```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd

# Load the dataset
# Using absolute path to ensure it works regardless of the current working directory
dataset = pd.read_csv(r"/Users/venkat/workspace/gitRepos/python-genAi-agenticAI/day-46-2026Feb2-ml-naive-bayes/logistic-classification.csv")

X = dataset.iloc[:, [2, 3]].values
y = dataset.iloc[:, -1].values

# Split the dataset into the training set and test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state = 0)

# Training the Random Forest Classification model on the Training set
from sklearn.ensemble import RandomForestClassifier
#classifier = RandomForestClassifier()
#classifier = RandomForestClassifier(random_state = 0)
classifier = RandomForestClassifier(n_estimators = 33, max_depth = 5, criterion = 'entropy', max_features = 2, random_state = 0)
classifier.fit(X_train, y_train)

y_pred = classifier.predict(X_test)

from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print("Confusion Matrix:", cm)
```

```
from sklearn.metrics import accuracy_score
ac = accuracy_score(y_test, y_pred)
print("Accuracy Score:", ac)

bias = classifier.score(X_train,y_train)
print("Bias:", bias)

variance = classifier.score(X_test,y_test)
print("Variance:", variance)
```