International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 06 | June - 2022 Impact Factor: 7.185 ISSN: 2582-3930

DATA MINING USING NAIVE BAYES IN E-COMMERCE

NANDA KUMAR K, ALAMMA B H

MCA Department at Dayananda Sagar College of Engineering ,

MCA Department at Dayananda Sagar College of Engineering

----------***----------

ABSTRACT

Data mining is extraction of information, data mining can be used to find solution to complex problems. Naïve Bayes theorem calculates the each possible hypothesis and outputs the most probable one. In this paper data mining techniques that can be applied in various computer system. Data mining is the extracting required data from large set of rawdata. Data mining is also used in multimediaAlgorithms and statistical methods are also used to find patterns in data . Marketing, risk management, fraud detection, cyber security, and medical diagnostics all use data mining techniques.

KEY WORDS: Data mining, naive bayes, classification, e-commerce.

1.INTRODUCTION

Data mining is extraction of information, data mining can be used to find solution to complex problems [1]. But it requires some experts to understand processed

information even the experts face some difficulty in understanding the knowledge of data mining [2]. In order to analyse or understand we include data mining in various computing system. Consider multimedia mining, which is divided into five data types: image data, audio data, video data, text data, and digital ink. Image data is pictures or photos [3].

Audio data is songs, music, mp3 and video data is videos from mobile or camera, text data is data used in messages, Digital ink is represented in 2d or 3d video or light pen.

Data mining algorithms are used for analysing data and developing data models to identify patterns some of the algorithms arere c4.5,k-means and naïve bayes.

We can learn this kind of algorithms with the help of probability theory, logic, control theory, principle of statistics.

2. CLASSIFICATION

1.C4.5 Algorithm

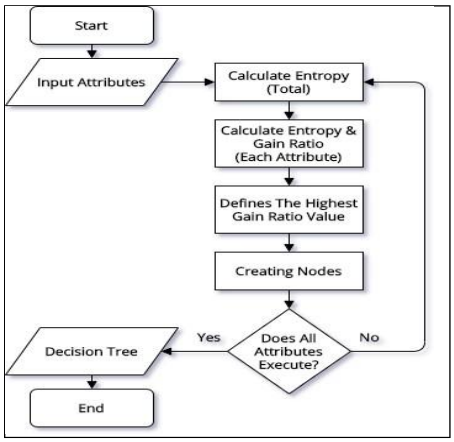

Fig 1: c4.5 flowchart

It was developed by Ros Quinlan. This approach is a Quinlan Id3 extension that is used to create a decision tree.. The main use of this algorithm is classification. It is also referred to statistical classifier. It creates decision trees from a collection of training data using the concept of information entropy.Predicting the information about an occurrence is what information entropy is all about. It can handle incomplete data as well.

2.K-means algorithm

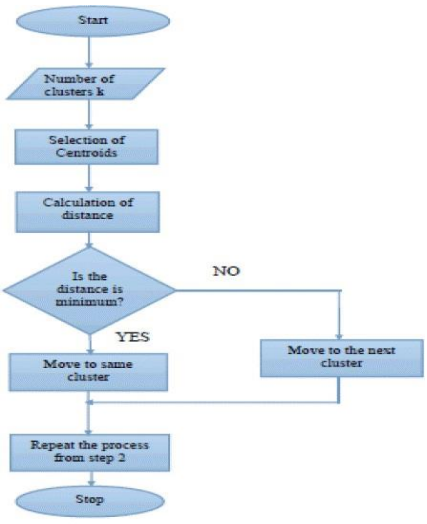

FIG. 1. **Flowchart for K-means clustering.**

Fig 2: k-means flowchart

In k-means k is unsupervised learning algorithm, In given a data set of items, the algorithm will categorize the items into k groups.

3.Naive Bayes algorithm

Naïve Bayes algorithm is combination of algorithm. The Bayes theorem is used. It's a learning algorithm that's supervised. The Bayes theorem assesses all potential hypotheses and returns the most likely.

3.IMPLEMENTATION

1.E-commerce

Our Shopping portal has consists of so many stores near our location. The customer can able to buy the products from our application and they can add the different products from different stores. Once the product added in the cart we are analysis the data with the various stores and trying to find the products which will be available for low cost. It will compare various stores product price for each product and it will analysis the product price and it will give the result where we are getting all the product with low cost. We are using sentimental analysis to find the low cost for the products added in the cart.We are implementing Navie Bayes Classifiers to analyze our data it will check each and every product in the dataset and used to identify the product which we are getting for the low cost. This algorithm will provide the accurate result to predict the low cost grocery. Our Main objective of the project is to save the customer money and make their purchase should be easy and simple.

4.PROPOSED SYSTEM

Most of the E-commerce application will be available to buy the product from the different store. In all the application we can get the product from the different stores and all the application is optimized for the customer and everybody wants to get everything Online. But the difficulty is where we will get all the products with low cost. Because all the people will prefer to buy the products with low cost. Our Proposed system will overcome those all the disadvantage and it will produce the accurate result to identify the product with low cost.

5.CONCLUSION

We have discussed several data mining techniques which can be helpful for extracting information. The data mining technique such as c4.5,k-means,naïve bayes can be used.

For mining information from data which later can be used for Various applications. So we use naïve bayes theorem in e-commerce. E-commerce plays an important role in real world so using data mining in e-commerce is very useful for marketing, business improvement, feedbacks, quality assurance, availability of stocks, price inflation. It is also helpful development in various e-commerce fields and also in decision making, helps in making .organisation more profitable. Several other factors can be determined using data mining.

REFERENCES

1.A study of data mining methods for intelligent computing systems by S. Iwin Thanakumar Joseph of Coimbatore, India.

2. A research study on data mining approaches in the pharmaceutical business. Nithya Jojen is a student at St. Joseph's College in Irinjalakuda.

3.A overview of data mining and machine learning approaches utilised in multimedia systems.