

Prediction of Student Enrolment Using Data Mining Techniques

Heena Sabnani¹, Mayur More², Prashant Kudale³, Prof. Surekha Janrao⁴

^{1,2,3} BE Engineering, Terna Engineering College, Nerul, Navi Mumbai.

⁴Professor, Dept. of Computer Engineering, Terna Engineering College, Maharashtra, India

Abstract - Data mining has been successfully implemented in the business world for some time now, its use in higher education is still relatively new, i.e. its use is intended for identification and extraction of new and potentially valuable knowledge from data. Using data mining the aim was to develop a model which can derive the conclusion on students' enrolment behaviour. Different methods and techniques of data mining were compared during prediction of students' enrolment. This paper contributes a technique that will help the institute to analyse the current scenario of admission by predicting the enrolment behaviour of student. It contributes a technique like Apriori which analyses the enrolment behaviour of the student considering the branch he is enrolled in and the branches he preferred to seek admission in. The system also proposes an enhanced data mining technique like Naïve Bayes algorithm which predicts to which course the student can enrol. This will help the student in admission procedure.

Key Words: Data Mining Techniques; Apriori; Branch Prediction.

1. INTRODUCTION

Data mining is the extraction of hidden predictive information from database. Generally data mining is the process of analysis of data from different perspectives and summarizing it into useful information used to increase profit. For higher education institutions whose goal is to contribute to the improvement of quality of higher education, the success of creation of human capital is the subject of a continuous analysis. In our daily life, we make our choices at most cases relying on recommendations from newspapers, people, or the Internet (e.g., book reviews, movie, restaurant rating, etc.). However, as the amount of information available on the Internet grows, searching for and making decisions about information becomes difficult.

Admission process is one of the most important processes when it comes to any educational institution. But this admission process is a very tedious and tiresome work. The aim of our system is to help the institute to analyse its admissions, based on the preferences considered by the student.

The major issue in the Student community who are at the stage of Higher Secondary is the selection of their career. It is mainly due to lack of information in the area which they want to choose. Ignorance is the first issue they face that blocks them from continuing to the right destination. Finally they choose some course and institution randomly after sacrificing their own dream. Because of this our country

loses many different potential students in various areas. India produces 3, 60,000 engineering graduate every year, only 25% of them are employable. One of the drawbacks which make them unemployable is improper selection of right branch during enrolment process. In our research work, we will propose a solution for this problem using knowledge based decision technique. Our motivation behind this work is that, if students enrol themselves in right branch, they will be able to perform in a better way.

Our system helps both college and student. It will help the college, regarding admission from student's perspective. Our system helps college in increasing the intake, developing labs, more facilities through the stream analysis which will give the prediction of related branches which students mostly prefer. Our system will also consist of a mock test, and by giving the mock test, the student will get the result to which particular branch he/she can enrol.

2. LITERATURE SURVEY

In this paper we have studied an association mining technique as Apriori that uses the partial information about the contents of a shopping cart for the prediction of products that the customers wish to buy or are more likely to buy along with the already brought products [1].

The proposed knowledge based decision technique will guide the student for admission in proper branch of engineering. Another approach is also developed to analyse the accuracy rate of for decision [2].

Previous research has focused predominantly on how to obtain exhaustive lists of associations. However, users often prefer a quick response to targeted queries. For instance, they may want to learn about the buying habits of customers. For this, paper proposes an approach that converts the market-basket database into an itemset tree [3].

This paper proposes the use of a recommendation system based on data mining techniques to help students to take decisions on their academics itineraries. More specifically, it provides support for the student to better choose how many and which courses to enrol on, having as basis the experience of previous students with similar academic achievements [4].

Right from making test papers, conducting test and giving results; everything has to be done manually. A lot of paper work is involved. A lot of time and energy is required to get results, but still no accurate suggestion could be provided to the student to choose the stream.

So far, no system has been developed to help the college to analyse its admissions so that it can develop its branches on the basis of preference's submitted by students.

3. PROPOSED METHODOLOGY

In our proposed framework Fig - 1, students have to complete the registration after they have authenticated themselves by username and password. When student logs in, then he/she can give test provided by the system. After completion of test, results will automatically be computed by system.

With the help of submitted answers and data mining techniques, the system provides a suggestion to the student, regarding which branch could be opted by him. This is done to solve the problems faced with manual examination writing, as there is need for a computerized system to handle all the work.

