

Original Research Paper

A Descriptive Analysis of Students Learning Skills Using Bloom's Revised Taxonomy

Joy Christy Antony Sami and Umamakeswari Arumugam

Department of Computer Science Engineering, School of Computing,
SASTRA Deemed to be University, Thirumalaisamudram, Thanjavur, Tamil Nadu, India, 613 401

Article history

Received: 06-07-2019

Revised: 30-07-2019

Accepted: 09-08-2019

Corresponding Author:

Umamakeswari Arumugam
Department of Computer
Science Engineering, School of
Computing, SASTRA Deemed
to be University,
Thirumalaisamudram,
Thanjavur, Tamil Nadu, India,
613 401
Email: umamakeswari.arumugam@gmail.com

Abstract: The academic committees worldwide suggest technical institutions to follow Revised Bloom's Taxonomy (RBT), a framework that helps to develop learning objectives. The model classifies a hierarchy of educational objectives such as cognitive, sensory and affective domains that are not only helping the students to evolve thinking abilities but also to identify the skills they are lacking with. Analysis of students RBT skills through data mining techniques is more valuable and is yet to be explored. This paper employs predictive and descriptive techniques of data mining to analyze the RBT level of each student. The methodology uses a classifier to classify the RBT level of questions under six levels such as remembering, understanding, applying, analyzing, evaluating, creating and performs clustering of students with respect to overall RBT level and lacking RBT skill of each student. The experimentation is carried out with university students. The results show that the proposed classifier is able to achieve 98% accuracy by correctly classifying RBT levels of input questions. The results also shows that the proposed work creates précised and meaningful clusters of overall RBT level/Lacking RBT skill of each student with precision 0.83 and 0.79 which could help the instructor to design different pedagogical approaches to improve students learning.

Keywords: Bloom's Taxonomy, Cognitive, Psychomotor, Affective, Clustering Techniques, K-Means Algorithm

Introduction

The objective of Revised Bloom's Taxonomy is to create achievable learning goals and a corresponding plan to meet them (Anderson *et al.*, 2005). The instructors witness the learning goals in behavioral terms. Well-defined goals enable the educators to effectively prepare lesson plans with appropriate content and sources according to the categories in the cognitive domain (Borg, 2003). Instructors can also design valid strategies and assessment tools to ensure whether the categories are met in line with the level of objectives. But, these tools do not contain any procedure for bridging the gap between students actual performance with cognitive domain and it is difficult for the instructors to analyze the performance of each student in mass groups. Moreover, there is no software tool available for describing the post facto effect with Bloom's learning domains. Bloom's Taxonomy verbs can be extended to determine the level of student's interests, attitude and expertise towards a subject (Adams, 2015).

Bloom's Taxonomy verb based questions help the instructors to analyze the level of each student under six

categories of cognitive domains such as remembering, understanding, applying, analyzing, evaluating and creating. These levels have a good impact in bridging the gap between education and employability. Mapping of RBT level motivates the instructors to understand whether a student simply by hearts the subject/interpolates problems in their own words/applies the learned concepts into new environments/distinguishes between facts and interference/makes judgments about the ideas on materials or builds a new structure or patterns based on the knowledge acquired. Not all students good at all categories of cognitive domains. There should be some measure to identify the RBT level and lacking skill of each student.

The scope of this work is to conduct predictive and descriptive analyses of students learning skills based on Revise Bloom's Taxonomy levels with the following objectives:

- To build an SVM based cognitive domain prediction model using trained RBT questions

- To use the model for predicting RBT levels of given set of questions
- To synthesize students marks based on each RBT level
- To cluster students with similar cognitive domains to analyze the lacking cognitive domain
- To analyze the post facto effect to improve students end semester exams

The paper analyzes the cognitive domain of students in higher education with respect to Bloom's Taxonomy. The impact of this proposed work would reflect in the students performance. The expected outcome of the paper is a software tool that analyzes the cognitive domain of students studying in the same class. The software clearly depicts the strength of the students within the six categories of remembering, understanding, applying, analyzing, evaluating and creating. The software also analyzes the lacking cognitive domains of the students within the same categories by considering the continuous internal assessment marks of each subject. For instance, if a student is able to create new concepts in one subject, but can only score marks in remembering and understanding categories of questions in other subject, the system will identify the discrepancy of the student's performance between the two subjects.

The paper ensures the application of clustering technique is quite helpful for the formation of homogenous clusters of learners. The clusters will allow the educator to design most effective teaching strategies that really cater to the needs of students, especially those lacking with RBT domain skills. This study is also intended to address the difficulties of the students in create level, which means the difficulties in building new solutions from the learned concepts, as it kindles research in students. Moreover, the analysis of Bloom's Taxonomy based student performance is yet to be explored nationwide.

The remaining sections of the paper are organized as follows: Section 2 reviews the existing literature in data mining with Revised Bloom's Taxonomy, Section 3 explains the proposed methodology, Section 4 discusses the experimentations and results and finally Section 5 concludes the social impacts and research findings of the work.

Review of Literature

Jayakodi *et al.* (2016) have automated the categorization of exam questions into RBT learning levels. The authors have used natural language processing to tokenize the words used in the questions and have compared the extracted verbs with Anderson Revised Bloom's taxonomy using Wordnet similarity comparison and cosine similarity module for generating rule set. The authors have claimed that the proposed rule set is able to achieve 70% accuracy in predicting the RBT level of questions. The authors have also suggested

evaluators to redesign the question papers based on the output of the classifier.

