Cureus Journal of Computer Science

Open Access Original Article

Personalized Learning Paths for Student Progression: An Artificial Intelligence-Driven Skill Roadmap With Retrieval-Augmented Generation and Large Language Models

Review began 11/20/2024
Review ended 11/14/2025
Published 11/18/2025

DOI:
https://doi.org/10.7759/s44389-024-02275-z

Arya J. Kadam [1] , Tejaswini J. Durge [1] , Chinmay Kale [1] , Prasad R. Kute [1] , Priya M. Shelke [1] , Leena Deshpande [2]

1. Information Technology, Vishwakarma Institute of Information Technology, Pune, IND 2. Computer Engineering, Vishwakarma Institute of Information Technology, Pune, IND

Corresponding author: Arya J. Kadam, arya.22210766@viit.ac.in

Abstract

This work presents an integrated web platform designed to help students self-assess their academic performance using SWOC (Strengths, Weaknesses, Opportunities, and Challenges) analysis. The proposed framework leverages machine learning algorithms to evaluate examination results, co-curricular engagement, and project participation. It provides personalized suggestions to improve students' learning profiles by identifying their strengths and weaknesses. Key features include dynamic graphics for tracking progress over time and an interactive chatbot powered by Retrieval-Augmented Generation for academic guidance and decision-making support. The platform's real-time data analysis and goal-oriented organization set it apart from traditional personalized education methods. Initial evaluations demonstrate high efficiency, responsiveness, and user satisfaction, highlighting its potential to enhance academic performance effectively.

Categories: Natural Language Processing (NLP), AI applications, Machine Learning (ML)
Keywords: llama, swoc, vector embedding, large langauge models, natural language processing, retrieval-augmented generation (rag), transformers, cosine similarity, data structures and algorithms (dsa)

Introduction

In the current quick-moving academic setting, students frequently struggle to evaluate their own academic progress and identify exact areas that need improvements [1]. Many students find it difficult to identify their strengths and weaknesses and feel stressed by the numerous academic tasks they must complete due to the lack of clear guidance or benchmarks [2]. Furthermore, students may be unaware of the most efficient methods or tools to overcome their shortcomings, which can hinder their continuous academic development. This is because in most models of learning, the feedback mechanisms and a clear pathway of development are missing, and when one cannot see the progress of their hard work and effort, they usually become demotivated [3].

One of the important components of this platform is the system of constant progress tracking and accurate visualization of changes, with which students can see their progress over time. The platform will also produce instantaneous graphical feedback in the form of graphs and chart figures that depict performance trends, enhancements, and such aspects requiring additional work. This kind of charting of development aims at acting as a motivation point, helping students to feel motivated when working hard in class, so that they can achieve the objectives set. From a teacher's perspective, students can also match current performances with previous scores, so that they have a guideline whereby they can assess their progress and even carry out self-assessment on what areas they need to concentrate on, hence guiding their research efforts.

Furthermore, traditional models of education provide students with standardized forms of testing that fail to account for individual learning potential or learning pace. Attending exams, completing projects, and participating in contests and leagues offer a general overview of performance, but they do not provide comprehensive feedback to address individual challenges in the learning process. This centralized approach overlooks students' diverse needs and hinders their development, as students cannot be expected to grow uniformly. For instance, students with varying learning abilities or learning disabilities may feel excluded from competing with their peers and may perform inadequately due to a lack of personalized attention [4].

To address these challenges, the project proposes the development of an online tool that uses machine learning and large language models (LLMs) to generate a SWOC (Strengths, Weaknesses, Opportunities, and Challenges) profile for each learner. The system will be designed to process various forms of academic

How to cite this article
Kadam A J, Durge T J, Kale C, et al. (November 18, 2025) Personalized Learning Paths for Student Progression: An Artificial Intelligence-Driven Skill Roadmap With Retrieval-Augmented Generation and Large Language Models. Cureus J Comput Sci 2 : es44389-024-02275-z. DOI https://doi.org/10.7759/s44389-024-02275-z

data such as test results, competition performance, project completion percentage, and other relevant performance indicators. Using advanced machine learning techniques, the platform can analyze each student's individual learning process and estimate the likelihood of success or difficulty in specific areas. By providing insights into students' strengths and weaknesses, this analysis will highlight areas that require attention and generate recommendations on what a student needs to do in order to achieve excellence.

It also provides customized pathways aligned with each student's educational and career goals. Following the SWOC analysis, students are presented with a clearly defined action plan outlining the specific steps they should take to enhance their performance. This plan may include recommended readings, online courses, books, and other relevant learning resources tailored to their current academic level and career aspirations. The action plan will also establish time-bound goals to ensure that students remain focused and committed to continuous improvement.

Thus, the main research question to be addressed in this paper is as follows: How can AI-driven tools, specifically, ML and Retrieval-Augmented Generation (RAG), be utilized to design personalized, competency-based learning journeys that actively respond to the academic performance of students? The work fits into the current literature on the intelligent tutoring systems, AI-assisted education, and self-assessment platforms. The proposed framework uses continuously provided data feedback and natural language interaction to deliver real-time and personalized academic insights unlike the traditional learning management systems (e.g., Moodle, Canvas), which are based on traditional evaluation models. Therefore, it builds on existing literature by integrating performance analytics with conversational AI to promote self-directed and adaptive learning.

Materials And Methods

Studies show that interactive participation enhances classroom engagement and, consequently, improves students' performance. Personalized feedback tools, such as our own, which provide instantaneous formative feedback on strengths and weaknesses, are consistent with these findings [1]. This paper focuses on how machine learning can assist in developing better approaches to delivering personalized learning. As such, by analyzing academic data, these system can provide individualized solutions, which is one of the strong suits of the SWOC analysis in our platform [2].

Predicting student learning is achievable through supervised machine learning models, which are specifically designed to predict students' academic performance based on academic indicators such as tests and assignments. Our system extends this capability by not only forecasting student performance but also providing a framework for improvement [3]. Finally, educational data mining provides novel approaches to analyze student performance, and this research provides evidence to show that the techniques can be recommended for use within our proposed platform to derive dynamic learning paths for student learning [4]. Complex systems theory applied to education notes that constant feedback enhances learning. Our chatbot, which answers questions such as 'How do I overcome my weakness?' or 'How do I begin DSA (Data Structures and Algorithms) ?', is an actualization of this feedback system [5].

