



UNIVERSITY OF INFORMATION  
TECHNOLOGY AND SCIENCES (UITS)  
DEPARTMENT OF INFORMATION TECHNOLOGY

LAB REPORT No.: 3

IT-452 : MACHINE LEARNING

---

## Naive Bayes Implementation

---

*Submitted To:*

Sumaiya Akhtar Mitu  
Lecturer,  
Department of IT ,  
UITS  
Email:sumaiya.akhtar@uits.edu.bd

*Submitted By:*

Name:Nazmul Zaman  
Student ID:2014755055  
Department of IT, UITS

## Contents

1	Abstract	2
2	Working Procedure	2
3	Advantages of Naïve Bayes Classifier:	2
4	Disadvantages of Naïve Bayes Classifier:	2
5	Applications of Naïve Bayes Classifier:	2
6	Formula	3
7	Code	3
8	Conclusion	5
9	References	5

## 1 Abstract

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.

## 2 Working Procedure

Convert the given dataset into frequency tables.

Generate Likelihood table by finding the probabilities of given features.

Now, use Bayes theorem to calculate the posterior probability

## 3 Advantages of Naïve Bayes Classifier:

Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.

It can be used for Binary as well as Multi-class Classifications.

It performs well in Multi-class predictions as compared to the other Algorithms.

It is the most popular choice for text classification problems.

## 4 Disadvantages of Naïve Bayes Classifier:

Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.

## 5 Applications of Naïve Bayes Classifier:

It is used for Credit Scoring.

It is used in medical data classification.

It can be used in real-time predictions because Naïve Bayes Classifier is an eager learner.

It is used in Text classification such as Spam filtering and Sentiment analysis

## 6 Formula

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Figure 1: Code for : x86 Code

## 7 Code

```

1  # -*- coding: utf-8 -*-
2  """lab3.ipynb
3
4  Automatically generated by Colaboratory.
5
6  Original file is located at
7      https://colab.research.google.com/drive/1
8      V_ljga8rq8Boe0bKa9Fy1tjg6l2xWQAd
9  """
10 import pandas as pd
11 # Load the dataset
12 url = "https://raw.githubusercontent.com/milaan9/
13       Python_Decision_Tree_and_Random_Forest/main/dataset/
14       playgolf_data.csv"
15 df = pd.read_csv(url)
16
17 # Display the first few rows of the dataset
18 print(df.head())
19
20 from sklearn.preprocessing import LabelEncoder
21
22 # Encode categorical variables
23 categorical_cols = ['Outlook', 'Temperature', 'Humidity', 'Wind
24                    ','']
25 encoder = LabelEncoder()
26
27 for col in categorical_cols:
28     df[col] = encoder.fit_transform(df[col])
29
30 # Separate features and target variable
31 X = df.drop('PlayGolf', axis=1)
32 y = df['PlayGolf']

```

```
31 from sklearn.model_selection import train_test_split
32
33 # Split the dataset into training and testing sets
34 X_train, X_test, y_train, y_test = train_test_split(X, y,
35                                                    test_size=0.2, random_state=42)
36
37 from sklearn.naive_bayes import GaussianNB
38 from sklearn.metrics import accuracy_score
39
40 # Create and fit the Gaussian Naive Bayes model
41 model = GaussianNB()
42 model.fit(X_train, y_train)
43
44 # Make predictions on the test set
45 y_pred = model.predict(X_test)
46
47 # Calculate the accuracy of the model
48 accuracy = accuracy_score(y_test, y_pred)
49 print("Accuracy:", accuracy)
50
51 # Define a new data point for prediction
52 new_data = [[2, 1, 0, 1]] # Example data: Outlook=Sunny,
53                            # Temperature=Cool, Humidity=High, Windy=True
54
55 # Make the prediction
56 predicted_class = model.predict(new_data)
57
58 # Map the predicted class to 'Yes' or 'No'
59 prediction = 'Yes' if predicted_class[0] == 1 else 'No'
60
61 # Print the prediction
62 print("Prediction: The person will", prediction, "PlayGolf.")
```

	Outlook	Temperature	Humidity	Wind	PlayGolf
0	Sunny	Hot	High	Weak	No
1	Sunny	Hot	High	Strong	No
2	Overcast	Hot	High	Weak	Yes
3	Rainy	Mild	High	Weak	Yes
4	Rainy	Cool	Normal	Weak	Yes

Accuracy: 1.0

```
Prediction: The person will No PlayGolf.
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does not have valid feature names, but GaussianNB was fitted with feature names
warnings.warn(
```

Figure 2: OUTPUT

## 8 Conclusion

The Naïve Bayes classifier is a supervised machine learning algorithm, which is used for classification tasks, like text classification. It is also part of a family of generative learning algorithms, meaning that it seeks to model the distribution of inputs of a given class or category.

It is based on the Bayes' Theorem for calculating probabilities and conditional probabilities. You can use it for real-time and multi-class predictions, text classifications, spam filtering, sentiment analysis, and a lot more

## 9 References

2. <https://www.javatpoint.com/linear-regression-in-machine-learning>
3. <https://www.javatpoint.com/machine-learning-polynomial-regression>