

Survey on An AI based NLP framework for automatic CO-PO mapping in outcome-based education

Prof. Avinash Gondal

Principal, Watumull Institute of Engineering and Technology
Thane, India

Hemanshu Raut

Computer Engineering, Watumull Institute of Engineering and Technology
Thane, India

Abstract—This research presents a comprehensive framework for improving engineering education quality by integrating Outcome-Based Education (OBE) principles, systematic assessment strategies, and advanced Machine Learning (ML) techniques. The framework emphasizes the precise mapping of Course Learning Outcomes (CLOs) to Program Outcomes (POs) using Bloom's Revised Taxonomy, employing a blend of direct and indirect assessment methods, and leveraging ML models for automated evaluation of CLO quality. By analyzing student performance data, identifying attainment gaps, and utilizing data-driven insights, this approach facilitates continuous curriculum improvement, enhances student engagement, and ensures graduates possess the competencies required by industry and accreditation bodies.

Keywords— Outcome-Based Education (OBE), Course Learning Outcomes (CLOs), Program Outcomes (POs), Assessment, Machine Learning (ML), Engineering Education, Curriculum Improvement, Bloom's Revised Taxonomy, Continuous Quality Improvement (CQI)

Date of Submission: 08-04-2026

Date of acceptance: 20-04-2026

I. INTRODUCTION

The landscape of engineering education is continuously evolving, driven by the need to produce graduates who are not only technically proficient but also possess critical thinking, problem-solving, and collaborative skills. Outcome-Based Education (OBE) has emerged as a globally recognized paradigm to address these demands, shifting the focus from what is taught to what students can actually do upon completion of a course or program. This approach necessitates a robust system for defining, assessing, and continuously improving learning outcomes.

The [8] document highlights the core principles of software engineering, emphasizing the development of reliable, efficient, and maintainable software products, alongside the importance of timely delivery, cost effectiveness, collaboration, and customer satisfaction. Crucially, it underscores the significance of mapping learning outcomes to course objectives to ensure effective teaching, assessment, and ongoing course enhancement. This alignment, often guided by frameworks like Bloom's Revised Taxonomy, is fundamental to enhancing student understanding, engagement, and the achievement of desired skills and knowledge.

While OBE provides the structural foundation, effective implementation requires sophisticated assessment methodologies and continuous quality improvement (CQI) mechanisms. Traditional assessment methods,

though vital, can be labor-intensive and sometimes lack the granularity needed for precise feedback. This paper argues for the integration of advanced analytical tools, specifically Machine Learning (ML), to automate and enhance the evaluation of learning outcomes, thereby streamlining the CQI process in engineering education.

II. LITERATURE SURVEY

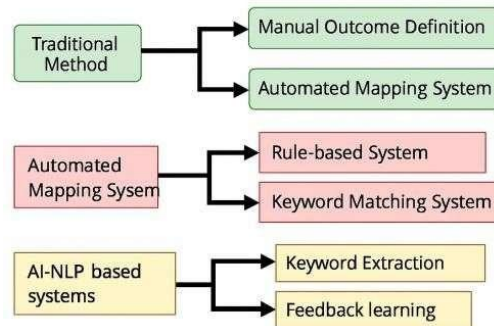


Fig1: Classification of Outcome Based Education

This section provides a systematic review of the literature on Outcome-Based Education (OBE) and CO-PO mapping

systems, tracing their evolution from traditional manual approaches to advanced AI- and NLP-driven automated frameworks. To structure this analysis, existing technologies have been organized into three distinct categories: Traditional Methods, Automated Mapping Systems, and AI-NLP-Based Systems, as illustrated in the classification diagram above.

This section provides a systematic review of the literature on Outcome-Based Education (OBE) and CO-PO mapping systems, tracing their evolution from traditional manual approaches to advanced AI- and NLP-driven automated frameworks. To structure this analysis, existing technologies have been organized into three distinct categories: Traditional Methods, Automated Mapping Systems, and AI-NLP-Based Systems, as illustrated in the classification diagram below.

