G H Raisoni College of Engineering
SY AI Semester-IV AY 2023-24 Division-A

UCAIP210: Machine Learning Algorithms Practicals Lab Manual

Practical Teacher : Dr Monika Y. Dangore

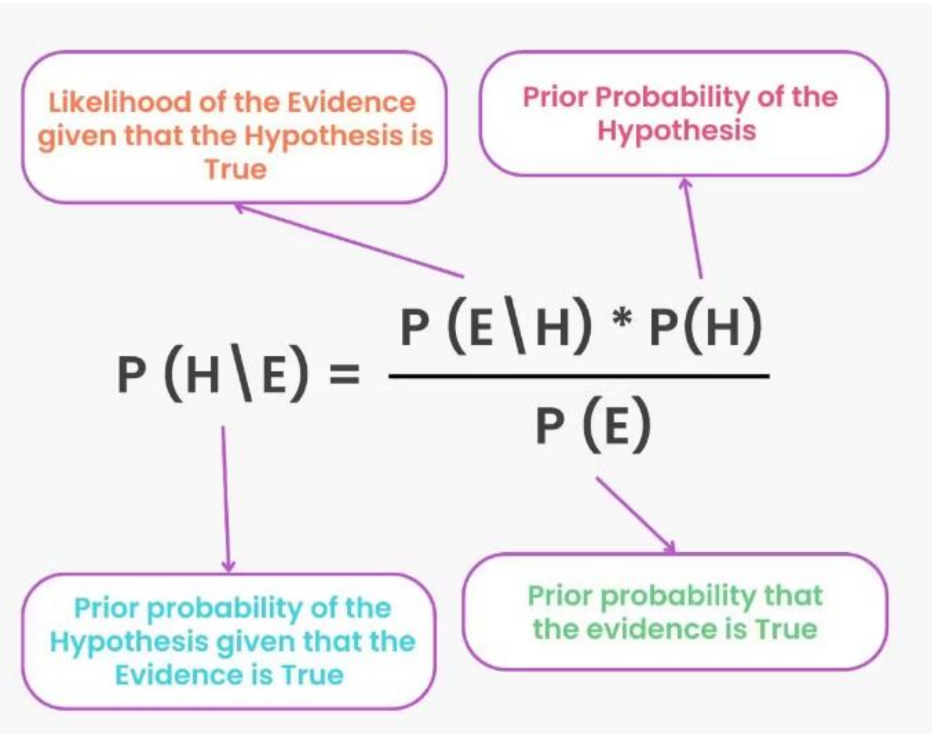
Experiment No: 2

Aim:

Implement Naïve Bayes Classifier on the dataset of your choice. Test and compare for accuracy and precision.

Introduction to Naïve Bayes Classifier:

The Naïve Bayes classifier is a popular supervised machine learning algorithm used for classification tasks such as text classification. It uses principles of probability to perform classification tasks.



It is an algorithm that learns the probability of every object, its features, and which groups they belong to. It is also known as a probabilistic classifier. The Naive Bayes Algorithm comes under supervised learning and is mainly used to solve classification problems.

For example, you cannot identify a bird based on its features and color as there are many birds with similar attributes. But, you make a probabilistic prediction about the same, and that is where the Naive Bayes Algorithm comes in.

Probability, Bayes Theory, and Conditional Probability

Probability is the base for the Naive Bayes algorithm. This algorithm is built based on the probability results that it can offer for unsolvable problems with the help of prediction.

Probability

Probability helps to predict an event's occurrence out of all the potential outcomes. The mathematical equation for probability is as follows:

```
Probability of an event = Number of Favorable Events/ Total number of outcomes
```

0 < = probability of an event < = 1. The favorable outcome denotes the event that results from the probability. Probability is always between 0 and 1, where 0 means no probability of it happening, and 1 means the success rate of that event is likely.

Bayes Theory

Bayes Theory works on coming to a hypothesis (H) from a given set of evidence (E). It relates to two things: the probability of the hypothesis before the evidence P(H) and the probability after the evidence P(H|E). The Bayes Theory is explained by the following equation:

```
P(H|E) = (P(E|H} * P(H))/P(E)
```

In the above equation,

- P(H|E) denotes how event H happens when event E takes place.
- P(E|H) represents how often event E happens when event H takes place first.
- P(H) represents the probability of event X happening on its own.
- P(E) represents the probability of event Y happening on its own.

The Bayes Rule is a method for determining P(H|E) from P(E|H). In short, it provides you with a way of calculating the probability of a hypothesis with the provided evidence.

Conditional Probability

Conditional probability is a subset of probability. It reduces the probability of becoming dependent on a single event. You can compute the conditional probability for two or more occurrences.

