**Naïve Bayes Classifier Implementation**

**Table of Contents**

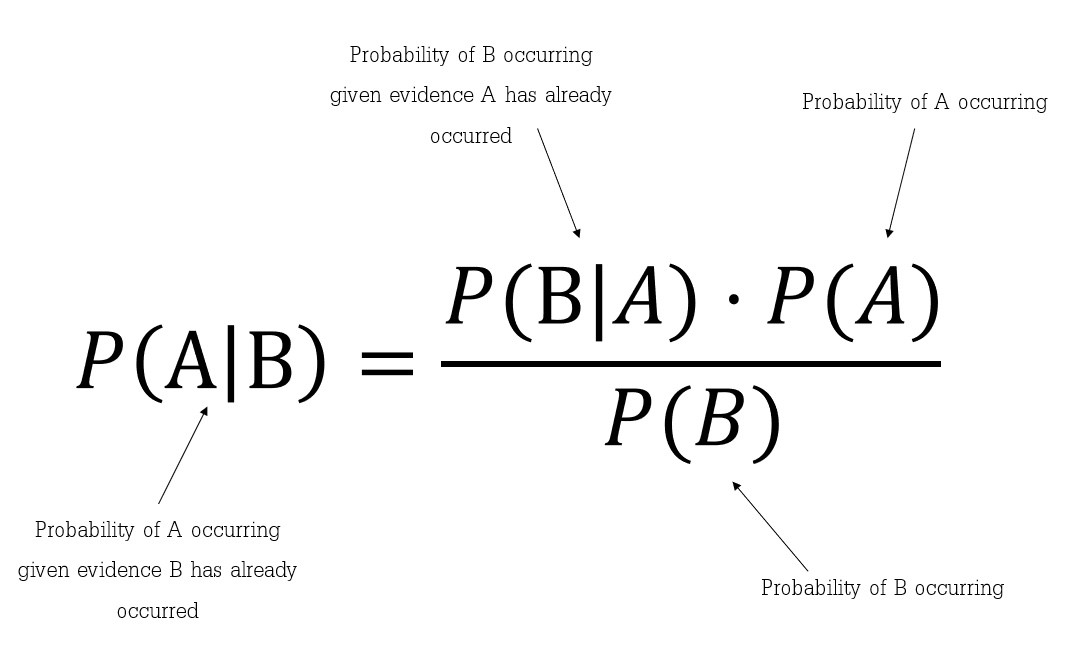
1. Introduction to Naïve Bayes Algorithm
2. Types of Naïve Bayes Algorithm
3. Application of Naïve Bayes
4. Model Training & Predicting Results
5. Result And Conclusion
6. Introduction to Naïve Bayes Algorithm

In machine learning, Naïve Bayes is a powerful algorithm for performing classification tasks on datasets.

This classification algorithm is based on Bayes theorem with strong independence assumption among the features available in the dataset. Naïve Bayes classification produces better results when we use it for textual data analysis.

Naïve Bayes classifier applies the Bayes theorem in practice. This classifier brings the power of Bayes theorem to Machine learning.

In practical, this classifier uses the Bayes theorem to predict availability probabilities of a given record such as if this record belongs to a particular class or not. The class with the highest probability is considered as the most likely class.



Naïve Bayes Classifier assumes that all the features are unrelated to each other. Presence or absence of a feature does not influence the presence or absence of any other feature.

In real world datasets, we test a hypothesis given multiple evidence on features. So, the calculations become quite complicated. To simplify the calculations, the feature independence approach is used to uncouple multiple evidence and treat each as an independent one.

