ISSN 2395-1621

A Business Analysis Application for Admission Process Using Data mining Algorithm

#1Karishma Tanpure, #2Pooja Jadhav, #3Ramesh Pawar, #4Durvankur Potnis

1karishmaptanpure1995@gmail.com
2jadhavpooja103@gmail.com
3rameshpawar1995@gmail.com
4durvankurp@gmail.com

#1234Department of Computer Engineering,

Zeal Education Society's
Dnyanganga College of Engineering and Research.

ABSTRACT

Business Intelligence is a broad category of applications which is analyzing business data to make better business decisions. The aim of this application is to reduce the paper work of human. In this paper, analysis of business for admission process to reduce system work and make better business decision for their improvement of business. In this application analyzing the student database and to observe the frequent patterns. The approach used for the naïve Bayes classification of data mining is to mine the well predicted output from the preprocess database and non-preprocessed database recursively. In this paper, an implementation of database analysis of student with the help of Naïve Bayes algorithm and this analysis is making with web application.

Keywords - Conditional pattern bases, preprocessing, classifier, prediction, educational data mining.

ARTICLE INFO

Article History

Received: 17th April 2017

Received in revised form : 17th April 2017

Accepted: 20th April 2017

Published online : 20th April 2017

I. INTRODUCTION

Data mining is the process of analyzing or extracting the patterns from different databases and summarizing it into useful information[3]. Some data mining algorithms are used to take out Predicted outputs such as naïve Bayes algorithm, k means algorithm. For business analysis this algorithms are very useful[6]. Data is the most important thing for any organization which is further processed to produce useful information. Data mining techniques are widely used for business to generate the useful pattern helpful for earning more profits and expand business. Since from many years, lots of research have been done by applying data mining algorithms on educational data for improvement in Education System. Data Mining can be useful for predicting the students' admissions. It is very time consuming and tedious task of processing these large volume of data[8]. In this we are implementing web application that includes student database with the help of data mining. The Web application which is going to be developed will be helpful for the college authorities as well as students[6]. This web application stores the records of previous year students admission in the engineering colleges database. The naïve Bayes algorithm is applying on student database which will gives the accurate prediction according to database. This data is again mined to maintain the records of district wise admissions, fees details. Range of marks, location, branches, marks, etc. The main principle behind the need of this application is taking forward step towards improvements in their strategies for their college admission and also used for the student's that they can easily get their predicted college. It can handle the details of previous marks of students, entrance exam mark, selected branch and all the details which filled on the admission form[8]. In this software, only admin or operator can view the details of students. So it is secure for both admin and student as well. In the admission management system of the colleges the

frequent items of students can be analyzed by referring the different attributes of the students. Predicting the admissions of the students is complex task for the colleges for that they required maintaining the students records manually by ink and paper. This is time consuming and very slow process. It is required to design easy admission management system to find out the frequent patterns of the student to improvement in the strategies of handle the database and also reduced the student efforts to find out their predicted college[3][8]. The main goal and objectives are :

- Predicting the admissions of the students is a complex decision for colleges or institutes.
- The main objective is to provide such patterns which will show that different categories of student who tent to take admission in which college.
- Provide ideas considering the patterns for easily analyzed data for different institutes.
- To reduce the efforts of institute's admin and student and their time.

II. PROBLEM DEFINITION

Today all the work at the time of admission of the students is done manually by ink and paper, which is very slow and time consuming much efforts and time[7].So it is required to Design an application for Student Admission System to speed up and make it convenient to use system.

III. BACKGROUND AND RELATED WORK

The main aim of our project is to develop the application which will be very useful in admission system. As any college management system includes the admission process of student from starting to ending till that student completes his course[8]. The requirement of all users is to:

1. Admin
2. Student
3. Institute

Login to the system through the first page of the application.

New student registration after log in into the system.

Student should get predicted college list according to marks.

Institute can see and analyze the data of any databases.

Admin login should be present who can read as well as remove any uploads files and student and institutes record.

In [1], How the data mining algorithm is used for the business development and analysis. FP tree algorithm is used for implement the business analysis tool. How the patterns are evaluate with the products and getting the frequent patterns among them. In this paper result of FP tree algorithm is implemented with minimum cost.

In [2], The naïve Bayes algorithm is implemented and how it is useful to find out well predicted outputs and accuracy.

In [3], A naïve Bayes algorithm is used to find out student performance by prediction and using this how this will performed accurately with large databases. Classification of student databases with yes and no form with prediction.

In [6], How the data mining can be use for student admission database or admission process is shown. One of the data mining technique is used in this application of data mining. How this is useful for education system is elaborated in this paper.

In [8], The student admission process using web application and how it could implement and beneficial for all users. How the human efforts could reduce is said in this paper.

