Bloom's Taxonomy and Rules Based Question Analysis Approach for Measuring the Quality of Examination Papers

Banage T. G. S. Kumara, A. Brahmana, and Incheon Paik

***Abstract*—Examinations play a major role in the teaching and learning. Questions are used to obtain information and assess knowledge and competence of students. Academics who are involved in teaching process in higher education mostly use final examination papers to assess the retention capability and application skills of students. Question writing process is a very challenging step for academics. Questions that used to evaluate different cognitive levels of students may be categorized as higher order questions, intermediate order questions and lower order questions. This research work tries to derive a suitable methodology to categorize final examination question papers based on Bloom's Taxonomy. The analysis was performed on computer science related papers. The study was conducted to check whether examination questions comply with the requirements of Bloom's Taxonomy at various cognitive levels. Natural language processing techniques were used to identify the significant keywords and verbs which are useful in the determination of the suitable cognitive level. A rule based approach was used to determine the level of the question paper in the light of Bloom's Taxonomy. The derived model introduced a quantitative approach to categorize undergraduate examination papers.**

***Index Terms*—Bloom's taxonomy, natural language processing, data mining.**

I. Introduction

Examinations play a significant role in teaching and learning process. Written examination is at effective way to evaluate the knowledge of students. Preparing a balanced question paper is a challenging task for lecturers. The situation is getting more challenging when lecturers try to produce good quality and fair questions to assess different level of cognitive.

There is not exact method to decide, which the level the examination paper belongs. Good and fair test must be in the middle level. All the questions should be analyzed to get the overall level of the question paper. It requires more time and effort to evaluate manually. Although if it is performed manually, the exact level of the question paper cannot be expected during a short time period. Collection of high level questions or collection of intermediate level questions or collection of low level questions does not represent a balanced examination paper. There must be an acceptable number of questions that represent all three levels in a balanced question paper. There are some standards to setting up examination question papers. It may depend on the institutional aspects or subject area that is going to test through the examination. But there is not an automated method to determine the exact level of the question paper. Without a balanced question paper cannot assess the learning outcome of learners with reference to the objectives.

To assess the exact knowledge level of students, the lecturers must have enough competence of setting up a balanced question paper. If there is a suitable methodology to determine the level of the paper academics can pay their full attention to the paper preparation process. A taxonomy has decided to determine the cognitive levels of knowledge based on the keywords included in the questions. But, it has not developed to determine the overall level of the question paper. Lecturer cannot decide whether that paper is good and fair test without determination the level of the test. They can do this process manually. But it is time consuming. Not only that, it is tedious and prone to mistakes. To overcome this difficulties in exam question paper setting, there must be an effective, automated methodology which follows international standards. The main purpose of this study is to derive a suitable methodology based on international standards to overcome those difficulties in question paper setting process to assess the knowledge of target student group.

The rest of this paper is organized as follows. In Section II, we discuss related work. Section III presents the proposed approach. Section IV describes our implementation and evaluation. Finally, Section V concludes the paper.

II. Related Works

There must be an effective examination methodology to assess the exact knowledge level of selected student group. Written examinations are the most common type among various examination types. Research work [1] had mentioned the importance of balanced question paper preparation to assess the learner's knowledge level. This study was conducted to evaluate the standards of classroom assessments in Virginia. According to the study, they have concluded that to increase the achievements and higher order thinking skills of students.

Research work [2] has conducted a study to classify questions using different approaches. The study described machine learning approach, statistical approach, structural approach and Rule based approach. This classification was

Manuscript received January 20, 2019; revised June 10, 2019.

Banage T. G. S. Kumara and A. Brahmana are with Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka, Sri Lanka (e-mail: btgsk2000@gmail.com, akilabrahmana@gmail.com).

Incheon Paik is with School of Computer Science and Engineering University of Aizu, Aizu-Wakamatsu, Fukushima, Japan (e-mail: paikic@u-aizu.ac.jp).

doi: 10.18178/ijke.2019.5.1.111

done according to a set of predetermined categories. Bloom's Taxonomy [3] has been used as a standard classification in education field. This taxonomy proved the supportive direction on education practice and teaching. Research work [4] conducted a study on Software Engineering assessment using Bloom's Taxonomy. They selected a sample of trial questions and specified and categorized them according to the appropriate Bloom's Taxonomy level. Researchers in [5] had done a study on analysis of examination papers in engineering courses to improve higher order thinking skills by getting the use of Bloom's Taxonomy. Engineering students must be capable of critical thinking skills and creative thinking skills to keep pace with rapidly changing engineering sector. The study selected end semester examination papers as the data set. The questions had categorized according to the levels of Bloom's Taxonomy. After the classification process, analysis was conducted to identify question categories that improve higher order thinking skills.

