A Comparative Study of Question Bank Classification based on Revised Bloom's Taxonomy using SVM and K-NN

Soumya K. Patil (soumyakpatil45@gmail.com), Shreyas M M (mm.shreyas@gmail.com)
Department of Computer Science and Engineering
PES University
Bangalore, India

***Abstract*— Education is not only about learning the facts, but the training of the mind to think. Based on the thinking ability of the person, the "Cognitive skill Domain" of the Revised Bloom's Taxonomy tries to distinguish the process of learning into six different levels. These levels are ordered in the increasing level of complexity. Lorin Anderson and David Krathwohl are the primary authors of the revisions to which is widely known as Bloom's Taxonomy that defines an ordering of cognitive skill and knowledge skills. Questions are the primordial component, it can be used to examine the skills and levels of the student. The primary aim of this paper is to make a comparative study of question bank classification based on Revised Bloom's Taxonomy using the machine learning techniques called SVM and K-NN in an attempt to achieve better performance and higher accuracy levels. In order to classify questions into the taxonomy the paper describes the various tests performed namely grammar check, in context check. This work helps the students to improve their thinking and questioning ability and the tutor for setting the objectives of the course and to make assessments.**

Keywords*— *Revised Bloom's Taxonomy, Question Bank, Machine Learning, Classification Algorithms, Grammar & In context tests

I. INTRODUCTION

Education is a process of modelling learner's behavior. Planning, teaching and assessment are the important stages of education which aims to make learner capable of decision making, good reasoning and critical thinking. Assessment [2] is the crucial step of determining learner's conceptual development. It helps to build cognitive skills of learner.

Question bank is an important component of the education system. Question bank [1] can be used for designing a more effective assessment by allowing a unique subset of questions to be chosen for each test, where the skills and levels of the students need to be examined. Hence, question bank needs to be managed properly which includes setting in context questions along with organization and classification of the questions.

Revised Bloom's Taxonomy is one such criteria to classify the questions into its six different cognitive skill levels. These levels are ordered in the increasing level of complexity and mainly focus on the cognitive skill domain or thinking ability of the person. Bloom's Taxonomy [3] is an essential concept that guides the tutor in preparing assessment, curriculum and setting objectives to the course. This can also be used to make performance analysis of the students based on the answers given or doubts asked by the students in the classroom.

Statistical approaches for question bank classification can be done by employing machine learning techniques SVM and K-NN. The text must be preprocessed before applying any machine learning techniques. The classifiers will be then trained by set of training examples for each category, which are predefined with the help of subject expert. Hence forth the classifiers will be used to classify set of questions.

Hence, the main aim of this paper is to make a comparative study using the machine learning classification framework for handling the problem of cognitive skill category determination (based on Revised Bloom's Taxonomy by Krathwohl) for the question bank questions in an attempt to achieve better performance and accuracy levels.

II. RELATED WORK

Many studies have sought to automatically classify questions based on Bloom's taxonomy. Therefore, some previous studies similar to the model introduced in this work are reviewed:

[1] The aim of the study is based on the process of examining the effectiveness of support vector machines (SVMs), So, we can overcome the question bank classification problem to classify into different levels of Bloom's cognitive skill domain. For that purpose, the dataset was collected which has pre-classified questions in it, where each question was processed through steps like removal of stop words, punctuations, tokenization, stemming, length normalization and term weighting. In approximation 70% and 30% of the dataset are used to build and evaluate the SVMs classifiers of type linear kernel respectively, SVM-Light software package was used and obtained preliminary results which showed a satisfactory effectiveness of SVMs with respect to classification accuracy and precision.

[2] The study has proposed a newest methodology that can be used in the automatic classification of the exam questions based on the cognitive skill levels of Bloom's taxonomy. By combining the three machine learning classifiers the study implements the voting algorithm based on the strength of these combination strategies. K-nearest Neighbor (K-NN), Support Vector Machine(SVM) and the Multinomial Naïve Bayes(NB) are the classifiers used in classification of the questions based on taxonomy without extracting any particular feature, later

they also used suitable feature selection methodologies like mutual selection techniques, chi-square test and odd ratio selection are used for classification.