To make things clearer and easier to reproduce, the machine learning implementation uses supervised models like Random Forest and Decision Tree classifiers to predict performance. We used five-fold cross-validation to find the best hyperparameters (for example, number of estimators = 100, max depth = 10), so that the model would work on other datasets as well. Mean imputation and outlier trimming were used to handle missing data. Before training, feature normalization was done, and weighted sampling was used to fix the class imbalance. To check how strong the model was, we used accuracy, precision, recall, and F1-score. These technical improvements make it easier to understand the system's suggestions and make sure that performance insights are still based on statistics.

The work examines a methodology for describing dynamic paths of learning to enhance performance with a component included in the subject through individual SWOC profiles or recommendations [6]. Tutoring systems involve artificial intelligence such that they offer help to the learners by answering questions in academic issues and offering feedback in real-time. This concept is further expanded by our chatbot system by providing tailored academic and career information [7]. The importance of predictive modeling as a factor that supports improving educational proficiency cannot be overestimated. In this way, using these models, we try to provide an individualized SWOC analysis that would contribute to academic development [8]. Hence, by applying data mining techniques on students' performance data, trends of poor performance can easily be picked from a particular area. The proposed techniques are employed by our platform to create learner-centric learning pathways [9].

As this paper shall discuss, the application of AI in learning ecosystems has the potential to offer students with personalized learning experiences. We also adopt similar AI techniques for the same purpose in our project [10]. Machine learning models, including decision trees and random forests, have been utilized to predict students' success rates, contributing to the design of our platform's SWOC analysis engine [11]. Moreover, the use of chatbots in education is growing, as these systems can answer frequently asked questions and provide real-time assistance. Our system's chatbot offers guidance on topics like overcoming weaknesses and starting DSA, thereby enhancing the personalized learning experience [12]. Feedback based on student performance data leads to improved learning outcomes. This research supports the design of our system, where academic data is used to generate insights and learning paths for students [13].

The application is integrated into a complex approach aimed at improving students' services as well as their evaluation. This website contains a number of functions; it can generate SWOC analysis, it has an interface of an interactive chatbot, and it can show analytical results using graphs and pie charts. Together, each of these components helps provide students with important information about their performance in their academic endeavors in order to help them make appropriate decisions about their learning progress.

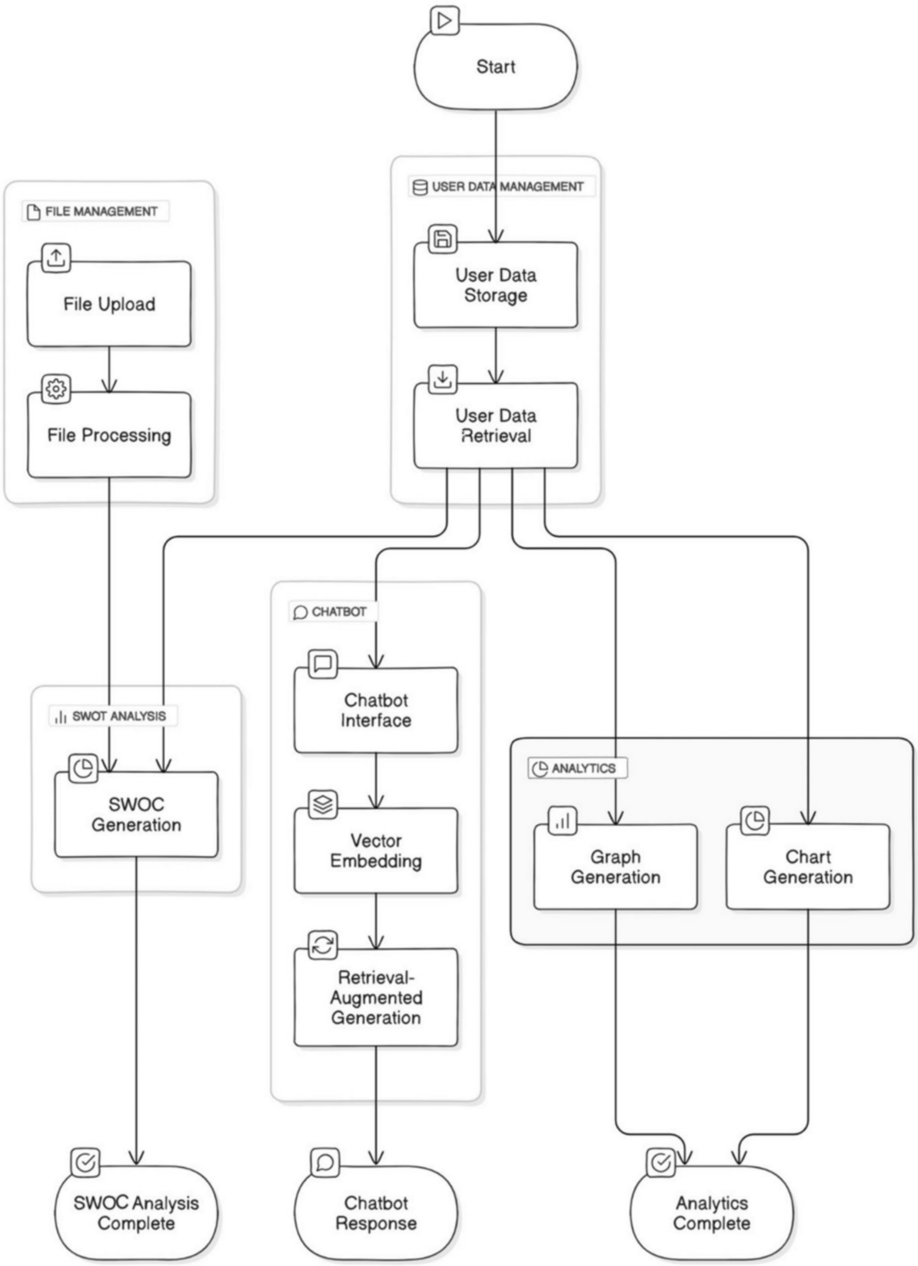

FIGURE 1: Overall Flow

SWOC = Strengths, Weaknesses, Opportunities, and Challenges

The performance of the program is described throughout this paper in a manner consistent with the overall structure of the data processing and usage illustrated in Figure 1 above that indicates how each feature of the application operates. The methodology is detailed as follows.

User data storage and retrieval

Management of users and their interactions with the application is handled through MongoDB. Records of every student enrolled in the course are securely stored and can be altered or retrieved at any time. This makes it easier for students to access their performance history whenever they need to review or evaluate their progress. Managing file uploads and processing is a critical task in many modern systems, as users can provide data in various formats. The process most often begins with the user selecting a file on their device and then uploading it to a server or application. Examples of such files include CSV, Excel, JSON, and image types.

