Journal of Theoretical and Applied Information Technology
31st August 2015. Vol.78. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EXAM QUESTIONS CLASSIFICATION BASED ON BLOOM'S TAXONOMY COGNITIVE LEVEL USING CLASSIFIERS COMBINATION

[1] DHUHA ABDULHADI ABDULJABBAR, [2] NAZLIA OMAR

[1, 2] Knowledge Technology Group, Center for AI Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

E-mail: [1]dhuhaftsm@yahoo.com, [2] nazlia@ukm.edu.my

ABSTRACT

Assessment through written examination is a traditional method but it is a universal test method practiced in most of the educational institutions today. Therefore, the question must be provided in accordance with the subject content learned by students to fulfil learning objectives. However, the process of questions writing is very challenging step for the lecturer. The situation is getting more challenging when lecturers try to produce good quality and fair questions to assess different level of cognitive. Thus, the Bloom's Taxonomy has become a common reference for the teaching and learning process used as a guide for the production of exam questions. Exam questions classification presents a particular challenge is the classification of short text questions due to short text involves text with less than 200 characters. In addition, the features of short text are very sparse and far. This study proposed a new method to classify exam questions automatically according to the cognitive levels of Bloom's taxonomy by implementing a combination strategy based on voting algorithm that combines three machine learning classifiers. In this work, several classifiers are taken into consideration. The classifiers are, Support Vector Machine (SVM), Naïve Bayes (NB), and k-Nearest Neighbour (k-NN) that are used to classify the question with or without feature selection methods, namely Chi-Square, Mutual Information and Odd Ratio. Then a combination algorithm is used to integrate the overall strength of the three classifiers (SVM, NB, and k-NN). The classification model achieves highest result through the combination strategy by applying Mutual Information, which proved to be promising and comparable to other similar models. These experiments aimed to efficiently integrate different feature selection methods and classification algorithms to synthesize a classification procedure more accurately.

Keywords: *Bloom's Taxonomy, Exam Questions, Voting Algorithm, Machine learning, Feature Selection*

1. INTRODUCTION

The objective of teaching and learning can be achieved by implementing assessments to measure students' cognitive levels. Therefore the written exam is a medium for educationists to confirm student knowledge and understanding, as well as, to assess the extent to which students are able to adapt a learning theory in a given situation [1].

Taxonomy produced by a group of educational psychologists [2] is a system based on educational objectives that classify the level of learning and understanding according to the levels of the taxonomy. There are three domains in Bloom's Taxonomy (BT): Cognitive, Affective and Psychomotor. At this moment, cognitive domain had received the attention of this study as it is closely related to the students' knowledge and understanding in the classroom. This domain is synonymous with the assessments used in any educational institution because it explains how cognitive response and attitude are interrelated. It consists of six levels to induce students to 'climb to a higher level, or step, of thought' starting with the simplest of knowledge, then comprehension, application, analysis, synthesis and, finally, evaluation. Table 1 shows each cognition level as well as illustrating verbs and simple examples that highlight the thought activity type of that level with examples of questions that agree with Bloom's Cognitive Level.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Six Categories Of The Cognitive Domain Based On Bloom's Taxonomy With Illustrative Examples

Category	Sample Keywords	Sample Behaviours	Sample Questions
Knowledge:	Arrange, recognize, relate, label, list, memorize, recall, define.	Students are able to define the 6 levels of Bloom's taxonomy	Define Inheritance concept.
Comprehension	Classify, explain, indicate, locate, recognize, describe, discuss, express.	Students can explain the purpose of Bloom's taxonomy	Explain the structure of a method in the program.
Application	Demonstrate, sketch, illustrate, operate, practice, schedule. predict, explain, interpret, employ, solve, use, write.	Students write an instructional objective for each Bloom's taxonomy level.	Demonstrate the relationship of all the packages, classes and methods of the program.
Analysis	Analyze, appraise, make a distinction calculate, categorize, differentiate, discriminate, distinguish, examine, list.	Students compare and contrast cognitive and affective domains.	List the advantages and disadvantages of using a container class such as ArrayList in place of an array.
Synthesis	Arrange, assemble, create, design, collect, compose, develop, set up, propose, write, organize, plan.	Students design a classification method for educational objectives that combine cognitive, affective, and psychomotor domains.	Write a JAVA program to show the Overloading concept.
Evaluation	Appraise, judge, predict, assess, attach, compare, defend choose, estimate, rate, compare, Justify.	Students judge the effectiveness of Bloom's taxonomy.	Justify the concept of inheritance and give the sample of code to illustrate your answer.

