



# Analysis and prediction of student placement for improving the education standards

S Anjali Devi \*, M Vishnu Priya, P Akhila, N Vasundhara

Department of Computer Science and Engineering, Koneru Lakshamiah Education Foundation,  
Vaddeswaram, Guntur, Andhra Pradesh, India

\*Corresponding author E-mail: [swarnaanjalidevi@kluniversity.in](mailto:swarnaanjalidevi@kluniversity.in)

## Abstract

Students' academic success can be evaluated based on their performance in the exams conducted by the institutions. In this paper, we propose a scheme where prediction of student final placement can be done based on the marks scored by them in the previous semesters. In order to predict the placement of the student we need some data to analyze. For this purpose we will supply students basic details and their previous academic information into the system which will be used to predict the placement of the student. This is done by generating association rules using apriori algorithm. Admin and user will use this system. Here user will be the student. Admin and user will use their login to access the system. Admin will add academic details of the students, like their SSC, HSC, Graduation marks (up to current semester, Back logs etc.). User will be the student. Admin and user will use their login to access the system. Admin will add academic details of the students, like their SSC, HSC, Graduation marks (up to current semester, Back logs etc.). This system can be used in schools, colleges and other educational institutions. This evaluation system is more accurate than other conventional methods. We are using a university data set to predict the placement of the student.

**Keywords:** J48 algorithm, Decision tree, Association rules, Academics, Prediction, Placement, Universities.

## 1. Introduction

Colleges have extensive quantities of module decisions, and it is trying for understudies to acquaint themselves with every one of the additionally vital to foresee general result and result in particular modules, given specific module decisions. This can be the reason for a recommender framework to help under studies in module choice. Recommender frameworks are at present considered as a prudent robotized answer for helping understudies with their decisions [5]. In this paper we endeavor to utilize information mining procedures to foresee understudies' results in view of early module execution and other understudy attributes. In the event that our techniques are effective for anticipating the more broad issue of understudy great respects execution, we would then be able to deliver more granular pre-word usages at the module level and those would shape the reason for a recommender framework. This paper reveals early pointers of dull execution that might be utilized to target medicinal activity for the concerned understudies. We expect to explore the accessible highlights that might be utilized for forecast, and in addition the kind of classifiers that may create the best outcomes. Understudies' scholarly achievement is evaluated based on the performance of the student in the exams conducted by the institutes or Universities[1]. In this paper, we propose a scheme where prediction of student final placement can be done based on the marks scored by them in the previous semesters.

### Classification:

Classification is done through

1. Classification rules
2. Decision tree
3. Classification formulae

This project is done by using J48 algorithm.[10].The following issues are faced by most the j48 algorithm.

- finding the splitting attributes
- Order of splitting attributes
- Total Number of the splits to be formed
- Balance of tree structure
- The ending criteria [10].

## 2. Literature survey

### 2.1 Data Mining Techniques to Predict Students at risk of poor performance

The accomplishment of good honors in Undergraduate degrees is critical with regards to Higher Education (HE), both for understudies and for the institutions that host them. Regardless of whether information mining can be utilized to feature execution issues at an early stage and propose therapeutic activities is clarified. A few data mining techniques were utilized to anticipate students 'outcomes in view of early module execution and other understudy qualities. The information was recovered from the data warehouse of a university, where data of the understudies and the results are gathered. Credits that identify with understudy demo designs and other preferences include[1]:

- gender;
- age during admission;
- over all score in year 1;
- over all score in year 2;

- over all score in year 3;

## 2.2 Predicting students' final passing results using the Apriori Algorithm

This paper predicts the final results of the students based on their performance in all subjects in the four semesters during their study period. In this research, a data mining task called an association was employed. This association is done through a technique called as the Apriori Algorithm. The rules were generated by SPSS Clementine 10.1 software.

## 2.3 Data mining: Concepts and techniques, in Cluster Analysis: Basic Concepts and Methods, (2011)

This paper explains about clustering and the need for clustering. The accumulation of information questions that are indistinguishable to another inside a bunch and non-indistinguishable to the articles in the other bunch called Cluster. This kind of procedure of collection comparable items into classes is called "Grouping". Common prerequisites of grouping include versatility, capacity to manage different kinds of information and characteristics, deciding information parameters, revelation of bunching oppressive shape and furthermore interpretability, convenience. The categories of fundamental clustering are: grid based methods, partitioning methods hierarchical methods and density-based methods[7]. K-Means clustering falls under this partitioning method. This creates k partitions as starting set; here k is the number of partitions to build. Objects are moved from one group to another to upgrade the partitioning. This technique is called Iteration relocation.