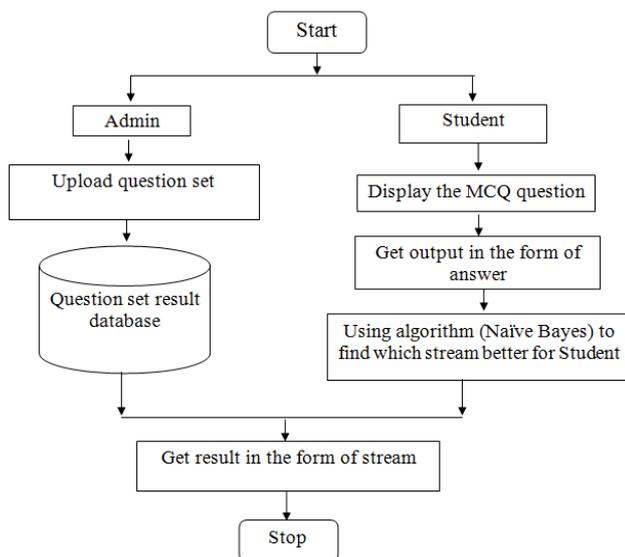


Fig -1: Student Branch Prediction.

In our framework Fig - 2, the stream which is most likely to be chosen is categorized by Apriori algorithm.

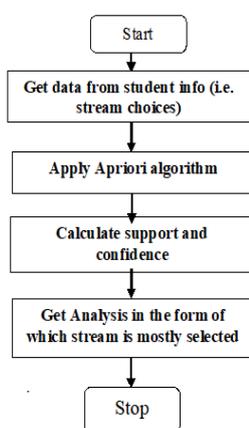


Fig -2: Branch-wise Analysis.

3.1 Apriori Algorithm

Apriori algorithm is used for Mining Association rules between Sets of Items in large databases.

Principle of Apriori Algorithm states that, if an item set is frequent, then all of its subsets must also be frequent.

Fig - 3, shows the preferences submitted by the students. By using the apriori algorithm, we find the related branches which are currently in trend or mostly preferred by the students. Using algorithm for this example, Computer, It, Extc are the related branches.

- 1.Comps 2.IT 3.EXTC 4.Civil 5.Mech
- 6. Elex

Student	Choice 1	Choice 2	Choice 3
Student 1	COMP(1)	IT(2)	EXTC(3)
Student 2	EXTC(3)	CIVIL(4)	COMP(1)
Student 3	EXTC(3)	MECH(5)	CIVIL(4)
Student 4	IT(2)	COMP(1)	EXTC(3)
Student 5	EXTC(3)	ELEX(6)	CIVIL(4)
Student 6	COMP(1)	EXTC(3)	IT(2)
Student 7	COMP(1)	ELEX(6)	EXTC(3)
Student 8	ELEX(6)	MECH(5)	CIVIL(4)
Student 9	IT(2)	MECH(5)	CIVIL(4)
Student 10	IT(2)	EXTC(3)	COMP(1)

Fig -3: Apriori Principle

Steps of Apriori Algorithm

Level-wise algorithm:

1. Let $k = 1$
2. Frequent item sets of length l is generated.
3. Repeat the above step until no new frequent item sets are identified
 1. Generate length $(k+1)$ candidate item sets from length k frequent item sets
 2. Prune candidate item sets containing subsets of length k that are infrequent
 - * How many k -item sets contained in a $(k+1)$ -item set?
 3. Count the support of each candidate by scanning the DB
 4. Eliminate candidates that are infrequent, leaving only those that are frequent

Note: steps 3.2 and 3.4 prune item sets that are infrequent.

3.2 Naive Bayes Algorithm

Naive Bayes model assumes that all variables contribute toward classification and that they are mutually independent. In other words, It assumes that variables are not correlated. This is unrealistic assumption for most of the datasets, however it leads to a simple prediction framework that gives surprisingly good result in many practical cases.

Total Questions given=30.

Branches	CIVIL		IT		COMPS		EXTC		MECH		ELEX	
Response	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Student	2	3	3	2	5	0	0	5	2	3	2	3

Fig -4: Naive Bayes

4. RESULTS AND IMPLEMENTATION

This system will help the institutions to make the necessary developments. Institutions can understand the admission scenario in a better way and can increase the intake of branch based on the prediction results. Student will be able to select the branch as per his/her abilities and field of interest.

