Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EDUCATIONAL DATA MINING FOR ENHANCED TEACHING AND LEARNING

[1]ATTA-UR-RAHMAN, [2]KIRAN SULTAN, [3]NAHIER ALDHAFFERI, [4]ABDULLAH ALQAHTANI

[1]Dep. of CS, [3,4]Dep. of CIS, CCSIT, Imam Abdulrahman Bin Faisal University, PO Box 1982, Dammam, KSA

[2]Dept. of Computer & Information Technology, JCC, King Abdulaziz University, Jeddah, KSA

E-mail: [1,3,4]{aaurrahman,naldhafeeri,aamqahtani}@iau.edu.sa, [2]kkhan2@kau.edu.sa

ABSTRACT

There are different classroom teaching techniques for effective teaching and learning like; teaching on black/white board, projectors, etc. Some of the students feel comfortable with one technique while others may be comfortable with some other. This study aims to discover trend of comfort level of students with respect to timetabling and teaching techniques. In this regard, a questionnaire is designed to acquire students' interest that will show what techniques are mostly liked or preferred by different types of students. Based on the feedback, machine learning algorithms are applied to extract the useful results. The analyses are conducted in WEKA and the proposed scheme is compared with other well-known techniques in the literature.

Keywords: *EDM, Machine Learning, Clustering, C-Mean, Apriori Algorithm, WEKA*

1. INTRODUCTION

Knowledge discovery in databases (KDD) is the way to find out hidden and interesting patterns from large data collections. KDD has been defined by Fayyad et al (1996) [1] as "The non-trivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data". KDD is an iterative process and have following steps also shown in Figure 1.

- Data selection
- Data cleansing
- Data reduction and projection
- Selection of data mining task and appropriate algorithm
- Data mining.

According to [2], in educational environment scheduling of courses is an important issue of NP-complete class of problems. This study focuses on generating the automatic timetable of two years degree program. This study completes its task in three steps as mentioned below:

- Two types of algorithms graph coloring and maximum matching are used in this work.

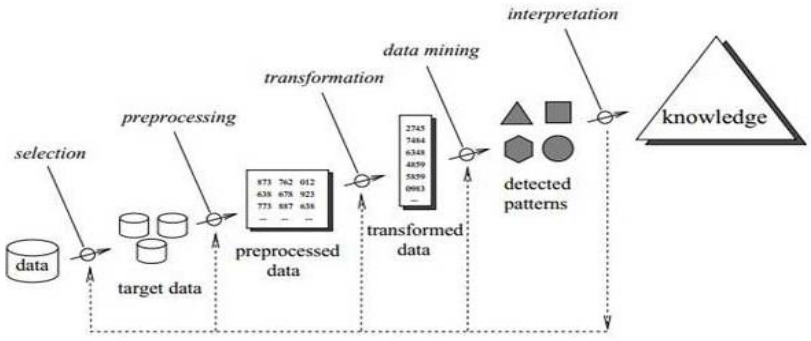


Figure 1: Data Mining Process

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Maximum software constraints are satisfied through simulated annealing.
- Through simulated annealing quality is enhanced by swapping events between different time slots.

In [3], the study solves the timetabling problem through auto generation of timetable by considering all the hard constraints and ignoring the soft constraints. In manual process soft constraints are considered while in auto generation method these constraints are ignored by the system. Automated system does not consider human feelings it is the main disadvantage of automated system. In order to build a system which works like humans it must understand the soft constraints of humans. This research used a method of rule mining to incorporate in the automated systems to get its attention towards soft constraints. Through the feedback that is obtained from the students it is observed that hardworking students prefer hard subjects in morning time. In other case if the hardworking student is able to get only second class, then he will be assigned some senior teacher to teach them in morning. If the student is not hardworking and average student then he would be assign three subjects in one day. The most useful observation in this study is that it is suggested to build the timetable so that during lab provide such student the opportunity to learn theory of those subjects which he had learnt that day.

This paper [4] investigates the problem related to learning and teaching methods by using two data mining techniques stepwise regression and decision tree. All the required data is collected from the students of Management Information System Department Bogazici University. Dimensions related to teachers and course characteristics are considered in this work. This research fin outs the factors which positively affects the learning and teaching methods. Teachers are suggested to encourage more students to attend the classes. It is observed that the courses which are taught by lower work load are more admirable by the students and the teachers behavior is also affect the learning process. Part time courses offered by teachers are preferred by students. Percentage of the enrollment questionnaire is also an important factor observed in this study.