Generally questions classification can be defined as the selection of a category for a given question from the predetermined questions categories based on their contents. The question classification is a unique form of text classification. A significant difference is that questions classification presents the issue of handling short sentences, unlike text documents, and hence, less information is available on each question that occurs [3] [4].

It would be inappropriate to use only statistical methods (such as TF-IDF, C-value, and N-gram) to classify the exam questions into Bloom's Taxonomy category [5]. This is due to the fact that statistical techniques require large data in each document to obtain high accuracy [6]. In fact, the accuracy will decrease when small data is used for a training document due to the lack of terms in text. In addition, the presence of terms that contribute to the density of text features is sparse and far. On the other hand, the rule-based approach uses rules determined manually by knowledgeable engineers, with the help of domain experts. However, developing such rules is tedious and time consuming [7]. Rule-based systems lack portability and robustness abilities. Additionally, the high cost of rules maintenance goes up, even when data is only marginally altered. In addition, the use of rules is not effective when large data are employed since a large set of rules must be developed. Over recent years, in an effort to overcome some of these difficulties, supervised machine learning approaches have been formulated for the classification of questions [8] [7] [9]. Machine learning approaches have a tendency to attain a quite high degree of accuracy for the task of questions classification in less possible effort.

To overcome the problem of exam question classification with a more effective solution, this study proposes a combination model which combines three machine learning approaches using a combination voting algorithm adopted to classify question items to agree with Bloom's cognitive levels.

2. LITERATURE REVIEW

Many studies have sought to automatically classify exam questions based on Bloom's taxonomy. Furthermore, although limited, research has addressed the use of Natural Language Processing (NLP) techniques to resolve this problem.

Therefore, some previous studies that proposed approaches similar to the model introduced in this work should be reviewed. Van Hoeij et al. (2004) developed a classification-based tool that uses Bloom's anatomy to evaluate the cognitive level of short essay questions.

Journal of Theoretical and Applied Information Technology
31st August 2015. Vol.78. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Experimental results revealed only moderate agreement on the general classification of the test items when a simplified classification tool based on Bloom's taxonomy was used. However, given more instructions and an enhanced classification algorithm, this approach may be useful in examination quality control. Therefore, this model requires more instructions and an improved algorithm for classification to be considered for application in a real-world scenario for the quality control of school exams.

Chang and Chung (2009) also applied Bloom's taxonomy to evaluate and classify English question item's cognition level. Their research included 14 general keywords for Bloom's taxonomy and considered 288 test items. They determined that the "knowledge" level of cognition has a 75% correct match. This result is relatively not high enough to be considered for enhancement.

Haris and Omar (2012) employed a rule-based approach for question classification using Bloom's taxonomy in NLP. A rule-based approach evaluates and classifies written examination questions for computer subjects. Experimental results confirm that the technique can support the automatic labelling and categorization of questions on the basis of Bloom's taxonomy. This technique is not as dynamic as machine learning techniques, but nonetheless exhibits satisfactory performance. Rule-based techniques are not only time consuming and expensive to implement but are also not as robust as other techniques, such as machine learning. In addition, this technique is only limited to computer subjects and works poorly in other domains.

Yusof and Hui (2010) created a model that categorized question items using an Artificial Neural Network (ANN) technique that used multiple feature methods. Of the three feature sets used, only DF reductions were performed efficiently with along with the addition of classification and convergence time. Meanwhile, the others did not perform adequately.

