Elevating Education Assessment: Machine Learning Empowered Classification of University Exam Questions via Bloom's Taxonomy

S.L.M.D.A. Umayanga, A. Mohamed Aslam Sujah, R.K.A.R. Kariapper, and A.R.M. Zulfi

Abstract This systematic literature review investigates the application of machine learning for the automated classification of university exam questions according to Bloom's Taxonomy. In the initial stage, 105 papers were identified, and after a rigorous screening process, 18 were selected for detailed analysis. The review focuses on several key themes, including Bloom's Taxonomy, Educational Assessment, Machine Learning in Education, Exam Question Classification, Automated Question Classification, and Taxonomy-Based Question Analysis. Our findings revealed significant advancements in the integration of machine learning for educational assessments, particularly in improving the accuracy and efficiency of classifying exam questions based on cognitive levels. However, challenges remain, such as the complexity of accurately aligning short-text exam questions with the different cognitive stages of Bloom's Taxonomy. These issues point to the need for further refinement of machine learning models and the datasets used for training. This review contributes to a deeper understanding of the technological advancements in automated educational assessment and highlights opportunities for further research to address current limitations, thereby enhancing the effectiveness of automated question classification in promoting comprehensive and balanced educational assessments.

Index Terms—**Bloom's Taxonomy, Educational Assessment, Exam Question Classification, Machine Learning in Education**

I. Introduction

BLOOM's Taxonomy was created by Benjamin Bloom in 1956 which is a classification system for educational objectives. It is organized into three domains: cognitive, affective, and psychosocial. Initially it has 6 categories and in 2001 Anderson and Krathwohl revised Bloom's Taxonomy, updating the category names and reordering the top two levels. Fig. 1 discloses the six categories of questions from bottom to top as lowest to highest cognitive level in Bloom's Taxonomy. Educational evaluation relies on accurately gauging students' learning outcomes, often accomplished through written exams, a common practice worldwide [1]-[2]. These exams aim to assess not only the depth of knowledge but also how students apply learned concepts in real-life situations. However, creating tests that genuinely represent various cognitive abilities poses a challenge for educators, especially when it comes to crafting short-text questions. Short-text questions, known for their brevity and lack of detailed information, make it challenging to align them accurately with Bloom's Taxonomy cognitive levels. This brevity complicates the identification of subtle cognitive abilities within these brief questions, making it difficult to create a well-rounded assessment tool. Consequently, ensuring fair and high-quality assessments across diverse cognitive domains remains an ongoing challenge in educational assessment practices [3]-[4].

S.L.M.D.A. Umayanga is a graduate from the Department of ICT, South Eastern University of Sri Lanka. (Email: anjulaumayanga047@gmail.com)

A. Mohamed Aslam Sujah is a Lecturer (Prob.) at the Department of ICT, South Eastern University of Sri Lanka. (Email: ameersujah@seu.ac.lk)

R.K.A.R. Kariapper is a Senior Lecturer at the Department of ICT, South Eastern University of Sri Lanka. (Email: rk@seu.ac.lk)

A.R.M. Zulfi is Senior Assistant Registrar at the Legal and Documentation, General Administration, South Eastern University of Sri Lanka. (Email: armzulfi@seu.ac.lk)

At the heart of educational assessment lies the goal of evaluating not just memorization but also understanding and practical application of learned material across different contexts. Designing exams that truly encompass this multifaceted aspect of learning continues to be a significant hurdle in educational teaching methods. The investigation explored the effectiveness of combining machine learning techniques with linguistically driven features, particularly those focused on syntactic text analysis, to classify questions according to Bloom's Taxonomy cognitive levels. While the

ISSN: 2961-5992

integration of these features resulted in only minor improvements in classification accuracies, their value in question classification was evident. Moving forward, future research aims to expand the database by evaluating a diverse range of questions across various academic fields. Additionally, the study exploring the implementation of TF-IDF, a well-established method in information retrieval and text mining, to enhance question classification within the framework of Bloom's Taxonomy. This exploration offers a promising avenue for refining and optimizing classification methodologies in this domain [4]-[5].