In [11], It is the combination of prediction of student databases and KDD technique using naïve Bayes Classification and how this will used for educational databases

IV. IMPLEMENTATION DETAIL

A. System architecture-

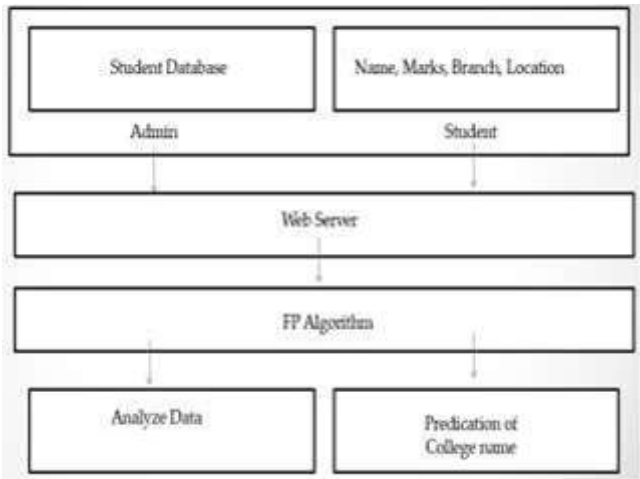

Fig.1 System Architecture

B. Algorithm-

Algorithm 1-
Step 1: Scan the student data set
Step 2: Calculate the probability of each attribute value. [n, n_c, m, p]
Step 3: Apply the formulae P(attribute value(ai)/subject value (vj))=(n_c + mp)/(n+m)

Where:
n = the number of training data item for which v = vj
nc = number of examples for which v = vj and a = ai
p = a priori estimate for P(ai,vj)
m = the parallel size of the sample
Step 4: Multiply the probabilities by p
Step 5: Compare the values and classify the attribute values to one of the predefined set of class.[3]

Naïve Bayes

$P(A|B)P(A)P(B)$……..(Eq. 1)

Where,
P(A) is the prior probability of A
P(B) is the prior probability of B
P(A|B) is the posterior probability of A given B
P(B|A) is the posterior probability of B given A

Since the denominator P(B) in Eq. 1 is the probability of the evidence without any knowledge of the event A, and since the hypothesis A can be true or false, Bayes' theorem can also be written as,

$P(A|B)=P(B|A)P(A)P(B|A)\times P(A)+P(B|\neg A)\times P(\neg A)$ (Eq. 2)

Where,
P(¬A) is the probability of A being false
P(B|¬A) is the probability of B given A is false[5]

Algorithm 2-
1. Variables are denoted in capitals such as Xi, and their values denoted by lower-case such as xi, and sets of variables are denoted by boldface letters such as X.

2. Let **X** = {X1,….,X□} be a finite set of observed random variables, called features, where each feature takes values from its domain Di. The set of all feature sets is denoted by Ω = D1 ×…× Dn. Let C, such that $c \in \{0,\ldots,u-1\}$, be an unobserved random variable denoting the class of a set of features.

3. A hypothesis $h: \Omega \rightarrow \{0,\ldots,u-1\}$, that assigns a class to any given set of variables is defined as a classifier. Each class c is assigned a discriminant function $fc(\boldsymbol{x}), c= 0,\ldots,u-1$. The classifier selects the class with the maximum

4. Discriminant function on a given set of variables, written as $h(\boldsymbol{x})= \arg max c \in \{0,\ldots,u-1\} fc(x)$.

5. The Bayes classifier $h*(\boldsymbol{x})$ uses the posterior probabilities given a set of variables as the discriminant function, i.e. $f*(\boldsymbol{x})=P(C=c\,|X=x)$. Applying Bayes' theorem from Eq. 1 to this function gives
$P(C=c \mid \boldsymbol{X}=\boldsymbol{x})= P(\boldsymbol{X}=\boldsymbol{x}\,|C=c)$
$P(C=c)P(\boldsymbol{X}=\boldsymbol{x})$. Since P(X = x) is the same for all classes it can be ignored. Hence, the Bayes' discriminant function can be written as

$f*(\boldsymbol{x})= P(\boldsymbol{X}=\boldsymbol{x}\,|C=c)\ P(C=c)$, where $P(\boldsymbol{X}=\boldsymbol{x}|C=c)$ $P(C=c)$ is called the class-conditional probability distribution (CPD) [5]. Thus the Bayes' classifier written as in Eq. 3 finds the maximum posterior probability hypothesis given x. $h*(\boldsymbol{x})=\arg max c$ $P(\boldsymbol{X}=\boldsymbol{x}\,|C=c)$
$P(C=c)$ (Eq. 3) Applying the assumption that features are independent given the class on Eq. 3, we can get the naïve Bayes classifier.

