Prediction of Course Selection by Student using Combination of Data Mining Algorithms in E-learning

Sunita B. Aher
M.E. (CSE) -II
Walchand Institute of Technology
Solapur University India

Lobo L.M.R.J
Associate Professor, Head, Department of IT
Walchand Institute of Technology
Solapur University India

ABSTRACT
Course recommender system aims at predicting the best combination of courses selected by students. Here in this paper we present how the combination of clustering algorithm- Simple K-means Algorithm & association rule algorithm- Apriori Association Rule is useful in Course Recommender system. If we use only the Apriori association rule algorithm then we need to preprocess the data obtained from Moodle database. But if we use this combination of clustering & association rule then there is no need to preprocess the data. So we present this new approach & also present the result. To test the result we have used the open source data mining tool Weka.

KEY WORDS
Simple K-means Clustering Algorithm, Apriori association Rule, Weka, Moodle

1. INTRODUCTION
The course recommendation system in e-learning is a system that suggests the best combination of courses in which the students are interested [9].

In this Course Recommendation System, we have considered the 13 course category. Under each category there will courses. So there are about 82 courses. The framework for Course Recommender System is given in the figure 1.
Student first logs in the learning management system e.g. Moodle & enrolled for those courses in which they are interested. This data is stored in the moodle database which we use to find out the best combination. After collecting the data from student which is stored in Moodle database, the next stage is to gather & prepare the data. In this step, first we select the data from database which is relevant. To test the result using Weka i.e. the best combination courses, first we need to preprocess the data & find out the result. The step-build the model, we directly select the relevant data from Moodle database. After collecting the data from Moodle database, we clustered the data using clustering algorithm e.g. Simple K-means algorithm. After clustering data, we use the Apriori Association Rule algorithm to find the best combination of courses. To find the result we need to preprocess the data from Moodle database but if we consider the combination of clustering & association rule algorithm then there is no need to preprocess the data. The preprocessing technique is explained in paper [10].