Looking ahead, there are numerous opportunities to further enhance exam classification systems through advancements in machine learning and Natural Language Processing (NLP) techniques. For instance, the use of deep learning models like Long Short-Term Memory (LSTM) networks and transformers could significantly improve the accuracy of classifying exam questions, particularly those requiring a deeper understanding of Bloom's Taxonomy. Moreover, the integration of multimodal data such as diagrams, charts, or even voice inputs present an exciting prospect for expanding the system's capabilities. This could enable more comprehensive classification across various question types, thus offering a broader range of assessment tools for educators. Additionally, incorporating real-time feedback mechanisms into these systems could provide immediate insights during exam creation, ensuring balanced question sets that align with intended learning outcomes [6]. However, the potential of current exam classification systems is not fully realized due to several challenges and limitations identified in existing studies. One major issue is the variability in the accuracy of these systems, with many studies reporting accuracy levels around 70%, which may not be sufficient for critical educational assessments. Moreover, many of the research studies informing these systems are outdated, relying on older datasets and machine learning techniques that are no longer at the cutting edge. This discrepancy in performance can also be attributed to the quality of the data used, the models applied, and the limited scalability of existing systems across different educational contexts. Furthermore, most systems still rely on isolated text-based data, failing to provide a holistic assessment tool that can handle the complexity of modern educational environments [7]-[9].

Given these challenges, there is a pressing need to conduct a systematic literature review (SLR) that examines the current state of research on exam classification systems. Such review would be crucial in identifying gaps in literature, understanding the limitations of current models, and proposing new directions for future research. By focusing on key factors such as model accuracy, the integration of diverse data sources, and real-time analysis capabilities, this review would provide a comprehensive and up-to-date analysis of the field, helping to drive improvements in educational assessments.

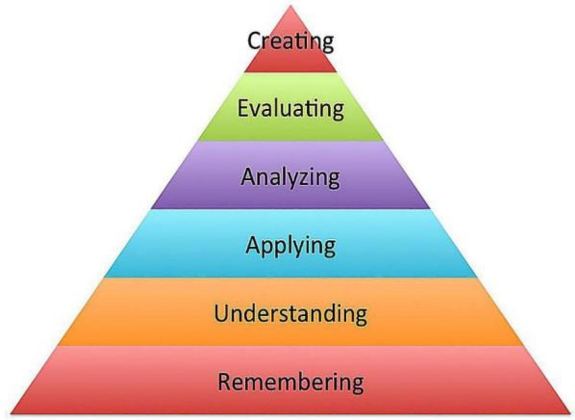

Fig. 1. Bloom's six categories of questions from lowest to highest cognitive level.

II. METHODOLOGY

A. Systematic Literature Review

The Planning phase involved developing a strategic approach to identify relevant literature. A broad search was performed using several electronic databases, including IEEE Explore, SpringerLink, Science Direct, ACM Digital Library, Scopus, and Emerald Insight. A total of 105 papers were initially downloaded from these sources. To ensure the relevance of the papers, duplicates were removed, leaving 80 unique papers. Further filtering was carried out to exclude non-English papers, resulting in 78 papers eligible for review. At this stage, specific search strings were also created to refine the search and retrieve the most pertinent studies.

During the Conducting phase, a systematic approach was employed to select the most relevant papers. Papers were assessed based on their titles, abstracts, and keywords, which narrowed the selection down to 50 papers. To enhance the comprehensiveness of the review, a snowballing technique was used to identify additional relevant studies. This technique led to the inclusion of two more papers from the references of the initially selected ones. After applying these methods, a total of 18 papers were finalized for the literature review, each carefully evaluated to ensure its relevance to the research objectives. Fig. 2 detailly explain the selection process flow of the research papers.

In the Reporting phase, the results were systematically documented and organized. The selected papers were compiled into a tabular format, detailing key aspects such as the title, abstract, keywords, research objectives, research questions, methodology, summary of results, threats to validity, and future perspectives. This comprehensive documentation was used to produce three main deliverables: the Systematic Literature Review (SLR), which includes all retrieved papers; the Research Overview, which summarizes the filtered papers; and the Detailed Literature Review, which provides an in-depth analysis of the 18 selected papers. This structured methodology called Systematic Literature Review (SLR) ensured a thorough and organized presentation of the literature, addressing the

ISSN: 2961-5992

research questions and offering valuable insights.

B. Research Questions

In this study, the primary focus is on the automated classification of exam questions using Bloom's Taxonomy, leveraging machine learning models, natural language processing, and other computational techniques. The research questions guide the exploration and analysis of these topics, aiming to address key aspects of automated educational assessment. TABLE I represents such research questions this study aimed to answer.