## 2.4 Improving Efficiency of Apriori Algorithm Using Transaction Reduction:

This paper explains about the experimental study on Apriori algorithm using transaction reduction by using association rules. The suggested system not just streamlines the calculation of lessening the measure of the hopeful arrangement of k thing sets, Ck, yet in addition limits the I/O spending by cutting down exchange records in the database. The fundamental favorable position of this paper is that the Apriori calculation can be used to find the frequent item sets and consequently association rules. It can be used to predict the customer purchase in supermarket.

The Apriori algorithm is "bottom up" approach but this implementation has a drawback .ie, expect Exchange database is memory inhabitant. Requires numerous database examines. Give us a chance to think about the case of market bin examination. It gives understanding into which items have a tendency to be bought together and which are most reasonable for advancement and noteworthy guidelines. Minor principles, individuals who purchase chalk-piece additionally purchase duster. Illogical, individuals who purchase portable additionally purchase pack.

## 2.5 Recommendation in Higher Education Using Data Mining Techniques

This paper describes about the enrolment of student course depends only on his/her decision. This are classified characterized into three primary classifications :content based ,shared construct and situated in light of recommender frameworks approaches .This paper portrays about different restrictions and describe various extensions that can make recommender systems more applicable in various fields in a better way with more accurate result [5]. This extensions includes various information regarding the students and their courses.

In this paper, they have researched about students academics and students to take right decision to select the courses. It provides the bolster for the understudies to pick better what number of and which courses to select on the premise of previous student data

with similar academic achievements. Based on the analysis of previous students data a recommendation system was developed.

## 2.6 Mining frequent patterns without candidate generation: A frequent-pattern tree approach (Apr 2004)

Mining incessant examples in time-arrangement database, exchange databases and numerous other various types of databases has been contemplated in information mining research. As, Apriori is exorbitant, particularly when there are enormous number of examples, here arrangement a model incessant case tree (FP-tree) structure, which is an offer prefix-tree enclosing for securing solidly stuffed, basic information about customary cases, and develop a beneficial FP-tree based trench procedure, FP-improvement, for trench the whole course of action of nonstop cases by adornment piece advancement space. The results which exhibits that the technique is productive and versatile for mining both long and short capacity enrichment, and is around a demand of size snappier than the Apriori count and speedier than a few as of late detailed new continuous enhancement mining strategies is FP-tree.

## 2.7 Bayesian Network for Data Mining

This paper explains Bayesian network that can be utilized to learn causal connections, additionally enables pick up to better comprehension about the issue area and to foresee the outcomes. This is a perfect portrayal for consolidating earlier learning and information. Likewise in this paper, there are distinctive strategies for the development of Bayesian systems from earlier information and get a thought of Bayesian measurable techniques for utilizing information to enhance these models.

## 2.8 The Third Information Systems International Conference A Review on Predicting Student's Performance using Data Mining Techniques.

This paper describes about the student's execution by utilizing information mining strategies to enhance the understudies' accomplishments. The fundamental target of this paper is to give a diagram on the information mining systems that have been utilized to foresee understudy execution. Understudy execution is anticipated in light of the last grades, evaluation check, last, most decisive test score and extracurricular exercises. This paper has assessed past examinations on anticipating student execution with different logical techniques. The vast majority of the specialists have utilized Total Review Point Normal (CGPA) and inward appraisal as informational indexes. Neural system and choice tree are two strategies exceptionally utilized by the analysts for foreseeing understudies execution. Based on prediction techniques students performance is evaluated.

## 2.9 Student Relationship in pursuing Higher Education Using Data Mining Techniques.

This paper describes about the students enrolment in the higher education's system to know them well, one approach to do this is by utilizing substantial administration and handling of understudies database. This may show at a beginning time which sort of understudies will possibly be selected and the regions to think upon advanced education frameworks for help. This depends on the choice tree as an order strategy. The created rules are examined and assessed utilizing diverse assessment strategies or techniques and the primary characteristics that may influence the understudy's devotion have been featured. This created rules is constructed utilizing VB.net programming dialect which enables the advanced education frameworks to foresee the steadfastness.

### 3. Overall survey

This survey helps to propose a model that analyzes student’s performance which helps in the prediction of student placement. To implement this model we need to collect the data of B.Tech prior placed student data(training set) and current pursuing student data(testing set). Then J48(Decision tree) algorithm is applied on the placed student data to generate a tree with specific splitting attributes and criteria. This tree generated from the training set help in identifying the placement of the B.Tech pursuing students. The tree generated can be verified using the test set given. This model identifies the students who are weak in their academic career and helps them in increasing their skills which helps them in getting placement.

