Performance Assessment and Classification of Students using CAMP Model with Data Mining Approach

Nayani Sateesh
Department of IT, CVR College of Engineering, Hyderabad, India
Email: nayanisateesh@gmail.com

Abstract - **With the advent of large number of professional colleges, especially with engineering colleges, the Quality of education is becoming compromised and unable to meet the demand vs. supply of Qualified and employable professionals in IT, ITES and other sectors. It needs lot of knowledge and skill based training programs to make the students employable and knowledgeable. In this paper the focus is on how best it can assess the students' performance and hence to provide the needful training and value added courses to bridge the gap to overcome the shortfalls. For this Assessment process CAMP (Conceptual, Analytical, Memory-Based, Practice-Based) model with data mining approach (EDM – Educational Data mining) is being utilized**

Index Terms—**Data Mining, Classification, Clustering, Rule based classification, Educational Data Mining (EDM).**

I. INTRODUCTION

The assessment process for evaluating the students' performance in the educational institutions for the past several years are based on the Quantitative rather than qualitative which is creating a gap between the supply vs. demand of employable graduates to the industries. This leads to a need for revising the assessment process and to train the graduates accordingly to make them employable. In this paper a mechanism is proposed to classify the students based on their performance using CAMP Model and hence to provide them additional support in the form of value added courses, refreshment courses and bridge courses etc.

II. RELATED WORK

A. Bloom's Taxonomy

During 1956, Benjamin Bloom headed a group of educational psychologists who developed classification of levels ofintell ectual behavior which is important in learning. This classification helps in an effective curriculum design and is presented as a pyramid [1]. This Taxonomy Originally developed as a method of classifying educational goals for student performance evaluation. By using Bloom's Taxonomy, can assess students on multiple learning outcomes. Bloom's Taxonomy is shown in Figure 1 and the subsequent of revisions made to this are show in Figure 2 and 3.

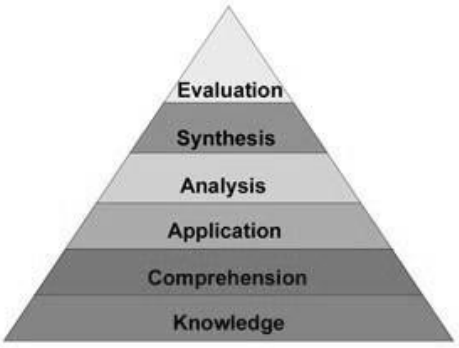

Figure 1: Bloom's Taxonomy (1956)

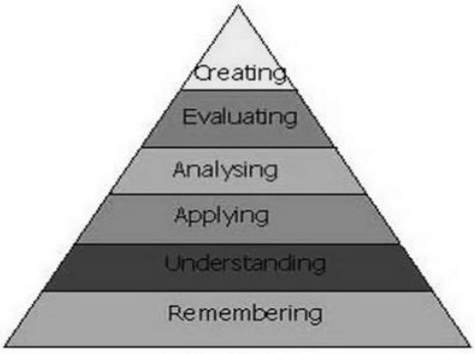

Figure 2: New version by Lorin Anderson (2001)

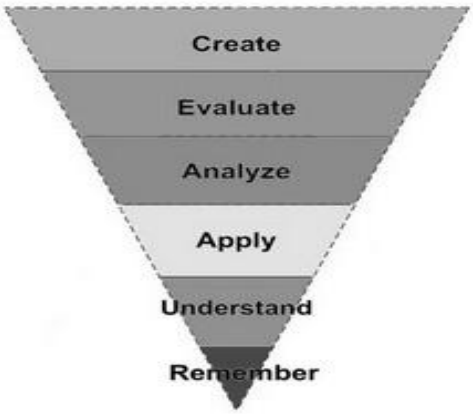

Figure 3: The Revised version by Lorin Anderson

The Taxonomy revised by the Lorin Anderson in figure 3 is the most adoptive taxonomy to the present need of the educational growth of the student in knowledge, creativity and employable dimension.

B. Graduate Aptitude Test in Engineering (GATE)

To assess the level students' competence at national level in core subjects, especially in engineering, GATE examination is being conducted in India. This test is becoming primary in screening the candidates in the process of recruitments in Government and PSU's .The pattern of evaluating the students' performance is in GATE is based the skills: Recall , Comprehension , application, analysis and synthesis. The following section represents what exactly these skills referring to.

Recall: Ability to memorize facts, principles, formulae or laws

Comprehension: Ability of understanding of the basics to draw simple conclusions from fundamental ideas.

Application: ability to apply knowledge either through computation or by logical reasoning.

