A Feasible Method for Blurbing Higher Education Institute

Dileep Kumar Yadav
CSE Dept. VBS Purvanchal
University, Jaunpur

Surjeet Kumar
MCA Dept. VBS Purvanchal
University, Jaunpur

Abstract: **In the last few decades, large number of Higher Education Institute grows. This creates a competition among these institutions to attracting the students for admission in their institutions. One of the best ways to reach the students is advertisement. In this paper, our focus is on advertisement which is economically feasible because the most of institutions are opened in self financial mode. Now a day a number of advertisement methods are available and our focus is on social networking sites like facebook, Ourkut, twitter and mobile application like WhatsApp. In this paper, we use Apriori Algorithm which explores support and confidence method to find out the best way of advertisement by using Weka Tools.**
 *Keyword***: Apriori Algorithm, Support, Confidence.**

1. INTRODUCTION

Data mining is a promising research field for educational data analysis. There are many data mining techniques such as clustering, classification, prediction, support, confidence and outlier analysis can be used for the purpose of analysis. Business world used data mining technique to find out the interest group for its product from a large number of consumers. No doubt they got satisfaction from data mining. This successful implication of data mining motivates researcher of education to use this technique due to rich data set.

In last few years Indian Higher Education Authorities focuses on opening self fund generating institution. In early days, these institutions easily attract students, who are seeking for admission; sometimes without any making special effort to attract them [1]. But increased number of self finance Higher Education Institute has faced a lot of trouble in attracting student. To solve this problem Higher Education Institute started thinking about some method to attract student by advertisement which is economically feasible.

2. DATA DESCRIPTION

The data is collected from different colleges situated in Jaunpur and Ghazipur district of U.P (India) .So data is first collected on Questionnaire paper i.e. Name, mobile no., facebook id, WhatsApp no., Gmail id and twitter id.Collected data is stored on database for further analysis. In this Questionnaire one section is very important which provides data for analysis i.e. advertisement method by which he or she came to know about college facilities, fee structure and other details?

Data Analysis

In last few years college uses traditional methods like hording, news paper, pamphlets radio of advertisement to attract students because most of the peoples are not using modern method of advertisement like WhatsApp, facebook, twitter, Messenger etc.

Table 1:Number of Occurrences of advertisement methods

Advertisement method	Code	Answer
WhatsApp	W	800
Facebook	F	650
Twitter	T	230
Messenger	M	435

3. DATA MINING

Data mining, knowledge discovery and machine learning are used in same context. They contains algorithm to find best pattern from an unstructured data with the help of computers. These algorithms attempt to fit a model to the data. The algorithms examine the data and determine a model that is closest to the characteristics of the data being examined [2].

Apriori Algorithm:

The Apriori algorithm is one of the most widely used tools for association rule mining. It uses priori knowledge of frequent item set property for association rule mining [3].

- In first step algorithm simply counts singleton items are the candidates and the items that has the support value less than a threshold value is eliminated from the candidate item list.
- In the second step algorithm the singleton items are combined to form two member candidate item set and support value less than threshold value is eliminated from the candidate item list.
- In the next step algorithm creates three member candidate item set and the process is repeat again. When all frequent item sets are accounted then the process stops.
- The frequent item sets are then used to generate association rules which have confidence values greater than or equal to the threshold value.

Support:

The support of an item (or a set of items) is the percentage of transactions in which that item occurs. It is defined as "the support (s) for an association rule $X{\rightarrow}Y$ is the percentage of transactions in the database that contain $X \cup Y$.

Support Count:
The occurrence frequency of an item set is the number of transaction that contains the item set.

Confidence:
Confidence is the ratio of the number of transaction that includes all items in the consequent as well as the antecedent to the number of transaction that include all items in the antecedent [3].

If $X \rightarrow Y$ then X is called antecedent and Y is called consequent and then

$$\text{Confidence} = \frac{\text{Support count of}(Y+X)}{\text{Support count of X}}$$

4. RELATED WORK

Data mining provides a great relieve to education by finding hidden pattern from unstructured student database. A lot of research is done in educational data mining area using data mining. Yadav and Pal [4, 5] make a comprehensive study on different research did in the area of education data mining. They also provide a list of data mining techniques.

Umesh kumar Pandey and Surjeet kumar yadav discussed data mining application to attract students in higher education institute using predictive models. In this paper Pandey and Yadav used traditional methods of advertisement like hoarding, news paper, pamphlets and radio and so on [1].

Bohannon T [6] discussed various application of predictive modeling in higher education. His paper concentrates on enrollment management, retention analysis and donor giving. He used a decision tree, backward regression, stepwise regression and a neural network data mining technique to complete his study.

Bharadwaj and Pal [7], applied the classification as data mining technique to evaluate student' performance, they used decision tree method for classification. The goal of their study is to extract knowledge that describes students' performance in end semester examination. They used students' data from the student' previous database including Attendance, Class test, Seminar and Assignment marks. This study helps earlier in identifying the dropouts and students who need special attention and allow the teacher to provide appropriate advising.