Suhaimi *et al.* (2016) have analyzed the effects of classifying the written exam question into six cognitive levels in Bloom's Taxonomy. The authors have tried to prove that the classification of written exam questions into correct cognitive level can generate a good set of exam questions. The authors have combined K-Means clustering algorithm and probability based classification approach to classify the class labels in blooms taxonomy and have obtained 50% accuracy. The authors have declared that some keywords may contribute to more than one class label resulting vague measurement that may impact with poor accuracy.

Sukajaya *et al.* (2015) have proposed a Bloom's taxonomy based serious game technology as an approach for measuring cognitive domain of learners. The authors have maximized the opportunities to analyze the potentials of all students based on their profile by collecting the learners data which are theoretically more representative to identify specific learner. The authors have conducted the analysis using Naïve Bayes and J48 using two types of procedure: raw data, transformed data into average value with consideration on cognitive domain of Blooms taxonomy and have obtained highest prediction accuracy with Naïve Bayes algorithm of 92.31%. The authors have claimed that integration of Bloom's taxonomy into serious game technology is potential in measuring learners' information in learning.

van Konsky *et al.* (2018) have explored an automatic identification of verbs and other parts of speech that impacts the semantic meaning of Bloom's taxonomy cognitive domains. The authors have classified Bloom's levels using a table lookup and machine learning approach. The authors have analyzed parts of speech in a training corpus with 13,189 learning outcomes at an Australian university. The authors have stated that their proposed approach assists human learning and teaching designers to write learning outcomes and also plays a vital role in automating Bloom's taxonomy classification. The authors have claimed that their method verb table lookup and machine learning approach has obtained 71% prediction accuracy by identifying same cognitive levels.

Gottipati and Shankararaman (2018) have presented an automated method to discover the impact of curriculum design and its competency using Bloom's taxonomy and Dreyfus model. The authors have developed a curriculum analytics tools that generates competency score for the entire curriculum with respect to cognitive and progression levels. The authors have tested the proposed method over 14 courses and the corresponding 578 competencies in information system management curriculum program and have claimed that their method performs in-depth analysis on the curriculum by discovering the cognition and progression statistics and has obtained 74.69% accuracy in predicting the competency of the curriculum.

Deshmuk *et al.* (2018) have implemented a student performance evaluation model for identifying the strengths and weaknesses of students for network analysis course studied by third semester Electronics and Communication Engineering students using fuzzy inference system. The authors have fuzzified the five inputs identify, understand, apply, analyze and design/create for the system to learn Bloom's Taxonomy cognitive domains. The rules are then applied to the questions to evaluate the critical thinking of students. The results evaluated are expressed in linguistic variables and compared with classical aggregate scores and the authors have claimed that the proposed model is flexible in terms of assigning importance to each criterion by modifying fuzzy rules.

Though the implications of data mining techniques in RBT analysis are enormous, the literature limits its contributions only with predicting RBT levels. Descriptive analysis of data mining can make obvious differences to understand the level of students in RBT standards. Thus, this work is not only trying to improve the prediction of accuracy of RBT levels but also to describe the learning ability of students.

Methodology

The component diagram of the proposed methodology shown in Fig. 1. denotes the major functionalities of the proposed system. The system builds a knowledge prediction model for classifying the cognitive levels based on Bloom's taxonomy. The model takes bloom's taxonomy based questions that cover all six RBT levels, as input. The model parses each question by tokenizing individual words and applies lemmatization techniques for removing the stop words in the question. The model also uses stemming algorithms to prune the most significant terms used in the question to its root word. The questions are finally converted as a dataset, where the significant terms of the question are independent variables and Bloom's taxonomy level for the question as dependent or class variable. The model uses support vector machine as it is a widely used classifier algorithm that often produces highest prediction accuracy compared to all other classifiers available in the market. The system implements the SVM classifier for predicting the RBT level of the questions prepared by the staff for internal and external assessment tests. Finally, the performance of the students has been analyzed by the system to map the cognitive level of the students. To accomplish this, the marks obtained by the students for all subjects in the same RBT level is averaged and normalized to logarithmic notation to put to a common form. The normalized marks are then inputted to K-means algorithm for the descriptive analysis. The step by step process of the proposed work is shown in Algorithm 1.

Algorithm 1: RBT Analysis

Input: CIA Question Paper, Students CIA Marks

Output: Students knowledge level with respect to RBT, Course Outcome Report

1. for each question Qp, p = 1, 2, 3.... n in the question paper do
2. predict RBTLevel using SVM Classifier
3. for each student Si, i = 1, 2, 3,... s in the database do
4. for each question Qp, p = 1, 2, 3,...n in the CIA Question Paper do
5. for each mark Mj,j = 1, 2, 3...n in the question do
 /*normalize marks*/
6. set Mj as Mj-mimmarks of Qj/maxmark of Qj-minmarks of Qj /*
7. for each student Si,i = 1,2,3....s in the database do
8. for each level Lx, x = 1,2,3...6 in RBT do
9. for each question Qp, p=1, 2, 3,...,n in the question paper do
10. if(Lx == Qp(RBTLevel))
11. set aggregatemarks[x]+=Mp
12. set number of clusters k as 3
13. choose the best set of centroids to cluster low, medium and average levels in overall RBT
14. for each cluster Ci, i = 1,2,3....k do
15. cluster students aggregate marks to the nearest centroids
16. if there is no change in the cluster group or cluster centroids
17. continue
18. interpret results
19. Set number of clusters k as 6
20. Choose the best set of centroids to cluster lacking skill of students in each level
21. Repeat steps 13 to 17

