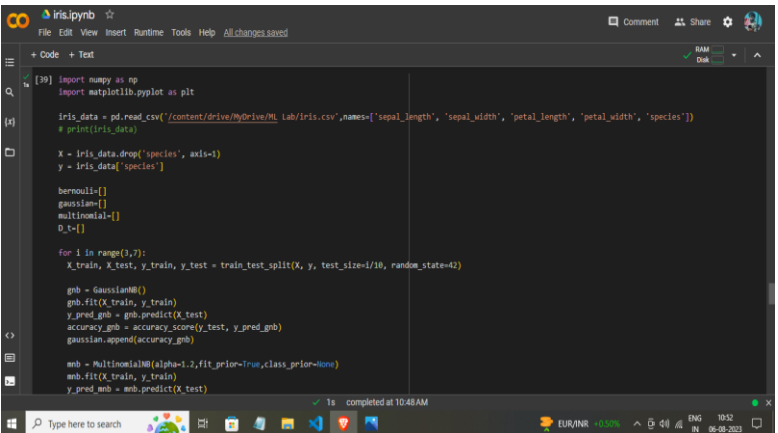
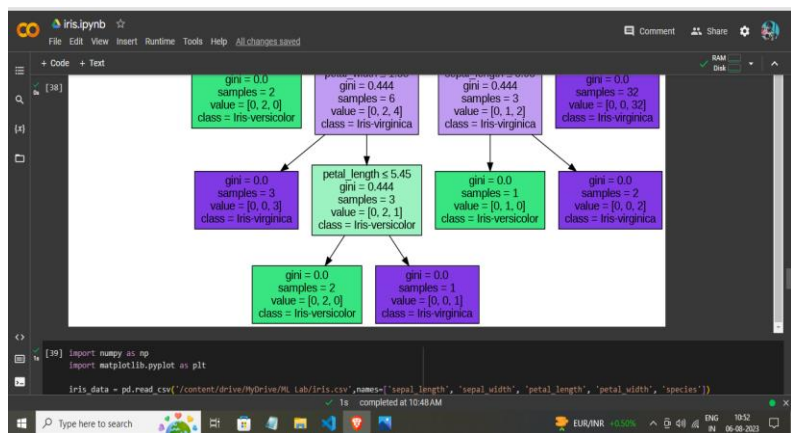
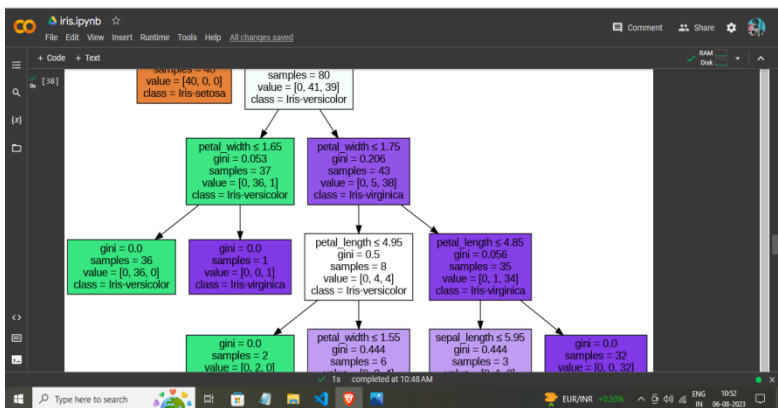
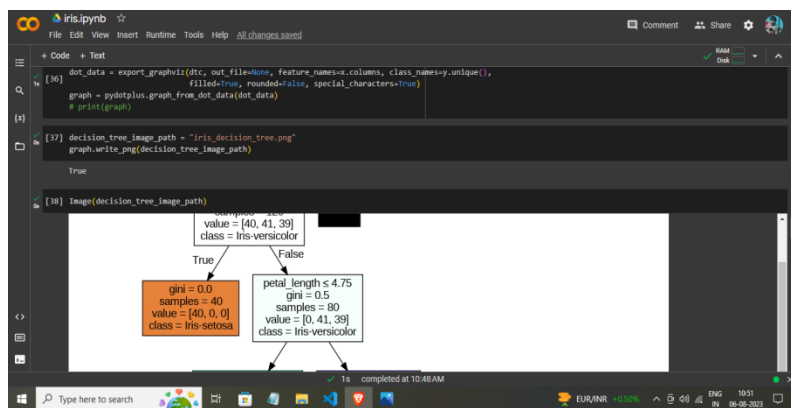
Decision Tree Classifier:
Accuracy: 1.00