Our system is divided into two models first is Student's Branch Prediction and second is Stream Analysis. Every student can register himself using Sign Up page, after successful registration he/she gets the login username and password. Using it, user can logged in the system. Our mock test will consist of 30 total questions. Using Naïve Bayes, we can predict which stream is better for student but the final decision will be his/her while choosing. Fig – 5, shows the demo result of the student, where student gives maximum right answers for computer branch.

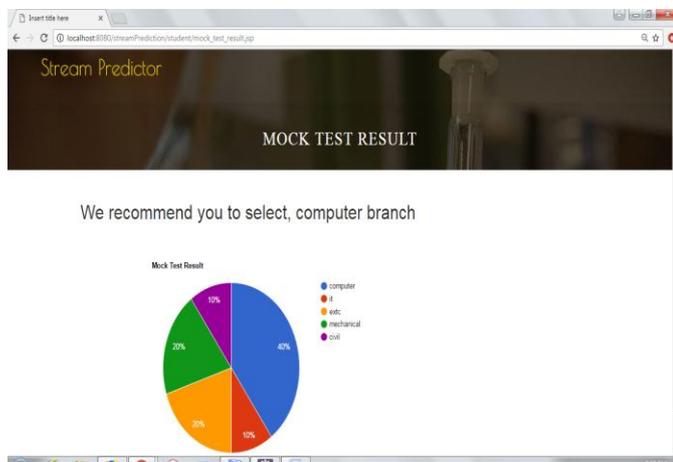


Fig - 5: Mock Test Result

Our second module consists of analysis part. Using apriori algorithm, we finds the related branches. Suppose maximum students prefer Computer as their first preference then it may possible that they prefer IT as their second preference, so using apriori we find the related branches which will help the institute to understand the current admission scenario in a better way. Fig – 6, consists of pie chart which shows the actual admissions happened in college.

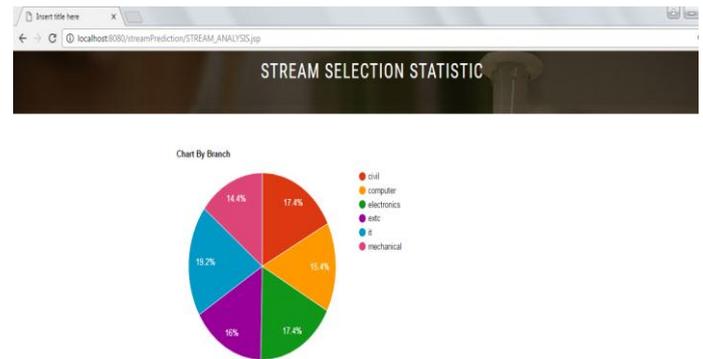


Fig - 6: Pie chart of actual admissions of every branch

Fig - 7 shows the result of apriori algorithm. Here we find the pairs of branches which is related to each other. That helps us to find out which branches are currently in trend and most preferred by students.

Fig – 8, shows the analysis of branch-wise admissions. It shows that the student already get admitted in one of the branch and our pie chart shows its most preferred second branch, which will help the institute to understand the admission scenario. So, that it will develop the infrastructure or increase the intake of second proffered branch along with the first.

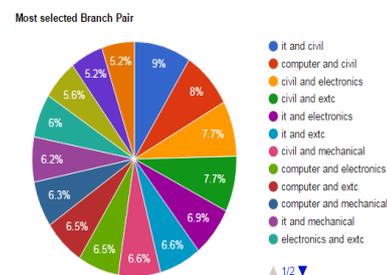


Fig - 7: Result of Apriori algorithm

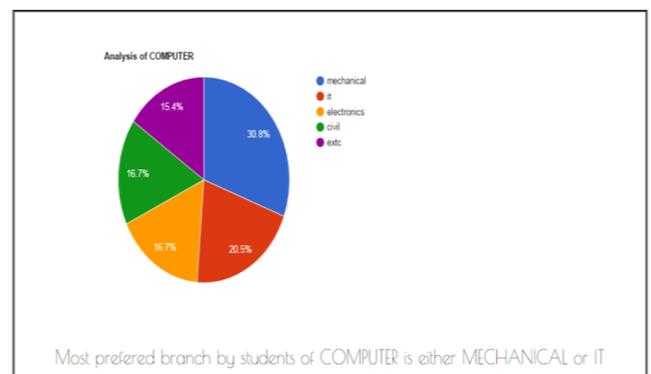


Fig - 8: Analysis of branch-wise admissions

Fig – 9, shows the mock test panel. It will consist of 30 basic questions, 5 from each branch (Computer, EXTC, IT, Electronics, Mechanical, Civil). This test will be a helping-hand to the student to find out the branch he/she is interested in and can pursue his/her career in it.

