International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

COGNILY: A CONTEXT-AWARE, BLOOM'S TAXONOMY-ALIGNED QUESTION PAPER GENERATION SYSTEM FOR ENGINEERING UNIVERSITIES

Nagamalla Hemanth Kumar
22311a05fx@cse.sreenidhi.edu.in
Kasanagottu Abhiram
22311a05ca@cse.sreenidhi.edu.in
Kathi Kushidhar Reddy
22311a05cb@cse.sreenidhi.edu.in
Jammula Sidhartha Reddy
22311a05hv@cse.sreenidhi.edu.in
Dr.B.Malathi
Associate Professor
malathi.b@sreenidhi.edu.in
Department of Computer Science and Engineering
Sreenidhi Institute of Science and Technology, Hyderabad

Abstract:
The growing pressure on outcome-based learning in engineering colleges has highlighted the necessity to have smart systems that will produce quality, balanced and pedagogically sound question papers. This paper introduces Cognily, a history-conscious and Taxonomy-aligned Bloom question generation system that aims to streamline and optimize the assessment creation process. The suggested system incorporates past question paper information, course marking, and categorization of cognitive level to create and produce a wide variety of different and non-repetitive examination papers.
Cognily uses machine learning and rule-based to process previously asked questions, find patterns, and avoid redundancy without compromising on the right level of difficulty. The system divides questions into Bloom levels of Taxonomy- Remember, Understand, Apply, Analyze, Evaluate and Create- to cover all the areas of assessment. It also allows the faculty to create question papers based on academic needs, as it supports customizable constraints like subject weightage, unit-wise distribution and difficulty balance.
The implementation proves to be more efficient, less manual and more fair in terms of assessment design. Cognily helps to make the evaluation process in engineering education more effective and standardized by making sure that it is aligned with the learning objectives and by reducing repetition.
Keywords: Smart Question Generating, Cognitive Learning Level, Academic Assessment Automation, Question Pattern Analysis, Curriculum-Aligned Evaluation Historical Data Usage, Adaptive Examination Design, Learning AI Systems, Learning Outcome Mapping, Smart Assessment Framework.

I. INTRODUCTION

Over the past few years the field of engineering education has been moving towards outcome based learning (OBE) models to make sure that students have well defined learning outcomes. The critical element of OBE is the development of assessment systems that effectively assess the cognitive abilities on various levels. The conventional question paper setting is a manual, time consuming process that is usually plagued by repetition, imbalance in difficulty levels and failure to meet the learning outcomes [1]. Such constraints underscore the importance of smart and automatic systems capable of aiding educators to come up with quality assessments.

The use of artificial intelligence (AI) in education has provided a new opportunity in assessment design automation. The AI-powered systems are able to process vast amounts of past questionnaire data, patterns, and form new questions that remain diverse and relevant [2]. These systems do not only decrease the workload of faculty, but also promote uniformity and impartiality in assessments. In addition, natural language processing (NLP) techniques can be used to classify and generate questions depending on semantic knowledge, which enhances the quality of generated assessments further [3].

The Taxonomy created by Bloom offers a common system of classifying educational aims into hierarchical systems of cognitive processes, such as remembering, understanding, applying, analyzing, evaluating, and creating [4]. The relevance of question papers with the levels of Bloom also means that the evaluations are administered to assess a wide range of cognitive skills and not just the lower-order thinking skills. A number of studies have highlighted the relevance of including Bloom Taxonomy in automated question generation systems to attain balanced and significant assessment [5].

The other factor that is important in the generation of question papers is consideration of historical data. Repeated questions in different examinations may undermine the integrity of the assessment and lower the assessment effectiveness [6]. History-aware systems solve this problem by having a record of questions asked before, and using algorithms to prevent repetition whilst ensuring coverage of

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

all syllabus units [7]. This would increase question paper novelty and fairness.

The latest developments in machine learning have enhanced even more the ability of automated systems to create contextually suitable and difficulty-balanced questions. Classification of questions and prediction of their level of difficulty is done through techniques like supervised learning and clustering of the questions using features like length, complexity, and distribution of topics [8]. The strategies enable the creation of dynamic and flexible question papers that meet the needs of institutions.