III. Background – Revised Bloom's Taxonomy, Support Vector Machines (SVM), K-Nearest Neighbors (K-NN)

A. *Revised Bloom's Taxonomy*

From late 90's to early 70's in U.S, the various domains of human learning were taken into consideration and attempts were made to categorize and classify these domains, 3 classes were discovered namely: Cognitive, Affective, Psychomotor.

Cognitive skill also known as head or knowing, Affective also known as heart, feelings or emotion and Psychomotor known as kinesthetic or doing. Further for each area of this human learning, the results yielded a series of taxonomies. These before mentioned taxonomies of human learning are arranged hierarchically from simple to more complex functions.

Cognitive Skill Domain: There exist 2 primary taxonomies of cognition:

1. Bloom's Taxonomy the original cognitive skill domain by Bloom which was described and published in 1956, Bloom is an educational psychologist from the University of Chicago. According to Bloom the thinking skills can be categorized into six different levels namely in the increasing order of complexity: knowledge, comprehension, application, analysis, synthesis and evaluation.
2. Lorin Anderson and David Krathwohl have proposed the Revised Bloom's Taxonomy in the year 2001, They also categorized cognitive skill domain into six levels namely: Remembering, Understanding, Applying, Analyzing, Evaluating, Creating as shown in figure 1 which usually describes an ordering of the intellectual abilities or knowledge and the cognitive skills

Fig. 1. Revised Bloom's cognitive skill levels [4]

Remember: Basically remember is to recall the knowledge from the memory. It also refers to the scenario when we use memory to make lists or to recite previously learned any type of information, to retrieve the facts and to define things or to retrieve the definitions. A few keywords that belongs to this domain are choose, define, find and how.

Understand: Determining and constructing the meaning of instructional messages through various forms such as orally, in the written form or through graphical communication or by using different materials. Activities also includes the various tasks such as to interpret, exemplify, classify, summarize, infer, compare, and explain. A few keywords of this domain are classify, compare, demonstrate and explain.

Apply: Using procedures on situations where learned material is applied. A few keywords of this domain are apply, build, plan and select.

Analyze: Breaking concepts and components into smaller units and also to determine how these units relate with one another and also with the overall structure built. A few keywords of this domain are analyze, assume, categorize and classify.

Evaluate: Composing judgments that are based on several similar or different criteria and standards. In order to demonstrate this generates specific reports and recommendations for the process of evaluation. As evaluation is more often a necessary part of one's precursory behaviour before they create something hence the evaluation level comes before Creation in this Lorin Anderson and David Krathwohl's newer taxonomy. A few keywords of this domain are agree, compare, conclude and decide.

Create: Put the elements all together to form a functional or a coherent whole product. This is the difficult level as it reorganizes elements to create new pattern. A few keywords of this domain are create, delete, design and develop.

In this revised Bloom's Taxonomy, cognitive skill levels have relationship among each other. Higher the level bigger will be complexity. Thus, assessment process starts from lower to higher level to encourage the learner for development of higher cognitive skills as shown in figure 1.

B. *Support Vector Machines (SVM)*

In the early nineties the Vladimir Vapnik and his team members discovered SVM as an addition to the machine learning techniques have developed the support vector machines SVM is a renowned supervised machine learning algorithms, which helps in successful application to many classification problems for e.g., to classify the text data and more recent problems like classification of the questions.

Based on the performance SVM are considered as the most effective methods for classification tasks, as confirmed by most of the researchers. From computational learning theory the principle basically revolves on the structural risk minimization techniques. The support vectors search of the decision surface and then separates the data points into several different or the multiple classes and later involves the decision making step.

C. *K-Nearest Neighbors (K-NN)*

The K-Nearest Neighbors is basically known as "lazy learner" – as it postpones the task of decision making of generalizations beyond the training data, the postponing is

done until every new query case has been located. It is the classifier that is depends on more number of examples. In a sequence to classify the text sample, the k-NN classifier approximately arranges the sample according to the training available training samples and then it uses the labels of the classes that belongs to the nearest k neighbors.