TABLE I
RESEARCH QUESTIONS

No.	Research Question
RQ 1	How can Bloom's Taxonomy be applied effectively in the automated classification of exam questions?
RQ 2	What machine learning models are most effective for classifying exam questions based on Bloom's Taxonomy?
RQ 3	How does the integration of NLP techniques improve the accuracy of automated question classification?
RQ 4	What are the challenges and limitations in automating the classification of exam questions using Bloom's Taxonomy?
RQ 5	How can the classification system be adapted to different educational contexts and assessment formats?

C. Conducting a Search for Primary Studies

Terms and Search Strings

To answer the research questions, a comprehensive search for primary studies was conducted using specific search strings and keywords related to the automated classification of exam questions using Bloom's Taxonomy. The search was performed across various scientific databases, including IEEE Xplore, SpringerLink, and Science Direct, and others. The search strategy was structured using the Population, Intervention, Comparison, and Outcome (PICO) framework, which guided the selection of relevant keywords and search strings.

TABLE II
SEARCH STRINGS

Area	Search Terms
Automated Classification	"Automated Classification", "Automatic Question Classification", "Question Categorization"
Bloom's Taxonomy	"Bloom's Taxonomy", "Cognitive Levels", "Educational Taxonomy"
Educational Assessments	"Educational Assessments", "Exam Questions", "Assessment Systems", "Question Papers"
Machine Learning Models	"Machine Learning", "LSTM", "Support Vector Machine", "Naive Bayes", "Learning Vector Quantization"
Natural Language Processing	"Natural Language Processing", "Text Mining", "tokenization", "Word Embeddings"

The search strategy involved using the above search strings to query scientific databases and repositories. This approach ensured the retrieval of relevant studies that addressed the application of Bloom's Taxonomy in automated question classification. The search was supplemented by manual browsing of reference lists and related papers to identify additional relevant studies. This comprehensive search process enabled the collection of a robust set of primary studies for systematic review and analysis that tabular in TABLE II.

Sources

This Systematic Literature Review was performed using IEEE Explore, SpringerLink, Science Direct, ACM Digital Library, Scopus and Emerald Insight electronic databases and considered the most relevant studies.

Inclusion and Exclusion Criteria

Inclusion Criteria: The inclusion criteria were meticulously chosen to ensure that the studies selected for this systematic literature review were relevant, high quality, and directly addressed the research questions related to the automated classification of exam questions using Bloom's Taxonomy. The selected studies included books, peer-reviewed papers, journals, and technical reports that focused on the application of natural language processing (NLP), machine learning, and other computational techniques in the domain of educational assessments.

ISSN: 2961-5992

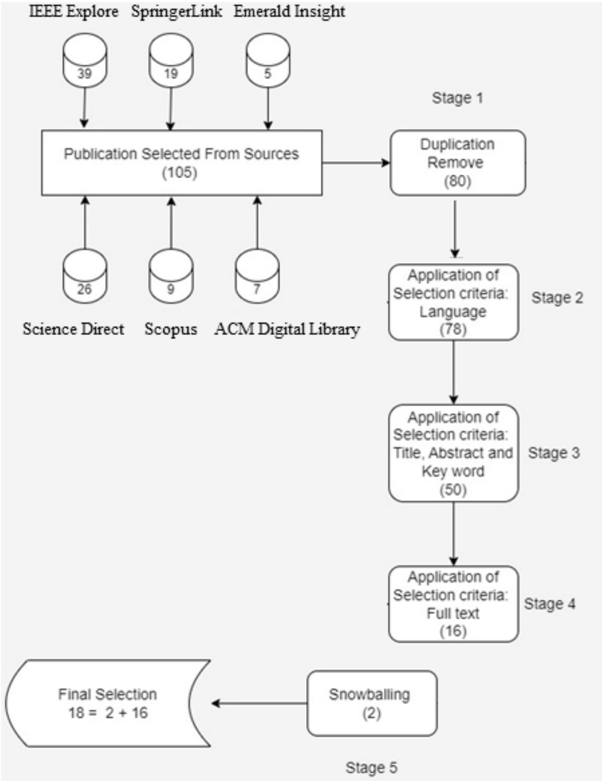

Fig. 2. Selection Process Flow

The inclusion criteria were designed to encompass studies that not only applied advanced computational techniques but also contributed to the theoretical understanding and practical application of Bloom's Taxonomy in educational settings. Studies that clearly articulated their methodologies, provided empirical evidence, and demonstrated the potential for replication or further research were prioritized. TABLE III lists out the three Inclusion Criteria for selection process.

