

Mining Exam Question based on Bloom's Taxonomy

Siti Hasnah Tanalol^a, Salmah Fattah^b, Rina Suryani Sulong, Mazlina Mamat

^aSchool of Engineering and Information Technology
Universiti Malaysia Sabah, 88999 Kota Kinabalu, Sabah
Tel : 04-928088-320 0004701, Fax : 088-320 348

E-mail : hasnah@ums.edu.my, salmahf@ums.edu.my, rinasuryani@gmail.com, mazlina@ums.edu.my

ABSTRACT

This paper discusses the implementation of Bloom's Taxonomy concept as a guideline in designing reasonable examination question paper. A good and reasonable examination paper must consist of various difficulty levels to tolerate the different capabilities of students. Here, the difficulty level of each question deposited in the question paper is determined from the criteria of keyword/s found in the question. A knowledge-based approach and text mining technique will be used to identify and extract information and keywords from textual content in the exam paper. Besides, using the prototype system developed, an illustration of the overall analysis for level of difficulty of examination question paper will be obtained. The outcomes from the system can be used as a guideline by the academician (exam question designer) to design/revise the exam paper accordingly.

Keywords

Text mining, Bloom's Taxonomy, text categorization

1.0 INTRODUCTION

Assessment plays an important part in the teaching-learning process at all levels of education. The main purpose of classroom assessment is to improve learning (Gronlund, 1968). Since assessment plays such an important and significant part in the future of students, there is no doubt that any assessment system will determine what students learn and the way in which they do this. Hence assessment will also determine the way in which we teach and what we teach. Education and assessment is related to each other. Through assessment, we can evaluate the significance or the effectiveness of learning experience to achieve the desired goal. Examination is one of the common methods to assess knowledge acceptance of the students. Based on the examination result, student's thoughts skills and behavioral can be developed.

Written examination is a conventional yet universal tool to evaluate the student's performance in education area.

Whether or not the written examination able to assess the student's ability is very much depends on the questions deposited in the examination paper. A good examination

paper must consist of various difficulty levels to tolerate the different capability of students. This is consistent with the objective of examination to categorize the students into groups of good, average and weak. One way to achieve this is to apply some guidelines into it. Therefore, this paper tries to adopt Bloom's Taxonomy concept in designing the good examination question paper.

2.0 BLOOM'S TAXONOMY

Bloom's taxonomy is a classification system of educational objectives based on the level of student understanding necessary for achievement or mastery. Educational researcher Benjamin Bloom and colleagues have suggested six different cognitive stages in learning from the simple recall or recognition of facts, as the lowest level, through increasingly more complex and abstract mental levels, to the highest order, which is classified as evaluation (Bloom, 1956; Bloom, Hastings & Madaus, 1971). There are six levels in the taxonomy as depicted in Figure 1, moving through the lowest order processes to the highest.

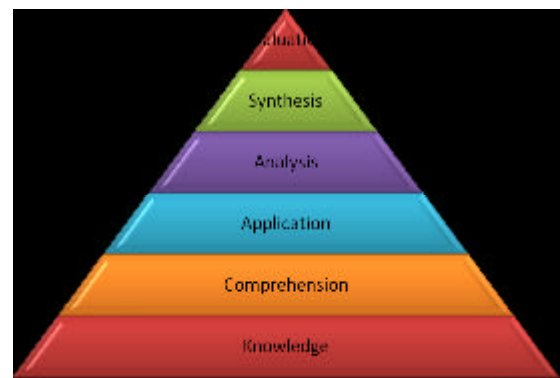


Figure 1: Bloom's Taxonomy Levels

The six developmental levels pertaining to the acquisition of knowledge and of intellectual analysis and skills are knowledge, comprehension, application, analysis, synthesis, and evaluation. Verb and question examples that represent intellectual activity on each level are listed as follow:

1. **Knowledge:** arrange, define, duplicate, label, list, memorize, name, order, recognize, relate, recall, repeat, and reproduce state.