Given an example which is a question indicated by d, first the system finds out the k nearest neighbors among the set of the training questions. Then each for nearest neighbor question with similar score and that is relative to the example question, is employed as the weight of the classes for the neighboring question in classification. The weighted sum in k-NN classification can be written as shown in equation 1

$$score(d,ti) = \sum_{dj=KNN(d)}^{ci} sim(d,dj)\partial(dj,ci) \quad \text{.....(1)}$$

For the examination question d the k-NN(d) indicate the set of the nearest k neighbors. If dj belongs to ci, δ (dj, ci) is equal 1; otherwise, it is equal to 0. The examination question d must be belonging to the class that has results in the highest weighted sum. In order to compute *sim (d, dj)*, where the Euclidean distance is used.

IV. Methodology

The proposed methodology is implemented in Python 3.5 Version. It is initialized through simple command line interface. Initially all the keyword lists, verb lists, questions are stored in CSV files and then operations are performed.

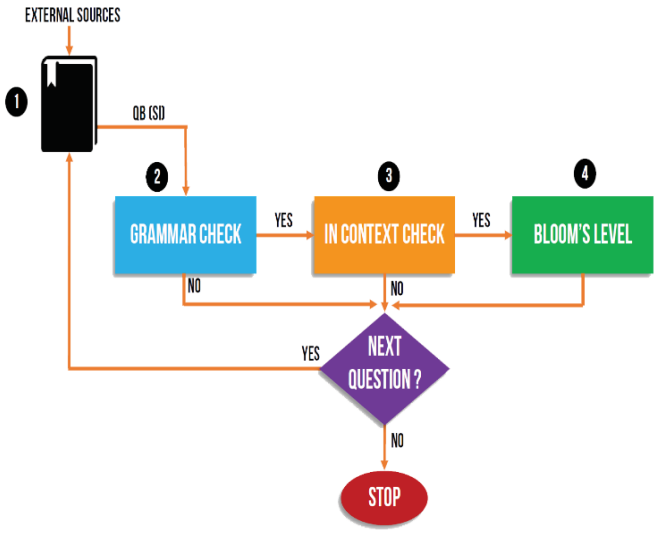

Fig. 2. Overview of Methodology

Step 1: Data set collection and pre-processing - The primary task is to collect a sufficient amount of data to make a comparative study as effective as possible. The data set consist of 1000 questions collected from various external sources like textbook, question papers, pedagogical activity and Q and A website. In this study we have taken questions related to operating system course. Apart from this, the catalog of keywords has to be collected, that is the list of action verbs pertaining to the cognitive skill domains of revised bloom's taxonomy was initially published by "Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing, Abridged Edition. Boston, MA: Allyn and Bacon" this list is enhanced by many others and is available in the public domain and we also take the synonyms of these key verbs to enhance the key wordlist. Along with this the keywords of the particular context/subject we are working on has to be collected, this in context keywords are taken from the index of the text book, as text book is our mini world. In the pre-processing phase initially we remove all the stop words, punctuation marks from the questions.

Step 2: Grammar checking phase - The data set is subjected to grammar checking phase to remove the grammatically incorrect sentences and ambiguous sentences. This is implemented using the English grammar checker Ginger API.

Step 3: In context checking - The data set is subjected to the context checking phase to remove the questions which are grammatically correct but does not belong to the subject context on which we are working. So, this step gives us the verified data set, which strictly belongs to our field of interest and domain. The question is said to be in context if it has at least one word that matches the keywords taken from the index of the text book, where text book is our mini world.

Step 4: Revised Bloom's level - The questions collected are then classified into one of the six levels of the Cognitive skill domain by utilizing specific key verbs. Also each of the levels is assigned a numerical value in increasing order starting from 1 for the "Remember" level and 6 for "Create" level. The classification is done by removing the punctuation symbol from the given question and splitting the stripped question on the space character. The individual words from the keyword list were matched with the words in the question to identify if they belonged to any of the levels in the domain. If they belong to the particular level, then the questions were assigned to that level with the numerical value assigned to that level. The keyword/verb list has some keywords which are ambiguous and belong to two levels at the same time, for example "choose" belong to both the levels "Evaluate" and "Create". In this situation we consider keyword in both the list but assign question to the higher level that is "Create", assuming one has to have higher level knowledge to answer the question.