TABLE III
INCLUSION CRITERIA FOR THE SELECTION PROCESS

No	Inclusion Criteria
IC1	Studies that employ Bloom’s Taxonomy for classifying or evaluating exam questions.
IC2	Research papers that utilize machine learning models such as SVM, LSTM, or Naive Bayes for classification.
IC3	Papers published in peer-reviewed journals or conferences within the last decade. (2015-2024)

Exclusion Criteria: The exclusion criteria were established to filter out studies that did not meet the necessary standards of relevance and quality for this review. This was done to maintain the focus on high quality research that directly contributes to the understanding and development of automated exam question classification using Bloom's Taxonomy.

The exclusion criteria effectively filtered out irrelevant or low-quality studies, ensuring that the review was based on the most current and applicable research. This process also excluded studies that did not employ the necessary computational methods, as well as those not aligned with the specific focus on Bloom's Taxonomy and educational assessments. Below TABLE IV lists out four Exclusion Criteria for the elimination process.

TABLE IV
EXCLUSION CRITERIA FOR THE SELECTION PROCESS

No	Exclusion Criteria
EC1	Studies not written in English.
EC2	Papers that only publish an abstract without full paper access.
EC3	Research outside the scope of Bloom’s Taxonomy or that does not focus on educational assessments.
EC4	Studies that do not utilize computational techniques for classification (e.g., manual classification only).

Keywording of Abstracts: Keywording of abstracts is a crucial step in the systematic literature review process, enabling the researchers to quickly identify the core themes and contributions of each study. This process involves two key steps: abstract analysis and keyword synthesis.

Initially, the abstracts of the selected papers were thoroughly reviewed to identify keywords and themes that reflected the main contributions of each study. The reviewers focused on extracting terms related to Bloom’s Taxonomy, machine learning models (e.g., SVM, LSTM, Naive Bayes), Natural Language Processing, and educational assessments. The context of each study was also assessed to understand the specific applications or educational settings being addressed.

Once the keywords from all studies were collected, they were synthesized into a comprehensive list, providing a high-level overview of the research landscape. This synthesis helped in categorizing the studies according to their primary focus, such as the application of Bloom’s Taxonomy in question classification, the use of specific machine learning models, or the integration of NLP techniques.

Data Extraction and Synthesis

During this study, a total of 105 papers were initially retrieved from six major databases: IEEE Explore, SpringerLink, Science Direct, ACM Digital Library, Scopus, and Emerald Insight. Specifically, 39 papers (37.14%) were

ISSN: 2961-5992

retrieved from IEEE Explore, 19 papers (18.10%) from SpringerLink, 26 papers (24.76%) from Science Direct, 7 papers (6.67%) from ACM Digital Library, 9 papers (8.57%) from Scopus, and 5 papers (4.76%) from Emerald Insight. In the first stage of the selection process, duplicates were identified and removed, leaving 80 unique papers for further assessment.

Next, the inclusion and exclusion criteria were applied to the selected publications to identify the most relevant studies. The inclusion criteria focused on papers that utilized Bloom's Taxonomy for classifying exam questions and incorporated machine learning models such as SVM, LSTM, or Naive Bayes for classification. Additionally, only peer-reviewed studies published between 2015 and 2024 were considered. The exclusion criteria filtered out non-English papers, studies without full paper access, and those that did not focus on educational assessments using computational techniques.

In the second stage, two papers were removed because they were not written in English. In the third stage, the titles, abstracts, and keywords of the remaining 78 papers were reviewed to assess their relevance to the research objectives. This resulted in the exclusion of 28 papers, which did not meet the necessary criteria, leaving 50 papers (47.62%) for a more thorough evaluation.

The fourth stage involved a detailed full-text analysis of the 50 papers. This stage applied a more rigorous set of inclusion and exclusion criteria to ensure that the studies were of high quality and directly addressed the research questions related to automated exam question classification. As a result, 34 papers were excluded due to their failure to meet specific inclusion criteria, such as focusing on manual rather than automated classification or not utilizing Bloom's Taxonomy. This left a final set of 16 papers (15.24%) for the review.

To enhance the comprehensiveness of the review, a snowballing technique was employed in the fifth stage. This technique involved reviewing the references of the selected papers to identify additional relevant studies. Two more papers were retrieved through snowballing, bringing the total number of papers included in the final synthesis to 18 (17.14%).

These 18 papers were thoroughly analyzed and synthesized. Key information, such as the title, abstract, keywords, research objectives, research questions, methodology, and results, was extracted and documented. This synthesis provided a clear understanding of the existing literature, identifying key trends, gaps, and opportunities for future research in the domain of automated classification of exam questions using Bloom's Taxonomy.