Some previous works have been reviewed, and their strengths and weaknesses have been considered. Hence, this study aims to model a machine learning classification framework on the basis of a classifier combination model for handling the problem of cognitive category determination for computer programming-related questions in an attempt to achieve better performance and accuracy levels than previously proposed models.

3. RESEARCH METHODOLOGY

The proposed method of this study incorporates a combined strategy utilising three machine learning approaches to classify the question components that match Bloom's cognitive level. To determine the question category, this method initially assigns three categories using three machine learning classifiers. Next, the final category was assigned to the question by combing the decisions from these approaches using Voting algorithm. Figure 1 shows the proposed combination model used to classify the question components into their corresponding Bloom's cognitive level.

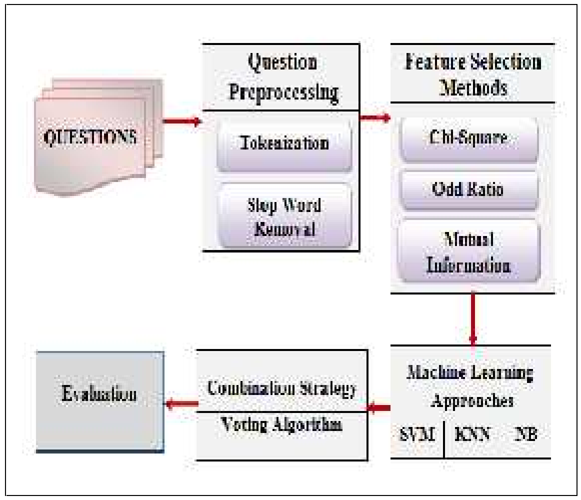

Figure 1: Proposed Combination Model for Classifying Exam Questions

As shown in Figure 1, constructing a Bloom's taxonomy question classification system using a machine learning approach requires many computational steps that include data planning, pre-processing, feature selection, classification and evaluation. The overall process of the machine learning approach to determine the Bloom's Taxonomy category of a certain question is described explicitly in the following:

3.1 Data Set Planning and Compiling Phase

The use of supervised machine learning techniques depends on the existence of training data. Such data is usually produced manually and labelled by experts in the field of relevance. Exam questions bank consisting of programming questions for the Bachelor of Information Technology Year 1 to 3 in the Faculty of Information Science and Technology / UKM from 2006 to 2011was utilised for the training and test data sets. This data set was obtained from [12]. In addition, more effort has been made

Journal of Theoretical and Applied Information Technology
31st August 2015. Vol.78. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to collect many questions from various web sites of Bloom's taxonomy works in programming subject to increase the training data sources.

3.2 Pre-Processing Modules Phase

The pre-processing phase is a highly important phase in any system that uses a machine learning framework. Before feature selection and classification, all questions in the system were subject to the following pre-processing steps:

- Tokenizing of words depending on white space and punctuation marks.
- Stopword removal task, including the elimination of non-letter, punctuation marks and stop words.

3.3 Feature Selection (FS):

FS is an important step in any type of classification, such as in question classification systems. FS improves the efficiency of classification tasks in terms of speed and effectiveness of learning. FS methods (FSM) minimise the range of data proportions. Furthermore, such methods eliminate superfluous, redundant, and noisy data. In this work, several FSMs were selected based on the subsequent aspects [14]: the data size and data consistency. We selected three feature selection methods, namely, Mutual information, Odd ratio and Chi-square.

3.3.1 Mutual Information

The efficiency of the classifier might also be enhanced by eliminating some of the less beneficial attributes. One of the popular feature selection measurements is Mutual Information (MI). The expected value of pointwise Mutual Information *MI*(*t*, *c*) between the term *t* and category *c* is defined on the level of their co-occurrence [15].

$$MI\,(t,c) = \log \frac{P_r\ (t \wedge c)}{P_r\ (t) \times P_r\ (c)} \qquad (1)$$

3.3.2 Chi-Square statistic (X^2)

The X^2 statistic is one of the most popularly used feature selection algorithm. X^2 statistics measure the lack of flexibility between the category and the term [16]. It can be defined by the following equation:

$$X^2\,(t,c) = \frac{N(AD-CB)}{(A+C)(B+C)(A+B)(C+D)} \qquad (2)$$

where A is the number of times t and c co-occur, B is the number of times t occurs without c, C is the number of times c occurs without t, D is the number of times neither c nor t occurs, and N is the total number of training questions.