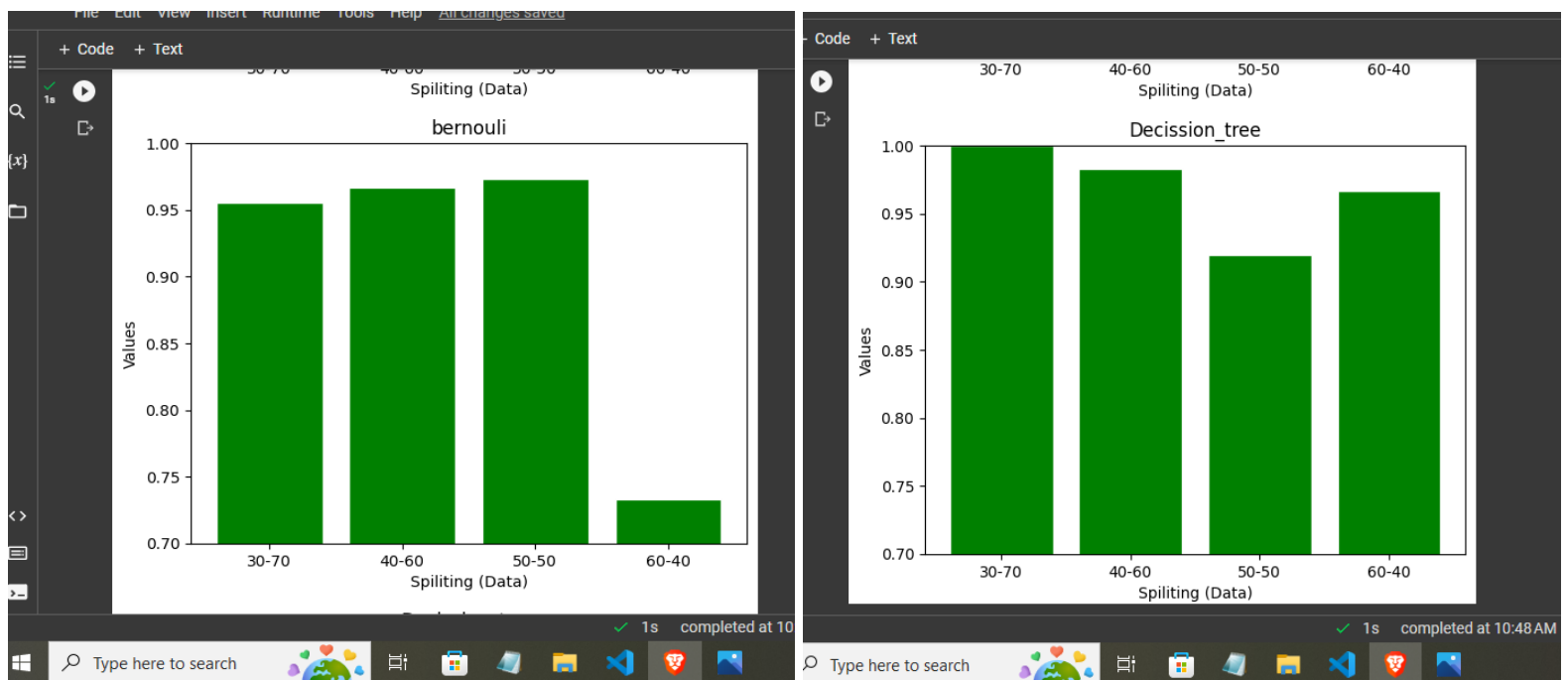
precision recall f1-score support
Iris-setosa 1.00 1.00 1.00 10
Iris-versicolor 1.00 1.00 1.00 9
Iris-virginica 1.00 1.00 1.00 11

accuracy 1.00 1.00 1.00 30
macro avg 1.00 1.00 1.00 30
weighted avg 1.00 1.00 1.00 30

