International Journal on Advanced Science Engineering Information Technology

Vol.8 (2018) No. 4-2
ISSN: 2088-5334

Question Classification Based on Bloom's Taxonomy Using Enhanced TF-IDF

Manal Mohammed[#], Nazlia Omar[#]

[#] *Faculty of Information Science and Technology, Center for Artificial Intelligence Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Malaysia*
E-mail: manal.altamimi@outlook.com, nazlia@ukm.edu.my

***Abstract*— Bloom's Taxonomy has been used widely in the educational environment to measure, evaluate and write high-quality exams. Therefore, many researchers have worked on the automation for classification of exam questions based on Bloom's Taxonomy. The aim of this study is to make an enhancement for one of the most popular statistical feature, which is TF-IDF, to improve the performance of exam question classification in accordance to Bloom's Taxonomy cognitive domain. Verbs play an important role in determining the level of a question in Bloom's Taxonomy. Thus, the improved method assigns the impact factor for the words by taking the advantage of the part-of-speech tagger. The higher impact factor assigns to the verbs, then to the noun and adjective, after that, the lower impact factor assigns to the other part-of-speech. The dataset that has been used in this study is consist of 600 questions, divided evenly into each Bloom level. The questions first pass into the preprocessing phase in which they are prepared to be suitable for applying the proposed enhanced feature. For classification purpose, three machine learning classifiers are used Support Vector Machine, Naïve Bayes, and K-Nearest Neighbour. The enhanced feature shows satisfactory result by outperforming the classical feature TF-IDF via all classifiers in terms of weighted recall, precision, and F1-measure. On the other hand, Support Vector Machine has superior performance over other classifiers Naïve Bayes, and K-Nearest Neighbour by achieving an average of 86%, 85%, and 81.6% weighted F1-measure respectively. However, these results are promising and encouraging for further investigations.**

***Keywords*— question classification; bloom's taxonomy; TF-IDF; support vector machine; naïve bayes; K-Nearest Neighbour.**

I. INTRODUCTION

The most traditional and classical way to evaluate students in educational institutes is by written examination. Therefore, many lecturers are trying to follow some framework such as Bloom's Taxonomy while preparing the exam questions to ensure the production of high-quality exams. The benefits of classifying questions regarding Bloom's Taxonomy Cognitive Domain (BTCD) are providing a suitable and appropriate way to measure students' intellectual abilities [1], and covering different thinking skills start from simplest to the most complex one. Therefore, the automatic classification of examination questions based on Bloom's taxonomy is highly required, especially in the educational environment [2], since the process of classifying exam questions manually is time-consuming. Furthermore, some academicians have no idea about Bloom's taxonomy [3], or have no ability to distinguish the difference between Bloom's taxonomy's levels which may lead to misclassification. Hence, this may lead to poor quality examinations [3][4]. Benjamin Bloom and his team introduced Bloom's taxonomy in 1956 which basically involves three domains. The domain that has developed for the purpose of assessing students' intellectual abilities and skills is known as Cognitive Domain [1]. Cognitive domain has a hierarchical structure which comprises six levels namely knowledge level, comprehension level, application level, analysis level, synthesis level and evaluation level. Knowledge level evaluates students' ability in memorizing facts and basic information e.g. *Label the parts of the microscope shown on the right*. Comprehension level measures students' ability in understanding ideas and topics based on previous knowledge e.g. *Describe in your own words what happens when a stream's velocity slows*. Application level evaluates students' skills in implementing acquired knowledge to new circumstances e.g. *Apply the storytelling technique here to a little story of your own*. Analysis level assesses students' ability in dividing information into pieces to classify them and find the relationship e.g. *Break down the main actions of the story*. Synthesis level evaluates students' ability to combine ideas together to create new solution e.g. *Create a set of guidelines to determine the points of a plant susceptible to localized corrosion*. Evaluation level measures students' ability to defend and judge issues based on some criteria e.g. *Assess the relative effectiveness of different* *graphical representations of the same data or biological concept*. Knowledge and comprehension level need lower thinking skills, whereas the rest of levels require higher thinking skills. Figure 1. demonstrates some verbs and keywords used in each level in BTCD [5].