There are various perspectives of academicians on categorization of examination question papers by getting the use of data mining. Research work [6] mentioned the importance of text mining in the extraction of factual information from a large data set. As well as they had mentioned the use of data mining, which is very useful to implement and solve different types of research problems. Not only that, the authors had mentioned that, text mining process is related to the data mining process which is used to derive information from unstructured or semi structured data sets such as emails, HTML files and full text documents. In this research work they had gone through the text mining preprocessing techniques.

Mishra *et al.* [7] had conducted a research work on question classification. This research work had discussed about the importance of question classification for question answering. Though they used support vector machine (SVM) concept. Research work [8] had conducted a research work on question classification. This work can introduce as another work related to the examination question classification using text classification techniques. This study was implemented on specific type of questions. The questions named as open ended questions. Questions can be separated into classes that enables to identify format and content of the expected response in this type. Open ended questions were successfully classified by using SVM concept.

III. Proposed Approach

Following Fig. 1. shows the steps of the proposed approach. Question preprocessing step is involved in extraction of keywords. Separation of words (Tokenization), white space removal, and removal of punctuation marks, stop word removal and elimination of non-letter characters were performed during the question preprocessing step [6].

The main rules were defined according to [9]. Rule for lower order questions was developed based on the keywords of knowledge and comprehension cognitive levels. Rule for intermediate order questions was developed based on keywords of analysis and application levels. Keywords of other two levels were used to define rule for higher order questions. These tasks performed during the step of applying rules. Question wise categorization is done according to the developed rules. Determination of the level of the question paper was done as the next step. According to the percentage of low level, intermediate and high level questions it was determined whether the examination paper was set in a balanced or imbalanced manner. This determination was performed according to [7].

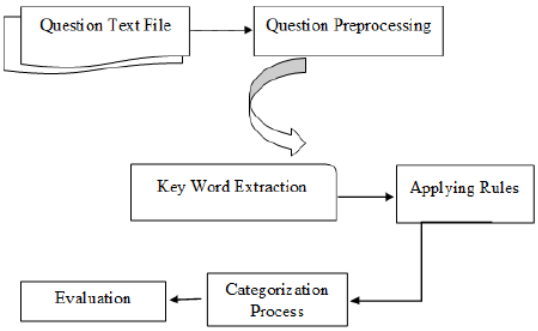

Fig. 1. Proposed model for question classification.

A. Text Preprocessing

Incomplete, noisy and inconsistent data must be corrected in preprocessing stage in order to obtain quality mining results. Text preprocessing can be divided in to two sub processes. Text cleaning and tokenization are those processes.

Examination papers may include various types of figures, images and formulas etc. Text questions were considered in this research. Therefore, text cleaning step was involved in removing tables, figures and other unwanted contents from text files. It was done manually. Tokenization is a step in text preprocessing which enables to split longer strings of texts into smaller pieces or tokens. Exploration of the words in a sentence is the aim of tokenization. Larger chunk of text, as paragraphs can be tokenized into sentences, sentences can be tokenized into words. After a piece of text has been tokenized appropriately, can perform further processing generally. Normalization was involved in removing punctuations from titles and contents, removing default stop words from title and contents, removing numbers, strip white spaces and removing non-letter characters.

B. Keyword Extraction

Benjamin Bloom developed a classification system of educational objectives based on the level of student understanding in 1950 [3]. This taxonomy contains six levels. This approach has been used by researchers, teachers, educationists, curriculum planners and examiners. The first three levels named Knowledge, Comprehension and Application known as LOTS (Lower Order Thinking Skills). Other three levels named Analysis, Synthesis and Evaluation known as HOTS (Higher Order Thinking Skills). To maintain the correct balance between questions in examinations, the academics must aware on standard classifications as Bloom's Taxonomy. The keywords from question paper were extracted to identify the level of a particular question.

Knowledge Level: Knowledge level defines as the lower level

or beginning level of the hierarchy. It also known as recalling of data. Some studies refer it as 'memorization'.

Computer Science related papers were selected in this work. Therefore, questions have involved in defining or describing computing terms, methodology and process. There are number of keywords associated with knowledge level. Define, List, State are some of them.

Examples:

a. Name the milestones in this project?
b. What are the three principle components of a data processing system?