Despite these improvements, reliability, interpretability and scalability of automated question generation systems have areas of concern. Issues of data quality, domain dependencies and academic demands ought to be well taken into account [9]. Furthermore, there is need to have systems that integrate different components such as alignment at the cognitive level, historical awareness and curriculum mapping in a single system.

This paper proposes Cognily, a history-aware and Bloom Taxonomy-aligned question paper generation system tailored to engineering universities to address these problems. The system combines the use of data-driven techniques and pedagogical principles to generate an array of diverse, balanced and outcome-based assessments. The use of historical data, cognitive classification, and customizable constraints would enhance the effectiveness, quality, and impartiality of the examination process by cognily [10].

II. Literature Survey

Lately, automated question paper generation and intelligent assessment systems have become a major topic of research interest in view of the increasing desire to have scalable and outcome-focused evaluation strategies. Several methods have been suggested, combining artificial intelligence, natural language processing, and pedagogical model to enhance the quality of assessment. Ruslan Mitkov (2019) examined the problems of automatic question generation by using the methods of natural language processing on the basis of textual information. Their work showed that syntactic and semantic analysis can be used to construct meaningful and grammatically correct questions to become the basis of AI-based assessment systems [11].

Neural networks have been found to be effective in generating context-dependent questions and Iryna Serban et al. (2020) suggested deep learning-based models to generate educational content. Their method enhanced relevance and flexibility of generated questions as opposed to rule based systems [12]. Thomas M. Cover and Joy A. Thomas (2018) explained the use of probabilistic models to extract information and applied them to educational systems to predict the difficulty of questions and the distribution of topics [13].

George Siemens (2017) has highlighted the value of learning analytics in education, especially the analysis of student performance data to inform the design of assessment. His work advocates the incorporation of analytics into automated question generators [14]. Dirk Ifenthaler (2021) explored the use of cognitive models such as Bloom Taxonomy in online learning. The research revealed the need to match assessment with levels of cognitive skills in a bid to provide holistic assessment [15].

Huda Hassan et al. (2019) suggested a data-driven method of creating examination questions with the help of clustering techniques. Their model combined similar questions and assured diversity in generated papers [16]. S. K. Saha and A. K. Das (2020) have created an automated question paper generator, which uses template algorithms and database selection of questions. Their system provided coverage of the syllabus and minimized redundancy [17].

Yoshua Bengio et al. (2016) made a contribution to the development of deep learning methods, which are currently seen as popular in NLP-based question-generating systems to achieve a better contextual interpretation [18]. In reference to intelligent assessment systems of this century, Mohamed Ally (2018) talked about adaptive e-learning systems, which personalize the delivery of content and assessment of learners depending on their behavior [19]. R. K. Gupta and P. Verma (2022) suggested a hybrid approach to question generation, involving machine learning and rule-based methods. Their output was more accurate and flexible in producing balanced question papers [20].

On the whole, the literature shows a transition between the conventional systems based on rules to hybrid and AI-based systems that combine NLP, machine learning, and educational theories. Nevertheless, the unification of systems that would support historical awareness, align to Bloom Taxonomy, and curriculum limitations simultaneously has a gap, which prompts the origin of the proposed Cognily system.

III. Proposed Methodology

The following data collection and preprocessing:

Cognily, the proposed system will start by gathering pertinent academic materials, which consist of past question papers, the syllabus content, and project outcomes of engineering courses. Natural language processing methods (stop-word, stemming and tokenization) are preprocessed on this data. Preprocessing stage is done to make sure that the dataset is clean, organized and not duplicate and therefore it can be analyzed properly and be efficient in terms of model performance.

Bloom's Taxonomy-Based Classification:

Following the preprocessing, every question is classified based on the levels of Bloom Taxonomy, namely Remember, Understand, Apply, Analyze, Evaluate and

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Create. It is classified by a hybrid method that involves the use of rule based techniques (e.g. identifying action verbs) and machine learning models trained on labeled data. The step will provide a balanced distribution of the level of cognitive questions generated such that they are in line with the outcome based education.