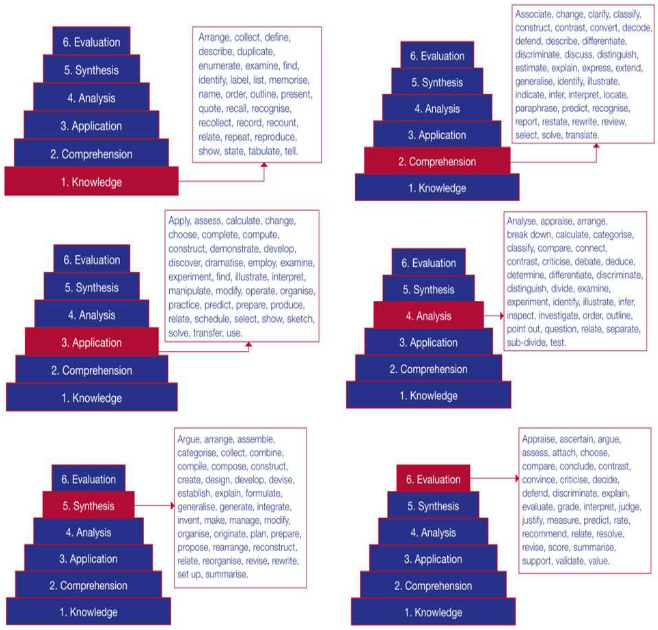

Fig. 1 Sample of verbs used in each level in BTCD [5]

II. MATERIAL AND METHOD

Bloom's taxonomy has become a trend in many aspects nowadays. Many researchers focused on building different applications based on Bloom's taxonomy like automating classification exams questions [6] [7] [8], or proposed assessment framework [9]. In addition, it is used in developing a paper generator system as in [10][11].

Researchers [12] [3] [6] [7] [8] [13] have put a lot of efforts and used different techniques to classify questions into Bloom's taxonomy cognitive domain. Some researchers used the rule-based technique while other used machine learning techniques, and even some of them used evolutionary algorithms. For example, the work of [7] proposed an online test system that classifies test question into Bloom's cognitive level. The system relied on the dataset which contains Bloom's verbs, test question fed into the system as an input, then the question is chunk into individual words, the extracted verb from question checked against the database to determine the level of this question, weight is applied in case of keyword overlap with other levels. The weight values are taken from several research papers related to Bloom's Taxonomy. The result shows that this method only efficient in classifying questions which belong to knowledge level. This approach depends on the database, which means it cannot handle new words that not stored in the database. In addition, the way of assigning a weight for the verb is inefficient.

Omar et al [3] proposed a rule-based approach, which applies NLP techniques, in order to extract important verbs and keywords and then assign weights to them to recognize the category of each question. In case of keyword overlapping the weight assign from experts' perspective, and the level of the question determined by selecting the higher weight. This technique is inefficient due to the requirement of writing several rules. In addition, the way of handling keywords weighting is exhausted, since many verbs need to be checked by experts, and inconsistent due to the various background knowledge of each expert.

The work of [8] categorize question into suitable Bloom's taxonomy level with a combination of rules and statistical approach (N-gram). The statistical technique used here in order to get over the weakness of rule, the result of average F1 is 86%. Similarly, the study of [4] classified question based into Bloom's taxonomy by rules. In this research, the rules generated automatically, WordNet used to get verbs from questions whereas cosine similarity used to identify the pattern of questions. As a result, 71% of classification matched experts annotation. Nevertheless, some studies show the success of rule-based approaches but still considered as exhausted and expensive since is not dynamic. There is need to write many rules manually to cover all types and domains of questions in order to increase the accuracy of the result. Thus, it is inefficient.

