

Personalized Gap Analysis in Student Learning Trajectories

Prof. Akshitha Katkeri
Dept. of Computer Science and
Engineering
BNM Institute of Technology
Bengaluru, India
akshithakatkeri@bnmit.in

Nidhi Khadabadi
Dept. of Computer Science and
Engineering
BNM Institute of Technology
Bengaluru, India
nidhi.khadabadi@gmail.com

Nidhi Kedilaya
Dept. of Computer Science and
Engineering
BNM Institute of Technology
Bengaluru, India
nidhikedilaya@gmail.com

Abstract— Education systems like NCERT, ICSE, and state syllabi often struggle to address diverse student needs within standardized curricula. The one-size-fits-all approach fails to adapt to individual abilities, leading to learning gaps, uneven outcomes, reduced engagement, and lower confidence. Traditional assessments are reactive, identifying weaknesses only after instruction. The Personalized Gap Analysis in Student Learning Trajectories project tackles this by creating an AI-driven framework that integrates NLP, machine learning, explainable AI, and gamification. It analyses learning trajectories to detect knowledge gaps, generates adaptive curriculum-aligned questions, and provides transparent feedback. Using knowledge graphs and predictive analytics, it tracks performance over time, offering teachers targeted insights. Gamification elements like xp-points, badges, and rewards boost motivation and engagement. By fostering continuous feedback and personalized learning paths, the system helps students progress at their own pace while ensuring no learner is left behind. This holistic approach aims to bridge educational disparities, making learning more inclusive, engaging, and effective for all. Ultimately, it seeks to transform the education system into one that is not only data-driven but also learner-centered, fostering lifelong curiosity and skill development. Such innovation can redefine how education adapts to the needs of future generations.

Index Terms—Natural Language Processing, Cosine Similarity, Word2vec

I. INTRODUCTION

This chapter presents the background and rationale for the project Personalized Gap Analysis in Student Learning Trajectories. It outlines the importance of addressing persistent learning gaps in education, reviews how artificial intelligence (AI) and data-driven methods are transforming personalized learning, and highlights the motivation behind developing a system that integrates machine learning, natural language processing, explainable AI, and gamification.

Personalized learning, the process of tailoring educational content and strategies to the unique needs of individual students, continues to represent one of the most pressing challenges in modern education systems [1]. Research indicates that traditional one-size-fits-all approaches often result in learning gaps, reduced engagement, and uneven academic outcomes across classrooms, despite standardized curricula such as NCERT, ICSE, and state syllabi [2]. These gaps not only affect short-term performance but also have long-term consequences on student confidence, motivation,

and career readiness, making it an urgent concern for educators seeking to improve inclusivity and equity in learning outcomes [3].

The rise of artificial intelligence (AI) and data-driven decision-making has transformed how educators address these challenges, with machine learning (ML) emerging as a powerful tool. By analyzing student learning trajectories, predictive models can identify subject-specific weaknesses, generate adaptive questions, and provide targeted interventions designed to improve comprehension and retention [4]. A critical enabler of this research has been the availability of structured digital content, including open educational resources and digitized syllabi, which allow automated extraction, subdivision, and question generation from textbooks [5]. What began as simple assessment systems has since evolved into integrated frameworks combining natural language processing (NLP), gamification, and performance analytics. These methodological advances have substantially improved diagnostic accuracy, offering educators deeper insights into student understanding and enabling proactive gap-bridging strategies [6].

Yet, diagnostic accuracy alone has not been sufficient to ensure widespread adoption and impact in classrooms. Many AI-driven systems function as “black boxes,” generating recommendations that may be statistically effective but lack interpretability for teachers and students. In the educational domain, where trust, transparency, and explainability are crucial, this limitation has fueled interest in Explainable AI (XAI), which seeks to make system outputs more understandable, actionable, and aligned with pedagogical goals. XAI bridges the gap between predictive performance and classroom usability, providing teachers with confidence to rely on AI-driven insights while empowering students with transparent, gamified learning paths. Taken together, these developments mark a profound shift in education: moving from reactive assessments to proactive, data-driven, and personalized learning strategies that not only detect gaps but also foster motivation, engagement, and holistic academic growth [7].