Analysis and synthesis: ability to analyze the data, facts and ability to compare the information to draw valid conclusions

C. Infosys Campus Connect – FP TEST

Infosys Campus Connect – Foundation Program is an initiative by the Infosys to connect with the campuses (engineering colleges/professional colleges) to train and bridge the gap in the student competency levels at their fundamental /core subjects and hence to make them employable. They use the model called CAMP [2] (Conceptual, Analytical, Memory based, Practice-Based) model to train and asses the students on fundamental courses. They made the Examination pattern in evaluating the students' performance with the CAMP model such that the components should be as follows:

C+M = 20-30% and A+P = 70-80%

D. EDM: Educational Data Mining

EDM is an emerging discipline (as an application of Data Mining) deals with developing the methods and technique to explore the data coming from the educational Institutions to understand the students' performance and hence to setup the teaching and learning practice in an effective way to make the students competent and employable [3,4]. It was initiated during 2000 year. This EDM approach can helps in making the operations fine-tuned and can be used as tool to asses at various levels of administration of an educational institute. It also helps in designing effective curriculum, mode of teaching and evaluation methods

III. PROPOSED WORK

All the existing methods that are discussed above are evaluating the overall performance of the students and hence classifying them into a certain class/grade. But there is a need to understand the student performance at minute level of the ability skills on which the student is good at and lagging even though overall performance of the student is good and clearing the assessment tests. This type of evaluation helps in making the student competent at each and every core competence level [5, 6, 7, 8, 9, 10, 11]. In this proposed work, the process is to evaluate the performance of the students based on CAMP model and classifying them in order to provide the value added courses or additional training to make the student more competent and employable on core domains.

A. SAMPLING & DATA COLLECTION

The assessment is done based on the student performance in their core subjects with reference to their abilities (CAMP skills) are done. For this, Purposive sampling method is used to conducted a survey with a structured questionnaire (An objective multiple choice type Test) administered with the respondents (students). The design of Questionnaire to assess the students is as follows with respect to each skill component. The scheme of the questionnaire is shown in detailed in Table 1.

In each of the subjects, 30% of the weightage is given to (Concept + Memory) based Questions and 70% of the weightage is given to (Analytical + Practice) based questions. If a student gets 40% of marks in a subject then that students is considered to be passed in that subject. If the student scores 50 % of the marks on overall then that student is considered to be passed in the examination.

TABLE I
EXAMINATION PASSED VS PLACEMENT

Core Subject	C+M	A+P	SUBJECT WISE PASS	TEST PASS SCORE
Operating Systems	30%	70%	40%	50%
C Programming	30%	70%	40%	
Data Base Management Systems	30%	70%	40%	
Object Oriented Programming	30%	70%	40%	
IT Fundamentals	30%	70%	40%	

B. DATA ANALYSIS AND CLASSIFICATION

Data being collected through Questionnaire is loaded to the SPSS Software to analyze and process the data. SPSS Stands for Statistical Package for Social Sciences. This Software helps us to load various types of data stored in the formats like excel, text, csv etc. To carry out this study,

 DOI:10.32377/cvrjst0717

Descriptive statistics like frequencies, Cross tabulation etc using SPSS. This Software allows filtering the datasets, helps in writing our own coding/scripting (i.e. syntax) to process and analyzing the datasets. Segments (clusters) are defined using the association rules [12, 13, 14, 15] based on the subject wise scores, overall score and number of subjects passed by the students in the examination. The association rules for the segments (clusters) are defined as follows in Table 2.

TABLE II
CLUSTER VS CLUSTER DEFINITION

Cluster/Segment	Cluster Definition/ Association Rule
PERFORMERS	[ALL SUBJECTS PASS] & [GOT >=50% SCORE in the Test]
MODERATORS	[AT LEAST PASSED IN HALF OF THE SUBJECTS] & [GOT >=50% SCORE in The Test]
UNDER PERFORMERS	[NOT PASSED IN AT LEAST HALF OF THE SUBJECTS]

IV. RESULSTS AND DISCUSSION

Based on the sample of 48 respondents, the data is analyzed and the results are described below. The results are verified with their (students) respective placement status. This comparison helps us in whether the clusters are defined correctly and is aligning with the placement status. Based on the cluster and the placement status , can introspect the result of students on which they got failed in securing the placement (job). Whether they failed to secure the job due to lack of CAMP Skills or any other like interpersonal skills, presentation skills etc. this helps us to conduct value added and additional courses to make them employable and to enrich their skills.

A. RESULTS

The results which are obtained from the data being collected and analyzed are shown in the following Tables 3 and 4.

TABLE III
EXAMINATION PASSED VS PLACEMENT STATUS

	PLACED	NOT_PLACED	Total
EXAM PASSED	54.17%	35.42%	89.59%
EXAM NOT PASSED	0%	10.42%	10.42%
Total	54.17%	45.83%	100%

TABLE IV
CLUSTER VS EXAMINATION PASSED AND PLACED

	EXAM_PASSED		Total
	PLACED	NOT PLACED	
Performers	47.92%	22.92%	70.84%
Moderators	6.25%	14.58%	20.83%
Under performers	0%	8.33%	8.33%
Total	54.17%	45.83%	100%

B. DISCUSSION

- BASED ON OVERALL RESULTS

1. In “Performers” cluster, students are good at CAMP skills but some of them are unable to get the placements. This may be due to the lack of interpersonal and presentation skills. So these students who are not placed under “Performers” should be given some training on interpersonal and presentational skills (soft skills).