3.3.3 Odd Ratio

In text classification, Odds Ratio (OR) was initially proposed for selecting relevant feedback terms. The general idea is that the distribution of relevant document features differs from the distribution of irrelevant document features [17]. Given a category $y_i \in Y$, a feature term t belongs to one or more documents in X. Let A equal how many times t is present in y_i, B is how many times t is present without y_i, C is how many times t is absent in y_i, D is how many times t is absent without yi and n is the size of the training set. OR methods compute the score of t belonging to y_i, as follows:

$$OR\,(t, y_i) \approx \frac{A\,D}{C\,B} \qquad (3)$$

After that calculating max score for each FS methods between term t and category c as:

$$FS_{max}\,(t) = \max_{i=1}^{m}\,(FS\,(t, c_i)) \qquad (4)$$

3.4 Classification

The following classifiers were used as the main classifiers in this work:

- Support Vector Machine (SVM) classifier.
- Naïve Bayes (NB) classifier.
- k-Nearest Neighbour (k-NN) classifier.

The classifiers are briefly described in the following:

3.4.1 Support Vector Machine (SVM)

Vladimir Vapnik and his team members developed the Support Vector Machine (SVM) in the early Nineties [18]. This is a well-known supervised learning algorithm that has been very successfully applied to many problems concerning classification tasks, such as text classification and, more recently, questions classification [9]. The basic idea of the SVM for a two-class learning problem is to find a linear decision boundary (also called a separating hyperplane) between two classes, which is as far as possible from the closest training instances within each class-the support vectors.

Although several variations of SVM have been created [19], this paper only discusses linear SVM; due to its popularity and excellent text classification [20]. SVM optimization SVM

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(dual form) reduces:

$$\vec{\alpha} = argmin\{-\sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (\vec{x_i}, \vec{x_j})\}$$

$$Subject\ to: \sum_{i=1}^{n} \alpha_i y_i = 0;\ 0 \le \alpha_i \le C \quad ....(5)$$

3.4.2 Naïve Bayes (NB)

The Naïve Bayes classifier has been used extensively in many applications involving classification tasks, such as text and question classification, where it achieves state-of-the-art results despite its simplified naive assumptions. It is the most common approach for questions classification [21]. With the help of a feature vector table, the algorithm is able to pull out the rear possibility, which is the word or term connected to numerous named entity (NE) classes, and assign it to the category that has the maximum rear possibility. Two major advantages of NB text classification algorithms are that they are easy to implement and often have superior performance. NB classifiers solve text classification problems as follows: given a document *d*, which is represented as a set of feature terms { t_k | k=1,2,…, |d| }, and c is a category in the category set C , where $|C| \succ 2$. Naive Bayes can be defined as the conditional probability of *c* given *d* constructed as follows:

$$P(c_i \mid d) = P(c_i \mid t_k, \ldots, t_{|d|}) = p(c_i) \prod_{k=1}^{d} P(t_k \mid c_i) \quad (6)$$

Where: $P(C_i) = \frac{N_i}{N}$ (7)

Where N_i is the number of documents associated with class C_i, and N the number of classes, and

$$P(t_k \mid C_i) = \frac{1 + n_{ki}}{1 + \sum_{h=1}^{l} n_{hk}} \quad (8)$$

Where n_{ki} is the total number of documents that contain feature t_k and belongs to class C_i, l is the total number of distinct features in all training documents that belong to class C_i [22].

3.4.3 k-Nearest Neighbour (k-NN)

The k-Nearest Neighbour, which is a classifier based on examples, is also described as 'lazy learning' because it postpones the decision making of generalizations beyond the training data until every new query case has been located. In order to classify a text sample, the k-NN classifier roughly arranges the sample according to the training samples available and then using class labels of the k most nearest neighbours.