### 4. Future Enhancement

Future research work includes improving the accuracy in prediction and providing the organization with the necessary information about the performance of the student. Accuracy includes decreasing the false positives and false negatives. This can be done by increasing the attributes. Then the details about the students such as in which subject he/she is weak should be provided to the organization. This helps the student in increasing his/her performance in the future. This model also helps the organization in improving their success rate.

### 5. Results

These are the results generated after doing in the weka tool[11]. In fig1 the details of students are represented. In the fig.2 these details are loaded to the weka tool [11].

	dept	cgpa	computer ssc	intermedi speaking	backlog	placemen	intership	higher studies		
1										
2										
3	cse	5	good	77	65	good	2	not placed	0	n
4	ece	6	good	76	88	good	1	not placed	0	n
5	cse	7	poor	87	87	avg	0	placed	1	n
6	eee	7.4	poor	73	78	good	0	placed	1	n
7	cse	8.6	good	79	89	avg	0	placed	1	n
8	eee	7.8	good	89	99	good	0	placed	0	y
9	cse	8.9	avg	67	76	avg	0	placed	1	n
10	it	9	avg	78	87	avg	0	placed	1	n
11	ece	6.87	good	87	78	good	1	not placed	0	y
12	cse	9.76	avg	99	98	avg	0	placed	1	n
13	ece	8.54	good	95	89	avg	0	placed	1	n
14	eee	7.67	good	67	76	good	1	placed	0	y
15	civ	7.567	avg	83	81	good	0	placed	1	y
16	eee	8.23	avg	81	86	avg	0	placed	1	n
17	ece	7.54	good	71	78	good	0	placed	1	n

Fig. 1: Student information

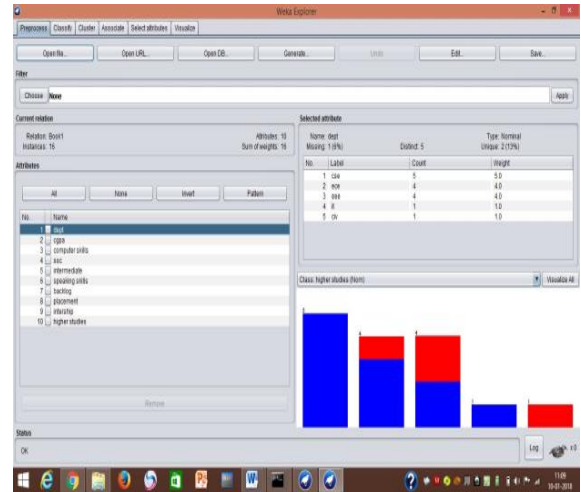


Fig. 2: Data is stored in training set.

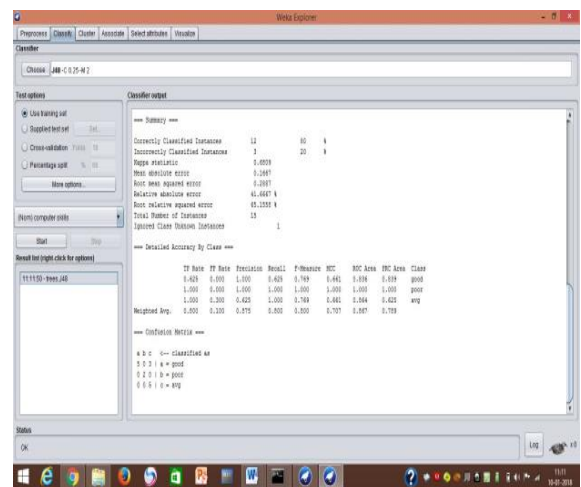


Fig. 3: Results of the test set.