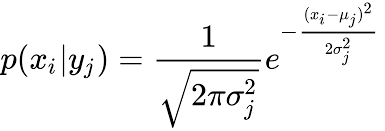
1. Types of Naïve Bayes

There are three types of Naïve Bayes Algorithm :-

* 1. Gaussian Naïve Bayes
  2. Multinomial Naïve Bayes
  3. Bernoulli Naïve Bayes

Gaussian Naïve Bayes Algorithm

When we have continuous attribute values, we made as assumption that values associated with each class are distributed according to Gaussian or Normal distribution. For eg. Suppose the training data contains a continuous attribute x. first, we segment the data by the class and then compute the mean and variance of x in each class. Let µi(mu i)be the mean of the value and let the σi (sigma i) be the variance of the values associated with the ith class. Suppose we have some observation value xi. Then, the probability distribution of xi given a class can be computed by the following equation: -



Multinomial Naïve Bayes Algorithm

With a multinomial Naïve Bayes model, feature vectors (samples) represents the frequencies with which certain events have been generated by a multinomial (p1,p2…..pn)where pi is probability that event i occurs. Multinomial Naïve bayes algorithm is preferred to use on data that is multinomially distributed. It is one of the standard algorithms which is used in text categorization classification.

Bernoulli Naïve Bayes

In the multivariate Bernoulli event model, features are independent boolean variables (binary variables) describing inputs. Just like the multinomial model, this model is also popular for document classification tasks where binary term occurrence features are used rather than term frequencies.

3. Application of Naïve Bayes Algorithm

Naïve Bayes is one of the most straightforward and fast classification algorithm. It is very well suited for large volume of data. It is successfully used in various applications such as :-

1. Spam filtering
2. Text classification
3. Sentiment analysis
4. Recommender systems

It uses Bayes theorem of probability for prediction of unknown class.

4. Model Training & Predicting Results

For for training & testing purpose we split dataset in a ratio of 70:30

Where, 70% data has been used for training

30% data for testing

For **Gaussian Naïve Bayes:**

1. **Classification Report:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | support |
| 0 [<=50K] | 0.89 | 0.78 | 0.83 | 4978 |
| 1 [>50K] | 0.50 | 0.70 | 0.58 | 1535 |
| Accuracy |  |  | 0.76 | 6513 |
| Macro avg. | 0.70 | 0.74 | 0.71 | 6513 |
| Weighted avg. | 0.80 | 0.76 | 0.74 | 6513 |

1. **Confusion Matrix:**

[[3899 1079]

[462 1073]]

1. **Accuracy:**

Accuracy is: 0.76

For **Multinomial Naïve Bayes:**

1. **Classification Report:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | support |
| 0 [<=50K] | 0.85 | 0.73 | 0.78 | 4978 |
| 1 [>50K] | 0.40 | 0.59 | 0.48 | 1535 |
| Accuracy |  |  | 0.69 | 6513 |
| Macro avg. | 0.63 | 0.66 | 0.63 | 6513 |
| Weighted avg. | 0.75 | 0.69 | 0.71 | 6513 |

1. **Confusion Matrix:**

[[3616 1362]

[625 910]]

True Positives(TP) = 3616

True Negatives(TN) = 910

False Positives(FP) = 1362

False Negatives(FN) = 625

The confusion matrix shows 3616 + 910 = 4526 correct predictions and 1362 + 625 = 1987 incorrect predictions.

In this case, we have

1. **Accuracy:**

Accuracy is: 0.69

For **Bernoulli Naïve Bayes:**

1. **Classification Report:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Precision | Recall | F1-score | support |
| 0 [<=50K] | 0.90 | 0.73 | 0.81 | 4978 |
| 1 [>50K] | 0.46 | 0.74 | 0.57 | 1535 |
| Accuracy |  |  | 0.73 | 6513 |
| Macro avg. | 0.67 | 0.74 | 0.69 | 6513 |
| Weighted avg. | 0.80 | 0.73 | 0.75 | 6513 |

1. **Confusion Matrix:**

[[3652 1326]

[402 1133]]

1. **Accuracy:**

Accuracy is: 0.73

5. Result & Conclusion

1. In this project, I build a Gaussian Naïve Bayes Classifier model to predict whether a person makes over 50K a year. The model yields good performance as indicated by the model accuracy which was found to be 0.76
2. I have compared the model accuracy with different types Naïve Bayes algorithms. Like:
3. For Gaussian Naïve Bayes Accuracy is 76%
4. For Bernoulli Naïve Bayes Accuracy is 73%
5. For Multinomial Naïve Bayes Accuracy is 69%