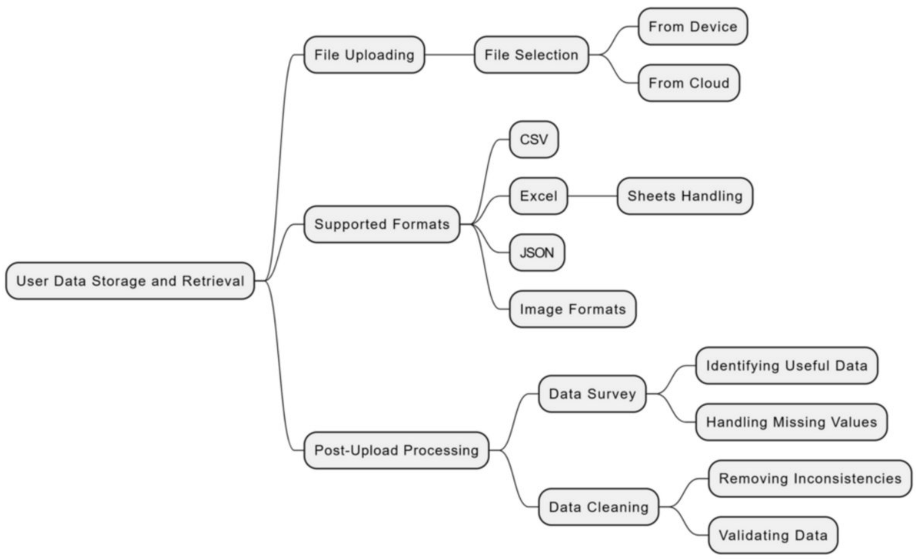

FIGURE 2: User Data Storage and Retrieval

The issue that arises is that once a file is uploaded, the application must be able to handle it well. This involves performing an initial scan of the file, identifying which data elements are useful for analysis, and, if necessary, further processing the data to make it easier to work with. For example, when working with spreadsheet files such as Excel, the application may need to read multiple sheets, manage different data formats, and address issues such as missing values or encoded data. In such cases, data cleaning becomes an essential step, during which inconsistent or invalid data is identified and removed to ensure the accuracy of subsequent analyses (see Figure 2).

Generating SWOC

Generating a SWOC analysis for students is an insightful process that leverages advanced language models, specifically LLAMA3.2:3b, and it can be implemented in real-time evaluations of learning environments (Figure 3). This model is selected because it meets the requirements for high-speed assessment without compromising the quality of the insights generated. The architecture of LLAMA3.2:3b has been designed under an educational settings, as it readily produces profile-specific feedback that reflects students' experiences.

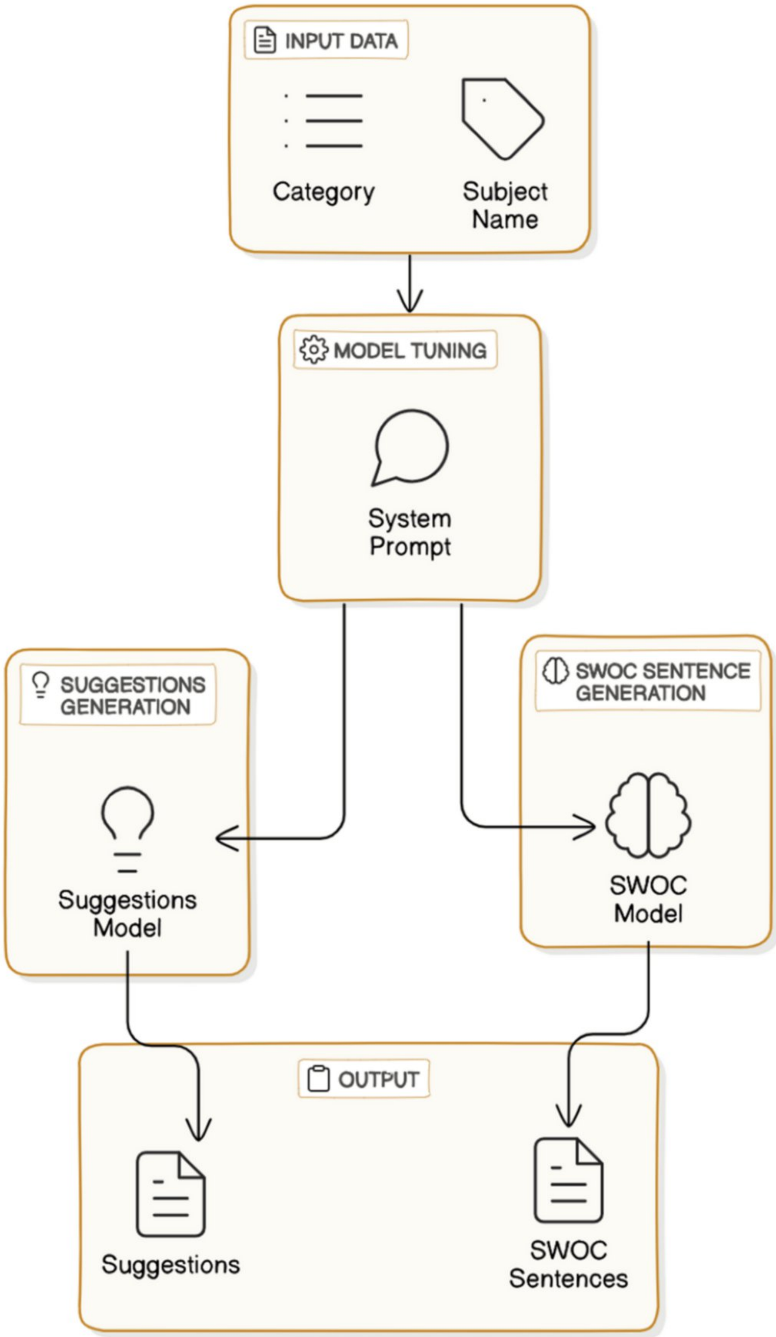

FIGURE 3: Generating SWOC

SWOC = Strengths, Weaknesses, Opportunities, and Challenges

The process starts with the organization of data collection procedure that involves information about students' performance and their activities. These are cumulative grade point average (CGPA), scores in various subjects, and amount of project work, internships, and activity in clubs. Once the data is accumulated, cutoffs are set by variance functions like means and standard deviation, allowing one to classify the performance measurement into assets and liabilities. For example, based on a threshold value, if a student's code yields a score above this value, it is considered a strong performance; the scores below this value indicate the student's areas of weakness. The use of statistical formulas like the mean (μ) and standard deviation (σ) is done for classification by creating thresholds. Refer to Equations (1) and (2), respectively.

$$\mu_p = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

$$\sigma_p = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \qquad (2)$$

Chatbot implementation

The integration of the chatbot uses the RAG model built to improve the specificity and correct contextuality of the answers. This model functions in two different phases; the first phase is the production of vector embeddings that act as replicas of the backend stored documents (Figure 4).