In [5], authors proposed teachers' assessment and profiling system (TAPS) using fuzzy rule based system and Apriori algorithm. System works as a decision support system to choose appropriate teacher for a subject etc. and to help the management, figure out the teaching deviancies for counsel. System does not provide any support related to timetabling. Similarly, in [6-7], authors tried to mine the educational domain data for finding teacher's performance in a particular course based on temporal students' feedback using association rule mining techniques and tried to build a decision support system for selection/rejecting a teacher for a specific term, course and class. In [8], authors applied Naïve Bayesian algorithm for effective Educational Data mining and covered various aspects. In [9-10], authors provide an overview of the work done in the field of EDM. Also categorized the work in terms of application areas in EDM and techniques applied. In [11-12], authors employed similar approach for network user's behavior analysis.

In this research, objective is to facilitate students using better timetabling and teaching methods through data mining techniques which will provide maximum flexibility of choice during enrollment, moreover, the choice of material, choice of timing, and choice of teachers etc. to improve the teaching methodologies which is missing in the literature.

Rest of the paper is organized as follows. Section 2 contains system model, section 3 introduces the tools and techniques used. Section 4 shows the results and discussion while section 5 concludes the paper.

2. SYSTEM MODEL

The system model is comprised of similar steps shown in Figure 1 with a little modification in our case. The steps are shown in Figure 2.

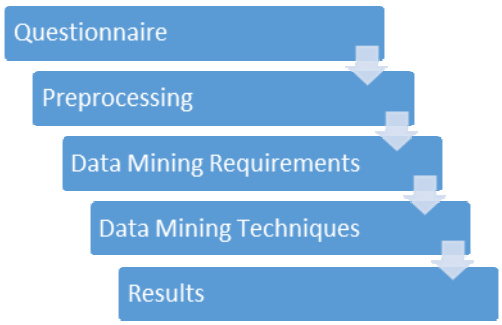


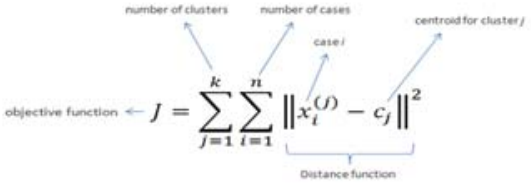
Figure 2: Work Flow

Through survey forms the data is collected from students about teaching techniques and suitable timetable and all the data is stored in MS Excel sheet for further processing. Student's data according to their semester, area, gender, and based on some other dimensions is saved in the sheet. Percentage of each question is calculated against the dimensions of students. From the Dataset some

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

attributes are used to analyze student's trends. The attributes used for this purpose are given in the Table 1. In the next step, the data is converted to CSV file and then to ARFF format to process in WEKS. An ARFF file is a text file that contains a list of instances having set of attributes.

The algorithms implemented in the WEKA can be applied on the data set for finding interesting relation and hidden pattern in dataset. The simple K-Means algorithm in WEKA automatically handles the categorical numerical attributes. During distance computations it automatically normalizes the numeric attributes. The distance measure is taken as Euclidean distance between instances and clusters. That can be written as:



The dataset used for this analysis is based on the student's data available in comma separated format (students-data.csv). The resulting data file is "students-data. Arff" and includes 1343 instances. We have used the K-Means algorithm to cluster the student's feedback against survey questionnaire in this data set in order to differentiate the resulting trends of students. The list of attributes and other parameters can be set after loading the file in WEKA.

In order to design a database for analysis, the primary data and some basic information about students are gathered from a survey of BIIT. The database is designed in MS Excel database management system. According to the required format the data is formed. In the next step, the data is converted to csv file and then to arff format to process in WEKA. An arff file is a text file that contains a list of instances having set of attributes. Figure 3 shows graphical representation of processed attribute.

3. TOOLS AND TECHNOLOGIES

For sake of data analysis two types of data mining techniques are used in WEKA. These techniques are most suitable for our case. That is K-mean clustering and Apriori algorithm. Here is a brief introduction to both pertaining to our system.

3.1 Clustering

Clustering is the type of unsupervised learning that allows a user to make groups of data in order to find the frequent patterns from the data. In classification only subset of attributes are used while in clustering it that every attribute is used in the data set to analyze the data. Clustering provides you benefit if the data set is defined and pattern needs to be determined from the data. Depending on the requirement of your task you can make a specific number of groups. The drawback of clustering is that without any knowledge about clustering user doesn't have idea about how many clusters he wants to make but in other case when user have an idea about it he can get benefits from it. WEKA quickly converts entire set of data into groups, so you can make conclusions fast. The steps followed during execution are:

1. Normalize each attribute in the dataset by dividing each value by the difference between the high value and the low value in the data set for that attribute.
2. Decide the number of desired clusters.
3. Compare the distance from each data sample in the cluster center by the least square method.
4. Compute the centroid.
5. By calculating the centroid distance if the clusters and its member do not change, then work is completed and clusters are formed. If the members are changed, then you need to repeat step three and start again until there is no change in clusters member.