III. Results and Discussion

The automation of exam question classification using Bloom's Taxonomy has been a focal point of research in the educational domain, particularly within engineering and computer science. Numerous studies have contributed to the development of methodologies and systems that improve the accuracy and efficiency of classifying exam questions according to the cognitive levels defined by Bloom's Taxonomy. This section provides an overview of the significant contributions from various researchers in this field.

In the study by Kithsiri et al. [1], an automatic classifier was developed specifically for engineering exam questions. The system utilized a combination of text processing techniques and Bloom's Taxonomy to categorize questions into different cognitive levels. The study highlighted the challenges of manual classification and demonstrated that the automated approach significantly reduced the time required for question classification while maintaining a high level of accuracy.

K. Jayakodi et al. [2] expanded on this work by integrating WordNet and cosine similarity measures into the classification process. Their approach involved using semantic analysis to understand the context of exam questions better, leading to more precise classification outcomes. The authors reported that this method outperformed traditional keyword-based techniques, particularly in handling questions that required higher order cognitive skills.

Another significant contribution was made by D. Uma et al. [3], who analyzed cognitive thinking in assessment systems using the revised Bloom's Taxonomy. They employed a detailed analysis of question phrasing and structure to improve the accuracy of cognitive level identification. Their results indicated that the revised taxonomy provided a more nuanced framework for evaluating higher education assessments, particularly in distinguishing between different levels of understanding and application.

Asshaari et al. [10] focused on applying the Rasch Measurement Model to engineering mathematics courses. Their research involved a detailed appraisal of how well exam questions aligned with Bloom's cognitive levels. By employing this model, they could quantify the separation between different cognitive levels, thus providing a more objective measure of question difficulty and alignment with learning outcomes.

In another study, Nazlia Omar et al. [4] automated the analysis of exam questions according to Bloom's Taxonomy. They introduced a machine learning approach that used labeled training data to classify questions. The study demonstrated that machine learning models could achieve high levels of accuracy, particularly when trained on a diverse dataset that covered a wide range of question types and cognitive levels.

The use of deep learning techniques was explored by Shaikh et al. [11], who applied LSTM networks and pretrained word embeddings to the task of classifying learning outcomes. Their approach leveraged the sequential nature of text data in exam questions, allowing for more accurate predictions of the cognitive levels. The study found that deep learning models could capture subtle differences in question phrasing that traditional methods often overlooked.

Anabela Gomes et al. [7] proposed a Bloom's Taxonomy based approach to teaching basic programming loops. They developed an educational tool that automatically generated questions aligned with specific cognitive levels. The tool was

ISSN: 2961-5992

particularly effective in reinforcing student learning by ensuring that questions were appropriately challenging and targeted the desired cognitive skills.

The integration of rule-based systems with Bloom's Taxonomy was explored by Kumara et al. [12], who developed an approach for measuring the quality of examination papers. Their system combined rule-based analysis with taxonomy-based classification to ensure exam questions were not only correctly categorized but also met the educational standards for cognitive skill development.

Cognitive classification was further enhanced by Eko Subiyantoro et al. [13], who used a learning vector quantization approach. This method allowed for more flexible classification by adapting to new types of questions that were not present in the initial training dataset. Their results showed that this approach could significantly improve the adaptability and robustness of automated classification systems.

K. Osadi et al. [14] introduced an ensemble classifier-based approach that combined multiple machine learning models to improve classification accuracy. By using a diverse set of classifiers, they were able to reduce the bias and variance in the classification process, leading to more reliable categorization of exam questions into Bloom's Taxonomy levels.

In their comprehensive review, Karima Makhlouf et al. [6] compared various approaches and techniques for classifying exam questions based on Bloom's Taxonomy. Their analysis highlighted the strengths and weaknesses of different methods, providing valuable insights into the most effective strategies for automating this process.

G.N.R. Prasad et al. [15] addressed the challenge of identifying Bloom's Taxonomy levels using NLP tokenization techniques. Their research focused on improving the accuracy of classification in languages with complex grammatical structures, demonstrating that NLP techniques could be effectively applied across different languages and educational contexts.

Erastus Karanja et al. [16] discussed the importance of aligning course learning outcomes with Bloom's Taxonomy in the context of project management education. Their study emphasized the need for a systematic approach to question classification to ensure that assessments are aligned with course objectives and learning outcomes.