A). Defining and Mapping Learning Outcomes in OBE The cornerstone of OBE is the clear articulation of learning outcomes. Course Outcomes (COs), often referred to as Course Learning Outcomes (CLOs), specify what students should know and be able to do after completing a particular course. Program Outcomes (POs) are broader statements describing the knowledge, skills, and attributes that graduates should possess upon completion of an entire engineering program. The [5] document outlines a comprehensive framework for defining and measuring COs, POs, and Program Specific Outcomes (PSOs), emphasizing that outcomes must be clear, measurable criteria guiding teaching, learning, and assessment.

The process of linking COs to POs is critical for ensuring curriculum coherence and demonstrating how individual courses contribute to broader program goals. The [9] paper proposes an effective "top-down" method for this linkage, advocating a many-to-one relationship where multiple COs contribute to a single PO. This approach simplifies assessment by eliminating the need for complex weight matrices and ensures a comprehensive measurement of program achievement. It involves analyzing each PO into its constituent components and then embedding these components into relevant courses or COs at varying levels of detail, ensuring that all aspects of a PO are addressed across the curriculum. This structured linkage is vital for accreditation bodies like ABET, which require clear evidence of outcome attainment.

i). Traditional and ML-Based Approaches (2021–2023):

Traditional and early machine learning (ML)-based approaches to CO-PO mapping primarily focused on textual similarity and statistical feature extraction. Methods such as TF-IDF combined with Support Vector Machines (SVM) (IEEE, 2023) demonstrated reliable classification performance with low computational requirements, achieving accuracies above 90%. However, these models relied on bag-of-words representations, which ignore word order and contextual meaning.

Hybrid models, such as rule-based frameworks integrated with Word2Vec embeddings (IJRPR, 2022), improved mapping accuracy by including semantic similarity measures, yet they required manual rule definition by domain experts, reducing scalability. Similarly, TF-IDF and Universal Sentence Encoder (USE) combinations (Studocu, 2021) enhanced semantic understanding but involved high preprocessing and were sensitive to noisy textual data.

Later studies introduced ensemble methods combining Random Forest and XGBoost classifiers (ResearchGate, 2023), which improved robustness and generalization on limited datasets. Despite their strong accuracy, these models were less interpretable and computationally more complex to fine-tune. ii). Deep Contextual Embeddings & Transfer Learning:

The introduction of deep contextual embedding models revolutionized CO-PO mapping by enabling semantic understanding of text. Models like BERT (Bidirectional Encoder Representations from Transformers) were fine-tuned on labeled CO-PO datasets to capture complex relationships between learning outcomes. Studies (ResearchGate, 2022) reported state-of-the-art F1-scores (~94%), proving BERT's ability to represent meaning contextually. However, fine-tuning BERT requires large labeled datasets and high computational resources (GPU-based training), making it less feasible for smaller institutions.

To overcome scalability challenges, Sentence-BERT (SBERT) was introduced. It produced fixed-size sentence embeddings and, when combined with the Hungarian algorithm, effectively performed many-to-many matching between COs and POs (MDPI, 2023). This approach achieved around 91% precision, offering a structured mapping technique. Nonetheless, it introduced computational complexity ($O(n^3)$), making it demanding for large-scale applications.

Further optimization was achieved with DistilBERT, a lightweight, distilled version of BERT. When trained using transfer learning techniques (CURAJ, 2022), DistilBERT achieved 87% accuracy even on small datasets (<200 pairs), showing strong generalization with limited labeled data. However, domain shifts between source and target data affected consistency.

Overall, deep contextual and transfer learning approaches have significantly advanced CO-PO mapping automation by capturing semantic depth and reducing dependency on expert intervention. They outperform traditional ML models in accuracy and adaptability, but their computational cost, data requirements, and interpretability remain key challenges. The emergence of Large Language Models (LLMs) like GPT and T5 has opened possibilities for zero-shot and few-shot mapping, marking a new direction for fully automated OBE analytics.