The proposed work employs SVM multiclass classifier with One against All approach. The method creates k SVM classification models where k represents the number of RBT Levels (Mustaqeem *et al.*, 2018; Wang *et al.*, 2019; Negri, 2018). The model trains all examples of m^{th} class with positive and negative labels such that the given l training data $(r_1, S_1), \dots, (r_l, S_l)$, where $r_i \in R_n, i = 1, 2, 3 \dots l$. and $S_i \in \{1, 2, \dots, k\}$ is the class of r_i , the m^{th} SVM solves the following Equation 1:

$$\min_{\omega^m, b^m, \epsilon^m} = \frac{1}{2} (\omega^m)^T \omega^m + C \sum_{i=1}^l \epsilon_i^m$$

$$(\omega^m)^T \phi(r_i) + b^m \geq 1 - \epsilon_i^m, \text{ if } s_i = m \quad (1)$$

$$(\omega^m)^T \phi(r_i) + b^m \leq -1 - \epsilon_i^m, \text{ if } s_i \neq m$$

$$\epsilon_i^m \geq 0, i = 1, 2, 3, \dots, l$$

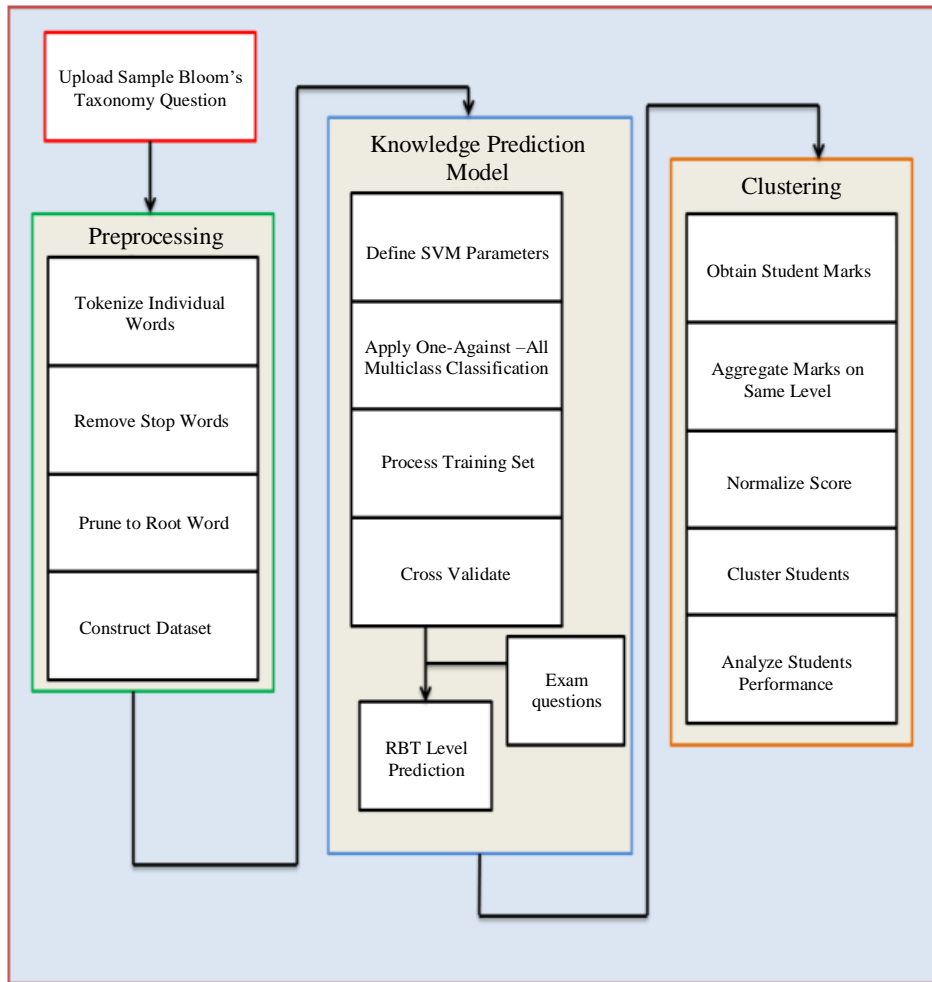


Fig. 1: RBT analysis component diagram

where, the training set s_i are mapped to a higher dimensional space by function ϕ is the penalty parameter. The margin between two groups of data is maximized to $2/\|\omega^m\|$ by minimizing the term $\frac{1}{2}(\omega^m)^T \omega^m$. The penalty term $C \sum_{i=1}^l \varepsilon_i^m$ is used to minimize the number of training errors, for nonlinear data. The basic notion of the proposed SVM is to balance and regularize $\frac{1}{2}(\omega^m)^T \omega^m$ and training errors.