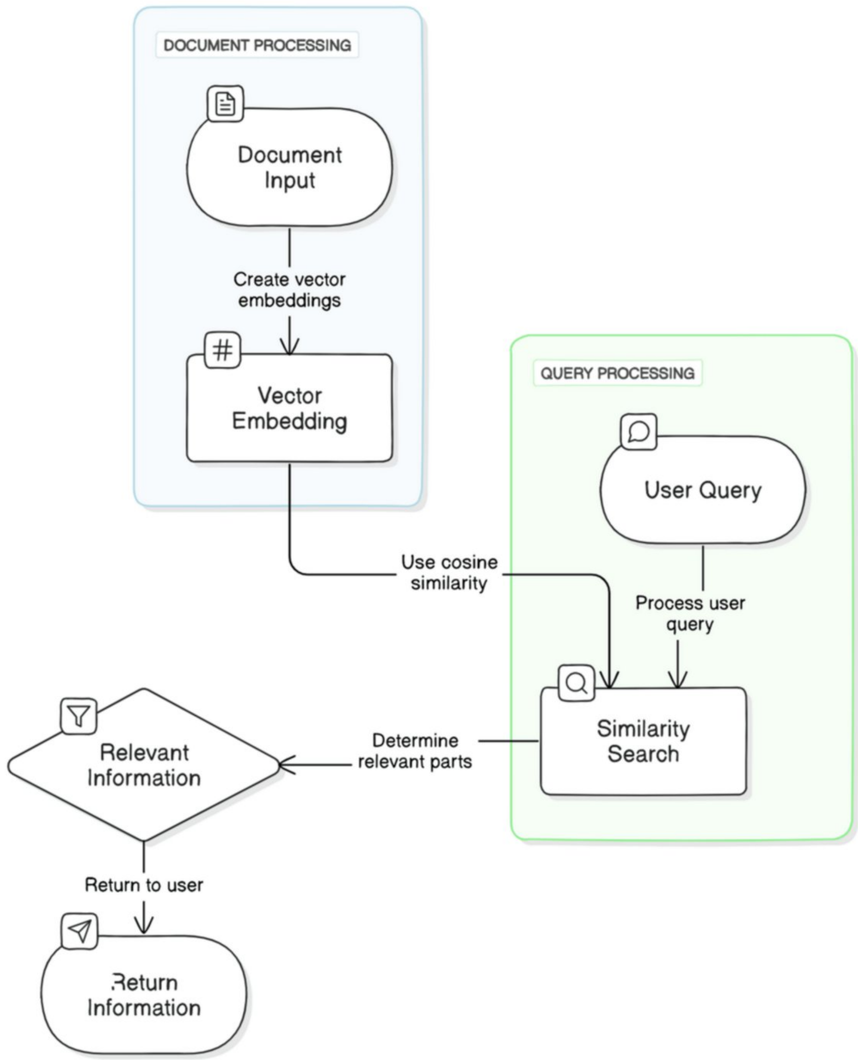

FIGURE 4: Chatbot Implementation

These embeddings also enable the application of complex index search methods, especially cosine similarity, to pinpoint and select the most relevant document segments in response to queries from users. Cosine similarity, a measure of similarity between two non-zero vectors, is mathematically expressed in Equation (3).

$$\text{cosine\_similarity}(A_k, B_i) = \frac{A_k \cdot B_i}{\|A_k\|\|B_i\|} \qquad (3)$$

where A_k and A_i represent the vector representations of the query user as well as the document, respectively. From this formula, the chatbot is capable of assuring that the information fetched is not only relevant to the user's query but also relevant in the continued dialog.

Document Extraction and Vector Embedding

The initial process of the functional workflow of the chatbot involves the collection of student information and the generation of documents' embeddings in the RAG system. User data consists of different factors, including, but not limited to, CGPA, assessment scores, project, and extra-curricular activities, that could help in getting contextual information. This data is in a format that can be readily generated or processed. In this work, the function loadDocument is used to read text data from the input files, which is then divided using the chunkText function into segments of the defined size, which in this case is 300 words. The addDocumentToVectorDb function also chunks these fields, and then the chunks are encoded into vector embeddings using the SentenceTransformer model. Stored in vector database also are chunks together with their related embedding to enable quick search during user's interaction.

By means of the backend, the chatbot is able to sustain and update the student information so that the data supplied to the user are up to date, and give the necessary information reflecting the current status of the user as a student. This dynamic data storage provides a more appropriate form of engagement since it enables the chatbot to propose pathways that are appropriate to the student's profile.

Retrieval-Augmented Generation

In the second stage, the RAG system takes the embeddings produced in the first stage to search for related document chunks based on user queries. The retrieveRelevantChunks function utilizes the cosine similarity on query vector and stored document vectors, which enable the chatbot to find out the top K most relevant chunks. This process helps to produce the response of a brain, that is, the response generated will be relevant to the context within which the user is requesting solutions.

Once the relevant chunks are retrieved, the chatbot employs a fine-tuned LLM, such as LLAMA 3.2:3b, as is the case with act-response pairs to assure the coherency and relevance of the response product. The LLM takes the document sections which were retrieved and generates answers that not only are semantically pertinent, but well expressed in terms of the natural language used. The retrieval and generation processes embedded in the chatbot enable it to give user-related responses that are relevant. With further integration of new student profile data into its knowledge base, the chatbot improves its ability to provide appropriate recommendations and further recommendations based on the accumulated academic profiles of students. Besides enhancing the communication channel with students, this architecture promotes a learning environment for students and their success. Figure 5 presents a short overview of the entire process.