On the other hand, machine learning techniques used widely in classification problems. The work of [14] used TF-IDF as feature extraction along with three classifiers KNN, NB and SVM to classify questions into Bloom's taxonomy. The dataset contains 100 questions for each level, divided into 420 questions as training set and 180 as the testing set. The experimental results show that SVM outperforms KNN and NB in term of accuracy and F1. The study proposed by [15], worked with the dataset that consists of 272 questions, 70% training set, and 30% test set which collected from several websites on Bloom's taxonomy. After pre-processed the questions, TF-IDF was calculated. The classification process handled by SVM is using linear kernel. This method shows satisfactory result regarding accuracy and precision, whereas it gave poor result regarding F-measure and recall. Using pure statistical approach TF-IDF requires a huge amount of data in order to return a reasonable result or some improvement in a way of calculating TF-IDF to enhance the outcome. Although [12] the size of dataset increased to 600 questions and used the same procedures in [15] except removing stop words, which cause to improve the result. Yet still, there is a need for extracting more other features.

Moreover, some researchers applied evolutionary algorithms, which usually used to solve optimization problems, to classify questions into Bloom's taxonomy. The researcher in [16] study applied classification and evaluation by particle swarm intelligence after pre-processing 1000 questions for each of Bloom's cognitive level from different courses in computer science major. The TF-IDF function was used in order to represent the question in form of term weights vector. However, this approach does not return good enough accuracy. Therefore, [17] enhance this method, by applying different feature selection techniques with particle swarm intelligence based on Rocchio algorithm.

In conclusion, a comprehensive review is performed by [18] in question classification. The review state that SVM classifier often outperforms other classifiers since it works well with unstructured text data.

The implementation of this study summarized in Fig. 2, which involves several steps to automatically classify question according to Bloom's taxonomy.

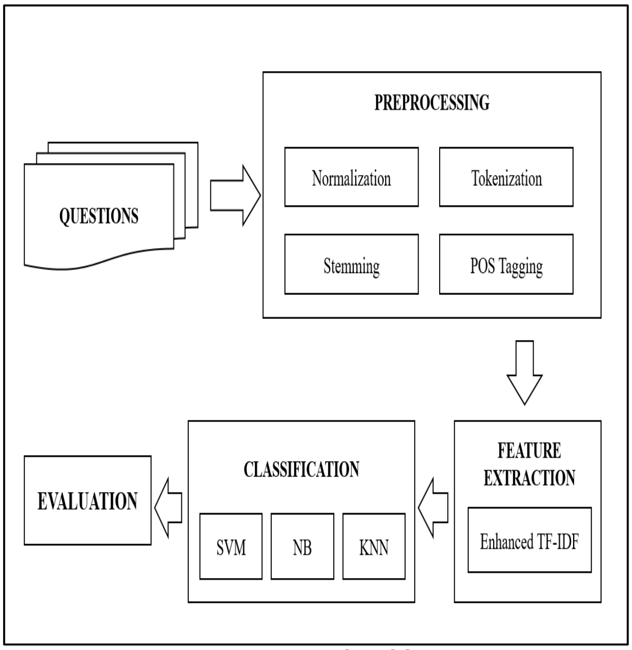

Fig. 2 Proposed Model

A. *Pre-processing*

Where the questions pass into several steps; normalization, tokenization, part-of-speech tagging, and stemming in order to refine and prepare them for next phase. In normalization process, the unwanted data will be eliminated such as punctuation marks, numbers, and stop-words. In addition, all words will be converted to lowercase. After that, the question will split into tokens. Finally, tagging process and stemming will occur.

B. *Feature Extraction*

The enhanced statistical feature, known as E-TFIDF, is an improved version of traditional TF-IDF feature. The proposed feature is inspired from this work [19]. TF-IDF is a very common weighting method used in information retrieval and text mining [20] which score the importance of the word in a document [21]. The higher TF-IDF value for the word shows the stronger relatedness to the document that appeared in. However, TF-IDF does not handle other information as an effect of word distribution among different classes [19]. Therefore, the impact factor is introduced. TF-IDF computed by multiplying the Term Frequency (TF) with the Inverse Document Frequency (IDF) which represented by the following equations:

$$TF(t,d) = \frac{c(t_d)}{T_d} \qquad (1)$$

where $c(t_d)$ indicates the number of term t appears in document d, and Td indicates the total number of terms in document d. The IDF of the term t calculated as:

$$IDF(t) = 1 + \log(\frac{D}{d_t}) \qquad (2)$$

where D is the total number of documents in the corpus, and dt is the number of the documents a term t appeared in. Finally, the TF-IDF represented by this equation:

$$TF-IDF(t,d) = TF(t,d).IDF(t) \qquad (3)$$

The enhanced TF-IDF will be calculated by multiplying TF-IDF with the impact factor, in which the impact factor is assigned to the word regarding its part-of-speech. Usually, verbs play an important role in order to determine the level of question. Thus, the impact factor of verbs will be higher than other words. Whereas noun and adjective are more important than other part-of-speech. Therefore, the impact factor for the term *t* calculated as follows:

$$IF(t) = \begin{cases} X(t) + 3, & if\ t\ is\ VB \\ X(t) + 1, & if\ t\ is\ NN\ or\ ADJ \\ 1 & ,\ otherwise \end{cases} \qquad (4)$$

where formula of X(t) is:

$$X(t) = \sqrt{\frac{1}{C} \sum_{i=1}^{C} (eq(t,c_i) - \frac{1}{C})^2} \qquad (5)$$

in which *C* denotes the total number of classes. $eq(t,c_i)$ refers to the number of documents that exist in class c_i and has term *t*, divided over the total number of whole documents. The enhanced TF-IDF can be represented by the following equation:

$$E-TFIDF(t,d) = TF-IDF(t,d).IF(t) \qquad (6)$$

C. *Classification*

In classification process, three of the most common machine learning classification algorithms are used:

- Support Vector Machine (SVM)
- Naïve Bayes (NB)
- K-Nearest Neighbour (KNN).

The first classifier used in this model is Support Vector Machine (SVM). The aim of SVM is to find the favourable hyperplane that separates two sets of data from each other, by maximizing the separation margin among the hyperplane and the set of data points closest to it. This study uses SVM with linear kernel since it is popular in text classification [22]. Fig. 3 demonstrate SVM method.

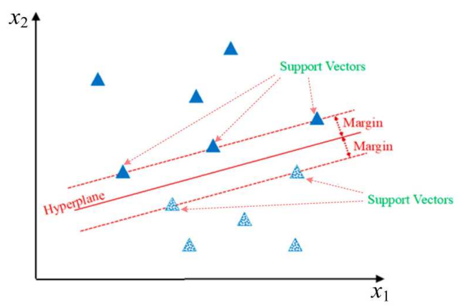

Fig. 3 Representation of SVM method [23]

The second classifier is Naïve Bayes (NB), which performs well in text classification [24]. NB is a probabilistic classifier that considers the existence of the specific feature in class is independent of any other existence feature [14]. In this study, the multinomial model of NB is used. According to [25], multinomial NB estimates the contingent likelihood of a specific term *t* given a class as the relative recurrence of term t in documents belonging to class *c* as:

$$P(t|c) = \frac{T_{ct}}{\sum_{t \in V} T_{ct}} \quad (7)$$

Consequently, this variety considers the number of frequencies of term *t* in training document from class *c*. Fig 4 shows the representation of multinomial NB algorithm.

```
TRAINMULTINOMIALNB(C, D)
  V ← EXTRACTVOCABULARY(D)
  N ← COUNTDOCS(D)
  for each c ∈ C
  do N_c ← COUNTDOCSINCLASS(D, c)
     prior[c] ← N_c / N
     text_c ← CONCATENATETEXTOFALLDOCSINCLASS(D, c)
     for each t ∈ V
     do T_ct ← COUNTTOKENSOFTERM(text_c, t)
     for each t ∈ V
     do condprob[t][c] ← (T_ct + 1) / Σ_t'(T_ct' + 1)
  return V, prior, condprob

APPLYMULTINOMIALNB(C, V, prior, condprob, d)
  W ← EXTRACTTOKENSFROMDOC(V, d)
  for each c ∈ C
  do score[c] ← log prior[c]
     for each t ∈ W
     do score[c] += log condprob[t][c]
  return arg max_{c∈C} score[c]
```

Fig. 4 Representation of NB (Multinomial model) algorithm [25]