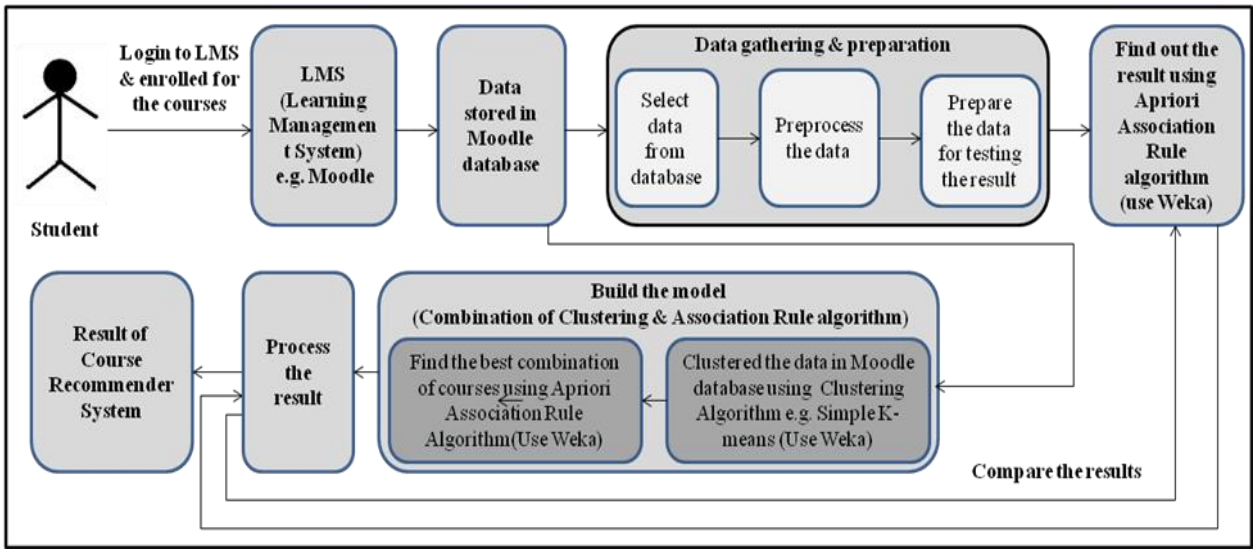

Figure 1: Course Recommender System

2. LITERATURE REVIEW

In paper [2], they explain Aspect oriented programming which offers a unique module, an aspect, to encapsulate scattered and tangled code, which makes it hopeful to solve the problem of crosscutting concerns. Identification and encapsulation of crosscutting concerns is the key problem in the migration from OO system to AO system. A novel aspect mining method which combines clustering and association rule technology is provided in this article. Clustering analysis based on the execution traces is provided to find out candidate aspects; while association rule mining based on the execution traces with ordered call is used to find out the crosscuts. Both the aspect code (advice body) and the crosscuts (pointcuts) are gotten after the above two processes, which constitute the aspect mining process.

In paper [3], they proposed a method for grouping and summarizing large sets of association rules according to the items contained in each rule. They use hierarchical clustering to partition the initial rule set into thematically coherent subsets. This enables the summarization of the rule set by adequately choosing a representative rule for each subset, and helps in the interactive exploration of the rule model by the user. Rule clusters can also be used to infer novel interest measures for the rules.

The main hypothesis discussed in the paper [4] is that Web content analysis can be used to improve Web usage mining results. They propose a system that integrates Web page clustering into log file association mining and uses the cluster labels as Web page content indicators. They experiment with several approaches to content clustering, relying on keyword and character n-gram based clustering with different distance measures and parameter settings.

In paper [5] they present a new approach to cluster web images. A tree-distance-based evaluation measure is used to evaluate the quality of image clustering with respect to manually generated ground truth. The experiments indicate that combining textual and content based features results in better clustering as compared to signal only or text-only approaches.

The main objective of the study [9] was to detect the patterns of use for additional services of mobile telecommunication through cluster analysis. The authors used empirical data of mobile telecommunication service collected from a company in Korea. As a result, it was possible to divide the customers into 8 clusters. After extracting the association rules for the eight clusters, these are the marketing implications.

The research [11] proposed a bracing approach for increasing web server performance by analyzing user behavior, in this pre-fetching and prediction is done by pre-processing the user access log and integrating the three techniques i.e. Clustering, Markov model and association rules which achieves better web page access prediction accuracy. This work also overcomes the limitation of path completion i.e. by extracting web site structure paths are completed, which helps in better prediction, decreasing access time of user and improving web performance.

The main objective of study [12] was to detect the patterns of use for additional services of mobile telecommunication through cluster analysis. The authors used empirical data of mobile telecommunication service collected from a company in Korea. As a result, it was possible to divide the customers into 8 clusters. After extracting the association rules for the eight clusters, these are the marketing implications.

3. DATA MINING ALGORITHMS

Here we consider the brief idea about the two data mining algorithm: Clustering algorithm & Association rule algorithm.

3.1 Clustering Algorithm

Clustering is finding groups of objects such that the objects in one group will be similar to one another and different from the objects in another group. Clustering can be considered the most important unsupervised learning technique.

3.1.1 Simple K-means Clustering algorithm

Simple K-means algorithm is a type of unsupervised algorithm in which items are moved among the set of cluster until required set is reached. This algorithm is used to classify the data set, provided the number of cluster is given in prior. This algorithm is iterative in nature.

The algorithm for Simple K-means algorithm is given in figure 2:

Algorithm: Simple K-means clustering algorithm
Input:
Set of Elements or Database of transaction
D= $\{t_1, t_2, t_3, \ldots, t_n\}$
Number of required Cluster k
Output:
Set of Cluster K
Method:
Make initial guesses for the means $\mathbf{m_1, m_2, ..., m_k}$;
Repeat
Assign each element t_i to the cluster having the closest mean.
Calculate the new mean for each cluster.
Until there are no changes in any mean

Figure 2: Simple K-means Clustering Algorithm

Consider the set of elements to be clustered as:
{2, 4, 12, 10, 3, 20, 31, 11, 25}
Suppose the number of cluster required (k) is 2 & initially consider the values of mean to the first two values: m1=2 & m2=4

m1	m2	K1	K2
2	4	{2,3}	{4,10,11,12,20,25,31}
2.5	16	{2,3,4}	{10,11,12,20,25,31}
3	18	{2,3,4,10}	{11,12,20,25,31}
4.75	19.8	{2,3,4,10,11,12}	{20,25,31}
7	25	{2,3,4,10,11,12}	{20,25,31}

So two clusters are K1= {2, 3, 4, 10, 11, 12} & K2= {20, 25, 1}

3.2. Association Rule Algorithm

Association rules are used to show the relationship between data items. Association rule generation is usually split up into two separate steps: First, minimum support is applied to find all frequent itemsets in a database. Second, these frequent itemsets and the minimum confidence constraint are used to form rules.