Question verbs: Define, list, state, identify, label, name, who? when? where? what?

2. **Comprehension:** classify, describe, discuss, explain, express, identify, indicate, locate, recognize, report, restate, review, select, translate,

Question verbs: Explain, predict, interpret, infer, summarize, convert, translate, give example, account for, paraphrase x?

3. **Application:** apply, choose, demonstrate, dramatize, employ, illustrate, interpret, operate, practice, schedule, sketch, solve, use, write.

Question verbs: How could x be used to y? How would you show, make use of, modify, demonstrate, solve, or apply x to conditions y?

4. **Analysis:** analyze, appraise, calculate, categorize, compare, contrast, criticize, differentiate, discriminate, distinguish, examine, experiment, question, test.

Question verbs: Differentiate, compare / contrast, distinguish x from y, how does x affect or relate to y? why? how? What piece of x is missing / needed?

5. **Synthesis:** arrange, assemble, collect, compose, construct, create, design, develop, formulate, manage, organize, plan, prepare, propose, set up, write.

Question verbs: Design, construct, develop, formulate, imagine, create, change, write a short story and label the following elements:

6. **Evaluation:** appraise, argue, assess, attach, choose, compare, defend, estimate, judge, predict, rate, core, select, support, value, evaluate.

Question verbs: Justify, appraise, evaluate, judge x according to given criteria. Which option would be better / preferable to party y?

As an Instructional Developer, this taxonomy will help us write assessment questions and excellent course. According to Bloom, et al. (1956), as teachers we tend to ask questions in the "knowledge" category 80% to 90% of the time. These questions are not bad, but using them all the time is. Try to utilize higher order level of questions. These questions require much more "brain power" and a more extensive and elaborate answer.

Susan Edington and Cathy Hunt (1996) discovered that testing should start with planning out the objectives to be tested-preferably at different levels of Bloom's taxonomy of cognition. Test questions should also be written for different levels of cognition. As a rule, include some questions from all six categories of Bloom's Taxonomy, but make sure no more than 40% are knowledge oriented.

3.0 TEXT MINING

According to Hearst (1999), text mining is also known as Text Data Mining. A number of researchers have described text mining as:

"Looking for patterns in unstructured text".

(Nahm and Mooney 2004)

"Text mining applies the same analytical function of data mining to the domain of textual information, relying on sophisticated text analysis techniques that distill information from free-text document"

(Doore et al. 1999)

"The process of extracting interesting and trivial pattern or knowledge from text documents"

(Tan 1999)

In this study, text mining technique will be used to extract questions and keywords from textual content in the exam paper. In this section we outline two text mining approaches known as text extraction and text categorization used based on Bloom's Taxonomy concept. We focus on the two approaches because in our current research we primarily rely on these fields for text mining.

3.1 Text Extraction

According to Jerome et al. (1999), term extraction involves two tasks: detecting "interesting" collocation of words (terms) and classifying them according to classes predefined by an expert. In this section, we discuss the details of our text extraction framework. This paper focuses on the detection of specific questions from textual exam paper and extracts them based on some criteria and the outcome appears as a collection of questions. We consider the task of constructing a database by applying an information extraction system to the exam paper documents. Figure 2 shows the overview of text extraction framework.

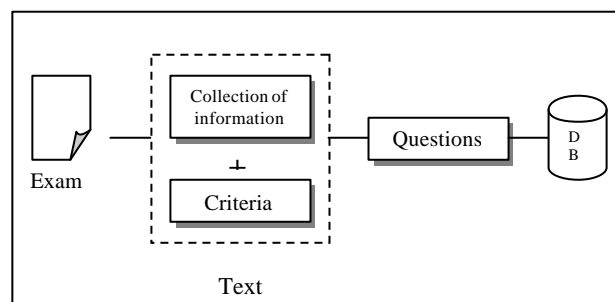


Figure 2: Overview of Text Extraction Framework

Next, we apply text categorization approach to the extracted data.