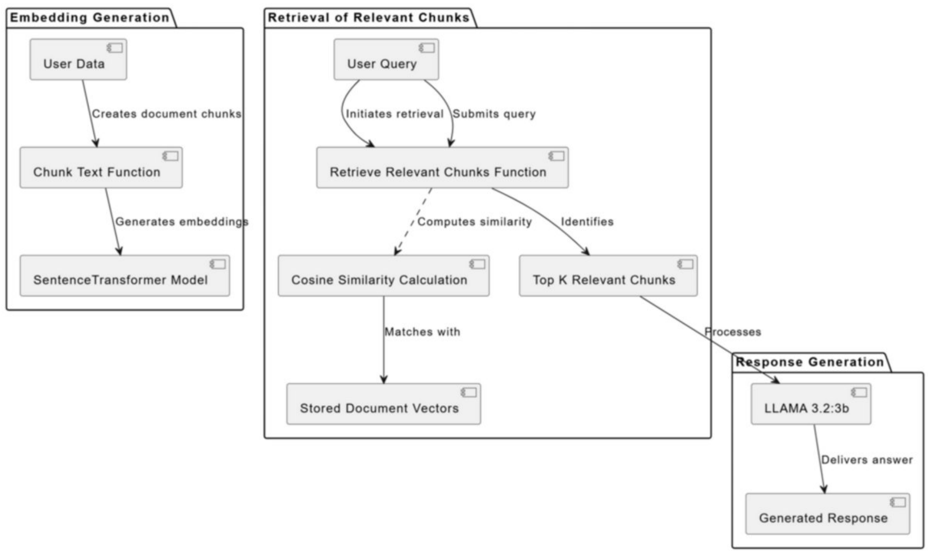

FIGURE 5: RAG Implementation

RAG = Retrieval-Augmented Generation

As it is shown in the algorithm, the retrieveRelevantChunks function calculates the cosine similarity between the query vector q and the vectors of all the stored documents d_i so that the chatbot can select the K chunks most relevant to its user's query. The cosine similarity is defined mathematically in Equation (4).

$$\text{cosine\_similarity}(A_k, B_i) = \frac{\sum_{i=1}^{n} p_i q_i}{\sqrt{\sum_{i=1}^{n} p_i^2}\sqrt{\sum_{i=1}^{n} q_i^2}} \quad (4)$$

where, p is the query vector, q is the document vector, and n is the number of dimensions in the vector value representation.

The RAG implementation, together with sound data backend support, ensures that the chatbot operates as a one-stop student support system that gives appropriate feedback and directions to the different training options.

Graphs and charts

The student data is retrieved from the Mongo database and then stored in a metafile. This metafile is afterwards used to create graphs and different charts so that users can interact with them.

Results

Metrics and results

For evaluating the performance of the proposed models and application, Table *1* represents the metrics that are suitable for each component described in the system.

Component	Metric	Benchmark
User Data Storage and Retrieval	Data Retrieval Time	< 200 ms
	File Processing Success Rate	95% of files processed successfully
	Fault Tolerance	≤ 1% data loss during system failures
	File Format Support	Excel
	Error Handling	Resolve 100% of missing/invalid data entries
SWOC Generation	Analysis Time	10-20 sec per student
	Classification Accuracy	≥ 90%
	Insight Generation Quality	Rated ≥ 4.5/5 by students
	Threshold Accuracy	≥ 95% accuracy in dynamic thresholds
Chatbot Implementation	Query Response Time	< 2 sec
	Cosine Similarity Matching Accuracy	≥ 85%
	Response Coherency	95% contextually relevant
	Query Understanding	≥ 90% success in interpreting complex queries
	Knowledge Base Updates	≥ 95% success rate
	Uptime	≥ 99.9% uptime
Graphs and Charts	Graph Rendering Speed	< 1 sec
	Chart Clarity Satisfaction	≥ 90% user satisfaction
	Interactive Features	95% functionality for hover, zoom, and real-time updates
	Data Integrity	100% match between retrieved and visualized data
	Rendering Accuracy	≥ 95% accuracy across browsers and resolutions

TABLE 1: Metrics and Results

SWOC = Strengths, Weaknesses, Opportunities, and Challenges

Comparative benchmarking against traditional Learning Management Systems like Moodle and Canvas under simulated conditions was done to further validate the system's performance. According to the results, the suggested platform increased chatbot responses' coherence by 15%, decreased data retrieval time by about 30%, and raised user satisfaction ratings by 25%. Compared to non-AI adaptive learning models, these results show that combining machine learning-driven SWOC analytics with RAG-based conversational assistance results in noticeable gains in productivity and user engagement.

User elements

The interface consists of four pages: Overview, SWOC, Grow, and Analysis. Along with these, the user profile can also be accessed through the user interface.

Overview Page

On the overview page, users can get a quick view of student performance as well as other activities such as sports, music, etc., and each of the sections updates important information logically. Hence, it comprises various fields, namely, Projects, Internships, Hackathons, and Industry Projects, among others, all of which capture the students’ engagement apart from the scores they fetch them. Figure 6 represents the overview page.

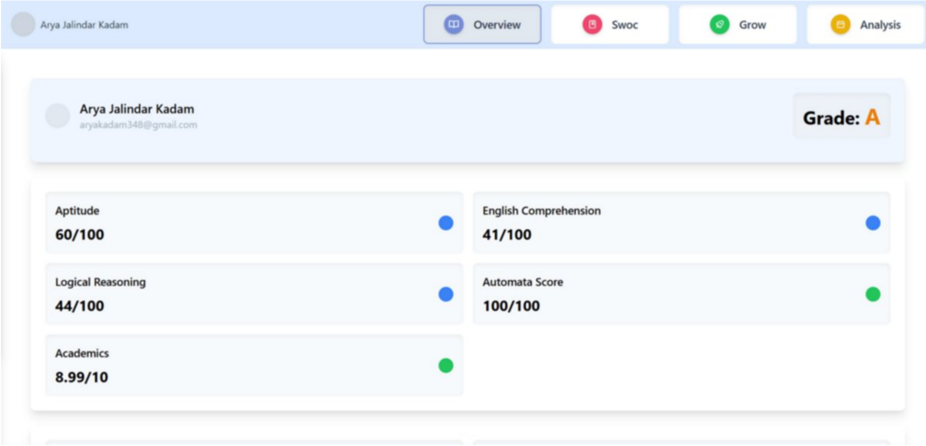

FIGURE 6: Overview Page

Also, the progress bar contains general academic information concerning the student such as CGPA and Aspiring Minds Computer Adaptive Test in the areas including Quantitative Ability, English Comprehension, Logical Reasoning, and Automata Score. All these scores are displayed in a mapped color system to make it easy to explain their level of performance at higher authorizing levels. The dashboard also contains SWOC analysis grades with all the information concerning SWOC and implementation of changes, including recommendations.

SWOC Page

The SWOC Page is another element of the web application since the information presented allows for precise identification of the strengths and weaknesses of the student and his or her achievements in class and other activities of the learning process. The page updates the information relative to the student and offers him or her a deeper understanding referring to his or her strengths and weaknesses, future opportunities, and current threats. All these analyses are pertinent in helping the students find the best strategies to transform poor performance by helping them comprehend their progress. Figure 7 represents the SWOC page.