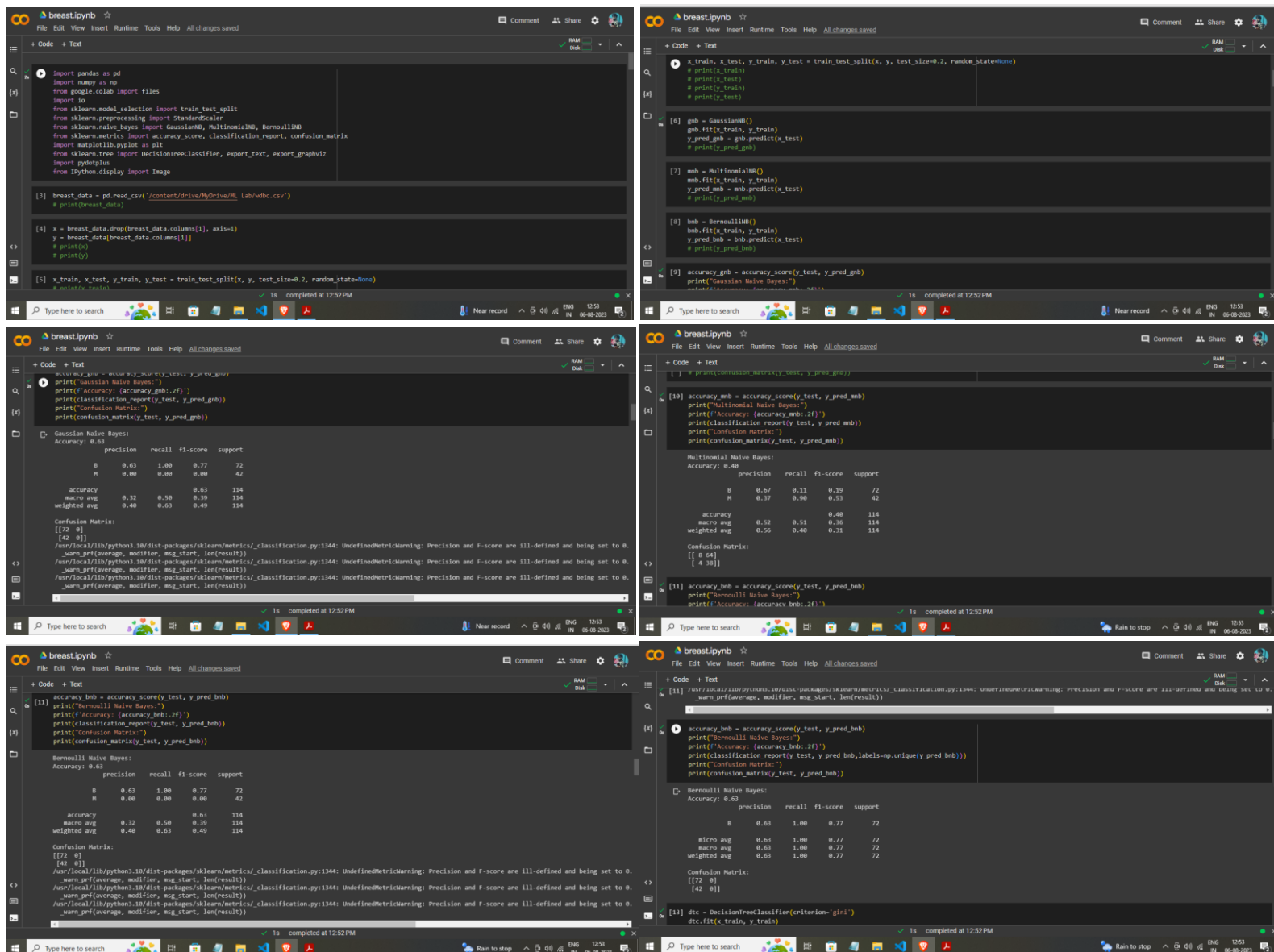
Confusion Matrix:
[[10 0 0]
 [0 9 0]
 [0 0 11]]

