AUTOEVAL - AUTOMATED CONTINUOUS EVALUATION FOR STUDENTS USING GENERATED FLASHCARDS, QUIZZES AND CUMULATIVE FEEDBACK BASED ON SYLLABUS AND SCHEDULE

Project Reference No.: *49S_BE_0838*

College : New Horizon College of Engineering, Bengaluru
Branch : Department of Computer Science and Engineering
Guide(s) : Prof. Mradula
Student(s) : Ms. Pallavishree S
Ms. Pavani A
Ms. Prabha Praveen Kini
Ms. Sharada Kumari V

Keywords:

Automated Evaluation, Natural Language Processing (NLP), Intelligent Content Generation, Adaptive Learning, Continuous Assessment, Schedule-Aware Feedback, Machine Learning, Flashcard Generation, Syllabus-Aligned Tools.

Introduction:

AutoEval is an innovative educational technology system designed to transform traditional student evaluation methods by implementing automated, continuous assessment mechanisms. The system leverages artificial intelligence and intelligent content generation to create personalized learning experiences through dynamically generated flashcards, adaptive quizzes, and comprehensive cumulative feedback systems. Currently, student assessment often relies on end-of-term examinations and manually created assignments, which creates several critical limitations.

First, the time-intensive nature of creating assessment materials places a substantial burden on educators, limiting their ability to provide regular feedback. Teachers spend significant time typing out quizzes and study materials—time that could be better spent on direct teaching. Second, traditional methods suffer from a gap between learning and evaluation; students typically receive feedback only after major assessments, missing opportunities for immediate course correction. Finally, there is often a disconnect between the academic schedule, syllabus progression, and evaluation content. Students may be assessed on topics not yet covered, or evaluation may occur too late to be effective for learning reinforcement. AutoEval solves these issues by providing intelligent, syllabus-aligned, and schedule-aware tools that automatically generate content based on the academic calendar.

Objectives:

1. Automated Content Extraction: To develop a robust NLP-based system capable of parsing complex academic documents (PDFs, DOCX, and slide decks) to identify and extract core concepts, definitions, and key learning points.
2. Dynamic Flashcard Generation: To implement an intelligent algorithm that transforms extracted concepts into personalized digital flashcards, facilitating active recall and spaced repetition for students.
3. Schedule-Aware Quiz Generation: To synchronize the assessment engine with the institutional academic calendar, ensuring that quizzes are generated and delivered precisely when the corresponding topics are covered in class.
4. Comprehensive Cumulative Feedback: To design a data-driven dashboard that aggregates student performance over time, providing longitudinal insights into learning trends rather than isolated test scores.
5. Educator-in-the-Loop Integration: To build a specialized interface for teachers that allows them to review, edit, and validate AI-generated content, ensuring that the automated outputs maintain high academic standards.
6. Adaptive Learning Reinforcement: To create a feedback loop where the system identifies a student's weak areas through quiz results and automatically prioritizes related flashcards for future study sessions.
7. Scalable Architecture: To develop a modular, API-first framework using Django and React that can be easily integrated into existing school and university Learning Management Systems (LMS).

Methodology:

The development of AutoEval follows a modular, pipeline-based approach that combines Natural Language Processing (NLP), Machine Learning (ML), and a modern web stack. The methodology is structured into five distinct phases to ensure the system is "schedule-aware" and "syllabus-aligned."

Phase 1: Data Acquisition and Pre-processing
The system begins by ingesting academic documents in various formats (PDF, DOCX, PPTX). Using libraries such as PyPDF2 and python-docx, the text is extracted and cleaned. Pre-processing involves noise reduction (removing headers/footers), tokenization, and sentence segmentation. For syllabus documents, a specialized parser identifies the structural hierarchy (Modules, Units, and Learning Objectives).

Phase 2: NLP Content Extraction Pipeline
This core phase utilizes Natural Language Processing to transform raw text into structured knowledge:
Keyword Extraction: Identifying high-frequency academic terms using TF-IDF and Rake-NLTK.
Named Entity Recognition (NER): Using SpaCy to identify specific dates, definitions, and technical concepts.
Semantic Analysis: Understanding the context of sentences to ensure that generated questions address core concepts rather than trivial data points.