The third classifier used is K-Nearest Neighbour, which is a type of lazy learning algorithm. The way that KNN classify object is by assigning the label to the test point that near to K nearest neighbour from training samples [26] as shown in Fig.5

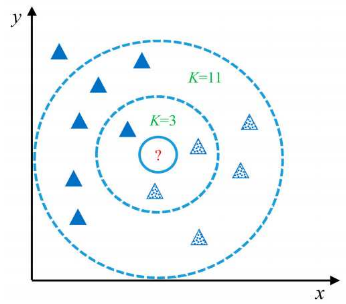

Fig. 5 Representation of KNN method [23]

D. *Evaluation*

The effectiveness of the proposed method is measured by calculating the following metrics: recall, precision, and f-measure, which are the most common metrics measurement in information retrieval. In order to define these metrics some terms must be introduced; True Positive (TP), False Positive (FP) and False Negative (FN). The term True Positive (TP) refers to the number of questions a classifier correctly classified to the appropriate BTCD. Whereas False Positive (FP) refers to the number of questions incorrectly classified to BTCD. Lastly, the term False Negative (FN) refers to the number of questions were not being classified.

Recall metric measures the perfection of classifiers, by computing this formula

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

Precision metric measures the fineness of classifiers, by calculating this equation

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

Finally, F1-measure that combine recall and precision, calculated by

$$F1 - measure = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (10)$$

III. RESULTS AND DISCUSSION

The experiments have been performed using the dataset that was introduced by [12], which consists of 600 labelled and open-ended questions, 100 question per each class. This dataset does not belong to a specific domain since the questions were collected from several resources. The performance of the proposed model evaluates via 5-fold cross-validation. The experiments were conducted to assess the improved TF-IDF versus the classical TF-IDF via three classifiers SVM, NB, and KNN. All the classifiers used with the default settings and trained using Scikit-learn library in Python. Moreover, for stemming the NLTK porter stemmer is used, and the part-of-speech tagging is performed with Stanford tagger (version 3.9.1).

A. *Evaluation of Each Individual Classifier*

This section discusses the results of classifying question into Bloom's Taxonomy regarding each classifier; SVM, NB, and KNN. The results display the performance of traditional TF-IDF and the enhanced E-TFIDF. The effect of impact factor in improving the quality of vector representation of questions can be observed. Since the basic idea in enhanced feature E-TFIDF is to focus on giving verbs higher value for impact factor compared to other words tags, which means give the verbs more attention in determining the class of the question.

1) *NB*: The experiment details for the classification by NB classifier with the classical feature TF-IDF is demonstrated in TABLE I. The result for the classification via NB using the enhanced feature E-TFIDF is showed in TABLE II.

TABLE I
RESULT OF NB WITH CLASSICAL FEATURE TF-IDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.900	0.868	0.884
Comprehension	0.861	0.897	0.876
Application	0.741	0.873	0.798
Analysis	0.950	0.725	0.821
Synthesis	0.770	0.889	0.825
Evaluation	0.809	0.857	0.830
AVG	**0.839**	**0.851**	**0.839**

TABLE II
RESULT OF NB WITH ENHANCED FEATURE E-TFIDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.969	0.874	0.919
Comprehension	0.861	0.877	0.867
Application	0.821	0.812	0.813
Analysis	0.931	0.849	0.884
Synthesis	0.789	0.806	0.795
Evaluation	0.741	0.939	0.823
AVG	**0.852**	**0.859**	**0.850**

As it can be observed from the result in tables that the enhanced feature outperform the classical feature. Fig 6 shows the comparison between TF-IDF and E-TFIDF using NB classifier.