Further, this process is continued for all the questions in the question bank. After step 2 if the question is grammatically incorrect we will directly go to the next question else will go for step 3 now if the question is in context the Bloom's level is assigned, if the question is not in context we will go to the next question and repeat step 2 to step 4 and if next question is last question will stop or else will again take a new question from the question bank and repeat from step 1 and the process continuous till the last question in the question bank.

V. Results and Discussion

This section includes the results which are obtained by following the proposed methodologies, on analyzing the results we can see the data and the output, which includes sample key verbs and their synonyms for each Bloom's level in figure 3, sample keyword list for in context checking in figure 4 and the set of questions with assignment in figure 5. Initially after execution of each step the minimum number of columns added to the output is 1 that is approved or not approved, which is got for grammar checking phase as shown in figure 5 and the maximum number of columns added are three that is approved, in context and Revised Bloom's taxonomy level. So, basically the question with single column are discarded for evaluation as they indicate questions are not approved. Later, questions with 2 columns indicate they are approved and are not in context of the textbook these questions are also discarded for assigning Bloom's level. Further, questions which are not- approved and which are not in context are not used for other calculations. Finally, we get complete classification shown as output with 3 columns for the proper questions that are grammatically correct, in context, with reading scores and are assigned with the Revised Bloom's level. The questions collected are restricted to be theoretical i.e., numerical problems are excluded and the key words for in context checking are limited to the text book which can be enhanced by using more textbooks.

REMEMBER	UNDERSTAND	APPLY	ANALYZE	EVALUATE	CREATE
choose	classify	apply	analyze	agree	adapt
define	compare	build	assume	appraise	build
find	difference	choose	categorize	assess	choose
how	demonstrate	construct	classify	award	change
label	explain	develop	conclusion	value	combine
list	extend	experiment with	contrast	choose	compile
match	illustrate	identify	discover	compare	compose
name	infer	interview	dissect	conclude	construct
omit	interpret	make use of	distinguish	consider	create
recall	outline	model	divide	criteria	delete
relate	relate	organize	examine	criticize	design
select	rephrase	plan	function	decide	develop
show	show	select	inference	deduct	discuss
spell	summarize	solve	inspect	defend	elaborate
tell	translate	utilize	list	determine	estimate
what	state	enumerate	motive	disprove	formulate
when			relationships	evaluate	happen
where			simplify	explain	imagine
which			survey	importance	improve
who			take part in	influence	invent
why			test for	interpret	make up
			theme	judge	maximize
			difference	justify	minimize

Fig. 3. Sample Key Verb List for Cognitive Skill Domain

As we execute our python script, For e.g. 1: question 0 is How is paging different from segmentation, which is grammatically correct, then for in context checking we pass the question through keyword list where the word "paging" matches "paging" in the operating system keyword list and question proves to be in context and for the cognitive category determination we check for the verbs which are present in the question by passing it through the list of key verbs, now in this key the verb "How" matches with the "How" of the list pertaining to the "Remember" level or level 1 of the Revised Bloom's Taxonomy, so we categorize this question into the level 1 of the taxonomy. In e.g. 2: question 20 is What are the differences between semaphore and monitors?, which is grammatically correct, then for in context checking we pass the question through keyword list where the word 'semaphore' matches and question proves to be in context and for the cognitive category determination we check for the verbs which are present in the question by passing it through the list of key verbs, now in this key the verb "What" and "difference" matches with the list pertaining to the "Remember" and "Evaluate" of the Revised Bloom's Taxonomy, In this situation we consider keyword in both the list but assign question to the higher level that is "Evaluate", assuming one has to have higher level knowledge to answer the question. So, we categorize this question into the level 4 of the taxonomy.

OPERATING SYSTEM CONTEXT KEY WORDS			
operating system	entry section	android	bankers algorithm
zones	mutual exclusive	android operating system	beladys anamoly
monitor	resouce allocation	apple ipad	best fit
limit register	allocation	apple ipad	worst fit
fat	buddy system	application containment	removable storage media
file-allocation table	indexed allocation	application interface	remote login
chrome	linked allocation	api	random access memory
faster read	contigous allocation	aqua interface	raw disk
operation	request edge	arm architecture	spooling
copy-on-write	assignment edge	arrays	backup and restore
thrashing	latency	assembly language	privilaged mode
windows nt	layered approach	asynchronous threading	multi-threaded
race	thrashing	atomic	hash maps
worm	direct access	authentication	hdfs
counter masures	device drivers	socket	hibernation
deadlock	interrupt	background processes	background
externel fragmentation	turn around time	bash	history
probing	response time	amd64 architecture	hybrid cloud
virtual memory	busy waiting	amdahls law	hybrid operating systems
main memory	spoofing	bitmaps	hybrid systems
secondary memory	spinlock	block	hypercalls