3.2.1 Apriori Association Rule

Apriori Association rule is used to mine the frequent patterns in database. Support & confidence are the normal method used to measure the quality of association rule. Support for the association rule X->Y is the percentage of transaction in the database that contains XUY. Confidence for the association rule is X->Y is the ratio of the number of transaction that contains XUY to the number of transaction that contain X.

The Apriori association rule algorithm is given in figure 3.

Algorithm: Apriori Algorithm
Purpose : To find subsets which are common to at least a minimum number C (Confidence Threshold) of the itemsets.
Input : Database of Transactions D={t_1,t_2, …, t_n}
Set if Items I= {I_1 ,I_2 ,….I_k}
Frequent (Large) Itemset L
Support,
Confidence.
Output : Association Rule satisfying Support & Confidence
Method :

C_k= Candidate itemset of size k
L_k = frequent itemset of size k
C_1= Itemset of size 1 in I
L_1 = {frequent items of size 1} i.e. L1 consists of the candidate 1-itemsets satisfying Minimum Support.
k=1;
repeat
  k=k+1;
  C_{k+1} = candidates generated from L_{k-1}
  for each transaction t in database D do
  begin
    Find out the support of candidate item C_{k+1}
    Compare candidate support count of C_{k+1} with minimum support
  Reduce the unfrequented k-itemsets from this set i.e. any k-itemset that is not frequent cannot be a subset of (k+1) itemset
  end
  L_{k+1} = candidates in C_{k+1} with Minimum support
until no more large itemset found.
Return $\cup_k L_k$.

Figure 3: Apriori association Rule Algorithm [10]

5. RESULT & IMPLEMENTATION

Here we are considering the sample data extracted from Moodle database as shown in Table 1. Here we consider 45 student & 15 courses. We consider the courses like C-programming (C), Visual Basic (VB), Active Server Pages (ASP), Computer Network (CN), Network Engineering (NE), Microprocessor (MP), Computer Organization (CO), Database Engineering (DBE), Advanced Database System (ADS), Operating System (OS), Distributed System (DS), Finite Automata System (FSA), Data Structure (DS-I), Software Engineering (SE), and Software Testing & Quality assurance (STQA). In this table yes represent that the student is interested in that course & no represent that student do not like that course .

In preprocessing step, we delete those rows & columns from sample table shown in table 1, having very less student count & less course count. After preprocessing of data we got 8 courses & 38 rows i.e. 38 students. These 8 courses are C-programming (C), Visual Basic (VB), Active Server Pages (ASP), Computer Network (CN), Network Engineering (NE), Operating System (OS), Distributed System (DS), Data Structure (DS-I). The graph for sample data before preprocessing & after preprocessing is shown in figure 4 & 5 respectively. In these graph the X-axis represents the courses & Y-axis represents the course count. The sample table after preprocessing of data is shown in table 2.

The result of applying Apriori association rule before & after preprocessing of data is shown in first & second row of table 4. Before preprocessing of data, we got the association rule containing "no" only. As we are recommending the course, we preprocess the data. The result after preprocessing of data is shown in second row of table 4. Now the association rule contains only "yes". The meaning of the association rule "DS=yes -> OS=yes" is that we can recommend to new student who has recently enrolled for DS course, the operating system as a course to be opted.

If we consider combination of clustering & association rule then there is no need preprocess the data. First we apply Simple K-means clustering algorithm to data selected from Moodle course. Here we consider the two cluster 0 & cluster 1. We apply the Apriori Association rule mining to both clusters. But using cluster1 we are getting the correct association rule as association rule contains only "yes". We ignore the cluster 0 as the association rule after applying the Apriori association rule algorithm contains "no" also. After clustering of data, we got the table 3. The result of applying the clustering algorithm-Simple K-means algorithm on sample data shown in table 1 is shown in third row of table 4. We consider two clusters as shown in table 4.

