IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, July 2022
Impact Factor 6.252

Classification of VTU Internal Questions Complexity Level using Cognitive Methods

Akhila U[1], Sushmitha K[2], Susheela M[3], Prof. Dr. H. Girisha[4]
Students, Department of Computer Science[1,2,3]
HOD, Department of Computer Science[4]
Rao Bahadur Y Mahabaleshwarappa Engineering College, Ballari, Karnataka, India

Abstract: *The accurate estimation of the difficulty position of the questions posed to scholars is essential to help them to learn more effectively and efficiently. Still, its agreed that preceptors generally fail to identify the correct difficulty position of the questions, according to the answers and final scores attained by their scholars. Therefore, this process examines the capability of preceptors for grading questions by difficulty position, comparing it with the scholars perception and the measures attained by an expert system of question automatic bracket. The results show that scholars perceptive questions more delicate than preceptors, except for the harder ones. In addition, preceptors are only smoothly more accurate (near to the expert system), in malignancy of the general scholars tendency to overrate the difficulty position of less delicate questions. Although no general conclusions can be attained about behaviour and delicacy of preceptors and scholars when they dissect the difficulty of literacy material, the handed analysis could be veritably precious for preceptors in order to descry unclear problem statements and scholars misconceptions.*

Keywords: Educational Technology, Automatic Question Classification, Expert Systems, Teachers Estimation, Students perception.

I. INTRODUCTION

It is very common that teachers have to estimate the difficulty level of learning material (questions, items, problems…) when they design and define assessment processes for their students. However, several studies indicate that teachers usually fail to identify the correct difficulty level of the questions, according to the answers and performance of their students [1]-[5]. Besides, if the aim of adapting assessment to students according to their knowledge level is to increase their motivation and efficiency, it is important to take into account their difficulty level perception. However, there are not many studies where the perception and estimation of difficulty level done by students and teachers are compared and analyzed in an objective way. Moreover, as it will be shown in next section, there are not definitive conclusions about the accuracy of the teachers' and students' estimations, since, beyond subjective impressions, it can depend on several factors like the type and representation format of the problem or the specific context. Thus, a more objective estimator, that specifically addresses the context under study, could be a powerful tool in such adaptive learning environments.

Bloom's taxonomy is set of hierarchical models used for classification of educational learning objectives into levels of complexity and specificity. It is very common that teachers have to estimate the difficulty level of questions when they design and define assessment processes for their students.However, several studies indicate that teachers usually fail to identify the correct difficulty level of the questions, according to the answers and performance of their students.Besides, if the aim of adapting assessment to students according to their knowledge level is to increase their motivation and efficiency, it is important to take into account their difficulty level perception.However, there are not many studies where the perception and estimation of difficulty level done by students and teachers are compared and analyzed in an objective way.

II. LITERATURE SURVEY

Education Taxonomy: Bloom's Taxonomy is widely accepted and used as an important framework to guide educators in developing a holistic assessment and promoting higher forms of thinking in education [6]. This Taxonomy was introduced by Benjamin Bloom and his research team in year 1956. The Bloom's Taxonomy consists of six cognitive

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, July 2022
Impact Factor 6.252

levels - Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. This framework was then modified by Anderson and team in year 2001. The revised Bloom's was changed from noun to verb forms [7]. The Bloom's verbs were reorganised also. The change of terminology was to indicate action because thinking implied active engagements. For example, the lowest level of original Bloom Taxonomy, "knowledge" inaccurately described a category of thinking. Thus, it was replaced with the verb "remembering". Besides, the top two January - February 2020 ISSN: 0193 - 4120 Page No. 4374 - 4385 4376 Published by: The Mattingley Publishing Co., Inc. levels were swapped. The revised taxonomy had swapped the "evaluation" stage down a level and the "creating" was revised to the highest level. In the "evaluating" level, students needed to defend, support, justify and evaluate their opinion, while at the highest level, the "creating", they needed to generate new ideas, create new products, or construct new points of view. It was revised in such a way because it was able to reflect the increase of complexity of thinking. Creative thinking was considered a more complex form of thinking comparing to the evaluating which was not necessarily involving creative thinking.

III. ARCHITECTURE DIAGRAM

Here, we are used question dataset. The dataset contains the information about the questions and different types of levels like L1, L2 and so on.,

The dataset is in the format as '.docx' or '.txt'**.**

The main assumptions of this process is

- To classify the different type of Levels effectively
- To implement the machine learning algorithm for better performance
- To implement the web application flask

It dependences on what type of the questions will be used in the docx file according to that the levels will be displayed.