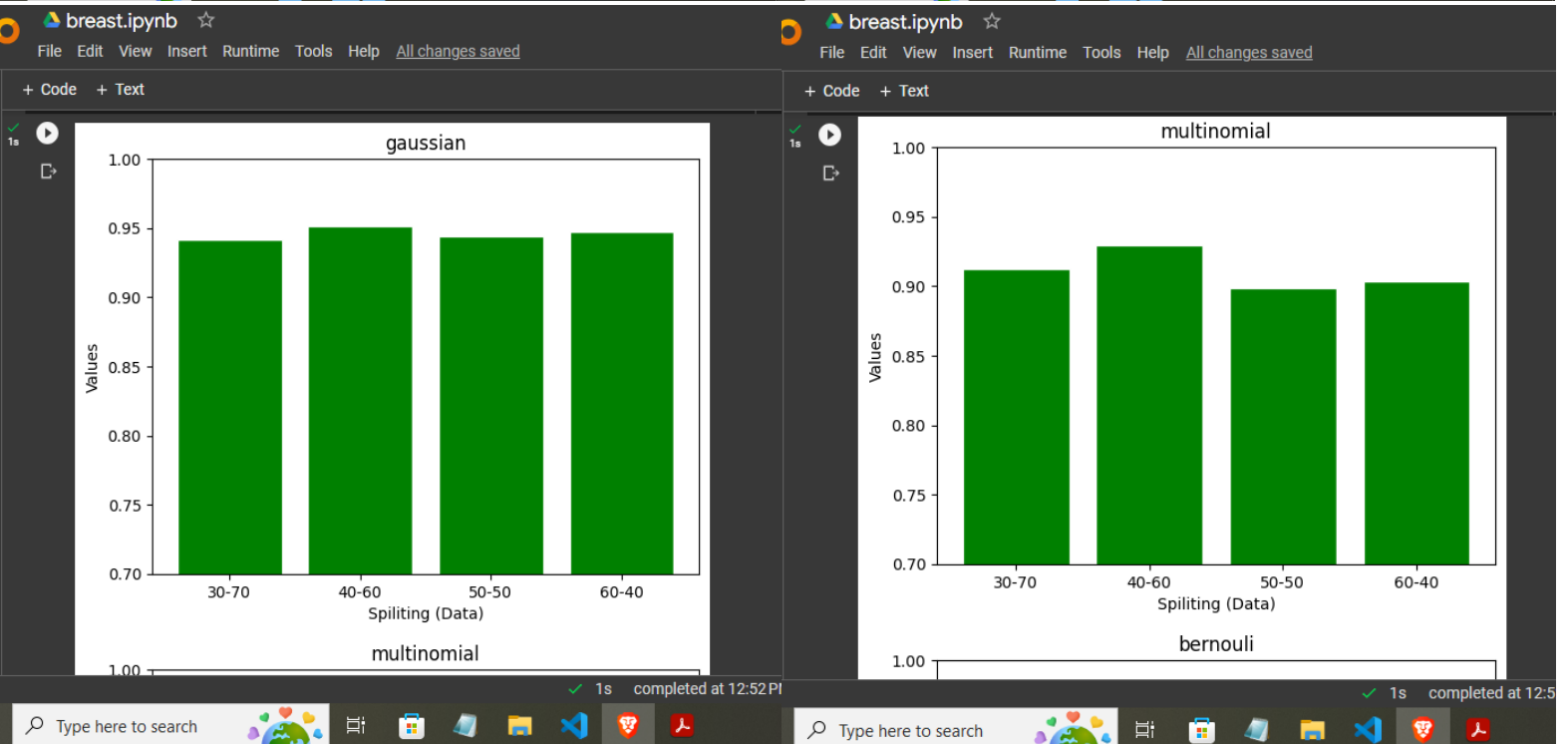
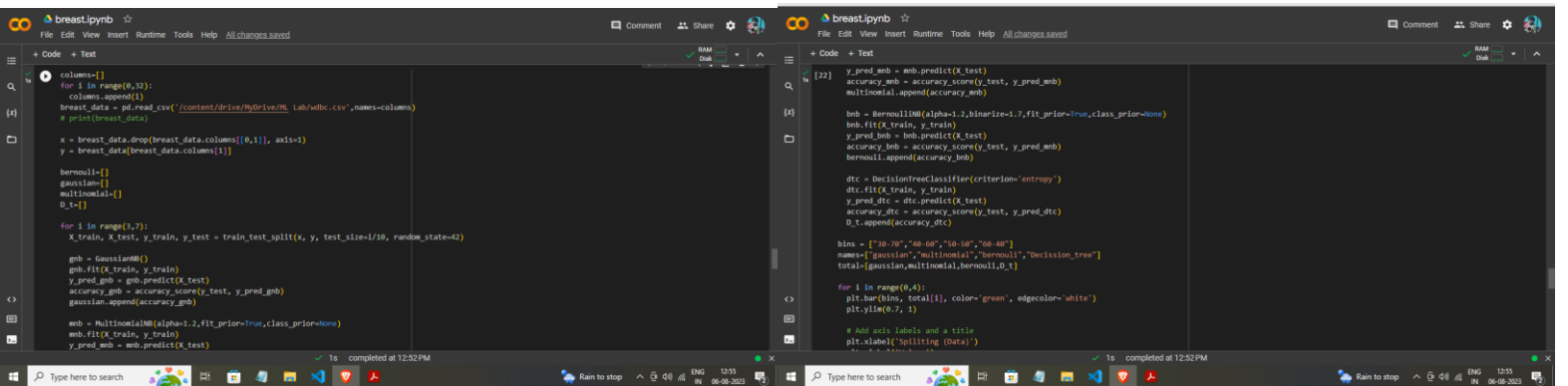
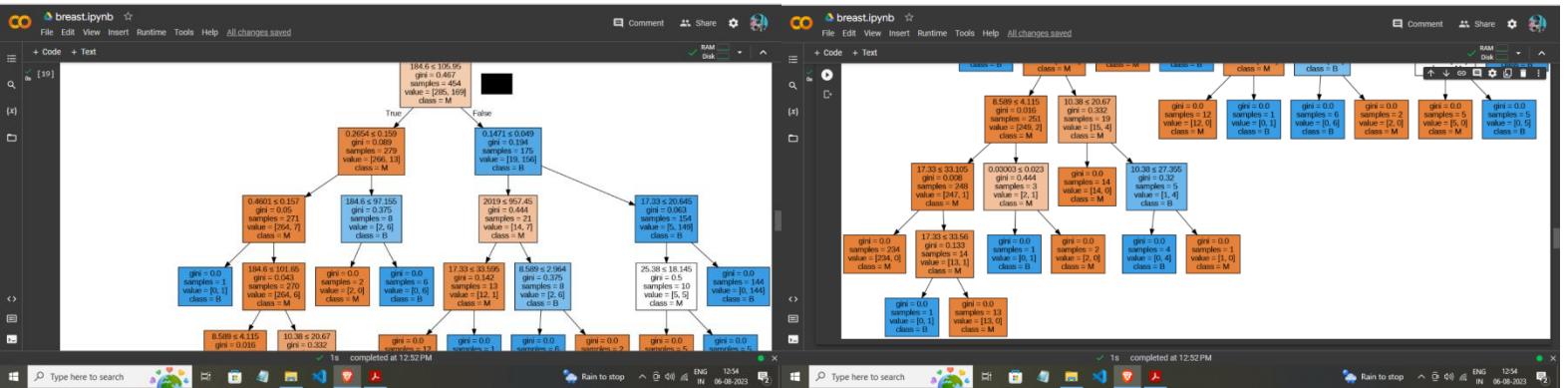
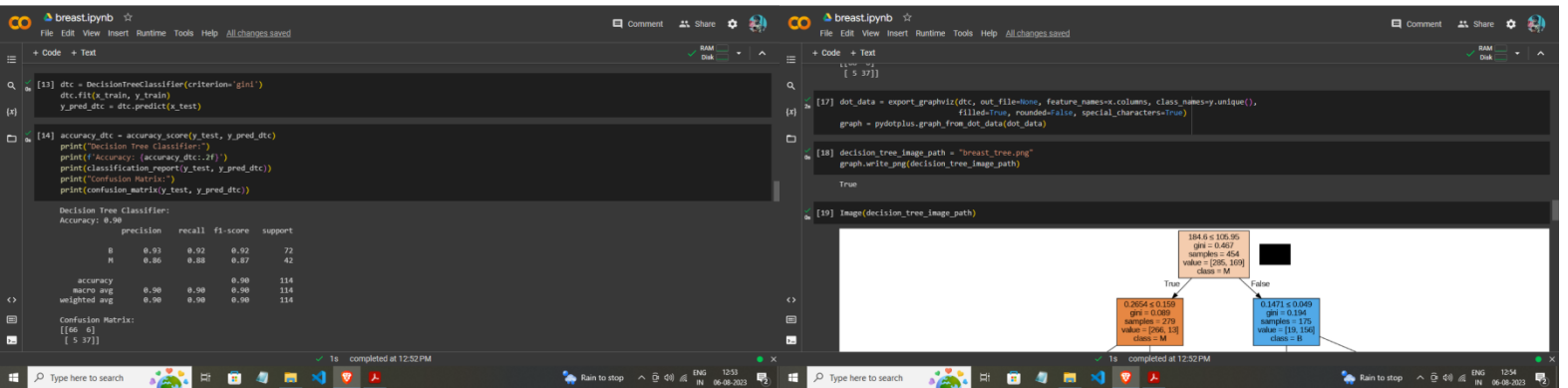
completed at 10:48 AM
```

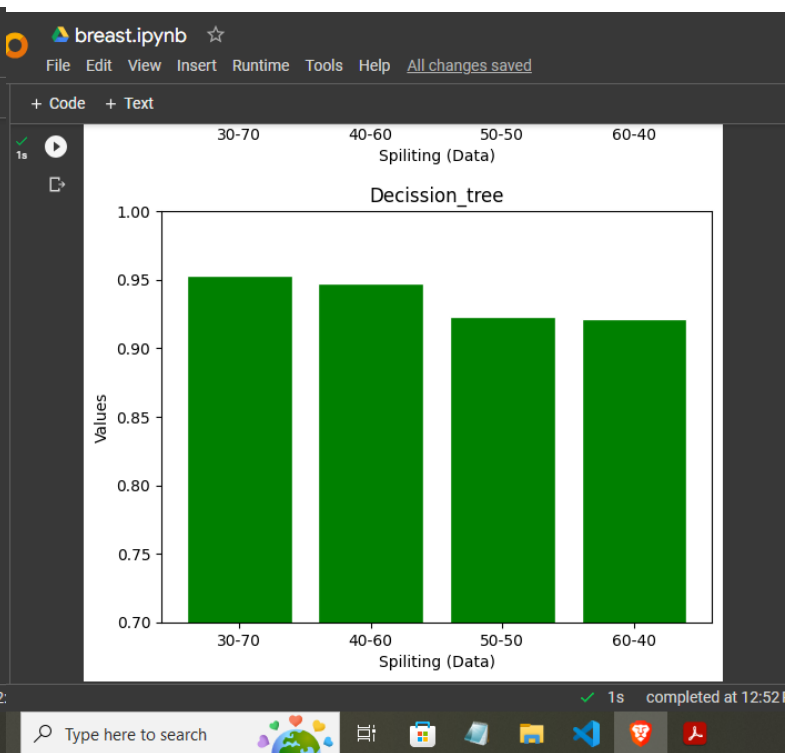
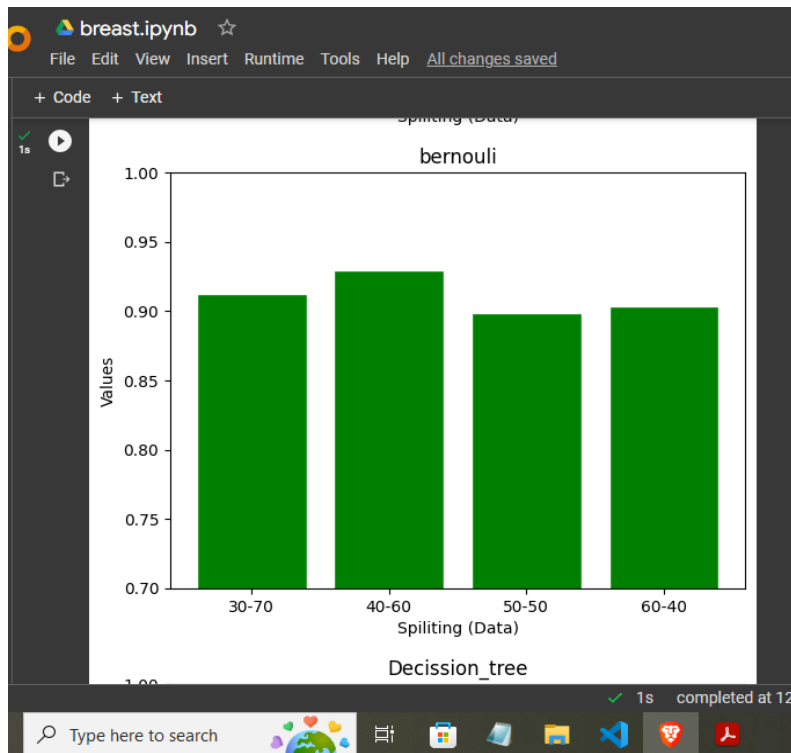




BREAST_CANCER : →







DIABETES : →

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
from sklearn.naive_bayes import GaussianNB, MultinomialNB, BernoulliNB
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier, export_text, export_graphviz
import pydotplus
from IPython.display import Image

# Load the diabetes dataset
diabetes_data = pd.read_csv('content/Drive/MyDrive/ML Lab/diabetes.csv')
print(diabetes_data)

x = diabetes_data.drop('Outcome', axis=1)
y = diabetes_data['Outcome']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(x, y, test_size=0.5, random_state=30)
```