Given the example test question d, the system identifies the k nearest neighbours amongst the training questions. Each nearest neighbour question similarity score, relative to the test question, is employed as the weight of the classes of the neighbour question classification. The weighted sum in k-NN classification can be written as:

$$score(d, t_i) = \sum_{d_{j=KNN(d)}} sim(d, d_j)\, \delta(d_j, c_i) \quad (9)$$

Where k-NN (d) indicates the set of k nearest neighbours to exam question d. If d_j belongs to c_i, δ (d_j, c_i) is equal 1; otherwise, it is equal to 0. Exam question d should belong to the class with the highest resulting weighted sum. In order to compute *sim(d, d_j)*, the Euclidean distance is used.

3.5 Combination Strategies Phase (Voting Algorithm)

A voting algorithm is used for the combination of the strength of all base classifiers. The label outputs of the participating classes of classifiers are used as features for the voting algorithm. To predict an unknown instance, the voting algorithm uses every classification model from its sub-process to determine the predicted class from the maximum votes given to the unknown test sample. This voting strategy determines test sample x class i with the most component predictions after counting the output of individual classifiers.

3.6 Performance Measures Phase

The performance of this study's exam question classification system is often gauged in terms of its precision (P), recall (R) and F-measure metrics. True Positive (TP) is the total number of questions correctly assigned to the category by a system. False Positive (FP) refers to the total number of questions incorrectly assigned to the category by a system. True Negative (TN) denotes the total number of questions correctly rejected by the system. False Negative (FN) refers to the total number of questions ignored by the system but belong to the category.

Precision is a percentage measure of the capability of a system to retrieve only relevant items.

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

Recall is a percentage measure of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

availability of all relevant items classified by the system.

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

For evaluating classification systems, F-measure is the most frequently used measure that combines precision and recall by function.

$$F_\beta = \frac{(1+\beta^2)*(precision * recall)}{(\beta^2 * precision + recall)} \quad (12)$$

For some values $0 \leq \beta \leq \infty$, usually, β is taken to be equal to 1, which means that the F_β function becomes

$$F_1 = \frac{2\ precision \times recall}{precision + recall}$$

Macro F-measure: first, the F_1-measure is evaluated locally for each category, these results are later summed up and the average of the F_1-measure is calculated.

$$F_1^{macro} = \frac{1}{m}\sum_{i=1}^{m} F_1(i) \quad (13)$$

4. RESULTS AND DISCUSSION

This section discusses the results from different classifiers of exam questions based on Bloom's Taxonomy. To examine the classifiers' performance without feature reduction, NB, SVM, and k-NN classifiers were primarily applied to the complete sample-term feature space. Experimental results for each Bloom's Taxonomy cognitive level using NB, SVM, and k-NN approaches are shown in Table 2 as F_1-measure. As shown, of the applications performed without any feature reduction method use, the highest performance was achieved using the k-NN approach on macro F_1-measure (80.82), and the worst performance was obtained with the NB classifier (74.95). Because k-NN is a simple non-parametric, the experiments showed that it is an effective classifier in the field of short text classification leading to a higher result than other approaches.

Table 2: The Performance (F_1-Measure) Of Bloom's Taxonomy Cognitive Level For Approaches Without Feature Selection

Cognitive Level	SVM	NB	KNN
Knowledge	79.9	66.66	79.9
Comprehension	80	80	80
Application	90	83.8	94.73
Analysis	72.71	72.72	83.33
Synthesis	79.9	79.9	80
Evaluation	66.66	66.66	67
Macro F_1	78.19	74.95	**80.82**

Different-sized sets of features were chosen from the feature space (50, 100, 150, 200 and 250) to test the efficiency of the three classifiers (NB, SVM and k-NN) with the three feature selection methods on the programming question and to evaluate the success of these methods individually or combined.

