

# Assignment 8

## Naive Bayes Classifier

- Aim: On the given dataset, perform the classification using Naive Bayes and performance measurement using accuracy, TPR, FPR, TNRR. Using weka Java API.

- Theory:

- Naive Bayes Algorithm.

It is a classification technique based on Bayes theorem with an assumption that independent among predictors. In simple terms a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

- Naive Bayes Model is easy to build and particularly useful for very large datasets. Along with simplicity, it outperforms many other sophisticated algorithms.



Naive Bayes model is easy to build and particularly useful for very large datasets along with sm

Bayes algorithm theorem provide a way to calculate posterior probability  $P(C|x)$  from  $P(C)$ ,  $P(x|C)$  and  $P(x)$

$$P(C|x) = \frac{\overset{\text{Likelihood}}{\downarrow} P(x|C) \overset{\text{prior probability}}{\downarrow} P(C)}{\underset{\text{Predictor Prior probability}}{\downarrow} P(x)}$$

class probability

### • Advantage.

→ It is easy & fast to predict class data set. It also performs well in multiclass prediction.

→ It performs well in case of categorical input variable compared to numerical variables.



## Disadvantages

- If categorical variables has a category which was not observed during training the model will assign 0 probability and will be unable to make a prediction.

## Application

- Real time Prediction
- Multiclass Prediction
- Recommendation System

## Conclusion

In this assignment we have implemented and understood the concept of Naive Bayes algorithm.