completed at 4:55PM

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[25] # Decision Tree Classifier
dtc = DecisionTreeClassifier(criterion='entropy') # You can also use 'entropy' for criterion
dtc.fit(X_train, y_train)
y_pred_dtc = dtc.predict(X_test)

# Evaluate Decision Tree Classifier
accuracy_dtc = accuracy_score(y_test, y_pred_dtc)
print("Decision Tree Classifier:")
print(f"Accuracy: (accuracy_dtc:.2f)")
print(classification_report(y_test, y_pred_dtc))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_dtc))

dot_data = export_graphviz(dtc, out_file=None, feature_names=X.columns, class_names=str(y.unique()),
                           filled=True, rounded=True, special_characters=False)
graph = pydotplus.graph_from_dot_data(dot_data)

# Save the decision tree image
decision_tree_image_path = "decision_tree.png"
graph.write_png(decision_tree_image_path)

# Display the decision tree image
Image(decision_tree_image_path)
```

completed at 4:55PM

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[25] # Save the decision tree image
decision_tree_image_path = "decision_tree.png"
graph.write_png(decision_tree_image_path)

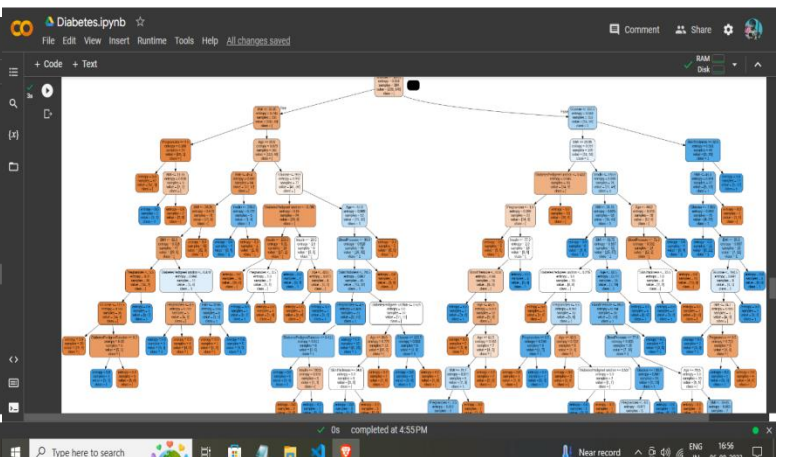
# Display the decision tree image
Image(decision_tree_image_path)
```

Decision Tree Classifier:
Accuracy: 0.74

	precision	recall	f1-score	support
0	0.82	0.81	0.81	262
1	0.60	0.01	0.60	122
accuracy			0.74	384
macro avg	0.71	0.71	0.71	384
weighted avg	0.75	0.74	0.75	384

Confusion Matrix:
[[212 50]
[48 74]]

completed at 4:55PM



Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[25]

```
# Gaussian Naive Bayes Classifier
gmb = GaussianNB()
gmb.fit(X_train, y_train)
y_pred_gmb = gmb.predict(X_test)

# Multinomial Naive Bayes Classifier
mnb = MultinomialNB(alpha=1.2, fit_prior=True, class_prior=None)
mnb.fit(X_train, y_train)
y_pred_mnb = mnb.predict(X_test)

# Bernoulli Naive Bayes Classifier
bnb = BernoulliNB(alpha=1.5, binarize=1.7, fit_prior=True, class_prior=None)
bnb.fit(X_train, y_train)
y_pred_bnb = bnb.predict(X_test)
```