Comprehension Level: Questions in this category must be capable of providing examples to illustrate a concept or an algorithm. Explain, Describe, Discuss, Convert are some main keywords in this comprehension level.

Examples:

a. Explain what you mean by 'Crashing' and 'Fast Tracking' tasks.
b. Describe SQL injection.

Application Level: This level is defined by applying the learnt concept to a certain scenario. Questions for Computer Science domain in this level have unique criteria. Questions in this type have involved in understand the concept and apply it to a new algorithm. Questions that include keywords like execute, solve, implement, use and operate fall in this level.

Examples:

a. How encapsulation helps in identifying the correct boundary of a class?
b. Modify the given 'while' loop into 'for' loop.

Analysis Level: This level requires to divide the information into smaller parts and analyze each of them. Compare, contrast, differentiate, examine are the keywords associated with the analysis level in Bloom's Taxonomy.

Example:

a. Compare and contrast the Kantianism and social contract theory by giving proper examples.
b. Differentiate logical modularity and physical modularity.

Synthesis Level: Synthesis level requires the ability of integrate and combine ideas or concepts by rearranging elements into a new whole. Questions associated with arrange, assemble, design, develop, and investigate fall into this level.

Examples:

a. Develop a network diagram using precedence diagramming method for the project.
b. Construct simple use case diagram to display the correct relationship.

Evaluation Level: This level is involved in judging, criticism, supporting or defending own stand. Questions in this level require the ability of students to make judgments about the value or merits of an idea, purpose, procedure method or product. Questions associated with verify, argue, appraise, judge, justify, evaluate and assess fall into evaluation level.

Examples:

a. 'Main memory usage is inefficient in fixed size partitioning.' Justify the given statement.
b. 'Attempting to break into a computing system without authorization should be illegal'. Justify given statement.

C. Rules Development

The algorithms developed according to the rules based on Bloom's Taxonomy [3]. This study developed new six rules based on the keywords in the six levels of Bloom's Taxonomy. The developed grammatical rules facilitate and improve the results of classification algorithm. Grammatical structure of the question analyzed to develop the rules based on Bloom's Taxonomy. The rules will distinguish the suitable keywords for each question. The syntactic patterns were utilized to identify rules.

A question may include some meaningful phrases (VB: Verb Phrase (base form), DT: Determiner (Stop words in English language), NN: Noun Phrase (singular or mass), JJ: Adjective Phrase, ADV: Adverb Phrase, PP: Prepositional Phrase). A question can divide into separate phrase to develop the rules. The rules were developed according to the phrases included in the question.

- **Rule 1**: Rule for Knowledge Level
  {(VB) (Knowledge Keyword) + [DT + NN + NN] + (PP)?}
- **Rule 2**: Rule for Comprehension Level
  {(VB) (Comprehension Keyword) + (DT) + [NP + PP] ?}
- **Rule 3**: Rule for Application Level
  {(VB) (Application Keyword) + (DT)? + (NP) + [DT + NP]}
- **Rule 4**: Rule for Analysis Level
  {(VB) (Analysis Keyword) + (NP) + [DT +NP]}
- **Rule 5**: Rule for Synthesis Level
  {(VB) (Synthesis Keyword) + (NP) + [(DT +VB) | (DT + VB)] + (NP) + (DT + NP)}
- **Rule 6**: Rule for Evaluation Level
  {(VB) (Evaluation Keyword) + (NP) + (DT) + (NP) + [(PP + NNP) | (DT + VB)] + (NP)?}

After the development of six rules, those rules were combined into basic three rules according to the study of [9]. They had presented three categories of questions as Higher Order Questions (HOQ), Intermediate Order questions (IOQ) and Lower Order Questions (LOQ). They have allocated six cognitive levels of Bloom's Taxonomy into these categories.

Higher Order Questions: Knowledge and Comprehension levels.

Intermediate Order questions: Application and Analysis levels

Lower Order Questions: Synthesis and Evaluation levels

Rule 1 and Rule 2 were used as one rule for LOQ. Rule 3 and Rule 4 were used as one rule for IOQ .Rule 5 and Rule 6 were used as one rule for HOQ.