Figure 4 shows a sample clustering window and Figure 5 provides with cluster visualization against attribute Q34 as an example. Where red color represents "Pure English" and "Blue" shows English-Urdu mixed medium of study. Figure 6 shows an example of students' trends towards class duration, discipline-wise, region-wise and gender-wise.

A. K-Means Method

- K mean is the type of unsupervised learning algorithms that is applied to solve the clustering problems.
- The procedure contains simple and easy way to classify the data set through a certain number of clusters.
- The basic step is to define k centroids for each cluster. The k centroids should be assigned in a cunning way because it causes different result by changing the position.

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Data Attributes

Attribute	Description	Values
Gender	Gender of students enrolled in institute	Male, Female
Area	Region of students where they belong	Rural, Urban
Semester	Current program of students	CS ,IT
Previous Qualification	Education history of students enrolled in institute	Intermediate, Graduate
Q1	Class Duration	60min,90min,180min
Q2	Class scheduling	M,T,W,T,F
Q3	Break duration of class	20min,40min
Q4	Class timings with respect to session	Morning/Evening
Q5	Rescheduling class in regular day with late timings	Yes/No
Q6	Combine class with other section	Yes/No
Q7	Consecutive classes	Left earlier/break
Q8	Learning material	White board/projector
Q9	Teacher's behavior	Friendly/serious
Q10	Class conduction by teacher	Yes/No
Q11	Oral questions by teachers answered orally by students	Yes/No
Q12	Discussion group by Students chairpersons	Yes/No
Q13	Lecture demonstration	Teacher/guest speaker
Q14	Class presentations	Student panel/Instructor
Q15	Debate on current issues	Yes/No
Q16	Audio tutorial lessons	Yes/No
Q17	Seminars and Tutorial	Yes/No
Q18	Labs and practical learning	Yes/No
Q19	Learning from tables, diagrams, charts	Yes/No
Q20	Peer teaching by assigned students to other students	Yes/No
Q21	Help by advisor about course difficulties	Yes/No
Q22	Suggestions for improving classroom assessments	Open book/home test
Q23	Online learning benefits students more than face to face	In long run/face to face
Q24	Feel comfortable when called by teacher on white board	Yes /No
Q25	Prefer surprise tests	Yes /No
Q26	Real time examples helps the students during study	Yes /No
Q27	Assignment submission by students	Soft copy/hard copy
Q28	Feel comfortable when same teacher take multiple courses	Yes /No
Q29	Multiple sections should taught same subjects by same teacher or multiple teachers	Single tech/multiple tech
Q30	Feel comfortable with more strength of students	Yes /No
Q31	Which type of paper pattern students like	Subjective/objective
Q32	Feel comfortable with co-education	Yes /No
Q33	Want to study from male teacher or female teacher	Male /female
Q34	With which language student feel comfortable in class	Mix / pure English