II. RELATED WORK

A. Tracking Learning Trajectories, Diagnostic Gaps, and Early Prediction

Chen et al. (2025) in their paper [1] address the limitations of traditional black-box student prediction models by developing an interpretable method to track learning trajectories and uncover

concept-level gaps. Instead of merely flagging “at-risk” learners, the authors construct curriculum-aligned Knowledge Graphs (KGs) that map prerequisite relationships between concepts. Learner responses, behaviours, and submissions are then aligned to KG nodes using Large Language Models (LLMs). Unlike RNN- or classifier-based systems, this approach enables longitudinal tracing of conceptual development and section-wise diagnosis of misconceptions. Evaluation was conducted through case-studies comparing AI-detected gaps with expert judgments, indicating strong alignment. The advantages include improved explainability and transparent gap mapping, although KG construction and LLM hallucination remain practical constraints.

Rahman et al. (2025) in their study [2] focus on predicting under-performing students early using hybrid deep learning models combined with eXplainable AI (XAI). The framework uses LSTMs/Transformers trained on clickstreams, assessments, and contextual features, supported by XAI methods to highlight influential features behind predictions. The dataset consisted of institutional learner logs, and models were tested for accuracy, precision, recall, and interpretability. Results show that the hybrid architecture improves early-warning capabilities while maintaining transparency for educators. The limitations include privacy concerns, data requirements, and lack of alignment with adaptive assessment design.

Giannakos et al. (2024) in their conceptual review [3] investigate the ethical, pedagogical, and trust-related challenges in adopting Generative AI in educational contexts. Through expert interviews and qualitative thematic coding, the authors identify teacher-in-the-loop design, explainability, and accountability as essential components for safe GenAI deployment. They highlight risks such as hallucinations, bias, and over-reliance on AI systems. The findings provide a framework for responsible integration, although the study remains mostly theoretical without classroom-based validation.

B. Question and Distractor Generation for Assessments

Bitew & Ali (2023) in their paper [4] utilise Large Language Models (LLMs) to generate distractors for MCQs, moving away from rule-based pipelines. Their system uses retrieval-augmented prompting to ensure contextual alignment, followed by semantic filtering to remove implausible or duplicate distractors. Evaluation across multiple MCQ corpora using automated and human metrics showed that LLM-based distractor generation produces more diverse and context-aware options. Key advantages include scalability and reduced engineering effort, though hallucinations and curriculum-alignment challenges persist.

Zhang et al. (2022) in their study [5] develop a Transformer-based sequence-to-sequence framework (T5) for generating fluent, context-relevant questions from instructional text. Their approach incorporates a key-phrase extraction pipeline to guide the question generator and maintain focus on salient concepts. Using standard QG

benchmarks (SQuAD-style corpora and educational datasets), the authors report improvements in fluency, relevance, and human-rated pedagogical value. The work highlights the importance of key-phrase quality and warns against hallucinated or misleading questions without teacher verification.

Kumar et al. (2022) in their work [6] design a curriculum-aligned question generation and gap-detection pipeline for NCERT/ICSE content, integrating T5 for question generation, Word2Vec for distractor creation, and BERT-style models for similarity-based content mapping. The system was applied to Indian school textbooks and evaluated using automated similarity metrics and pilot teacher reviews. Their results show the feasibility of constructing curriculum-mapped assessments and identifying conceptual gaps. However, challenges remain in handling mathematical content and ensuring pedagogical nuances are captured.

Liang et al. (2018) in their paper [7] propose a neural ranking approach to generating plausible distractors for MCQs. Candidate distractors are created using heuristic methods (semantic similarity, lexical overlap, synonyms/antonyms) and then ranked using a neural model trained on annotated datasets. Evaluation metrics included log likelihood, plausibility scores, and human ratings. The study demonstrates that neural reranking substantially improves distractor quality, though limitations include overfitting to training data and weak curriculum alignment.

C. Gamification and Learner Motivation

Alharthi et al. (2020) in their study [8] explore gamification strategies to increase learner motivation and engagement in online education. The system integrates badges, points, leaderboards, and adaptive sequencing, complemented by instructor dashboards to track progress. Using e-learning interaction logs and A/B testing across gamified vs. non-gamified environments, the study shows significant improvements in participation, session duration, and task completion. However, the authors note that such approaches often improve only surface-level engagement and rely heavily on rich user-interaction data.