Fig. 4. Sample Key Word List for in context checking

VI. Comparison of Revised Bloom's Level Classification Using SVM and K-NN

To validate our proposed methodology i.e., classification of questions into Revised Bloom's level and to make the comparative study using machine learning techniques namely, SVM and K-NN we are using an open source tool named orange: Orange is a free software for machine learning and data mining (package written in Python). It has a visual programming front-end for explorative data analysis and visualization and can also be used as a Python library. For this method, which gives the classification of questions into 6 different levels, the data set required is the question number and the verbs associated with that questions in case of test data (as we are validating our methodology the questions for test data are taken from the data collected in methodology). So, we use excel format to represent this data, where row represents the verbs and columns has question numbers and here the value 1 is used to denote the verb belongs to that particular question and 0 on the other columns and rows.

QUESTIONS	RESULT
how is paging different from segmentation?	Question 0 - approved in os context level 1
Some good	Question 1 not approved -
why should an operating system manage resources on behalf of user?	Question 2 - approved in os context level 1
when an operating system hangs what is the implementation with respect to hardware resources?	Question 3 - approved in os context level 1
what is page-fault in non-paged area?	Question 4 - approved in os context level 1
what is the role of operating system in case of ssd mass storage?	Question 5 - approved in os context level 1
at your dictionary we try to give you all of the tools you need to really understand what the word means.	Question 6 - approved not in os context
distinguish between node and file system block	Question 7 - approved in os context level 4
what happens to a process in the transient portion when operating system demands and no memory is free in user work area?	Question 8 - approved in os context level 6
how is virtual file system implemented internally with respect to i node?	Question 9 - approved in os context level 1
opearting system has three main function how are they scheduled?	Question 10 - approved in os context level 4
can a semaphore and a mutex be interchangable?	Question 11 - approved in os context level 3
do thread scheduling and process scheduling happen together?	Question 12 - approved in os context level 6
The close of this primal miscreant babble indicates strongly that species.	Question 13 not approved -
what happens if a system allows a file system to be mounted simultaniously at more than one location?	Question 14 - approved in os context level 6
in multi operating system implemented through virtual machine how does the guest operating system implement paging? does it get a raw disk?	Question 15 - approved in os context level 5
I dont like mangoes	Question 16 - approved not in os context
what are the difference between ntfs & fat?	Question 17 - approved in os context level 4
what should be done if a program needs more space than las?	Question 18 - approved not in os context
That medicine affects my sleep ability.	Question 19 not approved -
what are the differences between semaphore and monitors?	Question 20 - approved in os context level 4
how does operating system choose disk scheduling algorithms between single and multi users?	Question 21 - approved in os context level 6
seeing the word in a sentence can provide more context and relevance	Question 22 - approved not in os context
what happens transaction in log file when the system crashes?	Question 23 - approved in os context level 6
what is the difference between local and global page replacement ?	Question 24 - approved in os context level 4
in case solid state devices replace hard disk devices completely what are possible disk related enhancements?	Question 25 - approved in os context level 2

Fig. 5. Sample Dataset with Output Assignment

Similarly, for training data we take the key verbs which was initially published by Krathwohl and Anderson in their work - "A taxonomy for learning, teaching, and assessing, Abridged Edition. Boston, MA: Allyn and Bacon" and we also take the synonyms of these key verbs to enhance the key wordlist. where row represents the verbs and columns has numbers and here the value 1 is used to denote the verb with respect to that particular column number and 0 on the other columns and rows.