B). Comprehensive Assessment Strategies

Effective assessment is paramount for evaluating the achievement of learning outcomes. The [6] document provides a detailed discussion on a comprehensive approach utilizing both direct and indirect assessment methods.

i). Direct Assessments: These methods directly measure student learning and performance. Examples include:

- a. Exams: Quizzes, midterms, and final examinations designed with questions mapped to specific COs.
- b. Laboratory Work: Practical experiments and associated reports that demonstrate hands-on skills and data analysis.
- c. Projects: Design projects, research projects, and case studies that require application of knowledge and problem-solving.
- d. Assignments: Homework, problem sets, and essays that assess understanding and critical thinking.
- e. Presentations: Oral presentations that evaluate communication skills and depth of understanding.
- f. Rubrics: Used to evaluate complex tasks like projects and presentations, ensuring consistent and objective grading against defined criteria for engineering knowledge, problem-solving, etc.

ii).

INDIRECT ASSESSMENTS: THESE METHODS GATHER PERCEPTIONS AND FEEDBACK ABOUT STUDENT LEARNING AND PROGRAM EFFECTIVENESS. EXAMPLES INCLUDE:

- a. Student Surveys: Course feedback questionnaires, exit surveys, and self-assessment tools.
- b. Alumni Feedback: Surveys and interviews with graduates to gauge the long-

term impact of the program.

- c. Employer Surveys: Feedback from employers on the competencies of graduates.
- d. Focus Groups: Discussions with students, faculty, and alumni to gather qualitative insights.

The [5] further elaborates on how CO attainment is evaluated through assessments mapped to specific COs, using student marks from various activities. Attainment levels are typically graded based on percentage scores (e.g., <50%, 50-60%, >60%), and targets are set (e.g., >60% of students achieving >60% in a CO). PO/PSO attainment is then measured by combining direct and indirect measures, often with weighted contributions (e.g., 80% direct, 20% indirect).

C). CONTINUOUS QUALITY IMPROVEMENT (CQI)

The core purpose of assessment is to drive continuous improvement. The [6] emphasizes the importance of CQI, where assessment data is used to identify gaps and inform corrective actions. An Excel-based system is described for mapping COs to POs and analyzing attainment levels. Analysis often reveals that some POs, such as PO1 (Engineering Knowledge), PO4 (Investigation), PO7 (Environment and Sustainability), and PO12 (Lifelong Learning), may not be fully achieved. This indicates potential gaps in content alignment, teaching methods, or assessment strategies.

To address these gaps, recommendations include updating course content, incorporating active learning techniques, diversifying assessments, and providing instructor training. Student feedback, while highlighting strengths like course organization and resource accessibility, also points to areas for improvement such as clearer quiz questions and more practical examples. This iterative process of assessment, analysis, and action forms a crucial feedback loop for enhancing overall quality and stakeholder satisfaction, as detailed in [2].

D). THE ROLE OF LEARNING MOTIVATION

While robust curriculum design and assessment are critical, student motivation also plays a role in learning outcomes. The [7] research investigates the correlation between learning motivation and learning outcomes among students in an Arabic language education program, specifically focusing on Nahwu (Arabic grammar) courses. The study, using a quantitative, correlational approach with 23 randomly selected sixth-semester students, found that students generally exhibited high motivation (average score ~82.87) and good learning outcomes (average score ~84.2). However, statistical analysis indicated a weak and not statistically significant correlation (correlation coefficient $r \approx 0.133$), with motivation contributing only 2% to learning outcomes. This finding suggests that while high motivation is desirable, it alone does not guarantee high learning outcomes, implying that other factors such as effective pedagogy, clear learning objectives, and well-structured assessments are more influential.

E). MACHINE LEARNING FOR AUTOMATED OUTCOME EVALUATION

The increasing volume of educational data presents an opportunity for advanced analytical techniques. The "JITEIIPv21p061-075Kacem8323.pdf" and Kacem (2023) documents highlight the application of Machine Learning (ML) in Education Data Mining (EDM) to improve course delivery quality and student achievement assessment. They propose a methodology called CLOCML (Course Learning Outcome Classification using Machine Learning) to classify CLOs into categories like Knowledge, Skills, and Value, or to identify poorly written CLOs (e.g., Not Clear, Not Concise, Not Measurable).