3.2 Text Categorization

In 1999, Dirre et al. published a paper in which they described that the topic Categorization tool assigns documents to preexisting categories, sometimes called "topics" or "themes". The categories are chosen to match the intended use of the collection. In the Intelligent Miner for Text those categories are simply defined by

providing a set of sample documents for each category. All the analysis of the categories, feature extraction and choice of features, i.e., key words and phrases, to characterize each category, is done automatically.

These findings further support the idea of our research. In categorization, extracted data that gathered from text extraction approach are processed and grouped into categories that are predetermined based on Bloom's Taxonomy concept. A database consisting of keywords, arranged according to the six difficulty levels as proposed by Bloom. The outcomes then will be used by the system to make overall conclusion on the drafted examination question paper. Figure 3 illustrate the overview of text categorization framework.

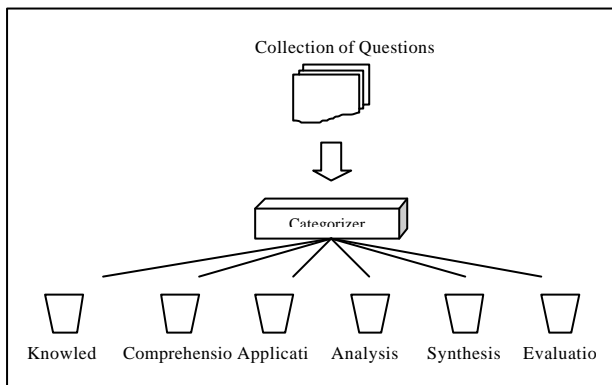


Figure 3: Overview of Text Categorization Framework

4.0 SYSTEM OVERVIEW

4.1 The Development Platform

The system will be developed on a Window-based system using Microsoft VisualBasic, C++ and MySQL. The Microsoft VisualBasic comes with all the tools required to develop the interface and the C++ as the engine to extract the keywords from the exam question and generate the difficulty level of examination question based on Bloom's Taxonomy. The system development will undergo several phases to implement its core modules that including Question Extraction, Level Categorization and Exam Question Analysis.

4.2 System Architecture

The system consists of several components including Database, Engine, and User interface. Figure 4 shows the relationships among the components.



Figure 4: The System Architecture

4.2.1 Database

The files of database are part of overall system that enables to efficiently store, search, sort and retrieve data.

The system uses MySQL database system and SQL (Structured Query Language). This database will store the examination questions, level of difficulties and keywords based on Bloom's Taxonomy.

4.2.2 Mining Engine

The engine will be written in C++. The function of this engine is to categorize the questions which will be retrieved from the Question Bank in the database. The categories are based on the first three level of Bloom's Taxonomy; knowledge, comprehension and application. The engine will retrieve the questions from the Question Bank and compare them with the keyword that kept in the database.

4.2.3 User Interface

The user interface acts as a medium of communication between users and the system. The user interface is developed using Microsoft VisualBasic.

4.3 Algorithm

In this particular research, document will be determined by its patent. A patent document contains dozens of items for analysis; some are structured, meaning they are uniform in semantics and in format across patents such as patent number, filing date, or assignees; some are unstructured, meaning they are free texts of various lengths and contents, such as claims, abstracts, or descriptions of the invention.

The patent of the document of the research is unstructured, which the document has to be extracted to the textbox before being processed. Since, the document patent is unstructured; the document must be split and indexed to be processed by frequent item set level wise search. Association rules are then extracted and interpreted with respect to set of quality measures and domain knowledge, under the control of an analyst. The text will be compare line by line in search of the particular characteristic. The text which has characteristics will be categorized before added into a database. Figure 5 and 6 exhibit the pseudo code for text extraction and text categorization.