Equation 1 can be solved for the 'k' decision functions Equation 2:

$$\begin{aligned}
 &(\omega^1)^T \phi(S) + b^1 \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 &(\omega^k)^T \phi(S) + b^k
 \end{aligned} \tag{2}$$

where, sample S is classified in one of the classes with the largest value of the decision function as shown in Equation 3:

$$\text{class of } S \equiv \operatorname{argmax}_{m=1, \dots, k} ((\omega^m)^T \phi(S) + b^m) \tag{3}$$

The basic notion of the paper is to check whether clustering techniques in data mining is useful to make a descriptive analysis of students performance based on RBT levels. This analysis is also intended to identify the lacking skills of students with respect to remember, understand, apply, analyze, evaluate and create. The analysis is carried out by generating two clusters outputs using k-means algorithm, where the former elucidates the overall analysis on students performance Vs. RBT and the latter denotes the lacking RBT skills of the students. K-means algorithm is one of the popular and widely used algorithms used for descriptive analysis (Dubey *et al.*, 2018). In this study, the cluster centroids are manually forced into k-means algorithm for getting

the desirable clusters of students. The distance between instances and cluster centroids are computed using Euclidean distance measures and the instances are grouped into the cluster with minimum distance (Yu *et al.*, 2018; Shyr-Shen *et al.*, 2018; Sardar and Ansari, 2018). The process gets repeated until there is no change in the cluster centroids or when no new cluster formation is possible to reduce root mean squared error. Equ.4. denotes the formulation of K-means algorithm (Singhal and Shukla, 2018):

$$RMSE = \sum_{x=1}^n \sum_{y=1}^m (\|I_x - C_y\|)^2 \quad (4)$$

Where:

$\|I_x - C_y\|$ = The difference between each instance and cluster centroid

n = The number of instances x^{th} cluster

m = The number of cluster centres

Experimentation and Result Discussions

The experimentation of the proposed work is split into two major segments such as predicting RBT level of questions and performing descriptive analysis of students. Predicting RBT level of questions starts with the construction of knowledge prediction model with support vector machine classifier. The model takes input as questions obtained from a corpus BCLsDataSet (Lashari *et al.*, 2012). The corpus contains six hundred questions,

100 for each Bloom’s Taxonomy level such remember, understand, evaluate, analyze, apply and create. The model uses NLP equivalent R programming methods to perform bag of words, part of speech and n-grams operations for stop words removal and root word extraction on the questions. Selected terms from questions with RBT level as class label is then built as a dataset and executed upon SVM classifier. Radial bias function is used as a kernel function for SVM classifier as it is a popular function used in kernelized algorithms. Table 1 shows the 6x6 confusion matrix of SVM classifier with the left axis showing the true class and the top showing the class assigned to an item with that true class.

Table 2 denotes the prediction accuracy and kappa statistics of SVM classifier for test set of BCLsDataSet. Accuracy denotes the percentage of correctly classified instances out of all instances. The SVM algorithm correctly predicts the RBT levels of questions upto 98%, another accuracy indicator is the kappa score, which is a measure of comparing the classification results to values assigned by chance. The kappa score closer to 1 reveals the classified results and the ground truth are identical and the kappa score of the SVM classifier for the experimental dataset is 0.972, which is close to 1, thus, can be used as a standard classifier for predicting the RBT levels of questions. The performance analysis of SVM classifier is compared with the algorithms presented in the literature and shown in Fig. 2.

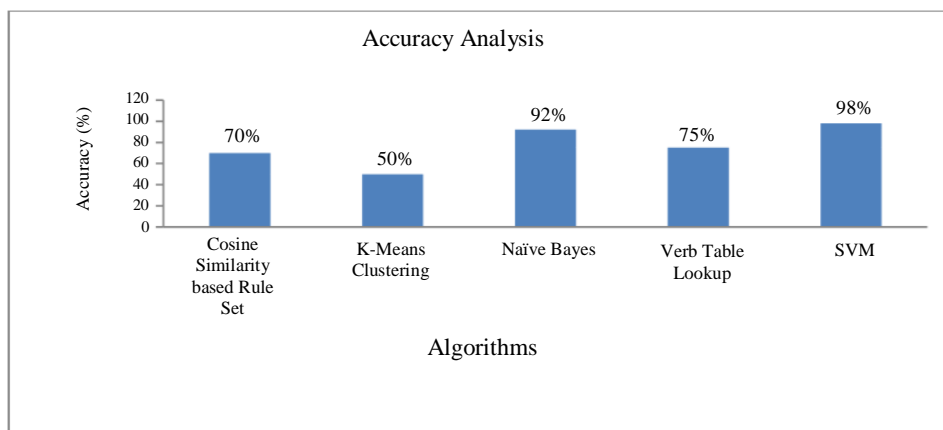


Fig. 2: Accuracy analysis of SVM classifier in RBT level prediction

Table 1: SVM confusion matrix

Actual number of instances	Predicted number of instances					
	Analyze	Apply	Create	Evaluate	Remember	Understand
Analyze (100)	99	1	3	0	0	3
Apply (100)	0	96	0	1	0	0
Create (100)	0	0	97	1	0	0
Evaluate (100)	0	3	0	97	0	0
Remember (100)	1	0	0	0	100	0
Understand (100)	0	0	0	1	0	97

Table 2: SVM accuracy and kappa statistics

Statistical measures	Value
Accuracy	0.9767
95% CI	(0.9612, 0.9872)
No Information Rate	0.1667
P-Value [Acc>NIR]	<2.2e ⁻¹⁶
Kappa	0.972

Table 3: Initial centroids for experiment 1

RBT	Remember	Understand	Apply	Analyze	Evaluate	Create
Centroid 1	0.3	0.3	0.3	0.3	0.3	0.3
Centroid 2	0.6	0.6	0.6	0.6	0.6	0.6
Centroid 3	0.9	0.9	0.9	0.9	0.9	0.9

Table 4: Cluster centroids for students overall RBT performance

RBT	Remember	Understand	Apply	Analyze	Evaluate	Create
Centroid 1	0.3639197	0.4955678	0.3450528	0.4692803	0.3138181	0.410300
Centroid 2	0.5872336	0.6679749	0.5890975	0.6697737	0.6726497	0.7112745
Centroid 3	0.7287417	0.8313449	0.8366759	0.7592867	0.7821078	0.7399388