Table 3 summarises the reported results of the NB, SVM, and k-NN with the feature selection methods at different feature sizes on Bloom's taxonomy question classification. The results of these classifiers are presented in terms of macro F_1-measure. As shown in Tables 2 and 3, the use of feature selection methods improved the result of SVM, NB and k-NN in the majority of feature sizes. The highest results for SVM, NB, k-NN among three FSM were obtained when applying MI at feature size 250, Chi at feature size 200, and MI at feature size 250 respectively; Figures (2, 3, and 4) show the result of each Bloom's category as an F_1-measure for these higher results.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: The Performance (Macro F_1-Measure) Of The NB, SVM, K-NN Classifiers With Feature Selection Methods At Different Sizes Of Features

Macro F_1-measure									
Feature Size	SVM			NB			KNN		
	Chi	MI	OR	Chi	MI	OR	Chi	MI	OR
50	70.47	74.48	72.74	66.40	61.49	73.26	89.5	63.29	87.14
100	86.20	81.15	69.64	71	68.43	79.79	86.28	84.77	87.14
150	87.08	86.20	72.60	83.12	70.48	79.79	84.77	84.77	87.14
200	86.20	86.20	70.40	**83.42**	83.12	79.79	84.77	84.77	87.14
250	86.20	**88.82**	75.05	83.12	83.12	79.79	84.77	**89.8**	87.14

Finally, to study the results of the combination framework that combined three machine learning approaches, FSMs at different feature sizes on Bloom's taxonomy question classification were applied. The combination model consists of two phases. In the first phase, a set of base-level classifiers, (SVM, NB, and k-NN) with executing FSM were generated. In the second phase, a combination strategy that combined the outputs of the base-level classifiers using a voting algorithm was conducted. When using a voting algorithm for combination, the outputs of all of the class labels of participating classifiers were used as inputs for the algorithm. The question was associated with several classes from base classifiers, and then designated to the final category with the maximum vote. Table 4 contains the reported results of the voting algorithm in term (macro F_1-measure) on Bloom's taxonomy question classification, when using feature selection methods in base classifiers.

Table 4: Results (macro F_1-measure) of Bloom's taxonomy classification for combination three approaches

Combination Approaches			
Feature Size	Chi	MI	OR
50	85.91	72.45	73.64
100	79.01	70.83	80.42
150	86.91	86.92	80.42
200	91.35	86.92	80.42
250	90.35	**92.28**	80.42

As shown in Table 4, the highest performance on the exam question's classification task was obtained with the voting algorithm when using MI in a weighted feature size equal to 250, because the performance of MI depends on its bias towards favouring rare terms; Figure 5 showed the result of each Bloom's category in term of F_1-measure for this result. This means the voting combination can integrate the overall strength of the base classifiers and remove different types of mistakes by single classifiers. Better results were obtained in those classes that had keywords that were not found in other classes. However, poor results of some of the base classifiers can be attributed to the similarity of terms for each class and occurring vague terms in some of the classification due to the relatively short length of questions.

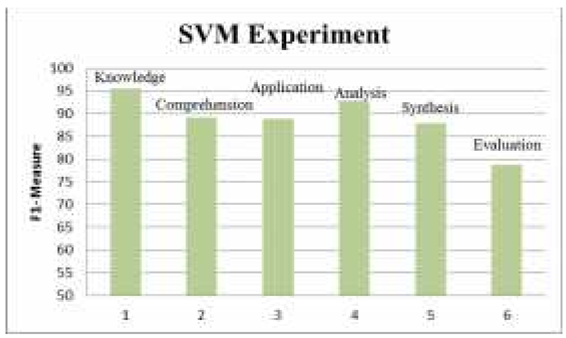

Figure 2: The Highest Result for SVM

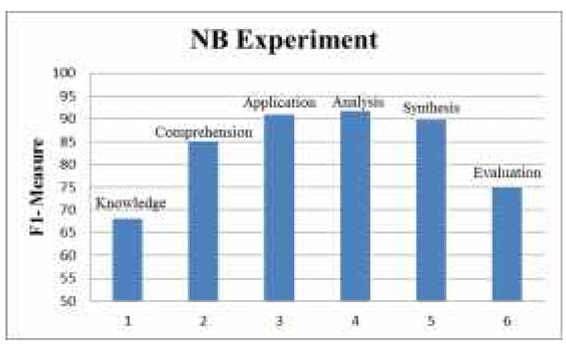

Figure 3: The Highest Result for NB

Journal of Theoretical and Applied Information Technology
31st August 2015. Vol.78. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