When you take events X and Y, the conditional probability of event Y is defined as the probability that the event occurs when event X is already over. It is written as P(Y|X). The mathematical formula for this is as follows:

```
P(Y|A) = P(X and Y) /P(X)
```

Bayesian Probability

Bayesian Probability allows to calculate the conditional probabilities. It enables to use of partial knowledge for calculating the probability of the occurrence of a specific event. This algorithm is used for developing models for prediction and classification problems like Naive Bayes.

The Bayesian Rule is used in probability theory for computing - conditional probabilities. What is important is that you cannot discover just how the evidence will impact the probability of an event occurring, but you can find the exact probability.

Types of Naïve Bayes Model:

Gaussian Naive Bayes

It is a straightforward algorithm used when the attributes are continuous. The attributes present in the data should follow the rule of Gaussian distribution or normal distribution. It remarkably quickens the search, and under lenient conditions, the error will be two times greater than Optimal Naive Bayes.

Optimal Naive Bayes

Optimal Naive Bayes selects the class that has the greatest posterior probability of happenings. As per the name, it is optimal. But it will go through all the possibilities, which is very slow and time-consuming.

Bernoulli Naive Bayes

Bernoulli Naive Bayes is an algorithm that is useful for data that has binary or boolean attributes. The attributes will have a value of yes or no, useful or not, granted or rejected, etc.

Multinominal Naive Bayes

Multinominal Naive Bayes is used on documentation classification issues. The features needed for this type are the frequency of the words converted from the document.

Precision Vs Accuracy:

In machine learning, precision and accuracy are two important metrics used to evaluate the performance of a model.

Precision refers to the ratio of correctly predicted positive observations to the total predicted positive observations. In other words, it measures the proportion of true positives (correctly predicted positive cases) out of all positive predictions made by the model.

Accuracy, on the other hand, measures the overall correctness of the model. It refers to the ratio of the total number of correct predictions to the total number of predictions made by the model, irrespective of whether they are positive or negative.

In simpler terms, precision measures how well a model can identify positive cases correctly, while accuracy measures how well a model can identify both positive and negative cases correctly.

For example, if a model predicts that there are 100 positive cases out of which 80 are actually positive, then the precision would be 80%, and if the model predicts a total of 200 cases out of which 180 are actually correct, then the accuracy would be 90%.

The micro average precision is the sum of true positives for a single class divided by the sum of predicted positives for all classes.

It is important to note that precision and accuracy are not always the best metrics to use, as they may not accurately reflect the model's performance in certain cases. Other metrics, such as recall, F1 score, and AUC-ROC, may also be used to evaluate a model's performance depending on the specific use case.

Program:

Attach the printouts of the program.

Result:

The concept of Naïve Bayes Classifier is studied and program is executed successfully.

To be skipped from Write-up:

Naïve Bayes Solved Example:

https://www.youtube.com/watch?v=XzSlEA4ck2I

Day	*Outlook*	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

$$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$$

D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

NAIVE BAYES CLASSIFIER

Example - 1

$$P(PlayTennis = yes) = 9/14 = .64$$

$$P(PlayTennis = no) = 5/14 = .36$$

Outlook	Y	N
sunny	2/9	3/5
overcast	4/9	0
rain	3/9	2/5
Tempreature		
hot	2/9	2/5
mild	4/9	2/5
cool	3/9	1/5

Humidity	Y	N
high	3/9	4/5
normal	6/9	1/5
Windy		
Strong	3/9	3/5
Weak	6/9	2/5

$$\langle Outlook = sunny, Temperature = cool, Humidity = high, Wind = strong \rangle$$

$$v_{NB} = \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \prod_i P(a_i | v_j)$$

$$= \underset{v_j \in \{yes, no\}}{\operatorname{argmax}} P(v_j) \quad P(Outlook = sunny | v_j) P(Temperature = cool | v_j)$$

$$\cdot P(Humidity = high | v_j) P(Wind = strong | v_j)$$

$$v_{NB}(yes) = P(yes)\ P(sunny|yes)\ P(cool|yes)\ P(high|yes)\ P(strong|yes) = .0053$$

$$v_{NB}(no) = P(no)\ P(sunny|no)\ P(cool|no)\ P(high|no)\ P(strong|no) = .0206$$

$$v_{NB}(yes) = \frac{v_{NB}(yes)}{v_{NB}(yes) + v_{NB}(no)} = 0.205 \qquad v_{NB}(no) = \frac{v_{NB}(no)}{v_{NB}(yes) + v_{NB}(no)} = 0.795$$