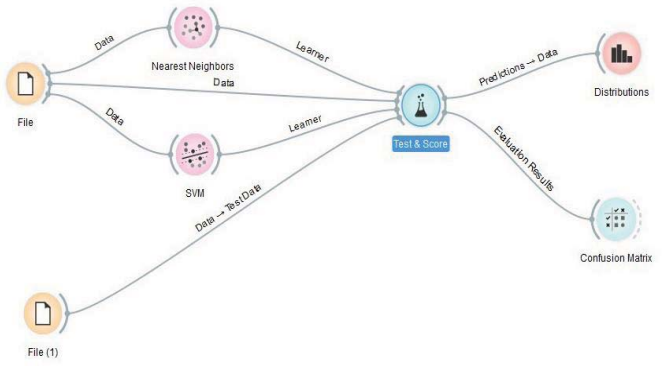

Fig. 6. Demonstration of orange tool

In the figure 6, the file indicates the train data and the file (1) indicates the test data, these data are connected to the 2 different classifiers namely SVM and K-NN. For SVM the parameters chosen are linear kernel and for K-NN the neighbors are set to 5 and Euclidean Distance is the metric used. And connected to the test and score module so this test and score model gives the values accuracy, precision, recall and f1 measure for the classifiers by evaluating the data. Which can be seen in figure 7.

METHOD	AUC	CA	F1	PRECISION	RECALL
KNN	0.605	0.338	0.252	0.272	0.338
SVM	0.690	0.362	0.496	0.990	0.362

Fig. 7. Test and score results (Case 1)

Later, this test and score data is connected to the distributions to get the distribution chart that represents the questions classified into 6 different levels which can be seen in figure 8.

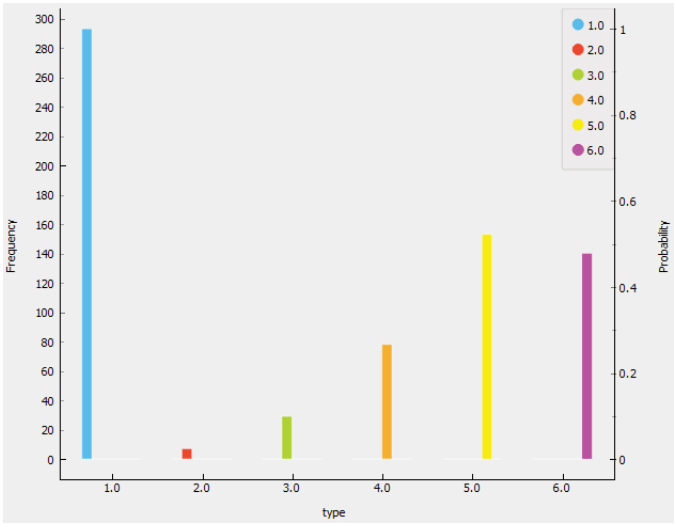

Fig. 8. Distribution results of test and score data (Case 1)

The test and score evaluated results are given to the confusion matrix which can be seen in figure 9 for K-NN and figure 10 for SVM. In the confusion matrix, the matrix shows the actual and predicted values for each class, i.e., the number of questions classified in that level. For example, in the SVM classifier in there are 422 questions in level 1, Out of these 100 are actually classified to class 1 and 322 are predicted to level 2. Based on this the accuracy for SVM is got to be 0.690 (69%) and for K-NN is 0.605 (60.5%), as discussed in section background, the K-NN is a 'lazy learning' because it postpones the decision making of generalizations beyond the training data until every new query case has been located and accuracy is bit higher for SVM and it is more effective approach (case 1).

Actual \ Predicted	1	2	3	4	5	6	∑
1	0	0	0	0	0	422	**422**
2	0	6	0	0	0	4	**10**
3	0	0	0	0	0	48	**48**
4	0	0	0	0	0	116	**166**
5	0	0	0	0	152	72	**224**
6	0	0	0	0	0	180	**180**
∑	**0**	**0**	**0**	**0**	**152**	**842**	**1000**

Fig. 9. Confusion matrix results for the test and score data of K-NN (Case 1)

Actual \ Predicted	1	2	3	4	5	6	∑
1	100	322	0	0	0	0	**422**
2	0	10	0	0	0	0	**10**
3	0	19	29	0	0	0	**48**
4	0	38	0	78	0	0	**116**
5	0	177	0	0	47	0	**224**
6	0	82	0	0	0	98	**180**
∑	**100**	**648**	**29**	**78**	**47**	**98**	**1000**

Fig. 10. Confusion matrix results for the test and score data of SVM (Case 1)

When these results are compared with results got from testing on train data we get higher accuracy as shown below (case2) shown in figure 11, Here the accuracy levels are 0.923 (92.3%) for SVM and 0.666 (66.6%) for K-NN. The distribution chart in figure 12 represents the questions classified into 6 different levels.