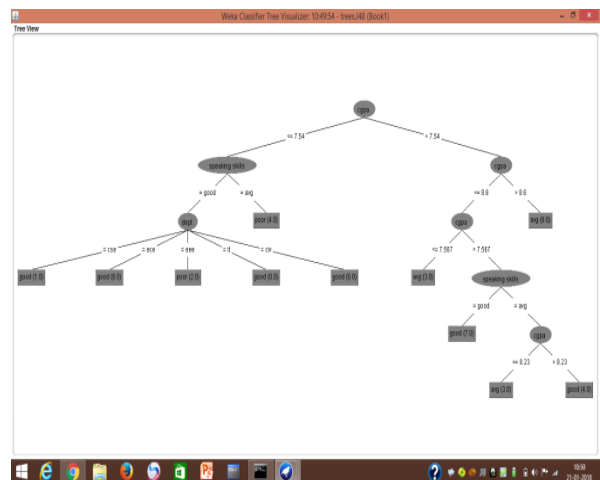


Fig. 4: Weka Classifier Tree

Table 1: Confusion matrix

	A	B	C
18	0	0	0
0	6	0	0
0	0	0	14

Here A= good, B= bad ,C=avg

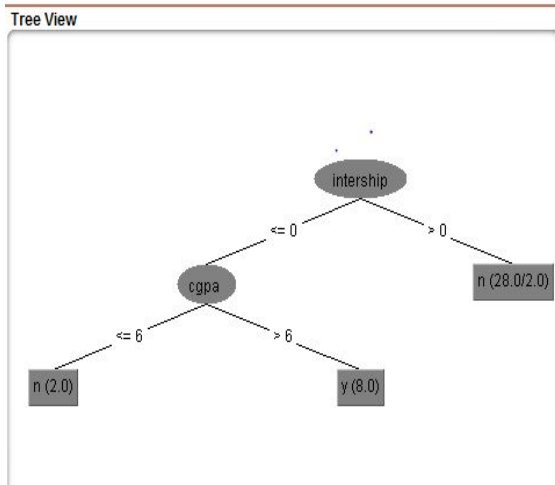


Fig. 5: Weka Classifier Tree

Table 2: Confusion matrix

	A	B
	28	0
	2	8

Here a = n b = y.

	TABLE1	TABLE2
Correctly Classified Instances	38	36
Incorrectly Classified Instances	0	2
Kappa statistic	1	0.855
Mean absolute error	0	0.0977
Root mean squared error	0	0.2211

## 6. Conclusion

World is becoming more competitive day-by-day. So, there is a need for evaluating one’s own performance frequently. So, through this paper we are proposing a model which helps in improving the students’ performance. By having an overview of all the models we propose that j48 algorithm suits best for the data set chosen for placement prediction.

## References

[1] Zahyah Alharbi, “Using Data Mining Techniques to Predict Students at risk of poor performance”. SAI Computing Conference 2016 July 13-15, 2016 | London, UK 978-1-4673-8460-5/16/\$31.00 2016 IEEE,pp 523-532.

[2] Julianti Kasih, Mewati Ayu† & Sani Susanto Transactions on Engineering and Technology Education Vol.11, No.4,2013.

[3] S. Shrimohammadi, A. Yassine and S. Singh, “Incremental mining of frequent power consumption patterns from smart meters big data”, in Proc. IEEE Electrical Power Energy Conf. (EPEC), Oct. 2016.

[4] Jaishree Singh, Hari Ram, Dr. J.S.Sodhi. International Journal of Scientific and Research Publications, Volume 3, Issue 1, January 2013 ISSN 2250-3153

[5] Recommendation in Higher Education Using Data MiningTechniques. Educational data mining 2009.

[6] T.Za\_bkowski and K. Gajowniczek, “Data mining techniques for detecting House hold characteristics based on smart meter data,” Energies, vol. 8, no.7, 2015.

[7] J. Han, J. Pei, and M. Kamber, “Data mining: Concepts and techniques,”in Cluster Analysis: Basic Concepts and Methods,

3rd ed. San Mateo, CA, USA: Morgan Kaufmann, 2011.

[8] Amirah Mohamed Shahiria., Wahidah Husaina, Nur’aini Abdul Rashida. Procedia Computer Science 72 (2015) 414 – 422.

[9] Boumedyen Shannaq, Yusupov Rafael, V. Alexandro. Vol. 10 Issue 11 (Ver.1.0) October 2010 Global Journal of Computer Science and Technology

[10] Hitarthi Bhatt et al. / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (5), 2015, 4785-4789.

[11] Dr. Seetaiah Kilaru, Hari Kishore K, Sravani T, Anvesh Chowdary L, Balaji T “Review and Analysis of Promising Technologies with Respect to fifth Generation Networks”, 2014 First International Conference on Networks & Soft Computing, ISSN:978-1-4799-3486-7/14,pp.270-273, August 2014.

[12] Meka Bharadwaj, Hari Kishore "Enhanced Launch-Off-Capture Testing Using BIST Designs" Journal of Engineering and Applied Sciences, ISSN No: 1816-949X, Vol No.12, Issue No.3, page: 636-643, April 2017.

[13] T. Padmapriya and V. Saminadan, “Distributed Load Balancing for Multiuser Multi-class Traffic in MIMO LTE-Advanced Networks”, Research Journal of Applied Sciences, Engineering and Technology (RJASET) - Maxwell Scientific Organization, ISSN: 2040-7459; e-ISSN: 2040-7467, vol.12, no.8, pp:813-822, April 2016.

[14] S.V.Manikanthan and K.srividhya "An Android based secure access control using ARM and cloud computing", Published in: Electronics and Communication Systems (ICECS), 2015 2nd International Conference on 26-27 Feb. 2015,Publisher: IEEE,DOI: 10.1109/ECS.2015.7124833.