Phase 3: Automated Content Generation Engine
The system employs two distinct generative models:
Flashcard Generator: Uses a "Definition-Concept" mapping algorithm to create active recall pairs.
Quiz Generator: Implements a Transformer-based model (such as T5 or BERT) to generate Multiple Choice Questions (MCQs) and Fill-in-the-blanks. The complexity is mapped to Bloom's Taxonomy to ensure a mix of "Remember," "Understand," and "Apply" level questions.

Phase 4: Schedule-Aware Integration & Database Modelling
To align with the academic calendar, the system utilizes a SQLite3 database where the syllabus is mapped against time-stamps. A scheduling algorithm monitors the current date and automatically triggers the release of flashcards and quizzes corresponding to the "Current Module" being taught in the classroom.

Phase 5: Evaluation and Feedback Loop
The system tracks student responses to build a Cumulative Feedback Profile. It uses statistical analysis to identify "knowledge gaps." If a student consistently fails questions on a specific topic, the Adaptive Learning module increases the frequency of related flashcards. Educators are provided with a Django-based dashboard to review AI-generated content before it is published to students.

Result and Conclusion:

The implementation of AutoEval is expected to yield a highly efficient, automated ecosystem that successfully bridges the gap between classroom teaching and student self-assessment. Based on the system design and initial testing phases, the project demonstrates that AI-driven content extraction can accurately identify core academic concepts from standard syllabi with high precision. The "schedule-aware" engine ensures that students are no longer overwhelmed by irrelevant material, instead receiving quizzes and flashcards that are perfectly synchronized with their weekly academic progress.

The primary results include:

1. Reduced Administrative Overhead: Automating the creation of quizzes and flashcards significantly reduces the time educators spend on manual content entry, allowing them to focus on high-level instructional tasks.
2. Improved Knowledge Retention: Using active recall (flashcards) and frequent, low-stakes testing (quizzes), the system promotes long-term memory consolidation as opposed to "cramming" for final exams.
3. Data-Driven Insights: The cumulative feedback system provides a clear visual map of student performance, enabling early identification of at-risk students who may need additional academic support.
4. Scalability: The modular architecture allows the platform to support a wide range of subjects, from technical engineering courses to humanities, by simply uploading different syllabus structures.

In conclusion, AutoEval represents a significant step forward in educational technology. By moving away from static, end-of-term evaluations toward a continuous, AI-assisted feedback loop, the system fosters a more responsive and personalized learning environment. It proves that integrating Natural Language Processing with academic schedules is not only feasible but essential for modernizing the digital classroom.

References:

- K. Koedinger, V. Aleven, and N. Heffernan, “Intelligent Tutoring Goes To School in the Big City,” *Proc. 13th Int. Conf. Artificial Intelligence in Education*, 1997.
- R. Ferguson, “Learning Analytics: Drivers, Developments and Challenges,” *Int. J. Technol. Enhanced Learn.*, vol. 4, no. 5/6, pp. 304–317, 2012.
- G. Siemens, “Learning Analytics: The Emergence of a Discipline,” *Amer. Behav. Scientist*, vol. 57, no. 10, pp. 1380–1400, 2013.
- H. Agrawal and H. Mavani, “Student Performance Prediction using Machine Learning,” *Int. J. Eng. Res. Technol. (IJERT)*, vol. 4, no. 3, pp. 724–727, 2015.
- J. Kim and B. Reeves, “Exploring the Effects of Content and Form on Student Engagement,” *Learn., Media Technol.*, vol. 44, no. 1, pp. 29–45, 2019.
- H. Foster, “The Impact of Formative Assessment on Student Learning Outcomes: A Meta-Analytical Review,” *Acad. Educ. Leadership J.*, vol. 28, no. S1, pp. 1–3, 2024.
- Y. Iwashita, A. Vargo, M. Iwata, and K. Kise, “Personalizing Augmented Flashcards Towards Long-Term Vocabulary Learning,” pp. 263–266, 2024, doi: 10.1145/3652920.3653045.
- P. Ana and P. T. Bukie, “Design and Implementation of Online Examination Administration System for Universities,” *Glob. J. Math. Sci.*, vol. 12, pp. 39–51, 2015, doi: 10.4314/gjmas.v12i1.16.
- M. Donmez, “AI-Based Feedback Tools in Education: A Comprehensive Bibliometric Analysis Study,” *Int. J. Assess. Tools Educ.*, vol. 11, pp. 622–646, 2024, doi: 10.21449/ijate.1467476.
- Z. Nasim, Q. Rajput, and S. Haider, “Sentiment Analysis of Student Feedback Using Machine Learning and Lexicon Based Approaches,” pp. 1–6, 2017, doi: 10.1109/ICRIIS.2017.8002475.
- B. S. Bloom, “Learning for Mastery,” *Eval. Comment*, vol. 1, no. 2, pp. 1–12, 1968.
- M. A. Hassan, U. Habiba, H. Khalid, M. Shoaib, and S. Arshad, “An Adaptive Feedback System to Improve Student Performance Based on Collaborative Behavior,” *IEEE Access*, vol. 7, pp. 107171–107178, 2019.
- W. Mutesi and J. Tumwine, “Sentiment Analysis of Student Feedback: An Implementation of a Natural Language Processing (NLP) Algorithm,” *Avance Int. Univ. Res. Repository Extension*, vol. 10, pp. 78–86, 2024.
- S. Selvakani, “Automatic Question Generation System Based Natural Language Processing Using Python,” *Int. J. Adv. Res. Comput. Sci.*, vol. 14, pp. 29–32, 2023, doi: 10.26483/ijarcs.v14i2.6961.
- M. Adarkwah, “The Power of Assessment Feedback in Teaching and Learning: A Narrative Review and Synthesis of the Literature,” *SN Social Sci.*, vol. 1, p. 75, 2021, doi: 10.1007/s43545-021-00086-w.

- T. Wu, "Fostering Inspirational Learning Through AI-Enhanced Formative Assessment," 2025, doi: 10.4018/9798-3693-6351-5.ch008.
- X. Li, H. Xu, J. Zhang, and H.-h. Chang, "Deep Reinforcement Learning for Adaptive Learning Systems," *J. Educ. Behav. Stat.*, vol. 48, 2022, doi: 10.3102/10769986221129847.
- C. Llurba, R. Palau, and J. Mogas Recalde, "Learning Analytics for Smart Classrooms," 2023, doi: 10.4018/978-1-7998-9220-5.ch103.
- Lin, M. McDaniel, and T. Miyatsu, "Effects of Flashcards on Learning Authentic Materials: The Role of Detailed Versus Conceptual Flashcards and Individual Differences in Structure-Building Ability," *J. Appl. Res. Memory Cogn.*, vol. 7, 2018, doi: 10.1016/j.jarmac.2018.05.003.
- M. El-Bishouty, T.-W. Chang, S. Graf, Kinshuk, and N.-S. Chen, "Smart e-Course Recommender Based on Learning Styles," *J. Comput. Educ.*, vol. 1, 2014, doi: 10.1007/s40692-014-0003-0.

Future Scope:

The future scope of this project includes:

1. Integration with Mainstream LMS: A primary goal for future iterations is the development of plug-and-play modules for popular Learning Management Systems such as Moodle, Canvas, and Google Classroom. This would allow institutions to adopt AutoEval without changing their existing infrastructure.
2. Support for Multi-Modal Content: While the current system focuses on text-based documents, future work will involve expanding the NLP pipeline to process educational videos and audio lectures. By using speech-to-text and visual recognition, the system could generate quizzes based on what a professor said during a live session.
3. Advanced Gamification: To further increase student engagement, future versions could incorporate competitive leaderboards, digital badges, and "streak" rewards. This would transform the evaluation process into a more interactive and motivating experience.
4. Deep Learning for Question Diversity: Exploring more advanced Generative AI models (like GPT-4 or specialized Llama variants) could allow the system to generate Higher-Order Thinking Skills (HOTS) questions. This would move beyond simple recall to testing critical thinking and problem-solving abilities.