Table 1: Sample data from Moodle database [9]

Courses → Roll_No | v	C	VB	ASP	CN	NE	MP	CO	DBE	ADS	OS	DS	FSA	DS-I	SE	STQA
1	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	yes	no	no
2	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
3	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	yes	yes
4	no	no	no	yes	yes	no	yes	no	no	no	no	no	no	no	no
5	yes	yes	yes	yes	yes	no	no	yes	no	yes	yes	no	yes	no	no
6	yes	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no
7	no	no	no	yes	yes	yes	yes	no	no	no	no	no	no	yes	no
8	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
9	no	no	no	yes	yes	yes	yes	no	no	no	no	yes	no	no	no
10	yes	no	no	no	no	no	no	no	no	no	no	no	no	no	no
11	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
12	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
13	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	yes	yes
14	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	no	no	no	no
15	yes	yes	yes	no	no	no	no	no	no	no	no	no	yes	no	no
16	no	no	no	yes	yes	no	no	yes	yes	yes	yes	no	yes	no	no
17	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
18	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
19	no	no	no	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
20	yes	no	no	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
21	yes	no	yes	no	no	yes	yes	no	no	yes	yes	yes	no	no	no
22	no	no	no	no	no	no	no	yes	yes	yes	yes	no	yes	no	no

23	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	no	yes	no	no
24	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
25	no	yes	yes	no	no	yes	yes	yes	yes	yes	yes	no	no	no	no
26	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
27	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
28	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
29	no	no	no	no	no	yes	yes	yes	yes	no	no	no	no	no	no
30	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	yes	yes
31	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no
32	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
33	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
34	yes	yes	yes	no	no	no	no	no	no	no	no	no	no	no	no
35	no	no	no	no	no	no	no	no	no	yes	yes	no	no	no	no
36	no	no	no	yes	yes	no	no	no	no	no	no	no	yes	no	no
37	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
38	no	no	no	no	no	no	no	no	no	yes	yes	yes	yes	yes	yes
39	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
40	no	no	no	no	no	no	no	no	no	no	no	no	no	yes	yes
41	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
42	no	no	no	yes	yes	no	no	no	no	no	no	no	no	no	no
43	no	no	no	no	no	no	no	no	no	yes	yes	no	yes	no	no
44	no	no	no	no	no	no	no	no	no	no	no	no	no	no	yes
45	no	no	no	no	no	no	no	no	no	no	no	no	no	no	no

Table 2: Sample table 1 after preprocessing of data [9]

Subject → Roll No of Student ↓	C	VB	ASP	CN	NE	OS	DS	DS-I
1	yes	yes	yes	yes	yes	no	no	yes
3	yes	yes	yes	yes	yes	yes	yes	yes
4	no	no	no	yes	yes	no	no	no
5	yes	yes	yes	yes	yes	yes	yes	yes
6	yes	yes	yes	no	no	yes	no	yes
7	no	no	no	yes	yes	no	no	no
8	no	no	no	no	no	yes	yes	yes
9	no	no	no	yes	yes	no	no	no
11	yes	yes	yes	no	no	yes	yes	yes
12	yes	yes	yes	yes	yes	no	no	no
13	no	no	no	no	no	yes	yes	yes
14	yes	yes	yes	yes	yes	yes	yes	no
15	yes	yes	yes	no	no	no	no	yes
16	no	no	no	yes	yes	yes	yes	yes
17	yes	yes	yes	no	no	yes	yes	yes
18	yes	yes	yes	no	no	no	no	no
19	no	no	no	yes	yes	no	no	no
20	yes	no	no	no	no	yes	yes	yes
21	yes	no	yes	no	no	yes	yes	no
22	no	no	no	no	no	yes	yes	yes
23	yes	yes	yes	yes	yes	yes	yes	yes
24	yes	yes	yes	yes	yes	yes	yes	yes
25	no	yes	yes	no	no	yes	yes	no
26	yes	yes	yes	no	no	yes	yes	yes
27	yes	yes	yes	yes	yes	no	no	no
28	no	no	no	yes	yes	yes	yes	yes
30	yes	yes	yes	yes	yes	no	no	no
32	yes	yes	yes	no	no	yes	yes	yes
33	no	no	no	yes	yes	yes	yes	yes
34	yes	yes	yes	no	no	no	no	no
35	no	no	no	no	no	yes	yes	no
36	no	no	no	yes	yes	no	no	yes
37	yes	yes	yes	yes	yes	no	no	no
38	no	no	no	no	no	yes	yes	yes
39	yes	yes	yes	yes	yes	yes	yes	yes
41	yes	yes	yes	no	no	yes	yes	yes
42	no	no	no	yes	yes	no	no	no
43	no	no	no	no	no	yes	yes	yes

