# Ex No: 5 IMPLEMENT NAIVE BAYES MODELS

**Program:**

#Import scikit-learn dataset library

from sklearn import datasets

#Load dataset

wine = datasets.load\_wine()

# print the names of the 13 features print("Features: ", wine.feature\_names)

# print the label type of wine(class\_0, class\_1, class\_2) print("Labels: ", wine.target\_names)

# print data(feature)shape wine.data.shape

# print the wine data features (top 5 records) print(wine.data[0:5])

# print the wine labels (0:Class\_0, 1:class\_2, 2:class\_2) Print(wine.target)

# Import train\_test\_split function

from sklearn.model\_selection

import train\_test\_split

# Split dataset into training set and test set

X\_train, X\_test, y\_train, y\_test = train\_test\_split (wine.data, wine.target, test\_size=0.3, random\_state=109)

# 70% training and 30% test

#Import Gaussian Naive Bayes model

from sklearn.naive\_bayes import GaussianNB

#Create a Gaussian Classifier

gnb = GaussianNB()

#Train the model using the training sets gnb.fit(X\_train, y\_train)

#Predict the response for test dataset y\_pred = gnb.predict(X\_test)

# Evaluating model

#Import scikit-learn metrics module for accuracy calculation from sklearn import metrics

# Model Accuracy print("Accuracy:",metrics.accuracy\_score(y\_test, y\_pred))

# Output:

**Display features and labels in the dataset:**

Features: ['alcohol', 'malic\_acid', 'ash', 'alcalinity\_of\_ash', 'magnesium', 'total\_phenols', 'flavanoids', 'nonflavanoid\_phenols', 'proanthocyanins', 'color\_intensity', 'hue', 'od280/od315\_of\_diluted\_wines', 'proline']

Labels: ['class\_0' 'class\_1' 'class\_2']

# Display the shape of the dataset:

(178, 13)

# Display the top 5 records in the dataset:

[[1.423e+01 1.710e+00 2.430e+00 1.560e+01 1.270e+02 2.800e+00 3.060e+00 2.800e-01 2.290e+00 5.640e+00 1.040e+00 3.920e+00 1.065e+03]

[1.320e+01 1.780e+00 2.140e+00 1.120e+01 1.000e+02 2.650e+00 2.760e+00 2.600e-01 1.280e+00 4.380e+00 1.050e+00 3.400e+00 1.050e+03] [1.316e+01 2.360e+00 2.670e+00 1.860e+01 1.010e+02 2.800e+00 3.240e+00 3.000e-01 2.810e+00 5.680e+00 1.030e+00 3.170e+00 1.185e+03] [1.437e+01 1.950e+00 2.500e+00 1.680e+01 1.130e+02 3.850e+00 3.490e+00 2.400e-01 2.180e+00 7.800e+00 8.600e-01 3.450e+00 1.480e+03]

[1.324e+01 2.590e+00 2.870e+00 2.100e+01 1.180e+02 2.800e+00 2.690e+00 3.900e-01 1.820e+00 4.320e+00 1.040e+00 2.930e+00 7.350e+02]]

# Display the labels in the dataset:

[0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2

2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2]

# Model Accuracy:

Accuracy: 0.9074074074074074