Table 5: Initial centroids for experiment 2

RBT	Remember	Understand	Apply	Analyze	Evaluate	Create
Centroid 1	0.3	0.6	0.6	0.6	0.6	0.6
Centroid 2	0.6	0.3	0.6	0.6	0.6	0.6
Centroid 3	0.6	0.6	0.3	0.6	0.6	0.6
Centroid 4	0.6	0.6	0.6	0.3	0.6	0.6
Centroid 5	0.6	0.6	0.6	0.6	0.3	0.6
Centroid 6	0.6	0.6	0.6	0.6	0.6	0.3

Table 6: Cluster centroids for students lacking RBT skill

RBT	Remember	Understand	Apply	Analyze	Evaluate	Create
Centroid 1	0.3597271	0.8201930	0.8106352	0.6531691	0.7754325	0.7092969
Centroid 2	0.7286630	0.3450528	0.8157800	0.6923217	0.8428593	0.6703103
Centroid 3	0.7112745	0.6652774	0.4955678	0.6192803	0.7687836	0.5510300
Centroid 4	0.6450619	0.7714665	0.8299177	0.4419030	0.6158362	0.4816480
Centroid 5	0.7151979	0.8766680	0.8498470	0.7962746	0.3138181	0.7676972
Centroid 6	0.8446824	0.8190486	0.7803017	0.7490387	0.6879762	0.3262793

The second segment of the experimentation is conducted over 56 undergraduate students studying Computer Science Engineering at SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India. The samples are asked to take an online exam in JAVA programming with sixty questions each of which represents one of the six Bloom’s Taxonomy levels. There are ten questions in each RBT category. Identification of Bloom’s taxonomy level of a question has been automated through software and the questions are dispersed to the students in random manner. The duration of the test is set to Thirty minutes. The scores of the students have been retrieved in CSV file format for the analysis. Table 3 denotes set of initial centroids that are given as an input to K-means clustering algorithm for conducting the first experimental analysis. The initial centroids for the first experimental analysis are carefully chosen in such a way that it should clusters poor, medium and good level students based on overall RBT. Table 4 denotes the modified

cluster centroids of k-means algorithm during the last iteration, followed by the cluster plots, Silhouette distance, and the graphical representation of cluster centroids in Fig. 3-5 respectively.

K-means algorithm creates three clusters with size 4, 14 and 38. From the results it has been observed that there are 4 students scored low marks in all RBT levels, 14 students have scored medium marks in overall RBT and 38 students have better skills in all RBT levels.

Table 5 denotes the new set of initial centroids that are given as an input to K-means clustering algorithm for conducting the second experimental analysis. The initial centroids are now carefully chosen to create clusters that elucidate the weaker students in each RBT level. Hence, the number of clusters is six. Table 6 denotes the cluster centroids generated during the last iteration of K-means algorithm, followed by the cluster plots, Silhouette distance, and the graphical representation of cluster centroids in Fig. 6-8 respectively.