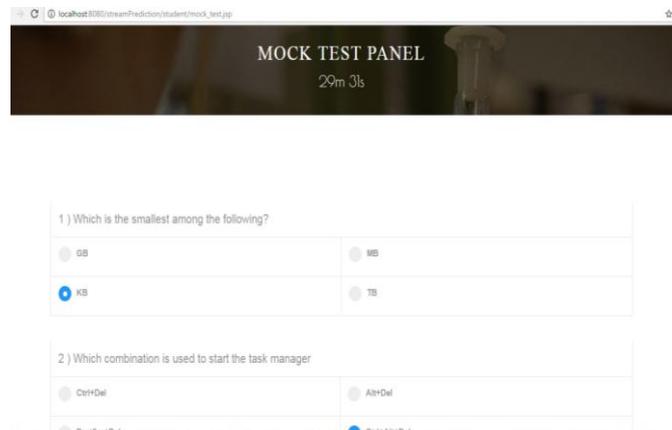


Fig - 9: Mock test panel

5. DISCUSSION

As a student perspective, by giving the test, student will able to select the branch as per his/her own interest or ability or field of interest. Our second module will generate the pie chart which shows the related branches which are currently in trend. Right from making test papers, conducting test and giving results; everything has to be done manually. A lot of paper work is involved. A lot of time and energy is required to get results, but still no accurate suggestion could be provided to the student to choose the stream. So far, no system has been developed to help the college to analyse its admissions so that it can develop its branches on the basis student preferences. This system will help the institutions to make the necessary developments. Institutions can understand the admission scenario in a better way and can increase the intake of branch based on the prediction results. Student will be able to select the branch as per his/her abilities and field of interest.

6. CONCLUSION

Information about test results is used to predict the suitable branch. This study helps to minimize the failure ratio and to take acceptable action for career. This study can facilitate the students, as it will guide them to take appropriate decision while choosing the stream as his/her career. This system will help the college to analyze the admissions and take the necessary actions depending upon the results.

7. FUTURE SCOPE

Our system will help the institutions to make the necessary developments. Institutions can understand the admission scenario in a better way. As the student's preferences would be taken into consideration, The colleges will be able to

increase the intake of branch based on the prediction results. In our system by giving the mock test, the student will be able to select the branch as per his/her abilities and field of interest. Our system will help the student in taking admission in particular branch in which the student is feasible to based on the results of mock test.

8. REFERENCES

- [1]Saroja T. V., Greeshma Nair, Kajal Nalawade, Yojana Prajapati, "Data mining technique to predict missing items and find optimal customer for beneficial customer relationship and management", Feb 2016, IJARCCCE(International Journal of Advanced Research in Computer and Communication Engineering)
- [2]Malaya Datta Borah, Rajni Jindal, Daya Gupta, "Application of knowledge based decision technique to predict student enrollment decision", International conference on recent trends in information systems.
- [3]Miroslav Kubat, Alaaeldin Hafez, Vijay V. Raghavan, Wei Kian Chen, "Itemset Trees for Targeted Association Querying", IEEE transactions on knowledge and data engineering.
- [4]Cesar Vialardi, Javier Bravo, Leila Shafti, Alvaro Ortigosa, "Recommendation in Higher Education Using Data Mining Techniques", Educational Data Mining 2009.
- [5]Kasun Wickramaratna, Student Member, IEEE, Miroslav Kubat, Senior Member, IEEE, and Kamal Premaratne, Senior Member, "Predicting Missing Items in Shopping Carts" IEEE
- [6]Narimel Bendakir and Esma "Using Association Rules for Course Recommendation" Departement informatique University Montreal
- [7]Pratish Guleria, Manu Sood "USAGE OF ASSOCIATION RULE MINING IN COURSE SELECTION FOR INDUSTRIAL TRAININGS" Department of Computer Science, Himachal Pradesh University,
- [8]Amer Al-Badarenah, Jamal Alsakran "An Automated Recommender System for Course Selection" Computer Science Department University of Jordan
- [9]<http://staffwww.itn.liu.se/~aidvi/courses/06/dm/lectures/lec7.pdf>
- [10]<https://www.slideshare.net/INSOFE/apriori-algorithm-36054672>
- [11]<https://www.analyticsvidhya.com/blog/2017/09/nave-bayes-explained/>