Manal Mohammed et al. [8] and Omar et al. [17] both explored the use of TF-IDF and word2vec techniques for classifying exam questions. Their research demonstrated that these text mining approaches could effectively capture the semantic meaning of questions, leading to more accurate classification results.

Anbuselvan sangodiah et al. [5] investigated the application of Support Vector Machines (SVM) for taxonomy-based features in question classification. Their study highlighted the high accuracy and computational efficiency of SVMs in handling large datasets, making them a suitable choice for automated classification tasks in educational settings.

In their work, Annisa Anindityya et al. [18] used a text mining approach with Naive Bayes classifiers to analyze exam questions based on Bloom's cognitive levels. Their results showed that Naive Bayes, despite being a relatively simple model, could achieve competitive accuracy when paired with effective feature selection techniques.

Finally, Mathews Zanda et al. [9] unpacked the revised Bloom's Taxonomy to develop case-based learning activities. Their research provided insights into how taxonomy-based classification could be integrated into curriculum design, ensuring that assessments not only measure cognitive skills but also enhance student learning experiences.

IV. Conclusion

This systematic review examined emerging field of automated exam question classification using Bloom's Taxonomy, analyzing 18 key studies that integrate machine learning and natural language processing techniques into educational assessment systems. The findings reveal significant progress in developing tools capable of improving the alignment between exam content and cognitive learning outcomes. However, despite the growing interest in this area, several limitations continue to hinder the widespread adoption and effectiveness of these systems.

Current challenges include the limited generalizability of models across subject areas, the absence of real-time classification features, and the reliance on small or domain specific datasets. These issues must be addressed to fully realize the potential of automation in educational assessments. To overcome these limitations, this review recommends the following:

Future research should focus on developing hybrid or ensemble classification models, combining approaches such as support vector machines (SVM), LSTM networks, and transformer-based architectures. These models are more likely to handle the variability in question structure and domain context effectively. Educational platforms should consider embedding real-time NLP tools into assessment design environments. This would enable instructors to receive immediate feedback on the cognitive levels targeted by their questions, thereby promoting balanced and pedagogically sound exam papers. There is a pressing need for the creation of large, open-access datasets that include a wide variety of question formats and academic disciplines. These datasets will support the training of more accurate and adaptable models. Classification systems should incorporate multilingual capabilities and context-aware analysis to accommodate diverse educational environments and curricula. Theoretical advancements must be tested in real-world classroom settings to evaluate their practicality, accuracy, and impact on teaching and learning outcomes. Developing intuitive, user-friendly interfaces and providing training resources will be essential to ensure that instructors can effectively use and trust automated classification tools. As the application of artificial intelligence

ISSN: 2961-5992

TABLE V
COMPARATIVE REVIEW OF AUTOMATED EXAM QUESTION CLASSIFICATION STUDIES BASED ON BLOOM'S TAXONOMY

Ref.	Title	Used Tools	Considered Features	Sector	Results	Limitations
[1]	An automatic classifier for exam questions in Engineering. A process for Bloom's taxonomy	Not specified	Classification of exam questions, Bloom's taxonomy levels	Engineering	Improved classification accuracy using Bloom's taxonomy	Questions and text in one subject domain.
[2]	An Automatic Classifier for Exam Questions with WordNet and Cosine Similarity	WordNet, Cosine Similarity	Semantic analysis, similarity measures	Education	Enhanced question classification accuracy	Limited to certain types of questions.
[3]	Analysis on Cognitive Thinking of an Assessment System Using Revised Bloom's Taxonomy	Assessment system tools, Revised Bloom's taxonomy	Cognitive levels, assessment	Education	Detailed analysis of cognitive levels in assessments	Accuracy is Low.
[4]	Automated analysis of exam questions according to Bloom's taxonomy	Automated tools	Exam question analysis, Bloom's taxonomy	Education	Effective categorization of questions	Limited to specific subject areas.
[5]	Taxonomy Based Features in Question Classification Using Support Vector Machine	Support Vector Machine (SVM)	Question classification, Bloom's taxonomy	Education	High accuracy with SVM	Computationally expensive.
[6]	Exam Questions Classification Based on Bloom's Taxonomy: Approaches and Techniques	Various classification techniques	Exam question classification, Bloom's taxonomy	Education	Comprehensive overview of techniques	General overview and lacks specific results.
[7]	Bloom's Taxonomy Based Approach to Learn Basic Programming Loops	Programming tools	Learning programming, Bloom's taxonomy	Computer Science Education	Enhanced learning outcomes for programming	Limited scope to basic loops.
[8]	Question Classification Based on Bloom's Taxonomy Using Enhanced TF-IDF	TF-IDF, Enhanced version	Question classification, Bloom's taxonomy	Education	Enhanced classification accuracy	Limited to text-based questions.
[9]	Unpacking the revised Bloom's taxonomy: developing case-based learning activities	Various classification techniques	Learning activities, Revised Bloom's taxonomy	Education	Enhanced learning activities using case-based methods	Not extensively validated.
[10]	Appraisal on Bloom's Separation in Exam Question of Engineering Mathematics Courses using Rasch Measurement Model	Rasch Measurement Model	Examination questions, Bloom's taxonomy separation	Engineering Education	Improved measurement accuracy	Requires further validation.
[11]	Bloom's Learning Outcomes' Automatic Classification Using LSTM and Pretrained Word Embeddings	LSTM, Word Embeddings	Learning outcomes classification	Education	High accuracy in automatic classification	Complex models require significant resources.