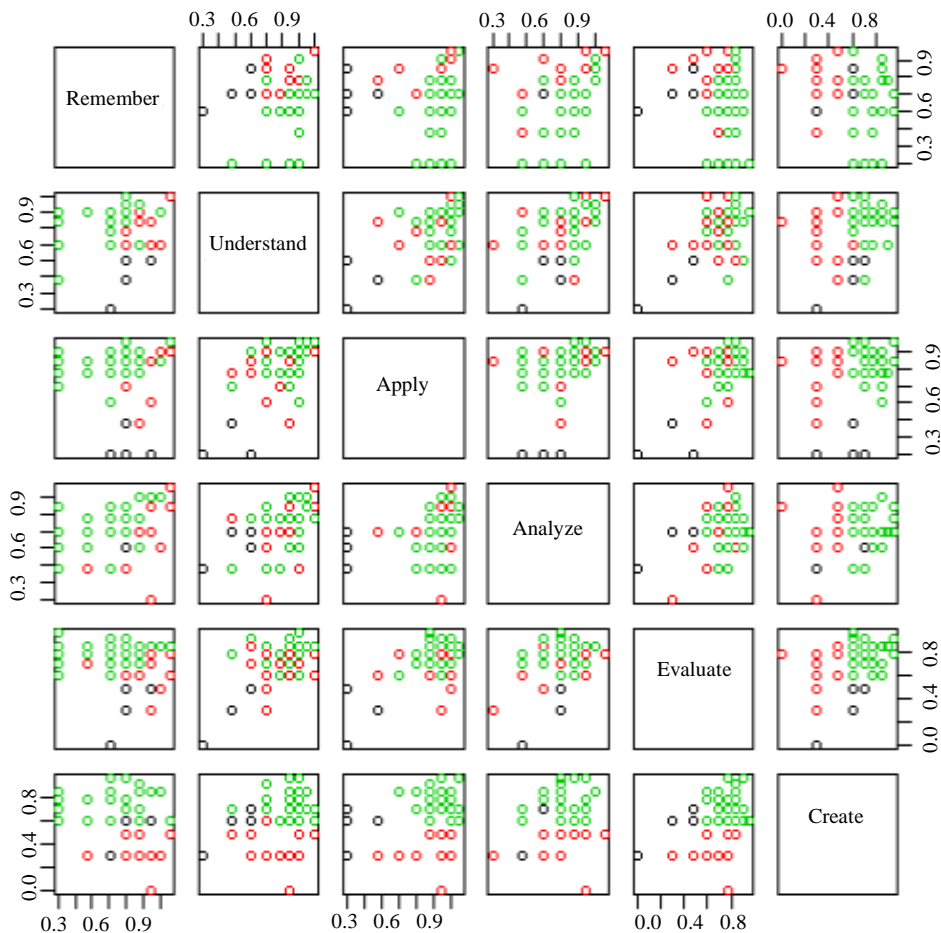


Fig. 3: Cluster results of students overall RBT performance

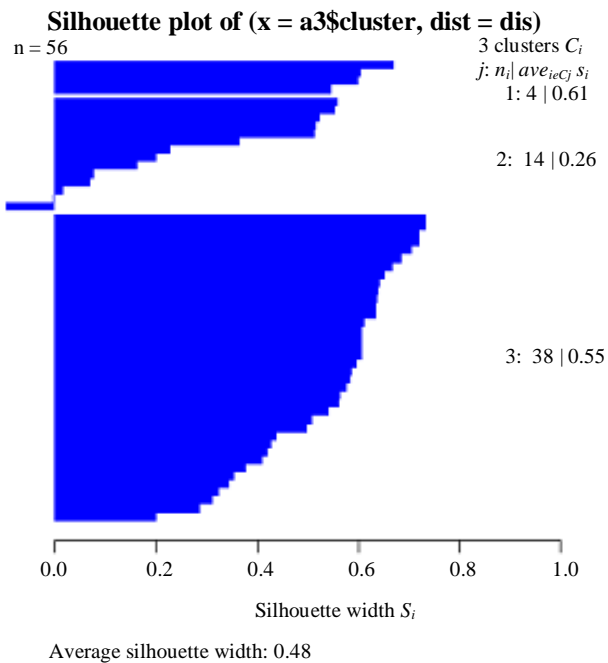


Fig. 4: Silhouette distance of students overall RBT performance

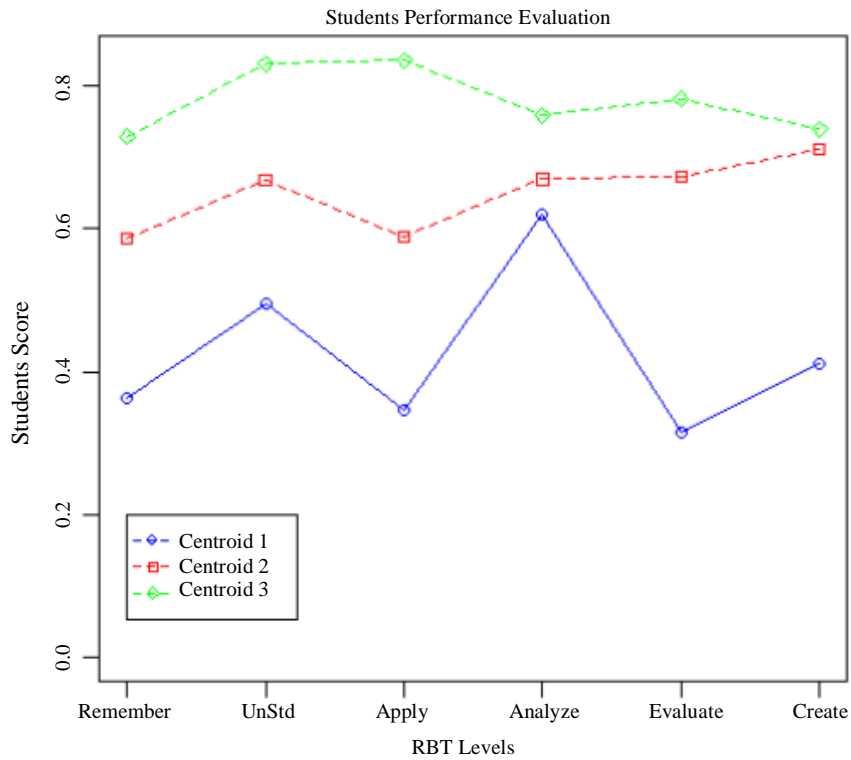


Fig. 5: Cluster centroids for students overall RBT performance

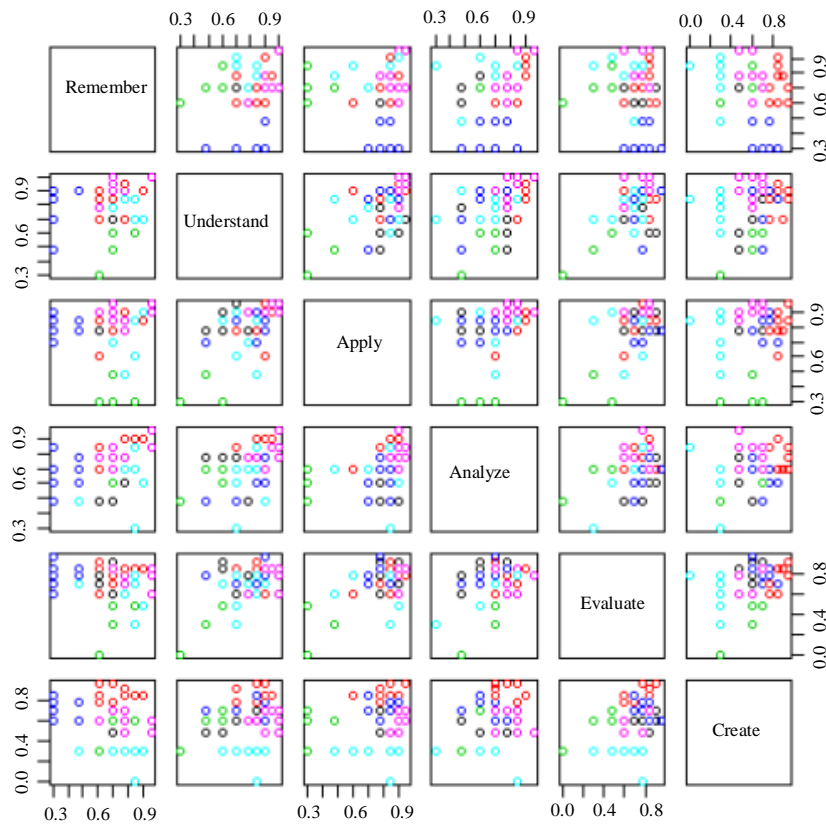


Fig. 6: Cluster results of lacking RBT skills

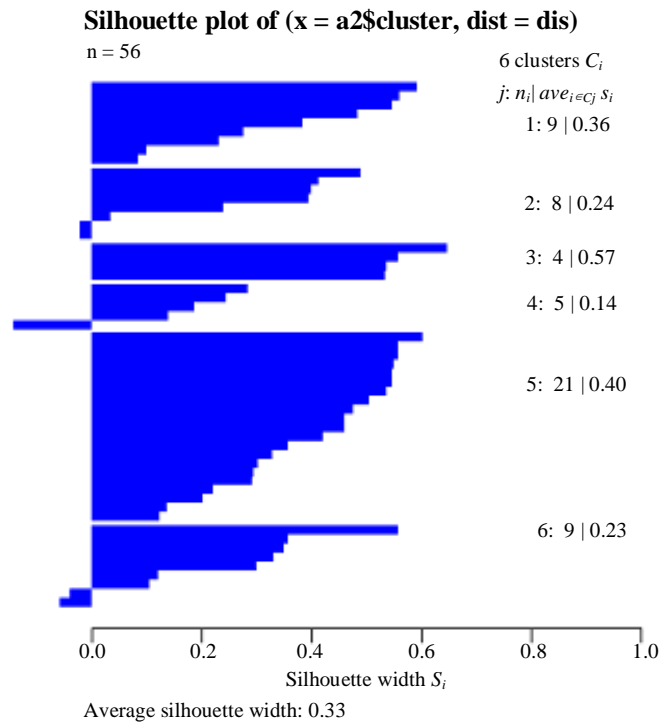