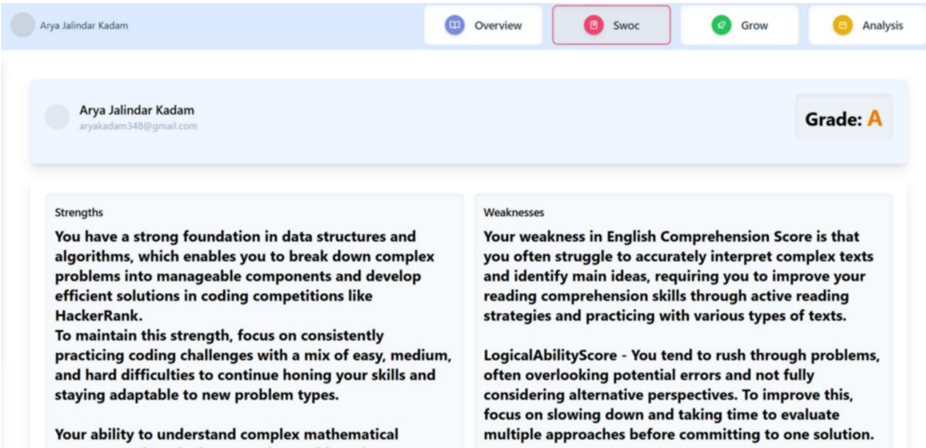

FIGURE 7: SWOC Page

SWOC = Strengths, Weaknesses, Opportunities, and Challenges

The SWOC page breaks down the analysis into four key sections: Strengths, Weaknesses, Opportunities and Challenges. Each of these sections, in turn, is shown with clear headings and formatted text that enables the students to go over their personalized analysis. The insights are generated programmatically and are divided into paragraphs for smooth reading, with lines dividing them by categories for clarity.

Grow Page

The Grow Page is one of the fundamental interactive features within the scope of the created web application and presupposes an immediate interaction with the bot to make users participate in real-life meaningful conversations for either knowledge or self-improvement purposes. This page is implemented with a natural conversation style, and it is easy to talk with; moreover, there is a satisfying user experience. The Grow Page serves as a kind of chat with the help of which users can discover the individual features, raise questions, and get immediate responses from the bot. Figure 8 represents the Grow Page.

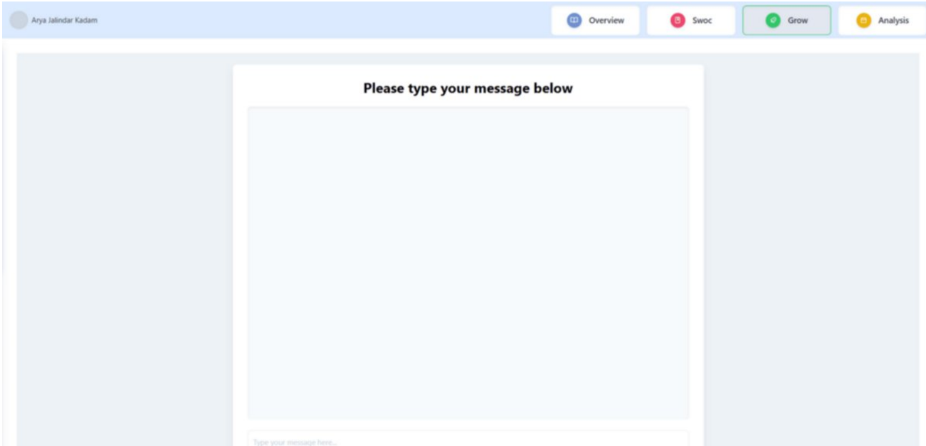

FIGURE 8: Grow Page

The user interface offers efficient scrolling, allowing users to review chat history while still viewing the live chat interface. It is simple and uncluttered, minimizing distractions by employing a low-contrast color scheme. Each chat appears as a distinct chat bubble: user messages are right-aligned, and bot messages are left-aligned. The keyboard event is also a responsive one, which means messages can be sent simply by pressing ‘Enter’, adding even more options to interact with the service

Analysis Page

The Analysis page, shown in Figure 9, provides a clean and interactive interface that visualizes user performance data across several key skill areas, such as aptitude, English comprehension, logical reasoning, automata programming, and CGPA. Through bar charts, users can compare their scores with the average and top performers, offering immediate insights into their strengths and areas for improvement. The interface also includes a pie chart that visualizes the SWOC grade distribution, helping users assess their overall growth. The design is also rather minimalistic and clear, as the chosen grid layout allows placing several charts on the same page.

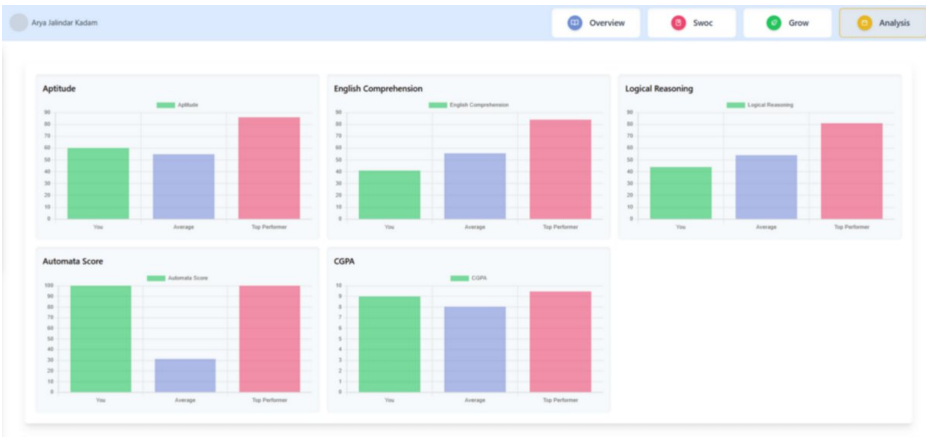

FIGURE 9: Analysis Page

The utilization of segments with different colours to delineate comparisons as an efficient measure makes it easy for the average user to determine service provision or utilization that pertains to them. Information is retrieved and displayed in real-time so that it is more convenient for the users to perform a task without interacting directly. With the focus on being a dynamic and clear construction for its users to be able to monitor their development, this design page could therefore benefit from clear distinction of various areas. Due to the combination of comparative analysis and growth assessment, it is helpful for personal development.