In the last few decades, number of Higher Education Institutions grows rapidly in India. This causes a cut throat competition among these institutions while attracting the student to get admission in these institutions. Most of the institutions are opened in self finance mode, so all time they feel short hand in expenditure. Therefore, institutions focused on the strength of students not on the quality of education. Indian education sector has a lot of data that can produce valuable information. Knowledge Discovery and Data Mining (KDD) is a multidisciplinary area focusing upon methodologies for extracting useful knowledge from data and there are several useful KDD tools to extract the knowledge.
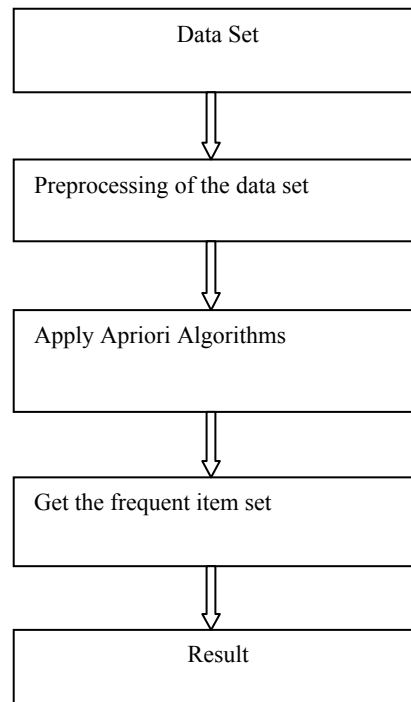
5. PERFORMANCE MEASURE

Figure.1: Methodology   Flow Chart

6. RESULT AND DATA ANALYSIS

In this paper, Apriori Algorithm with support and confidence is proposed to use for finding the best result of economically feasible advertisement methods. From the obtained results it is observed that the proposed method performs better and economically feasible advertisement in comparison to other methods.

There are different combinations   of advertisement to find the best combination of economically feasible ways from the given table:

Table.2 Different combination of advertisement

Sr. No.	Transaction id	Item set
1	T101	W,F,T
2	T102	F,M
3	T103	F,T
4	T104	W,F,T
5	T105	W,F
6	T106	F,T
7	T107	W,F
8	T108	W,F,T,M

Table 3.Result Analysis for the Confidence of the Data set

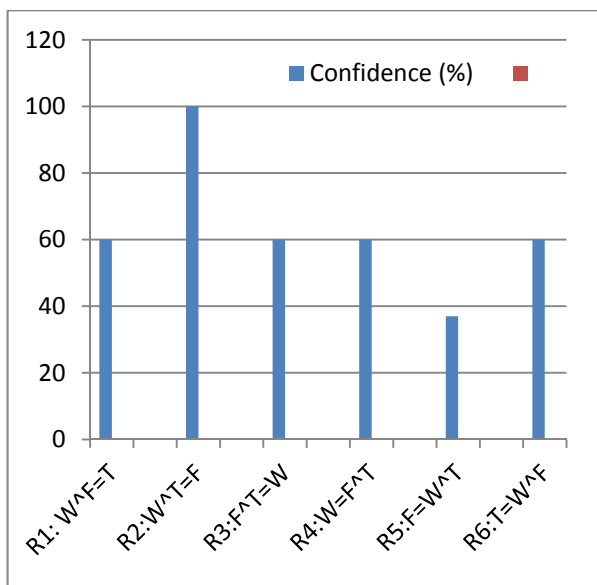|	Confidence (%)
R1: W^F=T	60
R2:W^T=F	100
R3:F^T=W	60
R4:W=F^T	60
R5:F=W^T	37
R6:T=W^F	60

Figure.2: Graphical Representation for the Confidence of the data set

RESULT

On table 2 by applying apriori algorithm on data set we conclude that most of the students who uses WhatsApp, facebook, twitter and messenger, the most frequent advertisement combination occurred is WhatsApp, facebook, twitter. From table 3 there are six relations found, out of six relation only relation R2 shows strong association which is 100%,this shows that the Confidence of WhatsApp, facebook and twitter is 100% of this advertisement. Similarly support and confidence level are showing for other relations also. From table 2 we can say that the entire advertisement mediums are connected to WhatsApp, facebook and twitter which is based on Internet.Today most of the student uses internet easily and access the WhatsApp, facebook and twitter.

7. CONCLUSION AND FUTURE WORK

In this paper, we tried to find out the best economically feasible advertisement methods. In table 2 most of the students have opted WhatsApp, facebook and twitter for advertisement is economically feasible way because in this method only internet connectivity required. In this era Internet users are increasing day by day. After performing data mining techniques like apriori algorithm, support and confidence on this data, we conclude that the WhatsApp, Facebook and Twitter is the best method for economically feasible advertisement.

REFERENCES

[1] Umesh Kumar Pandey, Surjeet Kumar Yadav, Saurabh Pal, "Data mining application to attract students in HIE", IJCSE, Vol.4 No.06 June 2012.
[2] M.H. Dunham, "Data Mining introductory and advanced topics", Pearson education, ISBN 978-81-7758-785-2.
[3] Sunita Tiwari, Neha Chaudhary, "Data Mining and Warehousing", Dhanpat Rai and Co.Educational and technical Publishers.
[4] Saurabh Pal, Surjeet Yadav, Mining Educational data to reduce dropout rates of engineering students", I.J. Information engineering and electronic business,Vol.4, No.2, 2012, pp 1-7.
[5] C.Romero and Ventura, "Educational data mining:A survey from 1995 to2005",Expert system with application 33,2007,pp 135-146.
[6] T.Bohannan, "Predictive Modeling in Higher Education",SAS global forum 2007,paper 074-2007.
[7] B.K.Bhardwaj and Saurabh Pal, Mining Educational data to analyze students performance",International Journal of Advance Computer Science and Application(IJACSA), Vol.2, No.6, pp 63-69,2011.