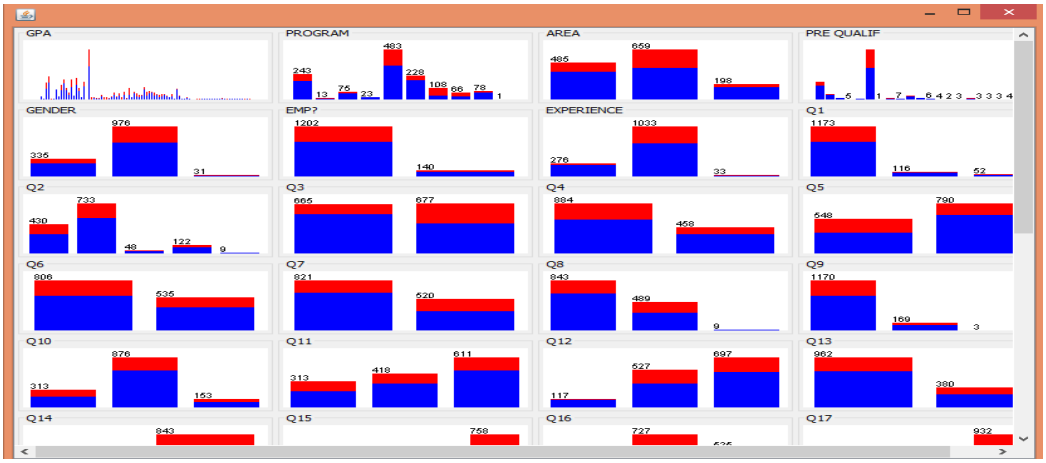


Figure 3: Graphical Representation of Processed Attributes

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

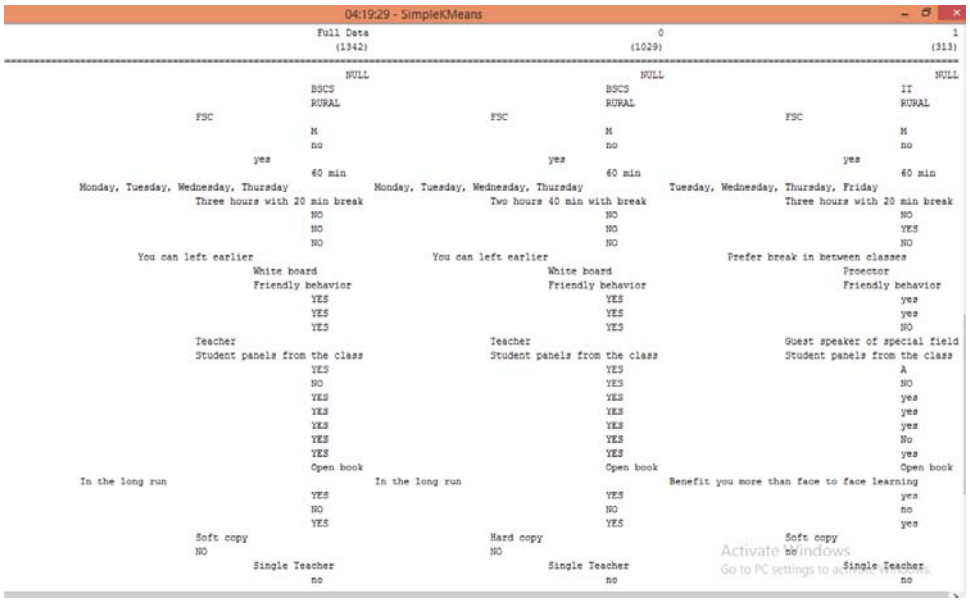


Figure 4: Sample clustering Window

- K-means clustering goal is to reduce intra cluster variance.

B. Example

Suppose we want to group the visitors to a website using just their age as given in the following way:
15,15,16,19,19,20,20,21,22,28,35,40,41,42,43,44,60,61,65
Initial clusters:
Centroid (C1) = 16[16] and centroid (C2)= 22[22]
Iteration1:
C1 = 15.33 [15,15,16]
C2 = 36.25 [19,19,20,20,21,22,28,35,40,41,42,43,44,60,61,65]
Iteration 2:
C1 = 18.56 [15,15,16,19,19,20,20,21,22]
C2 = 45.90 [28,35,40,41,42,43,44,60,61,65]
Iteration 3:
C1 = 19.50 [15,15,16,19,19,20,21,22,28]
C2 = 47.89 [35,40,41,42,43,44,60,61,65]
Iteration 4:
C1 = 19.50 [15,15,16,19,19,20,20,21,22,28]
C2 = 47.89 [35,40,41,42,43,44,60,61,65]

Results

- In iteration 3 and 4, no change has been occurred.
- Through clustering two groups have been defined.

The selection of centroids can affect the output of clusters, that's why the algorithm shows multiple results by changing the centroid position in order to get desired results.

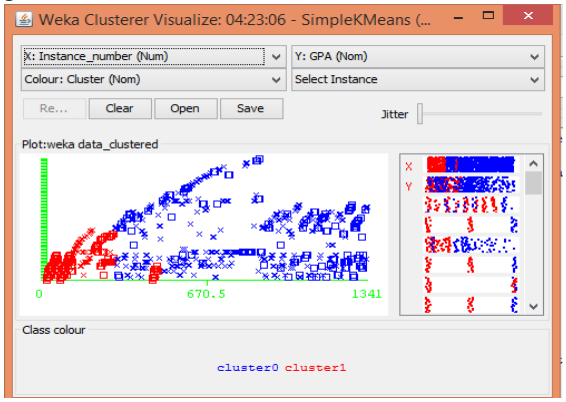


Figure 5: Cluster Visualization

3.2 Apriori Algorithm

Apriori algorithm is the algorithm of Boolean association rules. It is applied for mining frequent item sets, as described in [13] and [14]. It is used for the databases containing transactions. The approach used by Apriori is bottom-up approach. In which frequent subsets extended one item at a time. In this algorithm k item sets used to explore (k+1) item sets.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- Frequent subsets are extended once at a time
- Candidates of the groups are tested against the data
- It finds out the frequent item sets in the database and extend them into larger item sets till those item sets appear sufficiently often in the database
- The number of candidates being considered by only exploring the item sets whose support is greater than the minimum support.
- The item sets which are infrequent can be pruned.

3.2.1 Association rule mining

Association rule mining is the most specific method of data mining. Association rule mining is the process which is used to generate association rules that are built on the base of predefined mining support and confidence of a database. This method divides the problem into two sub problems. In first part it finds outs those item sets whose occurrence exceeds a predefined threshold in the database known as frequent or large item sets. In the second part it generates association rules for predefined large item sets with the constraints of minimal confidence. It is a two-step approach:

1. Frequent item set generation
2. Rules generation

The association rules are expressed as X→Y
This shows that whenever X appears Y also needs to appear. X and Y may appear as single item or set of items.