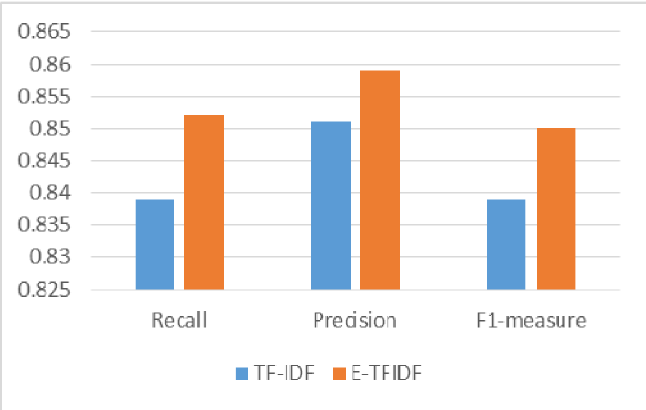

Fig. 6 Evaluation of TF-IDF vs. E-TFIDF via NB Classifier

2) *KNN*: The experiment details for the classification using KNN classifier with the classical feature TF-IDF is demonstrated in TABLE IIII. The result for KNN classification using the enhanced feature E-TFIDF is showed in TABLE IV. Moreover, Fig 7 shows the comparison between the two features with KNN classifier.

TABLE III
RESULT OF K-NN WITH CLASSICAL FEATURE TF-IDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.980	0.750	0.843
Comprehension	0.810	0.786	0.797
Application	0.660	0.790	0.711
Analysis	0.870	0.710	0.780
Synthesis	0.701	0.823	0.755
Evaluation	0.620	0.913	0.730
AVG	**0.774**	**0.795**	**0.769**

TABLE IV
RESULT OF K-NN WITH ENHANCED FEATURE E-TFIDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.948	0.796	0.863
Comprehension	0.829	0.852	0.838
Application	0.812	0.760	0.784
Analysis	0.900	0.832	0.862
Synthesis	0.761	0.802	0.779
Evaluation	0.661	0.919	0.769
AVG	**0.818**	**0.827**	**0.816**

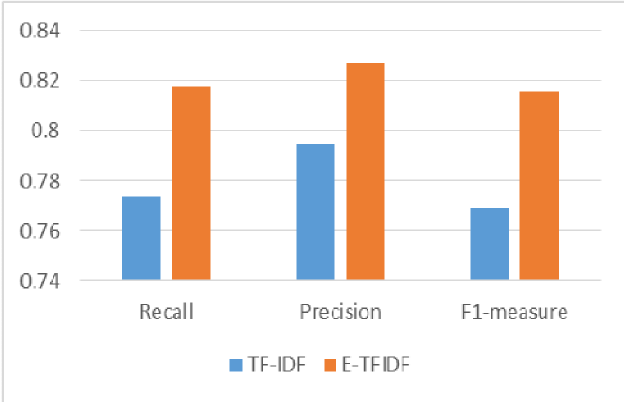

Fig. 7 Evaluation of TF-IDF vs. E-TFIDF via KNN Classifier

Obviously, the E-TFIDF achieves a good result and outperforms the performance of TF-IDF using KNN in terms of precision, recall and F-measure. In addition, it can be noticed that the enhanced feature improves the classification process strongly in KNN classifier by increasing the performance in term of F1-measure with 4.7% as shown in Fig 7.

3) *SVM*: The experiment details for the classification by SVM classifier with the classical feature TF-IDF is demonstrated in TABLE V. The result for the classification using SVM with the enhanced feature E-TFIDF is shown in TABLE VI. In addition, the result of the comparison between the two features with SVM is shown in Fig 8.

TABLE V
RESULT OF SVM WITH CLASSICAL FEATURE TF-IDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.979	0.862	0.916
Comprehension	0.830	0.955	0.885
Application	0.798	0.820	0.806
Analysis	0.919	0.815	0.862
Synthesis	0.751	0.894	0.810
Evaluation	0.811	0.830	0.811
AVG	**0.848**	**0.863**	**0.848**

TABLE VI
RESULT OF SVM WITH ENHANCED FEATURE E-TFIDF

Cognitive Level	Recall	Precision	F1-measure
Knowledge	0.969	0.900	0.933
Comprehension	0.851	0.894	0.868
Application	0.830	0.770	0.797
Analysis	0.910	0.927	0.918
Synthesis	0.818	0.843	0.824
Evaluation	0.790	0.864	0.821
AVG	**0.861**	**0.866**	**0.860**