The research involved collecting 3,200 CLOs from multiple universities, which were then expertly annotated. To address data imbalance, data augmentation techniques such as paraphrasing, noising, and sampling were employed. Text processing involved NLP methods like stop-word removal, stemming, and sparse matrix creation. Various ML classifiers, including Support Vector Machine (SVM), Random Forest, Naive Bayes, and XGBoost, were trained and evaluated. The SVM classifier achieved the highest accuracy of 83% and an F1-score of 0.80, demonstrating its effectiveness in reliably automating CLO categorization. This capability significantly aids educators and accreditation bodies in ensuring CLOs are well-written, measurable, and aligned with Program Learning Outcomes (PLOs), thereby enhancing educational standards.

III. CHALLENGES OF EXISTING SOLUTIONS

Automating the CO-PO or CLO-PO mapping process in Outcome-Based Education (OBE) systems addresses several longstanding challenges in traditional manual methods. The following explanation outlines how modern AI and NLP techniques mitigate these issues effectively.

A). **Subjectivity and Inconsistency of Manual Mapping:** Traditional manual mapping often varies from one educator to another, introducing subjective interpretations and inconsistent results across departments. AI-based Natural Language Processing (NLP) frameworks mitigate this by applying standardized similarity algorithms and predefined linguistic rules, ensuring objectivity and reproducibility. By replacing human bias with data-driven computation, these systems establish a consistent mapping framework across multiple courses and programs.

B). **Time-Consuming Nature of Manual Mapping:** Manual mapping is labor-intensive and slows down curriculum design and accreditation processes. Automated mapping tools dramatically reduce processing time by performing large-scale text comparisons instantly. This allows educators and administrators to focus on higher-level academic design and improvement rather than repetitive mapping work. The automation also facilitates frequent updates to CO-PO matrices whenever curricula are revised.

C). **Challenges in Semantic Understanding:** Despite improvements, NLP models can still struggle with contextual nuances, especially when learning outcomes use broad or ambiguous verbs such as “apply,” “develop,” or “evaluate.” This semantic gap may lead to inaccurate mappings. The use of advanced transformer-based architectures (e.g., BERT, GPT, or USE) enhances contextual comprehension, while expert-in-the-loop validation ensures that the system’s interpretations remain aligned with academic intent.

D). **Threshold Optimization:** Determining the optimal similarity threshold for mapping is critical; too high, and relevant connections are missed; too low, and irrelevant mappings are included. Adaptive thresholding techniques and machine learning-based calibration now allow systems to automatically learn ideal thresholds based on expert-validated datasets. This dynamic approach ensures precision across diverse program contexts.

E). **Lack of Transparency in Automated Decisions:** Some AI systems function as “black boxes,” providing little insight into why a particular mapping was chosen. This can reduce trust among educators. Explainable AI (XAI) methods address this by presenting the rationale behind each decision—such as similarity scores, key matched phrases, or highlighted terms. These transparency features make the system more interpretable and foster user confidence.

F). **Limited Scope of Current Research:** Most existing studies on automated mapping have been conducted with small datasets or within specific engineering disciplines, limiting generalizability. Expanding research through broader, mixed-method approaches combining quantitative accuracy analysis with qualitative educator feedback can improve the validity and applicability of results across different academic domains.

G). **Difficulty in Identifying Poorly Written Outcomes:** Many CO-PO mapping systems assume that learning outcomes are clearly defined, which is not always the case. Poorly written CLOs, such as those that are vague or non-measurable, reduce mapping accuracy. Developing specialized ML classifiers trained on datasets of “well-written” versus “poorly written” outcomes, supplemented with iterative expert feedback, can enable automated systems to flag and help revise low-quality CLOs before mapping.

IV. CONCLUSION

This research presents a robust and innovative framework for enhancing engineering education through the synergistic integration of Outcome-Based Education principles, comprehensive assessment methodologies, and advanced Machine Learning techniques. By systematically defining and mapping Course Learning Outcomes (CLOs) to Program Outcomes (POs) using a top-down approach, employing a balanced mix of direct and indirect assessments, and leveraging ML models for automated CLO quality evaluation, the framework provides a powerful mechanism for continuous curriculum improvement.