2. In "Moderators" cluster, students are moderate in CAMP skills. They need some additional training classes in the core subjects apart from soft skills.
3. In "Underperformers" cluster, students need to concentrate on the core subjects and special programs should be conducted to get the core domain knowledge apart from soft skills

- BASED ON SKILL LEVEL

1. Students that are not placed and have not answered 30% of Overall **C**oncept + Memory based questions should be given core subject knowledge by conducting the additional training programs on the core subjects along with the soft skills
2. Students that are not placed and have not answered 40% of Overall Analysis + Practice based Questions should be given practice based assignments, online tests and refresher courses apart from soft skills

CONCLUSIONS

In order to make the students employable, it is necessary to make the students more competent at the domain knowledge and also need to impart the soft skills and interpersonal skills to sustain in the competitive world. To make the students ready to compete, the curriculum and the teaching learning process should be fine-tuned in accordance with the industrial needs and the students' performance should be assessed and evaluated not only on the marks but also on their capability levels which helps in getting them placed in the industries

Since the study is purely empirical and on small size of data set, this study should be carried out on large size of data sets to know more insights of the data , though the results are on small data set resembles the reality. Secondly, the attributes that are considered for the classification under this study is among ability Skill set vs Placement status. This study further should be carried out with additional attributes like Academic background (rural/urban), Mode of instruction etc.

REFERENCES

[1] Bloom, B.S,"Taxonomy of Educational Objectives: The Classification of Educational Goals ", 1956, pp. 7-8.
[2] Campus Conncet ™ , An Infosys ® Industry-Academia Partnership Progam.
[3] K. Barker, T. Trafalis, and T. R. Rhoads, "Learning from student data," in Proceedings of IEEE Systems and Information Engineering Design Symposium, pp. 79–86, 2004.
[4] Dorina Kabakchieva, "Predicting Student Performance by Using Data Mining Methods for Classification", Cybernetics And Information Technologies, Volume 13, Issue 1, 2013, pp. 61-72.
[5] Kalpesh Adhatrao, Aditya Gaykar, Amiraj Dhawan, Rohit Jha and Vipul Honrao ,"Predicting Students' Performance Using Id3 & C4.5" , International Journal of Data Mining & Knowledge Management Process , Volume 3, Issue 5, September 2013, pp.39-52.
[6] Edin Osmanbegović, Mirza Suljić, "Data Mining Approach For Predicting Student Performance", Journal of Economics and Business, Volume 10, Issue 1, May 2012, pp.3-12.
[7] Samrat Singh, Dr. Vikesh Kumar, "Performance Analysis of Engineering Students for Recruitment Using Classification Data Mining Techniques", IJCSET, Volume 3, Issue 2, February 2013, pp.31-37.
[8] A. Merceron and K. Ycef, "Educational data mining: a case study", Proceedings of the 12th International Conference on Artificial Intelligence in Education (AIED '05), IOS Press, Amsterdam, The Netherlands, 2005.
[9] Romero C., Ventura S., Espejo, P.G. & Hervás C.,"Data mining algorithms to classify students", I International Conference on Educational Data Mining (EDM), Montreal, 2008, pp. 8-17.
[10] C. Romero, S. Ventura, P. Espejo, and C. Hervás."Data mining algorithms to classify students", Proceedings of Educational Data Mining, 2008,pp.20-21.
[11] Roxanne A. Ancheta, Rosmina Joy M. Cabauatan, Bartolome T. Tanguilig Lorena, W. Rabago, "Predicting Faculty Development Trainings And Performance Using Rule-Based Classification Algorithm" , Volume 2 , Issue 7, 2012, pp.203-209.
[12] Quang Hung Do and Jeng-Fung Chen, "A Neuro- Fuzzy Approach in the Classification of Students' Academic Performance", Hindawi Publishing Corporation Computational Intelligence & Neuroscience, Volume 2013.
[13] Jiawei Han, Michlelline Kamber and Jian Pei,"Data mining: concepts and techniques" , Morgan Kaufmann Publishers , 2008.
[14] Qasem A. Al-Radaideh, Eman Al Nagi , "Using Data Mining Techniques to Build a Classification Model for Predicting Employees Performance", International Journal of Advanced Computer Science and Applications, Volume 3, Issue 2, 2012.
[15] B. K. Bhardwaj and S. Pal, "Data mining: a prediction for performance improvement using classification" , International Journal of Computer Science and Information Security, volume 9, Issue 4, 2011, pp. 1–5.

 DOI:10.32377/cvrjst0717