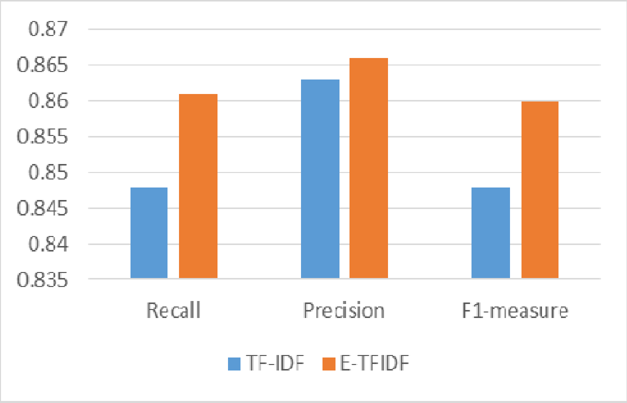

Fig. 8 Evaluation of TF-IDF vs. E-TFIDF via SVM Classifier

It is clearly, the E-TFIDF achieves a better result compared to TF-IDF using SVM in terms of precision, recall and F-measure.

B. *Evaluation Among Classifiers*

This section summarizes the results of TF-IDF and the enhanced feature E-TFIDF among all three classifiers SVM, NB, and KNN. This is due the value of the impact factor, which give a higher value for the relevant word in the document. Table VIIII represents an overall comparison between all classifiers. Clearly either by using TF-IDF or E-TFIDF the linear SVM has superior performance achievement among all other classifiers. Whereas NB comes at the second rank, then lastly KNN.

TABLE VII
OVERALL F1-MEASURE PERFORMANCE OF ALL CLASSIFIERS

	TF-IDF	E-TFIDF
NB	0.839	0.85
KNN	0.769	0.816
SVM	0.848	**0.86**

IV. CONCLUSION

The aim of the present study is to examine if the enhanced statistical feature can produce a reasonable result in classifying question in accordance to Bloom's Taxonomy. However, as shown in the previous section the outcome of the enhanced feature E-TFIDF produces satisfactory and promising outcome comparing to a traditional method.

These findings provide the following insights for future research: making more experiments with different sizes of corpus; using enhanced E-TFIDF along with other features; extraction of features that handle the order of words, such as N-gram; use of ensemble technique to improve the classification process.