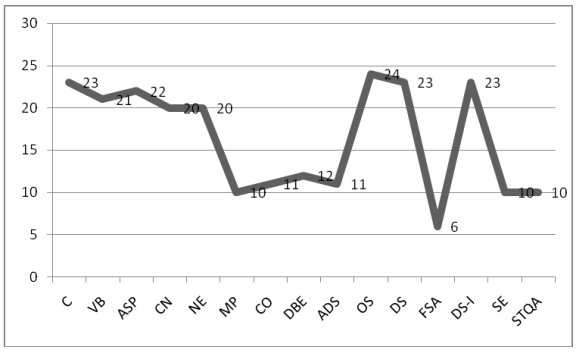

Figure 4: Sample table before preprocessing

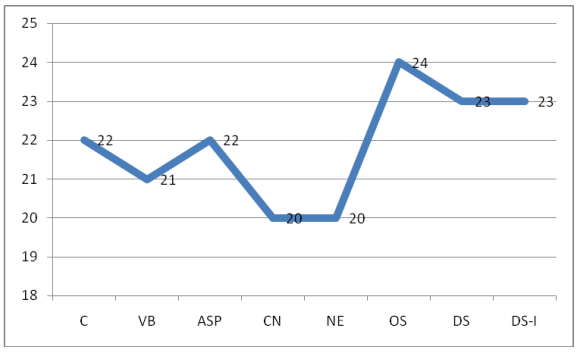

Figure 5: Sample data after preprocessing

Table 3: Table after application of clustering algorithm- Simple K-means algorithm

Courses → Roll_No | v	C	VB	ASP	CN	NE	MP	CO	DBE	ADS	OS	DS	FSA	DS-I	SE	STQA
1	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	yes	no	no
3	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	yes	yes
5	yes	yes	yes	yes	yes	no	no	yes	no	yes	yes	no	yes	no	no
6	yes	yes	yes	no	no	no	no	no	no	yes	no	no	yes	no	no
11	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
12	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
14	yes	yes	yes	yes	yes	no	no	no	no	yes	yes	no	no	no	no
16	no	no	no	yes	yes	no	no	yes	yes	yes	yes	no	yes	no	no
17	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	yes	yes
23	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes	no	yes	no	no
24	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
26	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no
27	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	no	no
28	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
30	yes	yes	yes	yes	yes	no	no	no	no	no	no	no	no	yes	yes
32	yes	yes	yes	no	no	no	no	yes	yes	yes	yes	no	yes	no	no
33	no	no	no	yes	yes	no	no	no	no	yes	yes	no	yes	no	no
37	yes	yes	yes	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
39	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
41	yes	yes	yes	no	no	no	no	no	no	yes	yes	no	yes	no	no

We use the combination of clustering algorithm- Simple k-means & association rule- Apriori Association Rule Algorithm to increase the strength of the association rule. Result of Apriori Association Rule algorithm after application of Simple K-means clustering algorithm is shown in fourth row of table 4. Cluster 0 result is incorrect as association rule contains "no". Result using cluster 1 contains "yes" only, so we these result to recommend the course to the student.