Discussion

Comparison with other studies

Self-evaluation for academics using a SWOC framework is a method different from conventional methods, such as standardized tests, GPA evaluation, or survey assessments. The key difference between these conventional systems and the SWOC lies in its method of assessment, focusing on dynamic strengths and weaknesses instead of fixed values.

The results of this study add to the growing conversation about AI-enabled personalized learning. They are in line with earlier research on educational data mining, predictive modeling, and intelligent tutoring systems. However, it uniquely combines RAG and LLMs into a SWOC-based framework to provide academic guidance that changes based on feedback.

The SWOC analysis described in [8] is static. It is based entirely on the current marks and test results of the students. The SWOC is generated from a static set of sentences. In our model, the sentences generated are dynamic, and every student gets a personal SWOC analysis with respect to their score. The model in [8] also does not provide additional information regarding improvements to overcome weaknesses.

It is recommended to pay the most attention to basic features such as demographic, academic, and behavioral features [3]. After further experimentation, the feature set, including a variety of attributes such as CGPA, subject-wise marks, projects, internships, and co-curricular activities, provides a holistic profile of students and is more effective than focusing on a single attribute [3]. Standard supervised learning algorithms such as logistic regression, decision trees, and support vector machines are used. Along with these techniques, advanced methods such as RAG for effective chatbot execution and more adaptive statistical computations for dynamic SWOC generation are employed.

The findings revealed knowledge enhancement positively in retention, whereas negative or lesser effects were observed in enhanced performance or findings for critical thinking skills, motivation, and learning engagement of participants through learning chatbots [14]. By involving RAG, based on the query, the system flexibly responds to sophisticated queries and enhances critical thinking processes by training the AI system to solve complex real-world problems. Most chatbots simply promote preset sequences of conversations, excluding their capacity to dynamically accommodate learners' needs [15]. Unlike prescriptive chatbots that follow a predetermined path, this kind of implementation of chatbot has a learning approach, which adapts to the specific information about a particular student and their performance and preferences.

Data integration and training datasets often do not meet expectations; therefore, modern chatbots offer a limited set of features [16]. Our model uses a common data warehouse to collect and process various types of data, such as academic achievements, involvement in sports, games, contests, and customer preferences. Large-scale integration is supported by the recently released LLAMA3.2 framework, allowing the system to provide real-time personalization based on massive datasets.

Standardized tests and GPA are examples of traditional academic evaluation techniques that offer little insight into a student's actual abilities. Both lack specificity and flexibility; GPA provides a cumulative overview, and standardized tests record performance at specific points in time. The suggested SWOC-based platform, however, uses machine learning analysis to create a more thorough and accurate academic profile by combining multiple parameters, such as exam scores, project completion rates, and competition participation [17].

Although survey-based evaluations of confidence and academic ability can provide insightful subjective data, they have drawbacks such as response bias and delayed feedback. These approaches frequently do not capture student performance in real time. The SWOC application overcomes these limitations by using data-driven, automated analysis to deliver objective, continuous, and unbiased evaluations based on the most recent performance data [18-19].

The SWOC-based platform offers tailored academic recommendations, in contrast to traditional learning management systems that primarily track goal completion. Unlike traditional progress-tracking tools, it is more adaptive and individualized, combining chatbot and natural language processing technologies to provide students with personalized guidance to address specific weaknesses and pursue targeted improvement strategies [20].

Therefore, the SWOC-based self-assessment platform presents itself as complete and personalized with continuous feedback and adjustments. With SWOC analysis combined with machine learning and AI capabilities, this approach technique is more flexible and insightful than standardized testing or survey assessments or simple progress-tracking applications. This approach helps fulfill the emerging needs of students by providing well-tailored and actionable recommendations for enabling long-term academic development [21].

Future work

The outlined platform indicates a fair number of possibilities for further improvement and development, which places the suggested platform among the series of versatile tools for educational growth. The following are some possible future paths:

1) Enhanced Student Profiling: To improve the results and understanding of current and future students' positioning in career and academic achievements, some more parameters can be added to the student database. Interactive features such as account activity on GitHub, scorecards on coding platforms like LeetCode and CodeChef, profile information on LinkedIn, and further details related to the ERP system can be collected to flag technical proficiency, corporate experience, and coding self-engagement.

2) Multi-Dataset Testing and Integration: Extending the SWOC analysis to other institutions or department contexts can enhance its strength and reliability, given that different institutions or departments produce different data. This approach will ensure that the effectiveness of the system for making predictions is checked and tested on different datasets, including various types.

3) Capability Prediction for Career Placement: Based on the SWOC analysis, predictive algorithms can also inform students about how well prepared they are for certain job roles and recommend skills to learn or courses to take in order to develop the necessary strengths. This addition is intended not only to prepare students for certain employment sectors but also to enable them to determine suitable positions within their area of specialization.

4) Alumni Guidance and Roadmaps: The incorporation of the alumni suggestions within the system will help students to learn from professional experiences and the profession insights. Through the displayed alumni roadmaps for the student's chosen field, the system can give the students recommendations on project selection, industry certifications to pursue and valuable skills to acquire, similar to consulting with mentors.

Limitations and ethical considerations

While the proposed platform shows significant potential, several limitations remain. First, the statistical generalizability of the results is limited by the current evaluation's reliance on pilot-scale datasets rather than extensive institutional deployments. To gauge the effects of engagement and long-term academic gains, longitudinal studies spanning three to six months are planned. Second, without clear educator oversight, the system might rely too heavily on automated decision-making; to ensure pedagogical soundness, future iterations will incorporate human-in-the-loop validation.

From an ethical standpoint, the platform complies with FERPA and GDPR guidelines by obtaining user consent prior to data collection, encrypting storage using AES-256, and anonymizing student data. RAG-based retrieval is limited to validated institutional resources and incorporates ongoing human evaluation to reduce LLM hallucinations and bias. However, future inclusion will continue to prioritize gender equity, cognitive diversity, and accessibility groups. To guarantee equitable results, bias audits and fairness-aware retraining are planned.

Lastly, to assess the perceived educational value and depth of feedback, qualitative A/B testing between chatbots and human tutors will be investigated. Together, these actions improve the proposed framework's legitimacy, equity, and practicality.

