

DEPARTMENT OF COMPUTER ENGINEERING

Instruction No. 01 and 02
LP-III/ML / Sr. No.01 and 02
Rev 00 Date: 27/12/17

Title: Assignment on Naive Bayes Classification Algorithm

Aim: Implement Naive Bayes Classification Algorithm on a given dataset.

Objectives: 1. Study the concept of Naive Bayes Algorithm
2. Implement the Naive Bayes Algorithm on Iris.csv dataset.

Input: Iris.csv

Theory:

Naive Bayes Classification Algorithm

Naive 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.

Naive 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 Naive Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

1. Bayes' Theorem:

Bayes' theorem is also known as **Bayes' Rule** or **Bayes' law**, which is used to determine the probability of a hypothesis with prior knowledge. It depends on the conditional probability. The formula for Bayes' theorem is given as:

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

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.

Algorithm:

- Data Preprocessing step
- Fitting Naive Bayes to the Training set
- Predicting the test result
- Test accuracy of the result(Creation of Confusion matrix)
- Visualizing the test set result.

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.

Disadvantages of Naïve Bayes Classifier:

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

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.

Output:

Correct predictions: 28

False Predictions: 2

Accuracy of the Naive Bayes Classification is: 0.9333333333333333

Conclusion:

Studied about Naive Bayes Classification and implemented it on Iris.csv dataset for classification.

Prepared by:

Dr.T.Bhaskar
Subject Teacher

Approved by:

Dr. D.B. Kshirsagar
Head, Computer Engineering