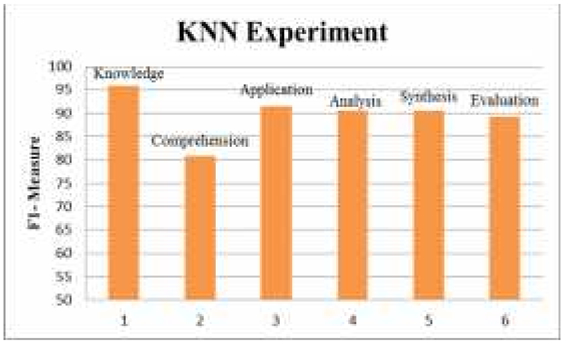

Figure 4: The Highest Result for k-NN

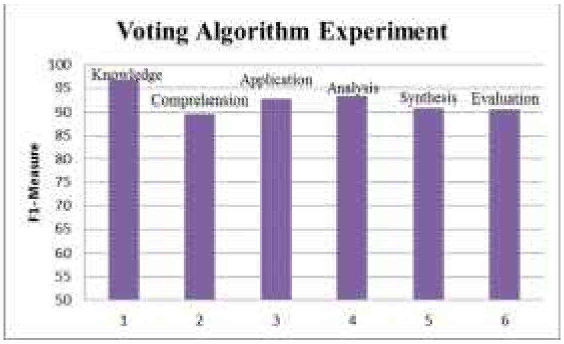

Figure 5: The Highest Result for Voting

As seen in Figures (2, 3, 4, and 5), the Voting algorithm achieves best result in each level of Bloom's taxonomy. Among the three machines learning the KNN achieves its best result in the Knowledge, Application, Synthesis and Evaluation meanwhile SVM achieves its best result in the Comprehension and Analysis. We can see that the highest performance is obtained when the feature selection methods made by Chi-square or MI. The research results were compared with the latest research on programming exam questions conducted by Haris and Omar, (2012) and proved to yield higher accuracy.

The next stage of this work will focus on experimentation of other feature selection methods to obtain better results. More effort will be taken on the experimentation of other question classification techniques to obtain more conclusive results in respect of the recall, precision, and F_1-measure

5. CONCLUSION

This study proposed a combination method based on a voting algorithm that combined machine learning approaches. The classification powers of all base classification models were combined. The results show that combination can outperform individual classifiers.

We conclude that the design of the proposed combination framework efficiently addresses the problem of cognitive category determination for programming questions and achieves satisfactory results. In our future work, this model will be supplemented through the use of other techniques to obtain enhanced results.

REFRENCES

[1] A. J. Swart, "Evaluation of Final Examination Papers in engineering: A case study using Bloom's taxonomy". Education, IEEE Transactions on, 53(2), 2010, pp. 257-264. DOI: 10.1109/TE.2009.2014221.

[2] B. S. Bloom, M. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl, "Taxonomy of educational objectives: Handbook I: Cognitive domain". New York: David McKay, 1956, 19, 56.

[3] X. Li, and D. Roth, "Learning question classifiers: the role of semantic information". Natural Language Engineering, 12(03), 2006, 229-249.

[4] Z. Hui, J. Liu and L. Ouyang, "Question Classification Based on an Extended Class Sequential RuleModel". The 5th International Joint Conference on Natural Language Processing, IJCNLP2011. 2011, PP. 938-946.

[5] X. H. Phan, L. M. Nguyen, and S. Horiguchi, "Learning to classify short and sparse text & web with hidden topics from large-scale data collections". In Proceedings of the 17th international conference on World Wide Web, 2008, pp. 91-100. ACM.

[6] J. Wang, L. Li and F. Ren, "An Improved Method of Keywords Extraction Based on Short Technology Text". International Conference on Natural Language Processing and Knowledge Engineering (NLP-KE). 2010, pp. 1-6.