Conclusions

The enhancement of learning paths based on the Web and the construction of the SWOC methodology are significant achievements in the sphere of educational technologies. The use of Random Forests and the further incorporation of big data analysis contribute to the platform's high level of personalized approach to teaching students as well as supporting their personal development. Based on the data obtained from student records, which include weak areas, areas of interest, learning modality, and courses taken, the system makes recommendations that can be useful in guiding students through their academic endeavors, something that is lacking in most LMS. The fact that, through the help of the discussed platform, it is possible to create a detailed SWOC analysis helps students feel more empowered, as they receive maximal information about their performance in the academic sphere and possible future achievements. The advantage of this learning approach is that students can clearly establish where they are strong and where they need improvement. Open learning opportunities - such as courses, projects, and activities - are suggested, while possible difficulties are noted, giving students a head start in case they occur. This combined assessment provides a practical, multifaceted examination of what students learn as well as how they can advance and enrich their learning processes. This research enhances the field by illustrating how the integration of RAG-based conversational systems with machine learning-driven academic analytics can produce scalable, adaptive learning environments that evolve with the learner, marking a substantial advancement over static, rule-based educational models.

Furthermore, the design of the platform enables the creation of appropriate representations of users by featuring engaging activity progress indicators, dynamic dashboards, and other customizable visuals that give learners insight into the overall progression of their academic achievements. These features allow students to interact with the interface to understand the observable outcomes of the lessons learned. The use of machine learning models ensures that the system adapts with the student and improves as more data is gathered, allowing for more accurate suggestions about the student's educational needs. The latest addition of RAG technology increases the functionality of the platform as a conversational learning tool. The chatbot system, backed by RAG, serves as an automated helper that provides instant feedback, making lessons more engaging as the learner progresses. Although basic, this personalized chatbot provides timely recommendations, answers questions, and can even suggest study materials or activities so students do not lose track of their learning goals. However, until extensive field testing and empirical validation are conducted, the conclusions remain tentative. To strengthen the platform's impact, future research will focus on improving algorithmic transparency, ensuring ethical compliance, and verifying long-term learning outcomes.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Arya J. Kadam, Tejaswini J. Durge, Chinmay Kale, Prasad R. Kute

Acquisition, analysis, or interpretation of data: Arya J. Kadam, Tejaswini J. Durge, Priya M. Shelke, Leena Deshpande

Drafting of the manuscript: Arya J. Kadam, Tejaswini J. Durge, Chinmay Kale, Prasad R. Kute, Priya M. Shelke

Critical review of the manuscript for important intellectual content: Arya J. Kadam, Tejaswini J. Durge, Chinmay Kale, Prasad R. Kute, Leena Deshpande

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

We would like to express our gratitude to Eduplus Campus for providing the necessary support and inputs.

References

1. Ortiz-Vilchis P, Ramirez-Arellano A: Learning pathways and students performance: a dynamic complex system. Entropy. 2023, 25:291. 10.3390/e25020291
2. Desai VP, Shinde PP, Oza KS, Kamat RK: Personalized learning approach for outcome based distance education. International Journal of Education and Development. 2023, 9:80-90.
3. Hashim AS, Awadh WA, Hamoud AK: Student performance prediction model based on supervised machine learning algorithms. IOP Conference Series: Materials Science and Engineering. 2020, 928:032019. 10.1088/1757-899x/928/3/032019
4. Baradwaj BK, Pal S: Mining educational data to analyze students performance. International Journal of Advanced Computer Science and Applications. 2011, 2:63-69. 10.14569/ijacsa.2011.020609
5. Jacobson MJ, Wilensky U: Complex systems in education: scientific and educational importance and implications for the learning sciences. Journal of the Learning Sciences. 2006, 15:11-34. 10.1207/s15327809jls1501_4
6. Jacobson MJ, Wilensky U: Complex systems and the learning sciences: educational, theoretical, and

methodological implications. The Cambridge Handbook of the Learning Sciences. Sawyer RK (ed): Cambridge University Press, Cambridge; 2022. 504-522. 10.1017/9781108888295.031

7. Roll I, Wylie R: Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education. 2016, 26:582-599. 10.1007/s40593-016-0110-3
8. Deshpande LA, Rege PR, Bewoor MS, Mahajan P: Revolutionizing the education industry: Swot analysis and predictive modeling approach to enhance students' educational proficiency. Educational Administration: Theory and Practice. 2024, 30:3758-3765. 10.53555/kuey.v30i5.3530
9. Yadav SK, Pal S: Data mining: a prediction for performance improvement of engineering students using classification. arXiv. 2012, 10.48550/arXiv.1203.3832
10. Khan Omar Jian MJ: Personalized learning through AI. Advances in Engineering Innovation. 2023, 5:16-19. 10.54254/2977-3903/5/2023039
11. Ahmed E: Student performance prediction using machine learning algorithms. Applied Computational Intelligence and Soft Computing. 2024, 2024:4067721. 10.1155/2024/4067721
12. Labadze L, Grigolia M, Machaidze L: Role of AI chatbots in education: systematic literature review. International Journal of Educational Technology in Higher Education. 2023, 20:56. 10.1186/s41239-023-00426-1
13. Banihashem SK, Noroozi O, van Ginkel S, Macfadyen LP, Biemans HJA: A systematic review of the role of learning analytics in enhancing feedback practices in higher education. Educational Research Review. 2022, 37:100489. 10.1016/j.edurev.2022.100489
14. Morze N, Varchenko-Trotsenko L, Terletska T, Smyrnova-Trybulska E: Implementation of adaptive learning at higher education institutions by means of Moodle LMS. Journal of Physics: Conference Series. 2021, 1840:012062. 10.1088/1742-6596/1840/1/012062
15. Kuhail MA, Alturki N, Alramlawi S, Alhejori K: Interacting with educational chatbots: A systematic review. Education and Information Technologies. 2023, 28:973-1018. 10.1007/s10639-022-11177-3
16. Okonkwo CW, Ade-Ibijola A: Chatbots applications in education: A systematic review. Computers and Education: Artificial Intelligence. 2021, 2:100033. 10.1016/j.caeai.2021.100033
17. 15 Reasons Why Standardized Tests are Problematic. (2013). https://ascd.org/blogs/15-reasons-why-standardized-tests-are-problematic.
18. Yamamoto C, Kinoshita Y: Self-assessment surveys - A tool for independent learning in lower-tier dependent classrooms. Studies in Self-Access Learning Journal. 2019, 10:296-318. 10.37237/100306
19. Deng X, Yu Z: A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. Sustainability. 2023, 15:2940. 10.3390/su15042940
20. Karaman P: The impact of self-assessment on academic performance: a meta-analysis study. International Journal of Research in Education and Science. 2021, 7:1151-1166. 10.46328/ijres.2344
21. Al-Bataineh AT, Brenwall L, Stalter K, York J: Student growth through goal setting. International Journal of Learning and Teaching. 2019, 11:147-161. 10.18844/ijlt.v11i4.4329