ISSN: 2961-5992

[12]	Bloom's Taxonomy and Rules Based Question Analysis Approach for Measuring the Quality of Examination Papers	Rules-based approach	Examination quality, Bloom's taxonomy	Education	Improved quality measurement of exam papers	Rule-based systems may lack flexibility.
[13]	Cognitive Classification Based on Revised Bloom's Taxonomy Using Learning Vector Quantization	Learning Vector Quantization	Cognitive classification, Revised Bloom's taxonomy	Education	Effective classification method	Requires further testing.
[14]	Ensemble Classifier based Approach for Classification of Examination Questions into Bloom's Taxonomy Cognitive Levels	Ensemble methods	Classification of exam questions, Bloom's taxonomy	Education	Increased classification accuracy	Computationally intensive.
[15]	Identification of Bloom's Taxonomy level for the given Question paper using NLP Tokenization technique	NLP, Tokenization	Question analysis, Bloom's taxonomy	Education	Accurate identification of Bloom's taxonomy levels	Limited by the quality of tokenization.
[16]	Improving project management curriculum by aligning course learning outcomes with Bloom's taxonomy framework	Curriculum tools	Project management, Bloom's taxonomy	Education	Improved alignment of curriculum with learning outcomes	Requires adaptation for different subjects.
[17]	Question classification based on Bloom's taxonomy cognitive domain using modified TF-IDF and word2vec	TF-IDF, word2vec	Cognitive domain classification, Bloom's taxonomy	Education	Improved classification performance	Requires careful parameter tuning.
[18]	Text Mining Approach Using TF-IDF and Naive Bayes for Classification of Exam Questions Based on Cognitive Level of Bloom's Taxonomy	TF-IDF, Naive Bayes	Text mining, cognitive level classification	Education	Effective classification using text mining	Limited to certain question types.

In education continues to evolve, automated classification systems based on Bloom's Taxonomy hold great promise for improving the quality and fairness of assessments. By aligning exam questions with intended cognitive outcomes more accurately and efficiently, these systems can support more meaningful evaluations of student learning. This review contributes to the growing body of research in this field and offers a framework for guiding future developments toward more reliable, scalable, and inclusive assessment technologies.

REFERENCES

[1] K. Jayakodi, M. Bandara, and I. Perera, "An automatic classifier for exam questions in Engineering: A process for Bloom's taxonomy," Proceedings of 2015 IEEE International Conference on Teaching, Assessment and Learning for Engineering, TALE 2015, pp. 195–202, Jan. 2016, doi: 10.1109/TALE.2015.7386043.

[2] K. Jayakodi, M. Bandara, and D. Meedeniya, "An automatic classifier for exam questions with WordNet and Cosine similarity," 2nd International Moratuwa Engineering Research Conference, MERCon 2016, pp. 12–17, May 2016, doi: 10.1109/MERCON.2016.7480108.

[3] D. Uma, S. Thenmozhi, and R. Hansda, "Analysis on cognitive thinking of an assessment system using revised Bloom's taxonomy," Proceedings - 5th IEEE International Conference on MOOCs, Innovation and Technology in Education, MITE 2017, pp. 152–159, Dec. 2018, doi: 10.1109/MITE.2017.00033.

[4] N. Omar et al., "Automated Analysis of Exam Questions According to Bloom's Taxonomy," Procedia Soc Behav Sci, vol. 59, pp. 297–303, Oct. 2012, doi: 10.1016/J.SBSPRO.2012.09.278.