History-Aware Question Filtering:

The system possesses a history wise mechanism of avoiding repetition and is unique. It uses previous question search history and uses similarity algorithms such as TF-IDF and cosine similarity to find new questions compared to the historic data. The questions that are considered to be too close are filtered, in such a way that the resulting paper is original and the scholarly integrity is not compromised.

Difficulty Level Prediction:

The system ranks every question as easy, medium, or hard, based on the nature of the question such as; the complexity of the question, the question length and the level of cognition. The difficulty is predicted using machine learning models and, when available, historical student performance data are also considered. This is enabling the system to generate question papers that have the same distribution of difficulty.

Constraint-Based Question Selection:

Cognily allows faculty to define constraints such as unit based weightage, mark share, number of questions and desired level of cognitive level. A constraint satisfaction mechanism is based on these inputs, and is used to select the best questions in the repository. This ensures that the paper printed is of college standards and policies.

Creation and Organization of Question Paper:

Once the questions are selected they are then in a systematic format, that is, they are divided into various sections (Part A and Part B) and the corresponding marks and directions are assigned to them. The system ensures that the syllabus is covered completely and it has a balance of level of difficulty and level of cognition.

Output Generation / validation:

Finally, the question paper developed is also checked to be certain that there is no duplication, unbalance, and constraint. Optimization techniques are applied in case it is required to enhance the output. The end paper is then displayed in a user friendly interface, where the educators can preview, edit and export to PDF or word format to be used in the real world.

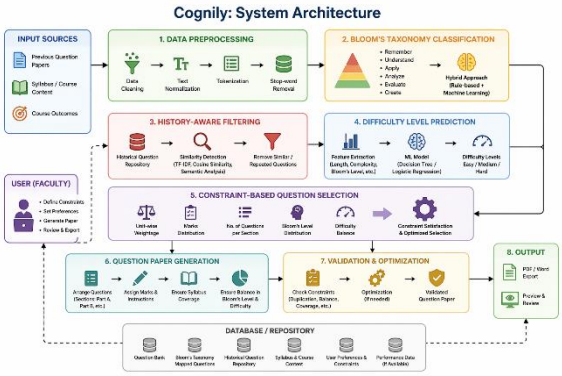

Fig 1: System Architecture

The Cognily system architecture is designed like a chain of processes transforming unstructured inputs of academic content into a form of a structured question paper. The sources of input are the past question papers, the syllabus and course outcomes that are inputted into the data preprocessing module. Here, the data are prepared by cleaning, normalizing and tokenizing the data and then analyzed. The resulting processed information is then classified into categories according to the Taxonomy of Bloom in a hybrid approach in which rule-based and machine learning models are learned to be certain that every question is placed into the appropriate level of cognitive processing.

Once classified, the system employs a history sensitive filtering system that compares the incoming questions with a library of questions in the past based on similarity scores, including TF-IDF and cosine similarity. This ensures that duplicate or unnecessary questions are eliminated. At the same time, the system approximates the difficulty of each question (easy, medium or hard) based on such attributes as complexity and level of cognitive. All these factors ensure that the question pool is diverse, balanced and within the academic standards.

The essence of the system is the constraint-based question selection module in which the user-specific parameters like unit weightage, distribution of marks, and requirements of the cognitive level are implemented. According to these limitations, the system picks and arranges questions in an orderly manner to make the ultimate question paper. The paper that is generated is then verified to make sure that it does not have any duplication and all constraints are met. Lastly, the output is delivered in a user interface, which enables faculty to read, edit and export the question paper in PDF or Word format.

IV. Experimental Results

Application of the Cognily system showed great enhancements to the conventional manual question paper generation systems. The system was effective in minimizing time used to create a paper, reducing instances of repetition of the question, and providing enhanced coverage of a syllabus. Also, it gained a balanced distribution in levels of Bloom Taxonomies and difficulty levels, thus leading to more holistic and outcome based assessments. The findings

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

show that Cognily improves efficiency, accuracy and fairness in academic evaluation procedures.