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[26]

```
# Bernoulli Naive Bayes Classifier
bnb = BernoulliNB(alpha=1.5, binarize=1.7, fit_prior=True, class_prior=None)
bnb.fit(X_train, y_train)
y_pred_bnb = bnb.predict(X_test)

# Evaluate Gaussian Naive Bayes
accuracy_gmb = accuracy_score(y_test, y_pred_gmb)
print("Gaussian Naive Bayes:")
print("Accuracy: (accuracy_gmb:.2f)")
print(classification_report(y_test, y_pred_gmb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_gmb))

cm = confusion_matrix(y_test, y_pred_gmb)
sns.heatmap(cm, annot=True, fmt="d", cmap="coolwarm")
plt.show()

# Evaluate Multinomial Naive Bayes
accuracy_mnb = accuracy_score(y_test, y_pred_mnb)
print("Multinomial Naive Bayes:")
print("Accuracy: (accuracy_mnb:.2f)")
print(classification_report(y_test, y_pred_mnb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_mnb))
```

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[27]

```
cm1 = confusion_matrix(y_test, y_pred_mnb)
sns.heatmap(cm1, annot=True, fmt="d", cmap="coolwarm")
plt.show()

# Evaluate Bernoulli Naive Bayes
accuracy_bnb = accuracy_score(y_test, y_pred_bnb)
print("Bernoulli Naive Bayes:")
print("Accuracy: (accuracy_bnb:.2f)")
print(classification_report(y_test, y_pred_bnb))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred_bnb))

cm2 = confusion_matrix(y_test, y_pred_bnb)
sns.heatmap(cm2, annot=True, fmt="d", cmap="coolwarm")
plt.show()
```

Gaussian Naive Bayes:
Accuracy: 0.77

	precision	recall	f1-score	support
0	0.83	0.84	0.84	262
1	0.65	0.62	0.64	122
accuracy	0.74	0.73	0.74	384
macro avg	0.74	0.73	0.74	384
weighted avg	0.77	0.77	0.77	384

Confusion Matrix:
[[221 41]
[46 76]]

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[28]

```
cm3 = confusion_matrix(y_test, y_pred_bnb)
sns.heatmap(cm3, annot=True, fmt="d", cmap="coolwarm")
plt.show()
```

Bernoulli Naive Bayes:
Accuracy: 0.67

	precision	recall	f1-score	support
0	0.68	0.97	0.80	262
1	0.36	0.84	0.07	122
accuracy	0.52	0.58	0.67	384
macro avg	0.52	0.58	0.44	384
weighted avg	0.58	0.67	0.57	384

Confusion Matrix:
[[253 9]
[117 5]]

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[29]

```
cm4 = confusion_matrix(y_test, y_pred_gmb)
sns.heatmap(cm4, annot=True, fmt="d", cmap="coolwarm")
plt.show()
```

Gaussian Naive Bayes:
Accuracy: 0.77

	precision	recall	f1-score	support
0	0.83	0.84	0.84	262
1	0.65	0.62	0.64	122
accuracy	0.74	0.73	0.74	384
macro avg	0.74	0.73	0.74	384
weighted avg	0.77	0.77	0.77	384

Confusion Matrix:
[[221 41]
[46 76]]

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[30]

```
cm5 = confusion_matrix(y_test, y_pred_bnb)
sns.heatmap(cm5, annot=True, fmt="d", cmap="coolwarm")
plt.show()
```

Bernoulli Naive Bayes:
Accuracy: 0.67

	precision	recall	f1-score	support
0	0.68	0.97	0.80	262
1	0.36	0.84	0.07	122
accuracy	0.52	0.58	0.67	384
macro avg	0.52	0.58	0.44	384
weighted avg	0.58	0.67	0.57	384

Confusion Matrix:
[[253 9]
[117 5]]

Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[31]

```
cm6 = confusion_matrix(y_test, y_pred_mnb)
sns.heatmap(cm6, annot=True, fmt="d", cmap="coolwarm")
plt.show()
```

Multinomial Naive Bayes:
Accuracy: 0.62

	precision	recall	f1-score	support
0	0.73	0.69	0.71	262
1	0.41	0.47	0.44	122
accuracy	0.57	0.58	0.57	384
macro avg	0.63	0.62	0.62	384

Confusion Matrix:
[[180 82]
[65 57]]