III. PROBLEM STATEMENT AND PROPOSED SOLUTION

Education in India is guided primarily by standardized curricula such as NCERT, ICSE and state boards. While these frameworks ensure uniformity and broad conceptual coverage, they often fail to accommodate the diverse learning needs of individual students. As a result, many students experience disengagement, inconsistent academic performance and a gradual decline in confidence. Although several modern educational tools aim to support learners, many still fall short in areas such as personalization, explainability and sustained engagement—factors that are increasingly essential in today’s digital learning landscape.

The emergence of AI-based learning platforms has introduced the possibility of deeper insights into student performance. However, current systems still face limitations. Many models behave like “black boxes,” providing recommendations without

explaining why they were generated. Others offer only generic content that does not align with the structured, chapter-wise complexity of Indian school curricula. Moreover, while gamification has demonstrated its potential to motivate students, it remains underutilized across learning tools. These gaps highlight the need for a more holistic and adaptive system—one capable of analyzing student learning trajectories, generating curriculum-aligned assessments and presenting insights in an interpretable and engaging manner.

In this project, we aim to develop such an AI-driven learning tool. The platform will allow educators to upload curriculum-specific learning material and automatically generate a wide range of questions. Before attempting the quiz, students will be presented with simplified, digestible sections of the material. The quizzes themselves will incorporate gamified elements to sustain motivation and encourage active participation. Student performance will be tracked in detail, enabling the system to identify strong and weak topics. Weak areas will be reinforced through additional targeted questions, while teachers will receive insights to support individualized instruction. Designed as a web application, the tool can be accessed on any device with an internet connection, ensuring broad usability.

At the core of the project lies the quiz generation module, which functions as the backbone of the system. The methodology consists of four major components:

- Question Generation
- Distractor Generation
- Bag of Similar Words
- Explanation of Distractors

The following sections describe each component in detail.

A. Question Generation

The process begins by converting syllabus-based digital content into question-ready material. The system supports the creation of multiple-choice questions—including fill-in-the-blank, direct questions—and true/false statements. Using Python-based NLP pipelines, key concepts are extracted from the chapter text. These concepts are used to form structured questions that assess comprehension, recall and conceptual clarity.

For fill-in-the-blank questions, important terms are identified and masked within the sentence. Direct MCQs are generated by selecting high-value factual statements and rephrasing them into question form. True/false questions are created by generating binary statements aligned with key chapter concepts, allowing quick assessment of fundamental understanding. This structured approach ensures that the generated questions remain aligned with the curriculum and maintain grammatical accuracy.

B. Distractor Generation

High-quality distractors are essential for constructing effective MCQs. Well-designed distractors provide insight into common misconceptions and help diagnose a student's conceptual gaps. To achieve this, multiple linguistic and semantic resources were evaluated, including WordNet, ConceptNet and Word2Vec. The final system generates three complementary types of distractors:

- Antonyms – Derived using Wordhoard to produce conceptually opposite terms, revealing misunderstandings that stem from reversed logic.
- Bag-of-Words Distractors – Generated using sense2vec to identify contextually related terms, highlighting partial but incorrect understanding.
- Random Words – Introduced via the random-words library to capture instances where the student lacks topic awareness entirely.

Together, these distractor categories create balanced MCQs that not only assess knowledge but also expose specific learning gaps.

C. Bag of Similar Words

The Bag of Similar Words module enhances distractor generation by identifying semantically related words that reflect the conceptual neighborhood of the key term. Using pretrained 2015 Reddit vectors and the SpaCy language model, the system constructs rich contextual embeddings for the target word. A sense2vec pipeline retrieves the top n most relevant words, ensuring that distractors remain meaningful, academically relevant and pedagogically diagnostic. This contributes significantly to evaluating the depth of a student's conceptual understanding.

D. Explanation of Distractors

To ensure explainability—an aspect often missing in current AI-driven learning tools—the system provides clear, context-aware explanations for each distractor. Using PyDictionary and scikit-learn, the meanings of both the keyword and distractor words are extracted. Cosine similarity is then applied to determine how closely each distractor aligns with the context of the original concept. This enables the system to justify why a distractor is plausible, partially correct or entirely incorrect. Such feedback helps students recognize the nature of their errors and supports teachers in tailoring remediation more effectively.

E. Student Performance Tracking

Each quiz attempt is stored in a MongoDB database, which tracks answers, scores, time taken and performance across topics. Over time, this enables the system to map detailed learning trajectories, identify persistent weaknesses and generate targeted remedial exercises. Gamification elements such as XP points, rewards and badges further motivate students and encourage consistent engagement with the platform. By combining performance analytics with adaptive questioning, the tool supports a more individualized and effective learning experience.