There are six major constraints, which are listed in order below, starting from the simplest behavior to the most complex.

- **Knowledge**: Recall data or information or specific items, remember definition of some terms.
- **Comprehension**: Recall but do a little more (e.g. paraphrase, define and discuss to some extend), understand the meaning, translation, interpolation, and interpretation of instructions and problems.
- **Application**: Do all of the above, but can take information of an abstract nature and use in a new situation or unprompted use of an abstraction. Applies what was learned in the classroom into novel situations in the work place.
- **Analysis**: Break down a communication into its constituent parts, revealing the relationships among them. Separates material or concepts into component parts so that its organizational structure may be understood.
- **Synthesis**: Pull together many disorganized elements or parts so as to form a whole. Builds a structure or pattern from diverse elements
- **Evaluation**: Makes judgments about the value of material or methods. Make judgments about the value of ideas or materials.

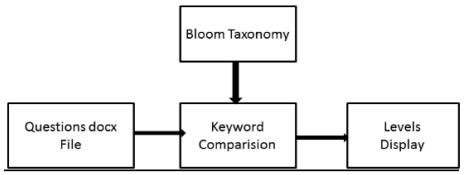


Fig 3.1: System Architecture

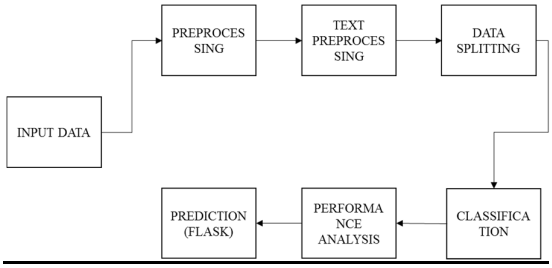


Fig 3.2: Sub System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, July 2022
Impact Factor 6.252

IV. PROPOSED SYSTEM

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.

4.1 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.

4.2 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.

4.3 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]}

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, July 2022
Impact Factor 6.252

- Rule 4: Rule for Analysis Level

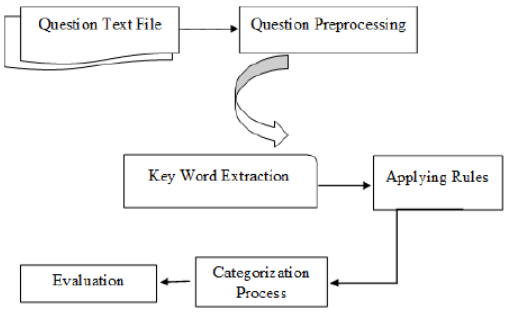
{(VB) (Analysis Keyword) + (NP) + [DT +NP]}

- 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)?}



Proposed model for question classification.

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 ofquestions 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.

V. METHODOLOGY

As aforementioned, when the students answered a question, they had to indicate their perception of difficulty according to a set of three values: 0, 1 and 2 (easy, moderate and hard, respectively). From these data, the average difficulty, as perceived by students, of the jth challenge dj has been calculated, as in (1).

$$d_j = \frac{1}{N_j}\sum_{i=1}^{3} n_{ij} \,.\, D_j$$

where nij represents the number of learners that give feedback responses belonging to the ith difficult level for the jth challenge, Nj is the total number of learners that rate the jth challenge, and Di is the challenge difficulty level quantified as 0, 1 and 2. On the other hand, teachers were also asked to indicate their difficulty level estimation, according to the same set of values: 0, 1 and 2. The average difficulty estimated by teachers has been calculated by using the same procedure shown in (1). Last, in order to compare teacher's and student's accuracy, the hybrid expert system that objectively measures the difficulty of the challenges has been used. Although the expert system works with three difficulty levels, it can obtain crisp numeric values that belong to two fuzzy sets in a considerable degree. Therefore, for a better analysis five levels have been used in this study, including those fuzzy areas found between two levels: easy, between easy and moderate, moderate, between moderate and hard, hard. Finally, the membership functions have also been used to interpret and translate the numeric values obtained from student's and teacher's input into the corresponding linguistic values.

VI. CONCLUSION

Examination question paper preparation process is a significant task in teaching process. In this paper, an automated model to categorize examination questions by performs well on question categorization. Teachers can easily analyze past

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, July 2022
Impact Factor 6.252

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

[2]. Bloom B. S., Taxonomy of educational objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc. 1956.

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

[4]. Burges, C. J. C., A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2:121–167, 1998.