METHOD	AUC	CA	F1	PRECISION	RECALL
KNN	0.666	0.590	0.530	0.551	0.590
SVM	0.923	0.904	0.902	0.913	0.904

Figure 11: Test and score results (Case 2)

In the confusion matrix shown in figure 14, the matrix shows the actual and predicted values for each class. For example, in the SVM classifier in there are 293 questions in level 1. Out of these 293 are perfectly classified to level 1. Similarly, in K-NN out of 293, 288 are actually classified to level 1 and 4 are predicted to level 5 and 1 is predicted to level 6. The accuracy for SVM is got to be 0.923 and for K-NN is 0.666.

Later the SVM and K-NN classification results in both Case 1 and Case 2 indicate the effectiveness of the machine learning techniques. When compared SVM proves to be better learner in both the cases. Using the proposed method, the question paper can be set by the tutor which usually covers all the type of questions and all the levels of the taxonomy.

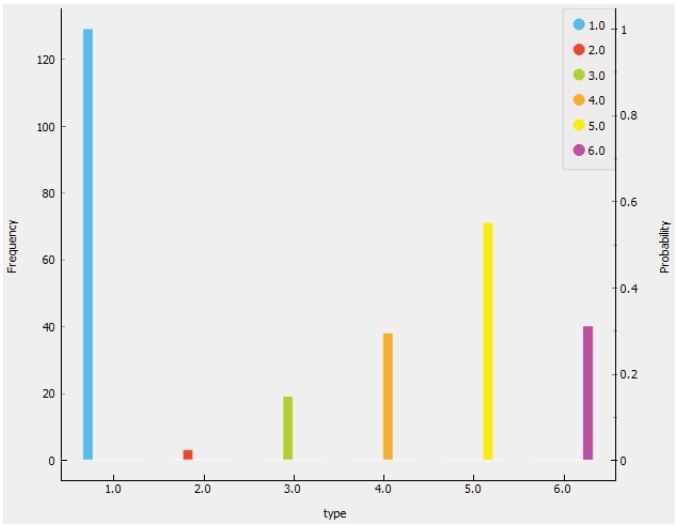

Fig. 12. Distribution results of test and score data (Case 2)

Actual \ Predicted	1	2	3	4	5	6	∑
1	288	0	0	0	4	1	**293**
2	4	0	0	1	1	1	**7**
3	18	0	0	1	5	5	**29**
4	62	0	0	5	8	3	**78**
5	70	0	0	11	70	2	**153**
6	39	0	0	12	39	50	**140**
∑	**481**	**0**	**0**	**30**	**127**	**62**	**700**

Fig. 13. Confusion matrix results for the test and score data of K-NN (Case 2)

Actual \ Predicted	1	2	3	4	5	6	∑
1	293	0	0	0	0	0	**293**
2	0	7	0	0	0	0	**7**
3	0	0	29	0	0	0	**29**
4	0	0	0	78	0	0	**78**
5	29	0	0	0	124	0	**153**
6	22	0	0	0	16	102	**140**
∑	**344**	**7**	**29**	**78**	**140**	**102**	**700**

Fig. 14. Confusion matrix results for the test and score data of SVM (Case 2)

A. Evaluation

The effectiveness of the SVM and the K-NN are done using these common measures:

1) Accuracy

The accuracy of a classifier is defined as the percentage measure of the capability of the system to classify all the questions into the proper category as denoted by equation 2.

$$Accuracy = \frac{(Tp + Tn)}{N} \quad \text{.....(2)}$$

2) Precision

The percentage measure of the capability of a system to retrieve only relevant Questions, as denoted by equation 3.

$$Precision = \left(\frac{Tp}{Tp + Fp}\right) \quad \text{.....(3)}$$

3) F-measure

The harmonic mean taken between the recall and precision, as denoted by equation 4.