Table 4: Result after application of data mining algorithms

Course considered	**Parameter Considered**	**Results**
Result of Apriori Association Rule before preprocessing & application of combination of Clustering & Association Rule		
C, VB, ASP, CN, NE, MP, CO, DBE, ADS, OS, DS, FSA, DS-I, SE, STQA	Minimum support: 0.7 Minimum metric <confidence>: 0.9	Best rules found: 1. CO=no → MP=no 2.DBE=no → ADS=no 3.CO=no FSA=no → MP=no 4.MP=no → CO=no 5.STQA=no → SE=no 6.SE=no → STQA=no 7.ADS=no → DBE=no 8.MP=no FSA=no →CO=no 9.FSA=no STQA=no → SE=no 10. FSA=no SE=no → STQA=no
Result of Apriori Association Rule after preprocessing & before application of combination of Clustering & Association Rule		
C, VB, ASP, CN, NE, OS, DS, DS-I	Minimum support: 0.5 Minimum metric <confidence>: 0.9	Best rules found: 1. DS=yes →OS=yes 2. VB=yes →ASP=yes 3. NE=yes → CN=yes 4. CN=yes → NE=yes 5. C=yes VB=yes → ASP=yes 6. DS=yes DS-I=yes → OS=yes 7. OS=yes →DS=yes 8. ASP=yes →C=yes 9. C=yes →ASP=yes 10.ASP=yes →VB=yes
	Minimum support: 0.6 Minimum metric <confidence>: 0.9	Best rules found: 1. DS=yes →OS=yes 2.OS=yes →DS=yes

Result After Application of Clustering algorithm-Simple K-means		
C, VB, ASP, CN, NE, MP, CO, DBE, ADS, OS, DS, FSA, DS-I, SE, STQA	Number of Cluster:2 Seed: 10	Cluster 0 Mean/Mode: no no no no no no no no no no no no no no Std Devs: N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A Cluster 1 Mean/Mode: yes yes yes yes yes no no no no yes yes no yes no no Std Devs: N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A N/A Clustered Instances 0 25 (56%) 1 20 (44%)
Result After Application of Clustering algorithm-Simple K-means & Association Rule-Apriori Association Rule		
C, VB, ASP, CN, NE, MP, CO, DBE, ADS, OS, DS, FSA, DS-I, SE, STQA	Minimum support: 0.85 Minimum metric <confidence>: 0.9	Cluster 1 Result (Correct Result) Best rules found: 1. VB=yes →C=yes 2. C=yes →VB=yes 3. ASP=yes →C=yes 4. C=yes →ASP=yes 5. ASP=yes →VB=yes 6. VB=yes →ASP=yes 7. VB=yes ASP=yes →C=yes 8. C=yes ASP=yes →VB=yes 9. C=yes VB=yes →ASP=yes 10. ASP=yes →C=yes VB=yes
	Minimum support: 0.7 Minimum metric <confidence>: 0.9	Cluster 0 Result (Incorrect Result) Best rules found: 1. ASP=no →VB=no 2. NE=no →CN=no 3. CN=no →NE=no 4. ADS=no →DBE=no 5. DBE=no →ADS=no 6. CO=no →MP=no 7. C=no ASP=no →VB=no 8. C=no VB=no →ASP=no 9. ASP=no FSA=no →VB=no 10. VB=no FSA=no →ASP=no

Figure 6 represents the graph for courses before & after application of Simple K-means clustering algorithm. In this graph X-axis represents the courses & Y-axis represents the course count for each count for each course. In this graph ST line represents line for the sample table. SKM1 Correct line represents the line for the cluster 1 after the application of Simple K-means clustering algorithm which is the correct cluster for the application of association rule. SKM0 line represents the line for cluster 0 which gives the incorrect association rule after applying the Apriori Association Rule Algorithm.