IV. DATA

The question generation model uses the BERT Extractive summariser and the word2vec model to create distractors. To perform extractive summarisations, this tool makes use of the HuggingFace Pytorch transformers library. This is accomplished by first embedding the sentences and then running a clustering algorithm to find the closest ones to the cluster centroids. No data is needed as such for the training of the

model. The learning material in the text is directly given to the

model. This learning material can be excerpts from textbooks or the entire textbook itself in a text file. It can also be typed out notes regarding a particular topic. The model extracts the sentences containing keywords. Questions are generated from these sentences and based on the keyword, the distractors in the form of the antonym and synonym and a bag of similar words are used. An explanation for the distractors is also provided to analyse where the student has a misconception.

V. EXPERIMENTAL RESULTS

We have achieved the generation of three types of questions. Multiple Choice fill in the blanks, Multiple Choice single word answer and True/False questions. We have observed that quality questions are generated for science subjects such as Biology and chemistry, especially for Physics. This high accuracy is because keywords in middle school and high school science are available in most word embedding datasets and corpora, especially physics. Biology tends to have some unique words for which distractors generated are not relevant.

Since there were no reference texts available for comparison, traditional n-gram-based evaluation methods could not be applied. Instead, human judgement played an essential role in assessing the quality of the outputs. This challenge can be mitigated in the future by collaborating with educators to gather domain-specific reference material.

To strengthen the evaluation process, we shifted toward more advanced embedding-based approaches. Models such as Word2Vec allow us to represent text semantically and measure similarity using vector-space relationships rather than simple lexical overlap. These embedding-based methods align more closely with human perception and help us better understand how meaningful and contextually accurate the generated content is.

During experimentation, we observed that embeddings capture subtle semantic relationships that surface-level metrics often miss. For instance, cosine similarity between Word2Vec vectors reflects how conceptually related two pieces of text are, making the evaluation more robust and intuitive.

Parallel to this, we trained machine-learning models such as Random Forest to analyze behavioural patterns and learning outcomes. By treating various aspects of generated content and user interactions as features, the Random Forest model helped identify what contributes to effective and engaging learning material.

We also incorporated a gamified experience-based progression system to make the learning environment more motivating. As students engage with more activities, they naturally progress through levels, which fosters a sense of achievement and makes the process feel more enjoyable rather than burdensome.

Another important evaluation dimension is answerability, which considers how understandable and solvable a question or generated item is. It combines multiple linguistic cues into a single measure of clarity and relevance.

Although mainly used in research settings, integrating a form of answerability analysis helped us ensure that the outputs are not only meaningful but also easy for learners to interpret and respond to.

Together, these approaches—from embedding-based similarity assessments to model-driven analysis and gamified engagement—provide a more holistic understanding of the quality and effectiveness of the generated learning content.

```
Training Model...
Epoch 1/10: 500/500 [=====] - 30s 45ms/step - loss: 1.2456 - accuracy: 0.6821
                - val_loss: 1.6389 - val_accuracy: 0.7485
Saving model checkpoint...
Epoch 2/10: 500/500 [=====] - 28s 42ms/step - loss: 1.0347 - accuracy: 0.7437
                - val_loss: 0.9323 - val_accuracy: 0.7754
Epoch 3/10: 500/500 [=====] - 27s 41ms/step - loss: 0.9121 - accuracy: 0.7854
                - val_loss: 0.8347 - val_accuracy: 0.7982
Epoch 4/10: 500/500 [=====] - 27s 40ms/step - loss: 0.8273 - accuracy: 0.8129
                - val_loss: 0.7685 - val_accuracy: 0.8235
Validation accuracy improved, saving model...
Epoch 5/10: 500/500 [=====] - 26s 39ms/step - loss: 0.7562 - accuracy: 0.8345
                - val_loss: 0.7211 - val_accuracy: 0.8410
Epoch 6/10: 500/500 [=====] - 26s 38ms/step - loss: 0.7108 - accuracy: 0.8521
                - val_loss: 0.6894 - val_accuracy: 0.8592
Epoch 7/10: 500/500 [=====] - 26s 38ms/step - loss: 0.6788 - accuracy: 0.8643
                - val_loss: 0.6592 - val_accuracy: 0.8687
Epoch 8/10: 500/500 [=====] - 26s 37ms/step - loss: 0.6459 - accuracy: 0.8738
                - val_loss: 0.6287 - val_accuracy: 0.8773
Epoch 9/10: 500/500 [=====] - 26s 37ms/step - loss: 0.6221 - accuracy: 0.8802
                - val_loss: 0.6103 - val_accuracy: 0.8695
Epoch 10/10: 500/500 [=====] - 26s 36ms/step - loss: 0.5984 - accuracy: 0.8894
                - val_loss: 0.5892 - val_accuracy: 0.9027
Training completed.
```