$$Fmeasure = \frac{2\lambda(Precision + Recall)}{Precision + Recall} \quad \text{.....(4)}$$

4) Recall

The percentage measure of the availability of all relevant Questions classified by the system, as denoted by equation 5.

$$Recall = \left(\frac{Tp}{Tp + Fn}\right) \quad \text{.....(5)}$$

These values of the percentage measures are used for the single class classification system, and the measurement of the effectiveness across the set of classes is done by the macro average. And, the terminologies used in equation 2 to 5 indicates these: True Positive –Tp gives the set of questions that are assigned to the correct category given. False Positive - Fp gives the set of questions that are assigned to the incorrect category given. False Negative Fn – has the set of questions that are not assigned to the incorrect category, and True Negative - Tn has the set of questions that are not assigned to the correct category.

VII. Further Enhancement

The future prospect of the paper will focus on the Knowledge level classification. More work will be done on experimenting with other question classification techniques. However, the concern will be in improving the accuracy of the results, and finally the question collection system will be made automatic as the present method is tedious and time consuming.

VIII. Conclusion

Questions are the primordial component. The proposed methodology is to make a comparative study of question bank classification based on Revised Bloom's Taxonomy using support vector machines and k-nearest neighbors, which helps to build cognitive skills of the learner. Here the grammatically proper and in context questions will be classified into different Bloom's cognitive skill levels. Proposed methodology also explores the effectiveness of the machine learning techniques which classifies the questions into Bloom's cognitive skill domain, which is of particular importance in question bank management systems. Hence our results and discussions show the classification of questions based on the proposed model and the comparison shows the verification of our model which indicates that our model is working. The Support Vector Machine and K-Nearest Neighbors are used for cognitive skill domain classification; their results indicate the effectiveness of the machine learning techniques. Using this method, the question paper can be set by the tutor which usually covers all the type of questions and all the levels of the taxonomy.

References

[1] Anwar Ali Yahya and Addin Osman, Automatic Classification of Questions into Bloom's Cognitive Levels Using Support Vector Machines, Faculty of Computer Science and Information Systems, Najran University, Najran, Kingdom of Saudi Arabia.

[2] Dhuha Abdulhadi Abduljabbar and Nazlia Omar, Exam Questions Classification Based On Bloom's Taxonomy Cognitive Level Using Classifiers Combination, Proceedings of Journal of Theoretical and Applied Information Technology Vol.78. No.3, 31st August 2015.

[3] Krathwohl, D. R.; Bloom, B. S.; Masia, B. B. -Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II: The Affective Domain. New York: David Mckay Company, 1964.

[4] Leslie Owen Wilson and ED. D, A Succinct Discussion of Anderson and Krathwohl – Bloom's Taxonomy Revised – The Second Principle and How to Use Them Effectively, Link - http://thesecondprinciple.com/teaching-essentials/beyond-bloom-cognitive-taxonomy-revised/, 16th September 2016.

[5] Committee of College and University Examiners. Taxonomy of educational objectives. Vol. 1. New York: David mckay, 1956.

[6] Wen Chih Chang and Ming Shun Chung, Automatic Applying Bloom's Taxonomy to Classify and Analysis the Cognition Level of English Question Items, 3-5 Dec. 2009 IEEE Conference at Taiwan, 25th February 2010.

[7] Syahidah Sufi Haris and Nazlia Omar, Bloom's Taxonomy Question Categorization Using Rules and N-Gram Approach, Journal of Theoretical and Applied Information Technology Vol.76. No.3, 30th June 2015.

[8] Anderson L.W., Sosniak L.A., Bloom's taxonomy: a forty-year retrospective. Ninety-third yearbook of National Society for the Study of Education, Pt.2., Chicago, IL., University of Chicago Press, 1994.

[9] Orange an Open source machine learning and data visualization for novice and expert. Interactive data analysis workflows with a large toolbox. It is maintained and developed at the Laboratory of Bioinformatics, Faculty of Computer and Information Science, University of Ljubljana. How does orange tool operate, link: http://orange.biolab.si?

[10] Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python. Journal of Machine Learning Research 14(Aug):2349−2353.

[11] Ginger grammar checking API, how does it work https://github.com/zoncoen/python-ginger