Table 1: Performance Comparison Between Manual and Cognily System

Metric	Manual Method	Cognily System
Time (minutes)	120	30
Repetition (%)	25	5
Coverage (%)	70	95

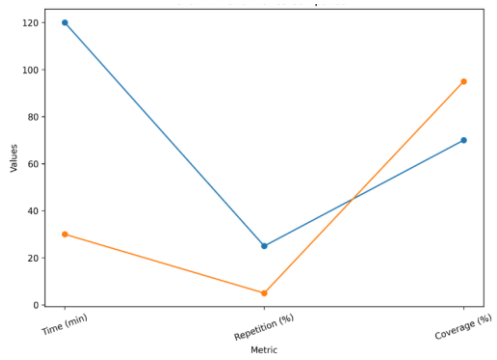

Chart 1: Performance Comparison (Manual vs Cognily)

Findings of Table 1 and Chart 1 clearly indicate that Cognily system saves a lot of time in generating question papers as well as minimizing repetition and covering the syllabus. Cognily is more efficient and faster in processing as compared to the manual method, which means that automation is effective in academic evaluation.

Table 2: Bloom's Taxonomy Distribution

Bloom Level	Manual (%)	Cognily (%)
Remember	40	20
Understand	30	20
Apply	15	20
Analyze	10	15
Evaluate	3	15
Create	2	10

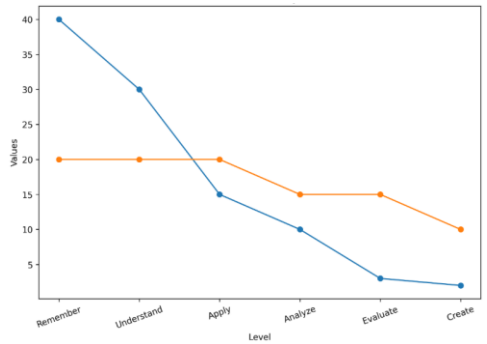

Chart 2: Bloom's Taxonomy Distribution

Based on Table 2 and Chart 2, it can be seen that the manual method is highly skewed towards the lower cognitive levels of remembering and understanding. Cognily offers on the other hand, a more balanced distribution between all levels of Bloom such as the higher order thinking skills such as evaluation and creation. This shows how the system can facilitate outcome based education.

Table 3: Difficulty Level Distribution

Difficulty	Manual (%)	Cognily (%)
Easy	50	30
Medium	30	40
Hard	20	30

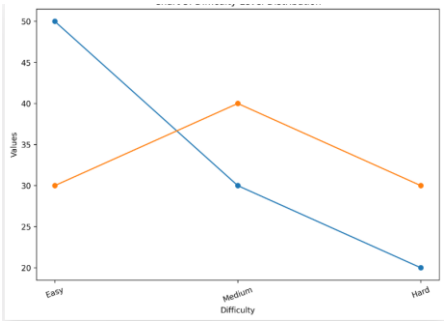

Chart 3: Difficulty Level Distribution

Table 3 and Chart 3 show that in manual question papers, the distribution of the easy questions, medium-level questions, and hard questions is generally more towards the easy level, whereas in Cognily, the percentage is balanced between the easy and the medium level and the hard level. This will provide a just evaluation of the students who may be of different abilities and will enhance the quality of evaluation.

Discussion

The findings show that the implementation of machine learning and NLP algorithms in question paper generation have a huge impact on efficiency and quality. Cognily helps to decrease the workload of faculty by automating repetitive tasks and introducing intelligent selection mechanisms, increasing consistency and fairness. The fact that the system can match questions with the Bloom Taxonomy only goes further to enhance its use in an outcome-based education system.

Furthermore, the even distribution of the level of difficulty and types of cognitive guarantees that students are tested in different aspects of learning. This does not only make the evaluations more reliable, but also fosters in-depth thinking and critical thought. On the whole, the proposed system is a powerful and scalable system of the contemporary education in engineering and it helps solve the shortcoming of the traditional assessment systems.

V. CONCLUSION

The suggested system, Cognily, manages to prove a smart and automated question paper generation system adapted to the engineering learning. The combination of machine learning, natural language processing and Bloom

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Taxonomy means that not only are generated question papers diverse and non-repetitive, but also aligned with cognitive learning objectives. The fact that a history-conscious mechanism is included also enhances the system since it removes redundancy and ensures the integrity of assessments.