Fig. 7: silhouette distance of lacking RBT skills

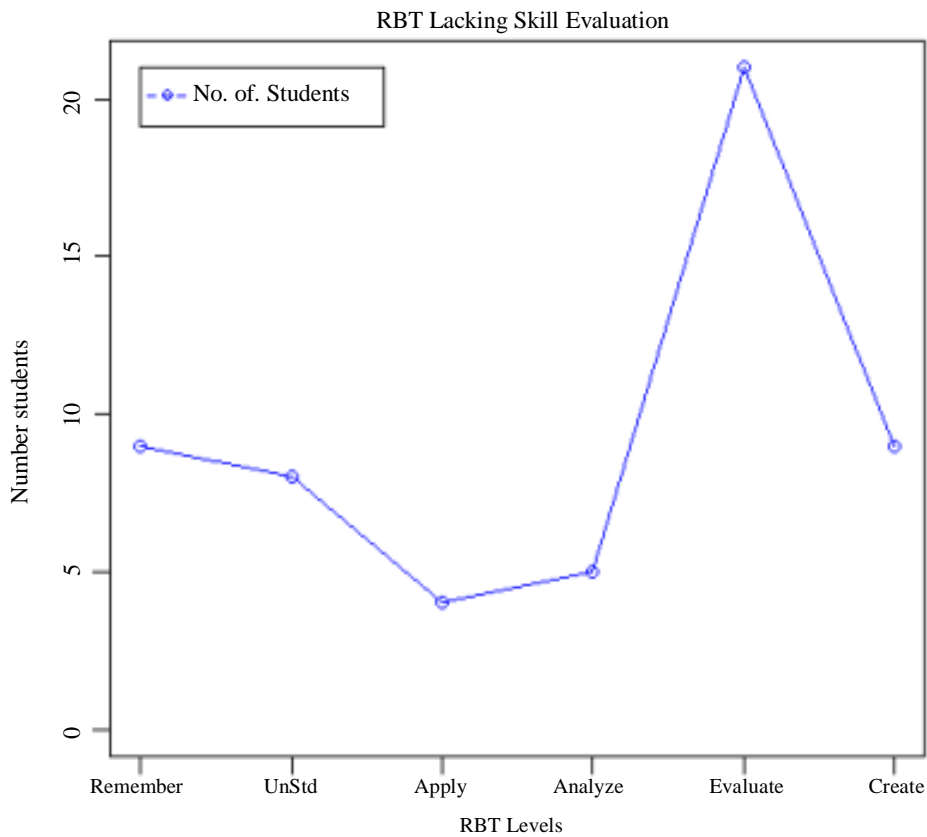


Fig. 8: Number of lacking students in Each RBT category

The number of elements in each K-means cluster is 9, 8, 4, 5, 21 and 9 respectively. The cluster results show that there are 9 students lacking with remembering, 8 in understanding, 4 in applying, 5 in analyzing, 21 in evaluating and finally 9 with creating ability.