[7] T. A. Rahman, "Question Classification Using Statistical Approach: A COMPLETE REVIEW". Journal of Theoretical and Applied Information Technology, 71(3), 2015. http://www.jatit.org/volumes/Vol71No3/9Vol71No3.pdf.

[8] K. Abidi, Z. Elberrichi, and Y.G. Tlili, "Arabic Text Categorization: A Comparative Study of Different Representation Modes", Journal of Theoretical and Applied Information Technology, Vol. 38, No. 1, 2012, pp: 1–5.

[9] D. Zhang and W. S. Lee, "Question Classification using Support Vector Machines". Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '03. 2003, pp. 26-32. DOI: 10.1145/860435.860443.

[10] M. J. Van Hoeij, J. C. M. Haarhuls, R. F. Wierstra, and P. van Beukelen, "Developing a classification tool based on Bloom's taxonomy to assess the cognitive level of short essay questions". Journal of veterinary medical education, 31, 2004, pp. 261-267. DOI: 10.3138/jvme.31.3.261.

[11] W.-C. Chang, and M.-S. Chung, "Automatic applying Bloom's taxonomy to classify and analysis the cognition level of English question items". Paper presented at the Pervasive Computing (JCPC), 2009 Joint Conferences on. DOI: 10.1109/JCPC.2009.5420087.

[12] S. S. Haris and N. Omar," A Rule-Based Approach In Bloom's Taxonomy Question Classification Through Natural Language Processing". 3rd International Conference on Advancements in Computing Technology (ICACT 2012), 2012, pp. 410-414. ISBN: 978-1-4673-0894-6.

[13] N. Yusof, and C. J. Hui, "Determination of Bloom's cognitive level of question items using artificial neural network". In Intelligent Systems Design and Applications (ISDA), 2010 10th International Conference on (pp. 866-870). IEEE. DOI:10.1109/ISDA.2010.5687152.

[14] Y. Liu, X. Huang, A. An, and X. Yu, "A Sentiment-Aware Model for Predicting Sales Performance Using Blogs". Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 2007, pp. 607-614. DOI: 10.1145/1277741.1277845.

[15] C. C. Aggarwal and C. Zhai, "A survey of text classification algorithms". In Mining text data, 2012, pp. 163-222. Springer US. DOI: 10.1007/978-1-4614-3223-4_6.

[16] L. Galavotti, F. Sebastiani, and M. Simi, "Experiments on the use of feature selection and negative evidence in automated text categorization". In Proceedings of ECDL-00, 4 th European Conference on Research and Advanced Technology for Digital Libraries (Lisbon, Portugal, 2000), 59-68. DOI: 10.1007/3-540-45268-0_6.

[17] Y. Liu, "A comparative study on feature selection methods for drug discovery". Journal of chemical information and computer sciences, 44(5), 2004, 1823-1828. DOI: 10.1021/ci049875d.

[18] C. Cortes, and V. Vapnik, "Support-vector networks. Machine learning". Boston: Kluwer Academic Publishers. 20(3), 1995, pp. 273-297. DOI: 10.1023/A:1022627411411.

[19] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features". Proceedings of the European Conference on Machine Learning (ECML), Springer, 1998, pp. 137-142. ISBN: 3-540-64417-2.

[20] Y. Yang and X. Lin, "A re-examination of text categorization methods". A CM Conference on Research and Development in Information Retrieval, 1999, pp 42-49. DOI: 10.1145/312624.312647.

[21] R. Yadav, M. Mishra, and S. S. C. E. T. Bhilai, "Question Classification Using Naïve Bayes Machine Learning Approach". International Journal of Engineering and Innovative Technology (IJEIT), 2(8). ISSN: 2013, 2277-3754.

[22] H. Alshalabi, S. Tiun, N. Omar, and M. Albared, "Experiments on the Use of Feature Selection and Machine Learning Methods in Automatic Malay Text Categorization". Procedia Technology, 11, 2013, 748-754.