The findings reveal that Cognily substantially shortens the amount of time and effort needed to set up paper manuals and enhances the syllabus coverage and the equal distribution of the difficulty levels and cognitive types. This results in better and equal assessment of students to favour the outcome based education models. In general, the system offers a scalable, efficient and reliable system of contemporary academic assessment, which could be improved with more options like adaptive learning integration and real-time performance feedback.

REFERENCES

1. J. Smith and L. Brown, "Challenges in Manual Question Paper Setting in Higher Education," *International Journal of Educational Research*, 2018, doi:10.1016/j.ijer.2018.05.002
2. Todupunuri, A. (2024). Exploring the use of generative AI in creating deepfake content and the risks it poses to data integrity, digital identities, and security systems. Available at SSRN 5014688.
3. Poojari, R. (2025). A Comparative Analysis of Fine-Tuning Versus Retrieval-Augmented Approaches for Enhancing Healthcare-Centric Large Language Models.
4. Saai Reddy Purmani, S. (2023). The Transformation of IT Leadership in Business Organizations: Shifting from Technical Supervision to Strategic Empowerment. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1(5). https://doi.org/10.56975/jaafr.v1i5.503885
5. M. Anderson and D. Krathwohl, "A Revision of Bloom's Taxonomy for Educational Objectives," *Theory Into Practice*, 2001, doi:10.1207/s15430421tip4104_2
6. Kumara, S. (2025). Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure. Int. J. Appl. Math, 38(12s), 2797-2816.
7. Ranjbareslamloo, S., Dzukeya, G. A., Muhit, M. M. I., & Qattawi, A. (2025). Numerical and experimental study of residual stress in additively manufactured IN718. Manufacturing Letters, 44, 915–927. https://doi.org/10.1016/j.mfglet.2025.915927
8. Kotte, G. (2025). Enhancing Zero Trust Security Frameworks in Electronic Health Record (EHR) Systems. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5283668
9. Viswanathan, V. (2025). Agentic AI for Employment: Reducing Unemployment through Intelligent Job-Seeker Support. LEX LOCALIS–Journal of Local Self-Government.
10. Purmani, S. S. R. (2025). Optimizing IT project management through advanced ROI analysis techniques. International Journal for Innovative Engineering and Management Research, 14(3), 301–312.
11. Agrawal, A. M., Gajula, S., Shinde, R. P., Shah, H., & Ghosh, H. (2025, July). Machine Translation for Long Sequences with Enhanced Attention Mechanisms. In 2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-6). IEEE.
12. Kalae, U. K. (2023). Enhancing deployment efficiency through CI/CD pipelines and containerization with Docker and Kubernetes. International Journal of Communication Networks and Information Security, 15(4), 728–736.
13. Poojari, R. Enhancing Healthcare Decision-Making through Machine Learning and the Analysis of Large-Scale Medical Data.
14. Mudusu, S. (2025). Health Insurance Fraud Detection: The Role Of Advanced It Systems In Preventing And Identifying Fraud. International Journal, 16(1), 3769-3777
15. D. Ifenthaler, "Cognitive Taxonomies and Digital Learning," Educational Technology Research and Development, 2021, doi:10.1007/s11423-020-09812-3
16. H. Hassan, M. Ahmed, and S. Khalid, "Clustering-Based Question Generation for E-Learning Systems," Expert Systems with Applications, 2019, doi:10.1016/j.eswa.2019.112879
17. S. K. Saha and A. K. Das, "Automated Question Paper Generation System Using Rule-Based Techniques," International Journal of Computer Applications, 2020, doi:10.5120/ijca2020912345
18. Mudusu, S. K. (2025). Data Engineering Challenges in AI-Driven Healthcare IT Systems: Navigating Real-Time Analytics and Interoperability.
19. M. Ally, "Adaptive E-Learning and Intelligent Assessment Systems," International Review of Research in Open and Distributed Learning, 2018, doi:10.19173/irrodl.v19i3.3432
20. R. K. Gupta and P. Verma, "Hybrid Machine Learning Approach for Automated Question Generation," International Journal of Advanced Computer Science and Applications, 2022, doi:10.14569/IJACSA.2022.0130456