D. Determine the State of the Question Paper

In this step, two algorithms to check the state (Balanced or Unbalanced) of the paper were defined. First, individual question categorization algorithm was defined to check the individual questions. Then, Paper State Check Algorithm was

defined to identify the final state. Following Fig. 2 and Fig. 3 shows the pseudo codes of the algorithms.

```
Begin {
Input: Document.txt   // Tokenized Text File without Stop words
readFile()
keyWordSelection()   {        // Select the exact Key word from the Document
        applyRulesToSelectTheLevel()
                        }
Print: Appropriate Level of the Question
}
End
```

Fig. 2. Individual question categorization algorithm.

Research paper [9] introduced percentages as shown in Table I that must be included in a balanced paper. In this work also, their percentage values were used. So, paper state check algorithm was developed according to the [9].

TABLE I: BALANCING THE EXAMINATION PAPER

Question Category	LOQ	IOQ	HOQ
Percentage distribution recommended	20-30%	40-50%	30-40%

```
Start {
Input: Document.txt   // Tokenized Text File without Stop words
readFile()
        for each main QuestionsInText :
        keyWordSelection()   {        // Select All Key word from the Document
                storeKeyworsInArray()
                getAllKeywordCount()
                applyRulesToSelectTheLevel() {
                        getLowLevelKeywordCount()
                        getIntermediateLevelKeywordCount()
                        getHighLevelKeywordCount()
        }

calculatePercentagesForEveryLevels()
applyBalancingRules() {       // Developed Rules are applied
        checkStateOfQuestion()
}
Print: High, Low and Intermediate Percentages
Print: Balance or Imbalance State of Question
}
End of Loop
getBalcePercentageOfMainQuestions()
getImbalcePercentageOfMainQuestions()
if balancePercentage  > givenValue // Value is given according to the rules
Print: Overall Paper is Balance
else
Print: Overall Paper is Imbalance
End
```

Fig. 3. Paper state check algorithm.

IV. EXPERIMENTS AND EVALUATION

The implementation was done on a computer running on Microsoft Windows 8.1, with an Intel(R) Core(TM) i3-3217U, a 1.80GHz CPU and 4 GB RAM. Java was used as the programming language to implement the program. The analysis was performed on computer science related end semester examination papers in the Department of computing and information systems of Sabaragamuwa University of Sri Lanka. Over 900 short essay questions which obtained from 30 question papers were allocated for the analysis.

Following Fig. 4 shows a sample output of Individual question categorization. Here, question is “Define the term IT Governance”. According to the output, it is a low level question.

```
run:
 [, , Define the term "IT Goverance"? ]
[,
,
Define
the
term
"IT
Goverance"?
]
[Define]
This is a low level question
BUILD SUCCESSFUL (total time: 1 second)
```

Fig. 4. Output after applying individual question categorization algorithm.

By applying the individual question categorization algorithm, the level of each and every individual question can be obtained. Generally examination papers may include full questions which include number of sub sections. To evaluate the overall question paper, the state of full questions that include sub sections was obtained.

```
run:
Percentage of Low Level Questions:      66.66666666666667%
Percentage of Intermediate Level Questions:     0.0%
Percentage of High Level Questions:     33.333333333333336%
Q1 is imbalance

Percentage of Low Level Questions:      66.66666666666667%
Percentage of Intermediate Level Questions:     33.3333333333333
Percentage of High Level Questions:     0.0%
Q2 is imbalance

Percentage of Low Level Questions:      40.0%
Percentage of Intermediate Level Questions:     60.0%
Percentage of High Level Questions:     0.0%
Q3 is imbalance

Percentage of Low Level Questions:      100.0%
Percentage of Intermediate Level Questions:     0.0%
Percentage of High Level Questions:     0.0%
Q4 is imbalance

Percentage of Low Level Questions:      100.0%
Percentage of Intermediate Level Questions:     0.0%
Percentage of High Level Questions:     0.0%
Q5 is imbalance

Percentage of Low Level Questions:      70.588235294117654%
Percentage of Intermediate Level Questions:     23.529411764705882%
Percentage of High Level Questions:     5.8823529411764714%
Overall Paper is imbalance
BUILD SUCCESSFUL (total time: 1 second)
```

Fig. 5. Output after applying paper state check algorithm.

There must be a balanced examination paper to access the exact knowledge of targeted student group. Algorithm represented in Fig. 3 returns the condition/state of the examination paper. It returns whether the question paper is balanced or unbalanced. The algorithm reads the whole question paper and checks the question paper with developed

rules. Then the user can get the actual condition of the examination paper. It will return the question wise analysis and finally returns the overall condition of the paper. Fig. 5 represents the output of the algorithm.