A. Confidence

Confidence denotes the strength of the association between X and Y.
Conf (X→Y) = SUPP (XUY)/SUPP(X)

B. Support

Support indicates the frequency of the pattern
Supp (X) = no. of transactions which contains the item sets X/ total no of transactions.

Example:
Consider, in a class
Total students: 100
Male: 20
So, 20/100*100= 20% which is support
In 20 Male students, 60 min class preference: 9 students
So, 9/100*100= 45% which is confidence

4. RESULTS AND DISCUSSION

Data from different sources is collected, reviewed, and then analyzed to form some sort of result or conclusion.

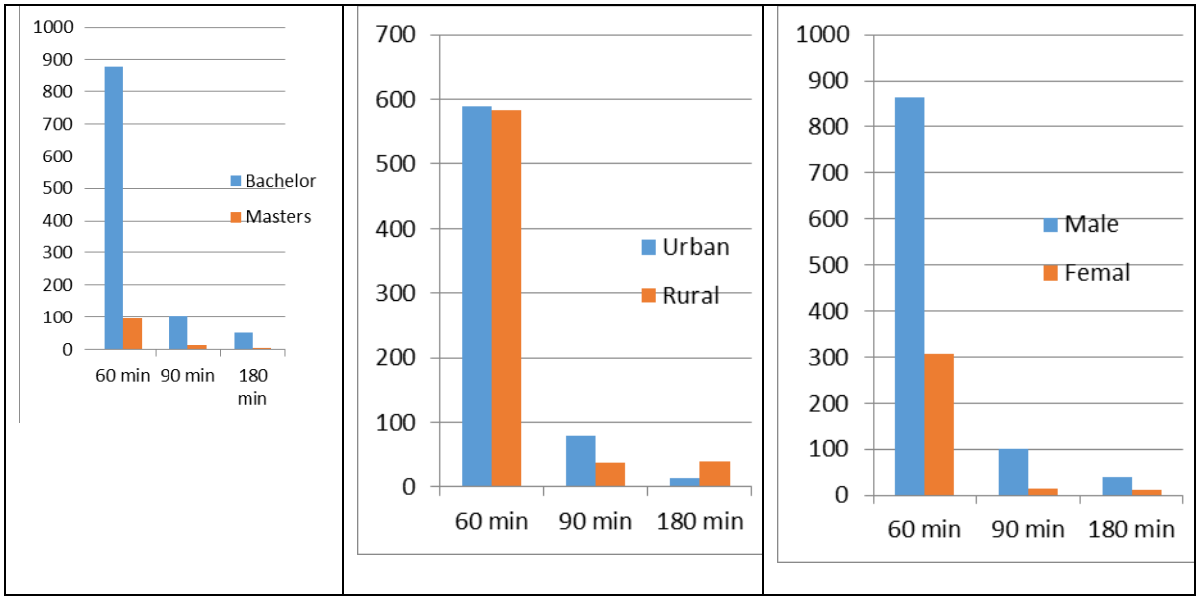


Figure 6: Trend of Students for Different Class Duration w.r.t. Discipline, Region and Gender

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

For this purpose, we show the graphical representation, Cluster Analysis by K-Means algorithm and Association Rule Mining for rules generation are used to show our results to understand the trends of students towards their timetable and teachers. Table-2&3 provide the trend analysis using K-Mean and Apriori algorithm respectively. The outcome is also discussed with in the table.

4.1 Trend Analysis by Clustering (K-Means Algorithm)

K-Means algorithm is used to cluster the student's feedback against survey questionnaire in this dataset, and to describe the resulting trends of students. Total instances were 1343.

Table 2: Trend analysis by K-Mean

Q. No.	Analysis	Outcome
1	Clustered Instances 0 442 (33%) 1 313 (23%) 2 587 (44%)	The resultant clusters shows the final trend analysis of students in which it is observed that 60 min (cluster-2) class tend to receive higher rating.
2	Cluster 0 ← Monday, Tuesday, Wednesday, Thursday (55%) Cluster 1 ← Tuesday, Wednesday, Thursday, Friday	The interpretation of above result shows that the combination of days of week (Monday, Tuesday, Wednesday, Thursday) is more suited to the students.
22	Cluster 0 ← Home tests (26%) Cluster 1 ← Open book (74%)	The cluster analysis shows that the students preferred open book test as compared to home test.
34	Cluster 0 ← Mixed 1029 (77%) Cluster 1 ← Pure English 313 (23%)	The cluster analysis shows that the students preferred mixed language (Urdu, English) in class.

