**Experiment No. – 1.3**

**Student Name: Milan Sharma UID: 23MAI10003**

**Branch: ME – CSE - AIML Section/Group: 23MAI – 1 (A)**

**Semester: 2nd Date of Performance: 07Feb2024**

**Subject Name: Machine Learning Lab Subject Code: 23 CSH 651**

Aim of the Experiment:

Develop an application for implementing the Naïve Bayes classifier.

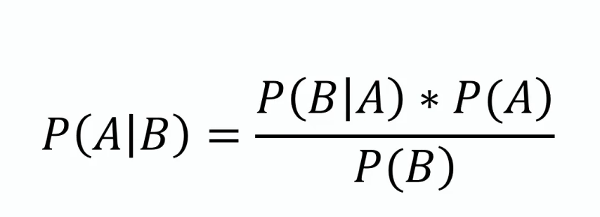
**Theory:**

Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in *text classification* that includes a high-dimensional training dataset.

Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object.

Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

Naïve Bayes is based on Bayes theorem :



Where,

P(A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P(A) is Prior Probability: Probability of hypothesis before observing the evidence.

P(B) is Marginal Probability: Probability of Evidence.

**Code:**

# Importing the libraries

import pandas as pd

# Read the dataset

dataset=pd.read\_csv("Downloads/Iris.csv")

print(dataset.head())

print(dataset.tail())

# Categorical to Numeric

dataset['Species'].replace(['Iris-setosa', 'Iris-versicolor', 'Iris-virginica'],[0, 1, 2],inplace=True)

# Assign Independent and Dependent Variables

x = dataset.iloc[:, [1,2,3,4]].values

y = dataset.iloc[:, 5].values

# Splitting 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 = 0.3, random\_state = 42)

# Fitting Naive Bayes to the Training set

from sklearn.naive\_bayes import GaussianNB

classifier = GaussianNB()

classifier.fit(x\_train, y\_train)

# Predicting the Test set results

y\_pred = classifier.predict(x\_test)

from sklearn.metrics import accuracy\_score

from sklearn.metrics import f1\_score

from sklearn.metrics import confusion\_matrix

# Calculate accuracy of the model

accuracy = accuracy\_score(y\_test, y\_pred)

f1 = f1\_score(y\_pred, y\_test, average="weighted")

cm = confusion\_matrix(y\_test, y\_pred)

print("Actual Values:")

print(y\_test)

print("\n")

print("Predicted Values:")

print(y\_pred)

print("\nNAIVE BAYES CLASSIFIER:")

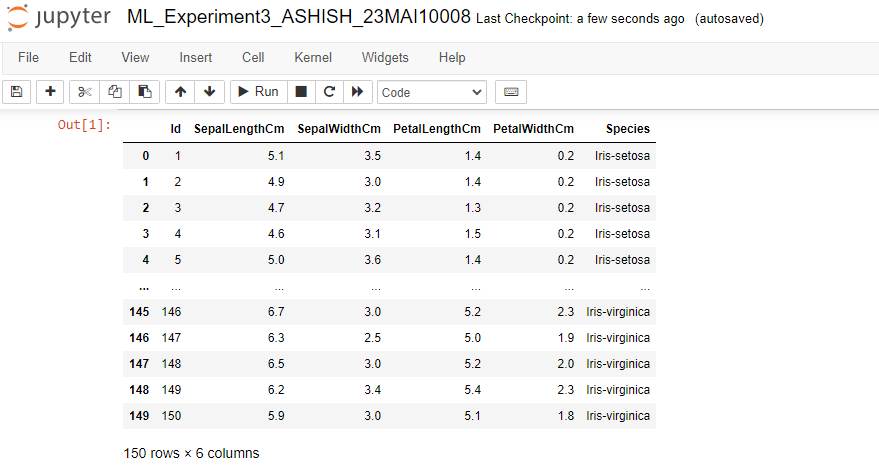
print("\tAccuracy:", accuracy)

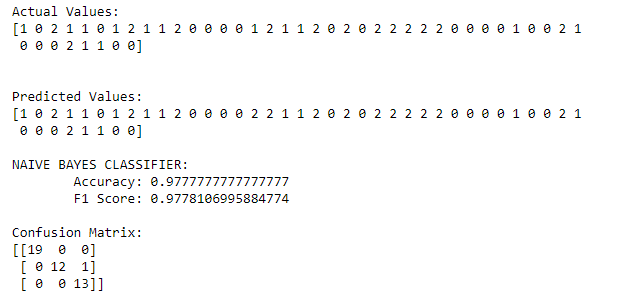
print("\tF1 Score:", f1, "\n")

print("Confusion Matrix:")

print(cm)

Output:





**Learning Outcomes:**

1. **I learnt about various python libraries like pandas, sklearn.**
2. **I learnt about the concept of Naive Bayes Classifier.**
3. **I learnt about how to read the dataset using pandas.**
4. **I learnt about the training and testing dataset for building the model.**