[5] A. Sangodiah, R. Ahmad, W. Fatimah, and W. Ahmad, "TAXONOMY BASED FEATURES IN QUESTION CLASSIFICATION USING SUPPORT VECTOR MACHINE," J Theor Appl Inf Technol, vol. 30, no. 12, 2017, Accessed: Dec. 11, 2025. [Online]. Available: www.jatit.org

[6] K. Makhlouf, L. Amouri, N. Chaabane, and N. El-Haggar, "Exam Questions Classification Based on Bloom's Taxonomy: Approaches and Techniques," 2020 2nd International Conference on Computer and Information Sciences, ICCIS 2020, Oct. 2020, doi: 10.1109/ICCIS49240.2020.9257698.

[7] "Bloom's Taxonomy Based Approach To Learn Basic Programming Loops | PDF | Teachers | Reading Comprehension." Accessed: Dec. 11, 2025. [Online]. Available: https://www.scribd.com/document/584123272/10-1109-FIE-2018-8658947

[8] M. Mohammed and N. Omar, "Question classification based on Bloom's Taxonomy using enhanced TF-IDF," Int J Adv Sci Eng Inf Technol, vol. 8, no. 4–2, pp. 1679–1685, 2018, doi: 10.18517/IJASEIT.8.4-2.6835

[9] M. Z. Nkhoma, T. K. Lam, N. Sriratanaviriyakul, J. Richardson, B. Kam, and K. H. Lau, "Unpacking the revised Bloom's taxonomy: developing case-based learning activities," Education and Training, vol. 59, no. 3, pp. 250–264, 2017, doi: 10.1108/ET-03-2016-0061.

[10] I. Asshaari, H. Othman, H. Bahaludin, N. A. Ismail, and Z. M. Nopiah, "Appraisal on Bloom's Separation in Final Examination

ISSN: 2961-5992

Question of Engineering Mathematics Courses using Rasch Measurement Model," Procedia Soc Behav Sci, vol. 60, pp. 172–178, Oct. 2012, doi: 10.1016/J.SBSPRO.2012.09.364.

[11] S. Shaikh, S. M. Daudpotta, and A. S. Imran, "Bloom's Learning Outcomes' Automatic Classification Using LSTM and Pretrained Word Embeddings," IEEE Access, vol. 9, pp. 117887–117909, 2021, doi: 10.1109/ACCESS.2021.3106443.

[12] B. T. G. S. Kumara, "Bloom's Taxonomy and Rules Based Question Analysis Approach for Measuring the Quality of Examination Papers," International Journal of Knowledge Engineering, pp. 20–24, 2019, doi: 10.18178/IJKE.2019.5.1.111.

[13] E. Subiyantoro, A. Ashari, and Suprapto, "Cognitive Classification Based on Revised Bloom's Taxonomy Using Learning Vector Quantization," CENIM 2020 - Proceeding: International Conference on Computer Engineering, Network, and Intelligent Multimedia 2020, pp. 349–353, Nov. 2020, doi: 10.1109/CENIM51130.2020.9297879.

[14] K. A., M. G., and W. V., "Ensemble Classifier based Approach for Classification of Examination Questions into Bloom's Taxonomy Cognitive Levels," Int J Comput Appl, vol. 162, no. 4, pp. 1–6, Mar. 2017, doi: 10.5120/IJCA2017913328.

[15] G. N. R Prasad Sr Asst professor, "Identification of Bloom's Taxonomy level for the given Question paper using NLP Tokenization technique," Turkish Journal of Computer and Mathematics Education, vol. 12, no. 13, pp. 1872–1875, 2021.

[16] E. Karanja and L. C. Malone, "Improving project management curriculum by aligning course learning outcomes with Bloom's taxonomy framework," Journal of International Education in Business, vol. 14, no. 2, pp. 197–218, 2020, doi: 10.1108/JIEB-05-2020-0038.

[17] M. Mohammedid and N. Omar, "Question classification based on Bloom's taxonomy cognitive domain using modified TF-IDF and word2vec," PLoS One, vol. 15, no. 3, p. e0230442, 2020, doi: 10.1371/JOURNAL.PONE.0230442.

[18] A. Anindditya, M. A. Hasibuan, and E. Sutoyo, "Text mining approach using TF-IDF and naive bayes for classification of exam questions based on cognitive level of bloom's taxonomy," Proceedings - 2019 IEEE International Conference on Internet of Things and Intelligence System, IoTaIS 2019, pp. 112–117, Nov. 2019, doi: 10.1109/IOTAIS47347.2019.8980428.

ISSN: 2961-5992