4.2 Analysis by Apriori Algorithm

The analysis is provided in Table 3.

4.3 Comparison

Here the proposed technique is compared with similar techniques in the literature based on following factors:

- Effective timetabling
- Effective teaching/learning methodologies (assessment, medium of instructions)
- Effective advisory

Based on student's type:

- Enrolled in degree
- Gender
- Area of residence
- Previous degree
- Working

Comparison of the proposed scheme with other techniques of educational data mining is given in Table 3. The comparison shows that the proposed scheme covers a wider range of the aspects of effective teaching and learning while most of the proposed schemes are limited in the scope. For example, either the techniques are student centric or teacher centric not both but the proposed scheme covers both aspects in addition to several attributes given in Table 1. Total number of attributes is significantly higher than [5-8]. Technique in [8] has three different attributes that are family background, marital status and student's financial status but timetable preference, class duration, assessment type and working students are missing. Nonetheless, in future extension of the work, these attributes may also be included to get a better insight of teaching and learning.

5. CONCLUSION

In this paper, a machine learning based education data mining technique is proposed and investigated to enhance teaching and learning quality in the educational institutes. A survey consisted of 38 questions, carefully selected to provide almost every aspect of teaching and learning in a typically educational institute. After preprocessing, machine learning techniques were applied and useful results/trends and preferences were obtained. This research could help the institute in selection of better teaching/learning practices in terms of classroom medium, effective timetabling etc. based on enrolled students, their area, educational background, gender and much more. Proposed scheme is compared with other schemes in the literature and it is observed that it covers wider aspects of educational data mining attributes compared to other schemes. Moreover, the proposed scheme is hybrid in terms of students and teacher point of view, while existing schemes are either student centric or teacher centric. In future, the scheme may be extended to incorporate questions regarding the subject areas, knowledge groups and other like attribute. Other machine learning techniques like support vector machines (SVM) etc. may be investigated for better classification and trend analysis in future.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] S. Fayyad, G.P. Shapiro and P.S. Smyth, "From Data Mining to Knowledge Discovery in Databases", *AI Magazine*, Vol. 17, No. 3, 1996, pp. 37-54.

[2] E.K. Burke and J.P. Newall, "A New Adaptive Heuristic Framework for Examination Timetabling Problems", *Computer Science Technical Report No. NOTTCS-TR-1*, 2002, pp. 1-2.

[3] N.P. Desai, "Preferences of Teachers and students for Auto generation of sensitive timetable: A Case Study", *Indian Journal of Computer Science and Engineering (IJCSE)* Vol. 2, No. 3, Jun 2011, pp. 461-465.

[4] A. Carlos, "Timetabling in Higher Education: Considering the Combinations of Classes Taken by Students", *In Proc. 10th International Conference of the Practice and Theory of Automated Timetabling*, 2014.

[5] Atta-ur-Rahman, "Teacher Assessment and Profiling using Fuzzy Rule based System and Apriori Algorithm", *International Journal of Computer Applications (IJCA)*, Vol. 65, No. 5, March 2013, pp. 22-28.

[6] Atta-ur-Rahman, S. Das, "Data Mining for Students' Trends Analysis using Apriori Algorithm", *International Journal Control Theory and Applications*, Vol. 10, No. 18, 2017, pp. 107-115.

[7] Atta-ur-Rahman, S. Das, "Big Data Analysis for Teacher Recommendation using Data Mining Techniques", *International Journal Control Theory and Applications*, Vol. 10, No. 18, 2017, pp. 95-105.

[8] A. Abusaa, "Educational Data Mining & Students' Performance Prediction", *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 5, 2016, pp. 212-220.

[9] B. Bakhshinategh, O.R. Zaiane, S. Elatia, D. Ipperciel, "Educational Data Mining Applications and Tasks: A Survey of the Last 10 Years", *Education and Information Technologies*, Vol. 23, No. 1, January 2018, pp. 537–553.

[10] A. Dutt, M. A. Ismail and T. Herawan, "A Systematic Review on Educational Data Mining", *IEEE Access*, Vol. 5, 2017, pp. 15991-16005.

[11] Atta-ur-Rahman, Zaidi D.N., Salam M.H., Jamil S., "User Behavior Classification using Fuzzy Rule Based System", 13th International Conference on Hybrid Intelligent Systems (HIS'13), pp. 118-123, December 04-06, Tunisia, 2013.

[12] Atta-ur-Rahman, S.A. Alrashed, A. Abraham, "User Behavior Classification and Prediction using FRBS and Linear Regression" Journal of Information Assurance and Security, vol. 12, no. 3, pp. 86-93, 2017.

[13] R. Agrawal, T. Imieliński and A. Swami, "Mining Association Rules between Sets of items in Large Databases", *in Proc. ACM SIGMOD Conf.*, Washington DC, 1993, pp. 1-10.