$fcNB(\boldsymbol{x})= \Pi P(Xj=xj\,|C=c)\ P(C=c) nj=1$ (Eq. 4)[5]

C. Set theory

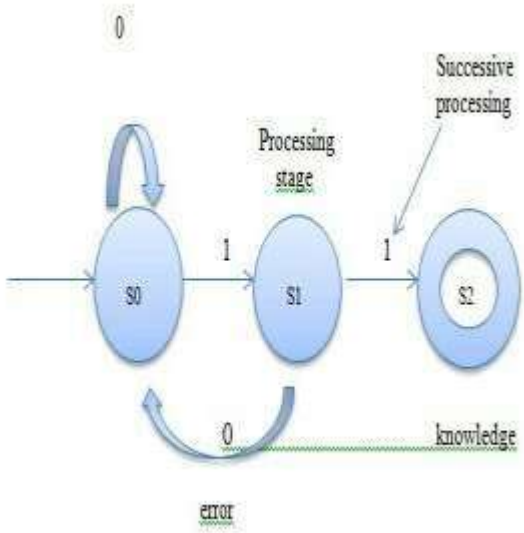

Fig. 2 State diagram S0: Initial state

1-Database input

0-Previous knowledge stat

S1: Processing state

1- Successful completion

0- Unexpected errors

S2: Represent knowledge to user

Let S be the system, S={ Q, ∑, ə, q0, F }

In this,

Q is the number of states that used in a proposed system.
{S0, S1, S2}
∑ is the transaction function that which function used at which states.
{0,1}
ə is the Data mining algorithm
{Naïve Bayes}
Q0 is the initial state of system.
{S0}
F is the final state of system.
{S2}

D. Mathematical model- From figure

State → Input	S0	S1	S2
0	S0	S0	-
1	S1	S2	

Fig. 3 State transition table

Input: Student database is given as a File
Output: Predicted item sets and analyze data or predicted college to student.

Success Conditions: Successfully generates the output according to that analyze data in simple manner

Failure Conditions: Fail to predict accurate data and analyze the data.

V. RESULT

In following figure marks of different branches with respected cut off list. All branches have their different cut off value. So here patterns from branches with cutoff with respect to student marks are elaborate in following figure.

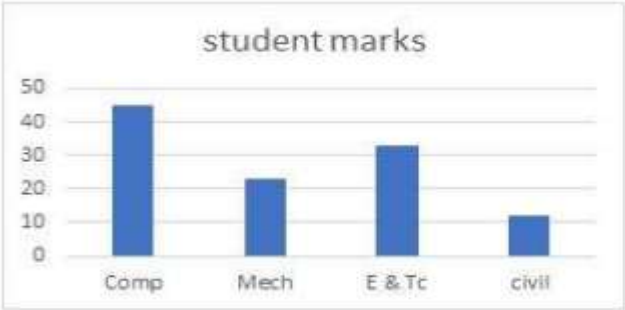

Fig. 4 Branch cut off graph

In this example the marks of computer branches with respective college cut offs. So here we can get result of this pattern like marks of subject and cutoff of college with respect to branch. So student can predict and analyze their predicted college according to this.

Fig. 5 Com cut off with diff. colleges

VI. CONCLUSION

In this paper, naïve Bayes algorithm of data mining approach for business analysis is used to make better business decisions. Data mining techniques can be useful in deriving patterns from student data, and this pattern can be useful to improve Admission System. So, the all hard work with use of ink and paper will reduce and all student can predict their college perfectly and all institutes can analysis their data to improve their strategies. This web application work can be helpful for any Educational System with reducing human efforts.

VII. ACKNOWLEDGMENT

The authors would like to thank the researchers as well as publishers for helping and guidance. We also thank the college authority for providing the required things and support systems. Finally, we would like to extend a heartfelt gratitude to friends and family members.

REFERENCES

[1] Mukund Prataprao Deshmukh, “Developing a business intelligence tool”, IEEE Conf., 2012.

[2] Priyanga Chandrasekar, “The impact of data preprocessing on the performance of naïve bayes classifier”, IEEE, 2016.

[3] Ms. Tismy, “ Devasia1 prediction of students performance using educational data mining, 2013.

[4] P. V. Praveen Sundar, "A comparative study for predicting student's academic performance using bayesian network classifiers", IOSRJEN, 2013.

[5]Bandana Garg, "Design and development of naïve bayes classifier", 2013.

[6]Dineshkumar B Vaghelaaand, Priyanka Sharma, "Students Admission Prediction using GRBST with Distributed Data Mining", Volume 2 – No.1, June 2015.

[7]Ashok Kumar Soni, "Student Admission System",2015.

[8]Jadhav Snehal Balasaheb, Supekar Bhagyashri Sitaram, "Web Based College Admission System", IJEDR Conf., 2014.

[9] M. Muchová, "An approach to support education of data mining algorithms", IEEE, 2017.

[10] Amjad Abu Saa, "Educational Data Mining & Students' Performance Prediction", IJACSA 2016.