TABLE II: PERCENTAGES AND BALANCE/IMBALANCE STATE OF QUESTION PAPERS

Subject Title	Low Level (%)	Intermediate Level (%)	High Level (%)	Balanced/ unbalanced
Human Resource Management	53.85	38.46	7.69	Unbalanced
Platform Technologies II	22.76	43.79	33.45	Balanced
Enterprise Resource Planning	25.00	41.00	34.00	Balanced
Software Engineering	78.79	15.15	6.06	Unbalanced
Software Quality Assurance	70.59	23.53	5.88	Unbalanced
IS Risk Management	22.72	43.00	34.28	Balanced
Business Intelligence	27.17	42.5	30.33	Balanced
Object Oriented Analysis	29.43	40.52	30.05	Balanced

Table II represents the obtained results after applying the algorithm in various question papers. It denotes the percentages of high, intermediate and low level questions and state of each and every question paper.

The person who wants to get the state of the question can store the question in a text file and run our algorithm. The algorithm returns the result as represented in Fig. 5. The academics can adjust their questions according to the correct percentages to get a balanced question. To maintain the correct balance between questions, users have to include low, intermediate and high level questions in appropriate percentages. It is very beneficial while allocating marks for the questions in the marking scheme.

V. CONCLUSION

Examination question paper preparation process is a significant task in teaching process. In this paper, an automated model to categorize examination questions by getting the use of cognitive levels described in Bloom's Taxonomy were developed. The combination of developed rules and Bloom's Taxonomy had proved that this model performs well on question categorization. Teachers can easily analyze past examination papers through this model and get ideas on future examination preparation process. As well as the model enables to adjust and modify the question paper in a quantitative manner.

In our future work, we plan to conduct a thorough evaluation of our proposed approach over existing methods. We also plan to increase the performance of the approach with the machine learning techniques.

REFERENCES

[1] K. W. Hawks, "The effects of implementing bloom's taxonomy and utilizing the Virginia standards of learning curriculum framework to develop mathematics lesson for elementary students," *Liberty University*, 2010.

[2] D. A. Abduljabbar, "Question classification using different approach: A whole review," *Middle-East Journal of Scientific Research*, vol. 24, no. 5, pp. 1724-1731, 2016.

[3] B. Bloom, *Taxonomy of Educational Objectives Book 1 Cognitive Domain*, New York: David, McKay Company, INC ,1956.

[4] V. N. Pawar, P. B. Kudal, C. D. Patil, and P. S. Gaidhani, "Software engineering assessments using blooms taxonomy," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 5, no. 11, November 2016.

[5] S. Narayanan, "Analysis of question papers in engineering courses with respect to hots (Higher Order Thinking Skills)," *American Journal of Engineering Education*, vol. 6, no. 1, June 2015.

[6] S. Vijayarani, J. Ilamathi, and Nithya, "Preprocessing techniques for text mining — An overview," *International Journal of Computer Science & Communication Networks*, vol. 5, 2016.

[7] M. Mishra, V. K. Mishra, and H. R. Sharma, "Question classification using semantic, syntactic," *International Journal of Web & Semantic Technology*, vol. 4, no. 3, 2013.

[8] J. Bullington, I. Endres, and M. Rahman, "Open ended question classification using support vector machines," *MAICS*, 2007.

[9] S. S. Ilango and D. Krishna, "Blooms taxonomy — Application in exam papers assessment," *Chemical Engineering (VITU)*, vol. 12, no. 12, 2015.

Banage T. G. S. Kumara received the bachelor's degree in 2006 from Sabaragamuwa University of Sri Lanka. He received the master's degree in 2010 from University of Peradeniya and Ph.D degree in 2015 from University of Aizu, Japan. His research interests include semantic web, web data mining, web service discovery and composition.

A.Brahmana received the bachelor's degree in 2019 from Department of Computing and Information Systems, Sabaragamuwa University of Sri Lanka. Her research interests include data mining, machine learning and software engineering.

Incheon Paik received the M.E. and Ph.D. degrees in electronics engineering from Korea University in 1987 and 1992, respectively. Now he is an associate professor in the University of Aizu, Japan. His research interests include deep learning applications in semantic web services, brain signal, natural language processing, and web data together with big data infrastructure and analytics.

Prof. Incheon Paik has organized numerous international conferences, and has served as a reviewer or editor of journals of JIPS, IEICE, Hindawi, IEEE. He is a member of IEEE, IEICE, IEIE, and IPSJ. Also, he is a director of IEIE Japan Branch, and served as a chair of IEICE Service Computing Technical Committee.