[14] R. Agrawal, and R. Srikant, "Fast Algorithms for Mining Association Rules", *in Proc. 20th Int. Very Large Data Bases (VLDB) Conf.*, Santiago, Chile, 1994, pp. 1-13.

Table 3: Analysis by Apriori Algorithm

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Attributes	Min-support, min-confidence	Association Rules Generated	Outcome and Discussion
Program, Gender, Area, Q1	(0.25, 0.6)	1. PROGRAM=BSCS, 483==> Q1=60min 438 conf:(0.91) 2. AREA=Urban, 485 ==> Q1=60 min 437 conf:(0.9) 3. AREA=RURAL, 659 ==> Q1=60 min 583 conf:(0.88) 4. GENDER=M, AREA=RURAL, 498 ==> Q1=60min 434 conf:(0.87) 5. GENDER=M, 976 ==> Q1=60 min 834 conf:(0.85) 6. AREA=RURAL, 659 ==> GENDER=M, 498 conf:(0.76) 7. AREA=RURAL, Q1=60min, 583 ==> GENDER=M 434 conf:(0.74) 8. PROGRAM=BSCS, 483 ==> GENDER=M, 348 conf:(0.72) 9. Q1=60min, 1173 ==> GENDER=M, 834 conf:(0.71) 10. AREA=RURAL, 659 ==> GENDER=M, Q1=60min 434 conf:(0.66)	The interpretation of the above association rules represents that 60 min class is highly preferred by the students having discipline BSCS and they are belonging to rural area and they are mostly Male.
Gender, Area, Q4	(0.2, 0.5)	1. AREA=RURAL 659 ==> GENDER=M 498 conf:(0.76) 2. AREA=RURAL Summer Timings =NO 415 ==> GENDER=M 311 conf:(0.75) 3. Summer Timings =NO 884 ==> GENDER=M 661 conf:(0.75) 4. Summer Timings =YES 458 ==> GENDER=M 315 conf:(0.69) 5. GENDER=M 976 ==> Summer Timings =NO 661 conf:(0.68) 6. AREA=Urban 485 ==> Summer Timings =NO 328 conf:(0.68) 7. AREA=Urban 485 ==> GENDER=M 311 conf:(0.64) 8. AREA=RURAL 659 ==> Summer Timings =NO 415 conf:(0.63) 9. GENDER=M AREA=RURAL 498 ==> Summer Timings =NO 311 conf:(0.62) 10. GENDER=M 976 ==> AREA=RURAL 498 conf:(0.51)	Best rules founds show that the rural area students cannot approach the university in summer season at 7:30am while some students of urban area admit the timings but they have low rating as compared to rural area students.
Program, Emp, Area, Q7	(0.2, 0.6) Consecutive Classes: CC Prefer Break (PB) Leave Early (LE)	1. CC =PB, 520 ==> EMP?=no, 480 conf:(0.92) 2. AREA=RURAL, CC =LE, 386 ==> EMP?=no, 349 conf:(0.9) 3. AREA=RURAL, 659 ==> EMP?=no, 592 conf:(0.9) 4. PROG=BSCS, CC=LE, 372==>EMP?=no, 333 conf:(0.9) 5. PROG=BSCS, AREA=RURAL, 306==>EMP?=no 270 conf:(0.88) 6. PROG=BSCS, 483 ==> EMP?=no, 426 conf:(0.88) 7. CS=LE, 821 ==> EMP?=no, 721 conf:(0.88) 8. AREA=Urban, 485 ==> EMP?=no, 425 conf:(0.88) 9. PROG=BSCS, EMP?=no, 426 ==> CS=LE 333 conf:(0.78) 10. PROG=BSCS, 483 ==> CC=LE 372 conf:(0.77)	Another important observation the association rules depict is that the students want to get free earlier and they did not preferred break between classes they admire consecutive classes and another interesting information is that they are mostly belonging to rural areas the reason to get free earlier can be their distance education and they want to go back to their homes at desired timings.
Program, Gender, Pre-qualification, Q8,Q9,Q13	(0.15, 0.9)	1. PROGRAM=BSCS PRE-QUALIF=FSC learning material =White board 258 ==> Guest spk, Techr =Teacher 249 conf:(0.97) 2. PROGRAM=BSCS PRE-QUALIF=FSC learning material=White board Attitude =Friendly behavior 222 ==> Guest spk, Techr =Teacher 213 conf:(0.96) 3. PROGRAM=BSCS GENDER=M Guest spk, Techr =Teacher 279 ==> learning material =White board 267 conf:(0.96) 4. PROGRAM=BSCS GENDER=M Guest spk, Techr =Teacher Attitude =Friendly behavior 240 ==> learning material =White board 228 conf:(0.95) 5. PROGRAM=BCS 228 ==> Attitude =Friendly behavior 216 conf:(0.95) 6. PROGRAM=BSCS learning material =White board 387 ==> Guest spk, Techr =Teacher 357 conf:(0.92) Attitude	The analysis of the above results shows the trend towards learning material and teachers behavior through which it is concluded that students like teacher's friendly behavior and they like to study through white board and these students of BSCS and have previous qualification FSC. Another observation is that they want lecture demonstration from teachers instead of guest speakers and their gender is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