REFERENCES

[1] B. S. Bloom, Taxonomy of Educational Objectives – Handbook 1 Cognitive Domain. London: Longman, 1956.
[2] A. Osman and A. A. Yahya, "Classification of exam questions using linguistically-motivated features: A case study based on bloom's taxonomy," in The Sixth International Arab Conference on Quality Assurance in Higher Education (IACQA'2016), 2016, vol. 467–474.
[3] N. Omar et al., "Automated Analysis of Exam Questions According to Bloom's Taxonomy," Procedia - Soc. Behav. Sci., vol. 59, no. 1956, pp. 297–303, 2012.
[4] K. Jayakodi, M. Bandara, and D. Meedeniya, "An Automatic Classifier for Exam Questions with WordNet and Cosine Similarity," in Moratuwa Engineering Research Conference (MERCon), 2016. IEEE, 2016, pp. 12–17.
[5] D. Kennedy, Writing and using learning outcomes: a practical guide. Cork, 2006.
[6] D. A. Abduljabbar and N. Omar, "Exam Questions Classification Based on Bloom ' S Taxonomy Cognitive Level Using Classifiers Combination," J. Theor. Appl. Inf. Technol., vol. 78, no. 3, pp. 447–455, 2015.
[7] W. C. Chang and M. S. Chung, "Automatic applying Bloom's taxonomy to classify and analysis the cognition level of english question items," 2009 Jt. Conf. Pervasive Comput. JCPC 2009, pp. 727–733, 2009.
[8] S. S. Haris and N. Omar, "Bloom's taxonomy question categorization using rules and N-gram approach," J. Theor. Appl. Inf. Technol., vol. 76, no. 3, pp. 401–407, 2015.
[9] A. Lajis, H. M. Nasir, and N. A. Aziz, "Proposed Assessment Framework Based on Bloom Taxonomy Cognitive Competency: Introduction to Programming," ICSCA '18, Kuantan, Pahang, Malaysia, 2018, pp. 1–5, 2018.
[10] A. Shaikh, A. Handore, and P. P. Kudal, "Quality Question Paper Generator on Role Based Hierarchy model," Int. J. Res. Advent Technol., vol. 5, no. 2, pp. 6–9, 2017.
[11] Rasim, A. Z.R. Langi, Y. Rosmansyah, and Munir, "Generation Quiz with Genetic Algorithm Based on Bloom ' s Taxonomy Classification in Serious Game Based Virtual Environments," 2016 Int. Conf. ICT Smart Soc. Surabaya, 20-21 July 2016 ISBN 978-1-5090-1620-4, no. July, pp. 20–21, 2016.
[12] A. A. Yahya, Z. Toukal, and A. Osman, "Bloom ' s Taxonomy – Based Classification for Item Bank Questions Using Support Vector Machines Bloom ' s Taxonomy – Based Classification for Item Bank," in Modern Advances in Intelligent Systems and Tools, SCI 431, no. April 2012, Springer-Verlag Berlin Heidelberg 2012, 2012, pp. 135–140.
[13] S. S. Haris and N. Omar, "A rule-based approach in Bloom's Taxonomy question classification through natural language processing," in International Conference on Computing and Convergence Technology (ICCCT),Seoul, 2012, pp. 410–414.
[14] A. A. Yahya, A. Osman, A. Taleb, and A. A. Alattab, "Analyzing the Cognitive Level of Classroom Questions Using Machine Learning Techniques," Procedia - Soc. Behav. Sci., vol. 97, pp. 587–595, 2013.
[15] A. A. Yahya and A. Osama, "Automatic Classification of Questions into Bloom's Cognitive Levels Using Support Vector Machines," in The International Arab Conference on Information Technology, Naif Arab University for Security Science (NAUSS), 2011, pp. 1–6.
[16] A. A. Yahya and A. Osman, "Classification of high dimensional Educational Data using Particle Swarm Classification," 2014 11th IEEE/ACS Int. Conf. Comput. Syst. Appl. AICCSA 2014, pp. 34–41, 2015.
[17] A. A. Yahya, "Swarm intelligence-based Approach for Educational Data Classification," J. King Saud Univ. - Comput. Inf. Sci., no. August, 2017.

[18] A. Sangodiah, M. Muniandy, and L. E. Heng, “Question classification using statistical approach: A complete review,” J. Theor. Appl. Inf. Technol., vol. 71, no. 3, pp. 386–395, 2015.

[19] L. Zhu, G. Wang, and X. Zou, “A Study of Chinese Document Representation and Classification with Word2vec,” in Computational Intelligence and Design (ISCID), 2016 9th International Symposium, 2016, pp. 298–302.

[20] K. Chen, Z. Zhang, J. Long, and H. Zhang, “Turning from TF-IDF to TF-IGM for term weighting in text classification,” Expert Syst. Appl., vol. 66, pp. 1339–1351, 2016.

[21] M. M. Altawaier and S. Tiun, “Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis,” Int. J. Adv. Sci. Eng. Inf. Technol., vol. 6, no. 6, p. 1067, 2016.

[22] M. Pota, M. Esposito, and G. De Pietro, A forward-selection algorithm for SVM-based question classification in cognitive systems, vol. 6. Cham: Springer, 2016.

[23] F. Wang, Z. Zhen, B. Wang, and Z. Mi, “Comparative Study on KNN and SVM Based Weather Classification Models for Day Ahead Short Term Solar PV Power Forecasting,” Appl. Sci., vol. 8, no. 1, p. 28, 2017.

[24] S. N. Das, M. Mathew, and P. K. Vijayaraghavan, “An Approach for Optimal Feature Subset Selection using a New Term Weighting Scheme and Mutual Information,” Int. J. Adv. Sci. Eng. Inf. Technol., 2011.

[25] C. D. Manning, P. Raghavan, and H. Schütze, “Text classification and Naive Bayes,” Introd. to Inf. Retr., no. c, p. 260, 2008.

[26] 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 Technol., vol. 11, no. Iceei, pp. 748–754, 2013.