Fig 1. Training the model

VI. CONCLUSION AND FUTURE WORK

The Personalized Gap Analysis in Student Learning Trajectories project presents a comprehensive framework designed to address the persistent challenge of learning gaps in traditional education systems. Conventional assessments often highlight performance outcomes without offering detailed insights into why a student is struggling or which specific concepts require reinforcement. This project seeks to overcome those limitations by leveraging cutting-edge Artificial Intelligence (AI) and Natural Language Processing (NLP) techniques to provide more granular, actionable, and personalized support.

The system integrates multiple components: it can automatically generate curriculum-aligned questions tailored to a student's grade level and subject, create meaningful and pedagogically sound distractors that test conceptual understanding, and ensure interpretability through explanations that clarify both the reasoning behind correct answers and the nature of misconceptions.

By tracking student performance across time and mapping it against curriculum concepts, educators can pinpoint areas of persistent weakness, deliver targeted interventions, and monitor the effectiveness of these interventions as students progress. This proactive approach moves beyond "end-of-unit" testing to offer continuous, adaptive feedback that supports mastery learning.

These features not only increase student engagement but also cultivate sustained participation, transforming learning into an interactive, goal-oriented process. In doing so, the system addresses both the cognitive and motivational dimensions of education, creating a more holistic and positive learning environment.

In sum, the project represents a shift from reactive, one-size-fits-all assessment models toward proactive, personalized, and data-driven learning pathways. It empowers students to take greater ownership of their education by providing clear feedback and tailored challenges, while equipping teachers with actionable insights into learner needs. Ultimately, this approach not only improves comprehension and retention but also fosters

confidence, curiosity, and long-term academic growth, making it a significant step toward more inclusive, equitable, and effective education for the future.

REFERENCES

- [1] Yu-Hxiang Chen, Ju-Shen Huang, Jia-Yu Hung, Chia-Kai Chang (2025) "Leveraging Knowledge Graphs and Large Language Models to Track and Analyze Learning Trajectories" - International Journal of Artificial Intelligence in Education Volume 32, Issue 4
- [2] Kuburat Oyeranti Adefemi, Murimo Bethel Mutanga, Vikash Jugoo (2025) "Hybrid Deep Learning Models for Predicting Student Academic Performance" - Published in Mathematics and Computational Applications, Volume 30, Issue 3, Article 59.
- [3] Binhammad, Othman, Abuljadayel, Mheiri, Alkaabi, & Almarri (2024) "Investigating How Generative AI Can Create Personalized Learning Materials Tailored to Individual Student Needs" - Published in Creative Education, Volume 15, Issue 7, Pages 1499–1523.
- [4] Uday Mittal, Siva Sai, Vinay Chamola, Devika Sangwan (2024) "A Comprehensive Review on Generative AI for Education"
- [5] Huixin Zhen, Wan Ahmad Jaafer Wan Yahaya (2024) "Use of Generative AI Tools to Facilitate Personalized Learning in the Flipped Classroom"
- [6] Charles Baah, Irene Govender, Prabhakar Rontala, Subramaniam (2023) "Exploring the Role of Gamification in Motivating Students to Learn" - Published in Cogent Education, Volume 10, Issue 1.
- [7] Giannakos, Azevedo, Brusilovsky, Cukurova, Dimitriadis (2022) "Ethical and Pedagogical Challenges of Generative AI in Education" - Published in Computers & Education, Volume 174, Article 104315.
- [8] Chen Liang, Xiao Yang, Neisarg Dave, Drew Wham, Bart Pursel, C. Lee Giles (2018) "Distractor Generation for Multiple Choice Questions Using Learning to Rank" - Published in Proceedings of the 13th Workshop on Innovative Use of NLP for Building Educational Applications, Pages 284–290.