		=Friendly behavior 336 ==> Guest spk, Techr =Teacher 306 conf:(0.91) 8. PROGRAM=BSCS PRE-QUALIF=FSC Attitude =Friendly behavior 264 ==> Guest spk, Techr =Teacher 240 conf:(0.91) 9. PROGRAM=BCS 228 ==> GENDER=M 207 conf:(0.91) 10. GENDER=M Guest spk, Techr =Guest speaker of special field 314 ==> Attitude =Friendly behavior 284 conf:(0.9)	Male.
Program, Gender, Area, Q33,Q34	(0.25,.5)	1. Teacher Gender =female 476 ==> GENDER=M 395 conf:(0.83) 2. AREA=RURAL 659 ==> GENDER=M 498 conf:(0.76) 3. AREA=Urban 485 ==> language =Mix Eng Urdu 361 conf:(0.74) 4. PROGRAM=BSCS 483 ==> GENDER=M 348 conf:(0.72) 5. language =Mix Eng Urdu 936 ==> GENDER=M 666 conf:(0.71) 6. Teacher Gender =female 476 ==> language =Mix Eng Urdu 336 conf:(0.71) 7. Teacher Gender =male 806 ==> language =Mix Eng Urdu 559 conf:(0.69) 8. GENDER=M 976 ==> language =Mix Eng Urdu 666 conf:(0.68) 9. GENDER= Teacher Gender =male 530 ==> language =Mix Eng Urdu 351 conf:(0.66) 10. Teacher Gender =male 806 ==> GENDER=M 530 conf:(0.66)	Above results show that if teacher is female and language which is mix Urdu English the rating is 64% and if the teacher is male and language is mix Urdu English then the rating is 94% and more interesting thing is that urban area students want to speak mix Urdu English and rural area students like to speak pure English which is point to be considered.
Program, GPA, Pre Qualif, Q20, Q21	(0.2, 0.5)	1. PROGRAM=BSCS PRE QUALIF=FSC 312 ==> Coaching=YES 282 conf:(0.9) 2. Peer Teaching =NO 458 ==> Coaching=YES 396 conf:(0.86) 3. PRE QUALIF=FSC Peer Teaching =YES 338 ==> Coaching=YES 292 conf:(0.86) 4. PROGRAM=BSCS 483 ==> Coaching=YES 408 conf:(0.84) 5. Peer Teaching =YES 569 ==> Coaching=YES 480 conf:(0.84) 6. PRE QUALIF=FSC 754 ==> Coaching=YES 540 conf:(0.72) 7. PROGRAM=BSCS Coaching=YES 408 ==> PRE QUALIF=FSC 282 conf:(0.69) 8. PROGRAM=BSCS 483 ==> PRE QUALIF=FSC 312 conf:(0.65) 9. Coaching=YES 878 ==> PRE QUALIF=FSC 540 conf:(0.62) 10. Peer Teaching =YES Coaching=YES 480 ==> PRE QUALIF=FSC 292 conf:(0.61)	The analysis of the above results shows the trend towards peer teaching and help by the advisor. Through the analysis it is concluded that most of the students are inclined to have coaching from an advisor rather than peer teaching. However, there also exist students that want both.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4: Comparison

Analysis parameter	Per [5]	Per [6]	Per [7]	Per [8]	Proposed
Timetable	No	No	No	No	Yes
Medium	No	No	No	Yes	Yes
Class duration	No	No	No	No	Yes
Assessment type	No	No	No	No	Yes
Gender wise	No	No	No	Only female	Yes
Area wise	No	Yes	No	Yes	Yes
Working students	No	No	No	No	Yes
Previous degree	No	Yes	No	Yes	Yes
Family background	No	No	No	Yes	No
Financial Status	No	No	No	Yes	No
Marital Status	No	No	No	Yes	No
Total attributes	5	6	15	21	38
Student centric or teacher centric	Teacher centric	Student centric	Teacher centric	Student centric	Student + Teacher centric (Hybrid)
Algorithm/s used	Fuzzy System, Apriori Algorithm	Apriori Algorithm	Apriori Algorithm, Weighted average	Naïve Bayesian, Decision Trees	Apriori Algorithm, K-Mean
Dataset	Student surveys	Student surveys	Student surveys	Student surveys	Student surveys