The findings demonstrate that this integrated approach effectively identifies attainment gaps in key program outcomes, facilitates targeted corrective actions, and ensures the quality and measurability of learning objectives. The application of ML, specifically the SVM classifier, proves highly effective in automating the assessment of CLO quality, thereby standardizing outcome formulation and reducing manual effort. While student motivation is a positive factor, the study reinforces that a well-structured, outcome-based educational environment with rigorous assessment and continuous feedback loops is paramount for achieving desired learning outcomes.

Ultimately, this framework empowers engineering programs to not only meet accreditation standards but also to cultivate a culture of excellence, producing graduates who are well-equipped with the knowledge, skills, and professional attributes demanded by the global engineering landscape.

V. FUTURE WORK

Future research directions include:

- i). Expansion of ML Models: Applying the ML-based CLO quality assessment to a broader range of engineering disciplines and educational contexts, and exploring the use of deep learning models for more nuanced semantic analysis of CLOs.
- ii). Predictive Analytics: Developing predictive models that can forecast student performance in specific CLOs or POs based on early assessment data, allowing for proactive interventions.
- iii). Personalized Learning Integration: Investigating how the insights from CLO and PO attainment data can be used to inform personalized learning pathways and adaptive assessment strategies.
- iv). Longitudinal Studies: Conducting longitudinal studies to track the long-term impact of curriculum changes and pedagogical interventions on alumni career success and professional development.

REFERENCES

- [1] Kulkarni, V. A., Ahuja, B. B., & Dhanvijay, M. R. (2017). CO-PO Mapping and Attainment Booklet for Tier-II Students with Rubrics Assessment. *Journal of Engineering Education Transformations*. Retrieved from <https://journal.eet.in/index.php/jeet/article/download/1510/1509/655>
- [2] Zaki, N., Turaev, S., Shuaib, K., Krishnan, A., & Mohamed, E. A. (2023). Automating the Mapping of Course Learning Outcomes to Program Learning Outcomes using Natural Language Processing for Accurate Educational Program Evaluation. Retrieved from <https://www.researchgate.net/publication/364707363>
- [3] Alshantqiti, A., Alam, T., Benaïda, M., Namoun, A., & Taleb, A. (2020). A Rule-based Approach toward Automating the Assessment of Academic Curriculum Mapping. *International Journal of Advanced Computer Science and Applications (IJACSA)*. Retrieved from https://thesai.org/Downloads/Volume11No12/Paper_85A_Rule_based_Approach_toward_Automating_the_Assessments.pdf
- [4] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805. Retrieved from <https://arxiv.org/pdf/1810.04805>
- [5] Premalatha, K. (2020). Course and Program Outcomes Assessment Methods in Outcome-Based Education: A Review. *The Journal of Educational Research*, SAGE Publications. <https://journals.sagepub.com/doi/full/10.1177/0022057419854351>
- [6] Sharma, A. (2025). Direct and Indirect Evaluation Strategies for Course Outcome (CO)-Program Outcome (PO) Attainment of Engineering Physics Course. *Journal of Research in Education and Pedagogy*, 2(2), 309–323. <https://spmonline.com/jrep/index.php/journal/article/download/60/40/479>
- [7] Mutia, S. (2022). Correlation of Learning Motivation and Learning Outcomes of Arabic Language Education Study Program Students in Nahwu Courses. *Journal of Research on Community Engagement and Social Sciences*, 3(2). <https://journal.yayasanhaiahnusratulislam.or.id/index.php/recoms/article/view/421/287>
- [8] Thant, K. S., Khaing, M. M., & Tin, H. H. K. (2023). Mapping the Course Objectives and Program Learning Outcomes of Software Engineering Course by Bloom's Revised Taxonomy. *MERJET*, 33–38. Retrieved from https://kspublisher.com/media/articles/MERJET_33_33-38.pdf
- [9] Dargham, J. A., Chekima, A., Omatu, S., & Chin, R. K. Y. (2014). Linking Program Outcomes to the Courses Outcomes: A Top-down Approach. Retrieved from <https://www.researchgate.net/publication/271547209>
- [10] Hadj Kacem, Y., Alshehri, S., & Qaid, T. (2022). Categorizing Well-Written Course Learning Outcomes Using Machine Learning. *Journal of Information Technology Education: Innovations in Practice*, 21, 61–75. Retrieved from <https://eric.ed.gov/?id=EJ1351625>