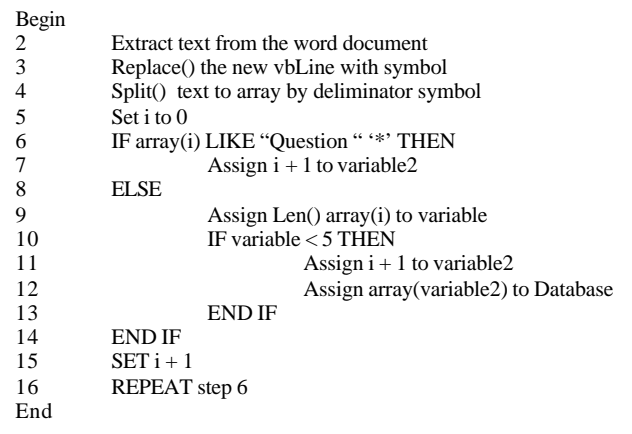


Figure 5 : Text Extraction Pseudo Code

```

START
1  Set k to 0
2  Assign array (i) to variable
3  Assign keywords.level to variable2
4  Assign InStr (variable, variable2) to varResult
5  IF varResult <> 0 THEN
6      Assign keyword.level to variable3
7      Assign InStrRev (variable, variable2) to varReverse
8      IF varReverse <> 0 THEN
9          Assign keyword.level to variable4
10         IF variable4 > variable3 THEN
11             Assign variable4 to Database
12             ELSE
13                 Assign variable3 to Database
14             END IF
15         ELSE
16             Assign variable4 to Database
17         END IF
18     ELSE
19         Delete record from Database
20     END IF
21 Set k + 1
22 Repeat Step 1
END

```

Figure 6 : Text Categorization Pseudo Code

5.0 CONCLUSION AND FUTURE WORKS

This paper discussed the implementation of Bloom's Taxonomy concept as a guideline in designing reasonable examination question paper. The system will be able to extract the questions and then categorize them into difficulty level based on Bloom's Taxonomy. With the implementation of text mining, the question will be categorized according to the level of difficulty. The outcomes from this system can be used as guideline by the academician (exam question designer) to design/revise the exam paper accordingly. Finally, the lecturer will be able to make analysis of the examination questions set and produce a quality exam paper based on the Bloom's Taxonomy.

REFERENCES

- Tan. (1999). "Text mining: The state of the art and the challenges". In Proceedings of the Pacific Asia Conference on Knowledge Discovery and Data Mining PAKDD'99 workshop on Knowledge Discovery from Advanced Databases., pages 65-70.
- Bloom, B.S. (Ed.). (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook 1. Cognitive domain. New York: McKay.
- Bloom, B.S., Engelhart, M.D., Furst, E.D., Hill, W.H. & Krathwohl, D.R. 1956. *Taxonomy of Educational*. New York: Longman Group Limited.
- Bloom, B.S., Hastings, J.T., & Madaus, G.F. (1971). Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill.

- Gronlund, E.N. 1968. *Constructing Achievement Test*. Englewood Cliffs, N.J.: Prentice-Hall.
- Jérôme Azé, Mathieu Roche, Yves Kodratoff, and Michèle Sebag. (2004). "Learning Interestingness Measures in Terminology Extraction" Proceedings Of World Academy Of Science, Engineering And Technology Volume 1.
- Jochen Dijkstra, Peter Gerstl, Roland Seiffert. (1999). Conference on Knowledge Discovery in Data archive Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining table of contents San Diego, California, United States. Pages: 398 – 401.
- Jochen Doorre, Peter Gerstl, and Roland Seiffert. (1999). "Text Mining: Finding Nuggets in Mountains of Textual Data". In Knowledge Discovery and Data Mining, pages 398-401.
- M. Hearst. (1999). "Untangling text data mining". In Proceedings of ACL'99: The 37th Annual Meeting of the Association for Computational Linguistics.
- Phye, G.D. (1997). Handbook of classroom assessment: Learning, adjustment, and achievement. San Diego, CA: Academic Press.
- Susan Edington and Cathy Hunt. 1996. Teaching Consultation Process Sourcebook. Classroom Evaluation and Assessment. *Stillwater, OK: New Forums Press*, page 57-66
- Un Yong Nahm and Raymond J. Mooney. (2004). "Text mining with information Extraction".