Conclusion

RBT aids the instructors to perform formative and summative assessments of students so as to map the objectives and motivation of a course in an academic curriculum. RBT is a powerful mechanism that assists students to learn at higher levels. But, the real challenge for the instructors is to make a descriptive analysis of students performance through RBT levels. There is only little/no effort been taken to group the students to find the number of students with poor, good and medium levels in all RBTs and the lacking skills of each student. This work is carried out to analyze students learning ability through predictive and descriptive analyses in data mining. Predictive analysis employs SVM classifier to obtain the RBT level of questions and descriptive analysis uses K-means clustering algorithm to group students based on learning ability. SVM classifier is able to correctly classify the RBT level of questions with the accuracy of 98% and found to be the best classifier for obtaining RBT levels of questions. The results of K-means clustering correctly group poor, good and medium level students and the lacking skill of each student with precision values 0.83 and 0.79. The cluster results help instructor to use different teaching methodologies and follow ups to ensure better learning of students.

Acknowledgement

The authors would like to thank SASTRA Deemed to be University for the facilities used in this work.

Author's Contributions

All authors contributed to the final version of the manuscript

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and there are no ethical issues involved.

References

Anderson, H.M., D.L. Moore, G. Anaya and E. Bird, 2005. Student learning outcomes assessment: A component of program assessment. *Am. J. Pharm. Edu.*, 69: 256-268. DOI: 10.5688/aj690239

- Adams, N.E., 2015. Blooms taxonomy of cognitive learning objectives. *J. Med. Library Associ. JMLA*, 103: 152-135. DOI: 10.3163/1536-5050.103.3.010
- Borg, S., 2003. Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teach.*, 36: 81-109. DOI: 10.1017/S0261444803001903
- Deshmukh, V.B., S.R. Mangalwede and D.H. Rao, 2018. Student performance evaluation model based on scoring rubric tool for network analysis subject using fuzzy logic. *Proceedings of the International Conference on Electrical, Electronics, Communication, Computer, Mechanical and Computing*, Dec. 15-16. IEEE Xplore Press, Mysuru, India. DOI: 10.13140/RG.2.2.32124.44169
- Dubey, A.K., U. Gupta and S. Jain, 2018. Comparative study of K-means and fuzzy C-means algorithms on the breast cancer data. *Int. J. Adv. Sci., Eng. Inform. Technol.*, 8: 18-29. DOI: 10.18517/ijaseit.8.1.3490
- Gottipati, S. and V. Shankararaman, 2018. Competency analytics tool: Analyzing curriculum using course competencies. *Educ. Inform. Technol.*, 23: 41-60. DOI: 10.1007/s10639-017-9584-3
- Jayakodi, K., M. Bandara, I. Perera and D. Meedeniya, 2016. WordNet and cosine similarity based classifier of exam questions using Blooms taxonomy. *Int. J. Emerg. Technol. Learn.* 11: 142-149. DOI: 10.3991/IJET.V11I04.5654
- van Kinsky, B., S. Huband and D. Gibson, 2018. Parts of speech in bloom's taxonomy classification. *Proceedings of the 35th International Conference of Innovation Practice and Research in the use of Educational Technologies in Tertiary Education*, (TTE' 18).
- Lashari, T., A. Maizam, A. Zainal, A. Akasah and M.J. Kesot, 2012. An affective cognitive teaching and learning framework in engineering education. *J. Eng. Educ.*, 1: 11-24. DOI: 10.11120/ened.2013.00011
- Mustaqeem, A., S.M. Anwar and M. Majid, 2018. Multiclass classification of cardiac arrhythmia using improved feature selection and SVM invariants. *Comp. Math. Methods Med.*, 1: 1-10. DOI: 10.1155/2018/7310496
- Negri, P., 2018. A MATLAB SMO implementation to Train a SVM classifier: Application to multi-style license plate numbers recognition. *Image Proc. Line*, 8: 51-70. DOI: 10.5201/ipl.2018.173
- Sardar, T.H. and Z. Ansari, 2018. An analysis of MapReduce efficiency in document clustering using parallel K-means algorithm. *Future Comp. Inform. J.*, 3: 200-209. DOI: 10.1016/j.fcij.2018.03.003

- Shyr-Shen, Y., S.W. Chu, C.M. Wang, Y.K. Chan and T.C. Chang, 2018. Two improved k-means algorithms. *Applied Soft Comp.*, 68: 747-755.
DOI: 10.1016/j.asoc.2017.08.032
- Singhal, M. and S. Shukla, 2018. Centroid selection in kernel extreme learning machine using k-means. *Proceedings of the 5th International Conference on Signal Processing Integrated Networks*, Feb. 22-23. IEEE Xplore Press Noida, India. pp: 708-711.
DOI: 10.1109/SPIN.2018.8474055
- Suhaimi, N.S., N. Arbin and N.N. Zulkifli, 2016. Studying the effects of performing text mining to improve classification of clustered questions based on bloom taxonomy. *Indian J. Sci. Technol.*, 9: 1-5.
DOI: 10.17485/ijst/2016/v9i28/97818
- Sukajaya, I.N., I.K.E. Purnama and M. Hery Purnomo, 2015. Intelligent classification of learner's cognitive domain using bayes net, naïve bayes and j48 utilizing bloom's taxonomy-based serious game. *Int. J. Emerg. Technol. Learn.*
DOI: 10.3991/ijet.v10i2.4451
- Wang, R., W. Li, R. Li and L. Zhang, 2019. Automatic blur type classification via ensemble SVM. *Signal Proc. Image Commun.*, 71: 24-35.
DOI: 10.1016/j.image.2018.08.003