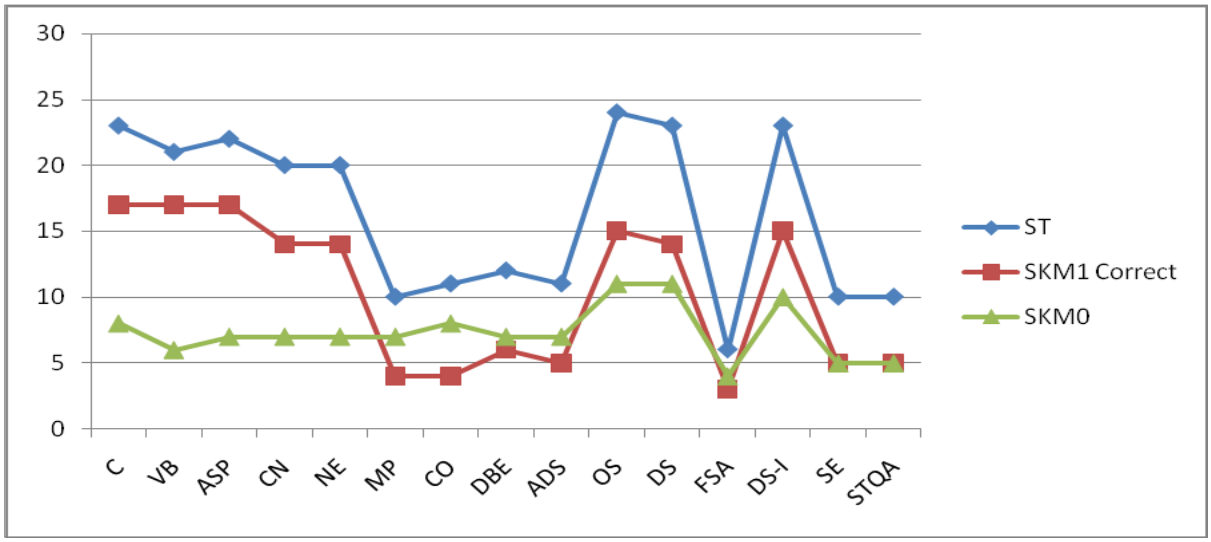

Figure 6: Graph for courses before & after application of Simple K-means Clustering algorithm

6. CONCLUSION AND FUTURE WORK

In this paper, we try to present how the combination of two data mining algorithm i.e. simple K-means clustering algorithm & Apriori association rule algorithms are useful in recommending the courses to the student. If we consider only Apriori association rule mining then we need to preprocess the data from Moodle database. But if we use the combination of clustering & association rule algorithm then there is no need to preprocess the data. Also we find that the combination of these algorithm increase the strength of association rule. Future work includes the atomization of this combination algorithm & testing the result on huge amount of data.

7. REFERENCES

[1] Castro, F., Vellido, A., Nebot, A., & Mugica, F. (in press). Applying data mining techniques to e-learning problems: A survey and state of the art. In L. C. Jain, R. Tedman, & D. Tedman (Eds.), Evolution of Teaching and learning paradigms in intelligent environment. Studies in Computational Intelligence (Vol. 62). Springer-Verlag.

[2] Lili He, Hongtao Bai:"Aspect Mining Using Clustering and Association Rule Method" IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.2A, February 2006

[3] Al¶³pio Jorge:"Hierarchical Clustering for thematic browsing and summarization of large sets of Association Rules" Supported by the POSI/SRI/39630/2001/Class Project

[4] Jiayun Guo, Vlado Keˇselj, and Qigang Gao:"Integrating Web Content Clustering into Web Log Association Rule Mining?" supported by NSERC

[5] Hassan H. Malik, John R. Kender:"Clustering Web Images using Association Rules,Interestingness Measures, and Hypergraph Partitions" *ICWE'06*, July 11–14, 2006, Palo Alto, California, USA. ACM 1-59593-352-2/06/0007.

[7] "Data Mining Introductory and Advanced Topics" by Margaret H. Dunham

[8] Sunita B Aher and Lobo L.M.R.J.. Data Mining in Educational System using WEKA. *IJCA Proceedings on International Conference on Emerging Technology Trends (ICETT)* (3):20-25, 2011. Published by Foundation of Computer Science, New York, USA (ISBN: 978-93-80864-71-13)

[9] Sunita B Aher and Lobo L.M.R.J. Article: A Framework for Recommendation of courses in E-learning System. *International Journal of Computer Applications* 35(4):21-28, December 2011. Published by Foundation of Computer Science, New York, USA **ISSN 0975 – 8887**

[10] Sunita B Aher and Lobo L.M.R.J.: "Preprocessing Technique for Association Rule Based Course Recommendation System in E-learning" selected in ICECT-12, proceeding published by IEEE

[11] Silky Makker, R.K Rathy: "Web Server Performance Optimization using Prediction Prefetching Engine" *International Journal of Computer Applications (0975 – 8887) Volume 23– No.9, June 2011*

[12] So Young Sohn, Yoonseong Kim" Searching customer patterns of mobile service using clustering and quantitative association rule" Published in journal Expert Systems with Applications: An International Journal, Volume 34 Issue 2, February, 2008