Journal of Information Technology Education: Innovations in Practice

An Official Publication of the Informing Science Institute
InformingScience.org

JITEiip.org

Volume 25, 2026

SHELF: A GENDER-SENSITIVE HYBRID AI FRAMEWORK FOR ADVANCING EDUCATIONAL EQUITY IN UNDERSERVED REGIONS

Ayesha Nawaz	Department of Computer Science, Government College Women University, Sialkot, Pakistan	cheemaayesha34@gmail.com
Muhammad Hamid*	Department of Computer Science, University of Kamalia, Pakistan	mhamid@ukm.edu.pk
Aisha Shabbir	Faculty of Computing and Information Technology (FCIT), Sohar University.	ashabbir@su.edu.om
Aisha Blfgeh	College of Computer Science and Engineering, Department of Computer Science and Artificial Intelligence, University of Jeddah, Jeddah, Saudi Arabia	aablfgeh@uj.edu.sa
Muhammad Saleem	Department of Industrial Engineering, Faculty of Engineering, King Abdulaziz University, Rabigh, Saudi Arabia	msaleim1@kau.edu.sa

ABSTRACT

Aim/Purpose	AI-based recommendation systems frequently fail to serve women in low-resource educational environments effectively, as they neglect critical realities, such as time constraints, limited connectivity, and socio-cultural barriers. This paper proposes a hybrid ensemble AI framework that directly addresses these contextual realities by integrating algorithmic fairness, gender-aware personalization, and low-bandwidth adaptability into a unified, women-focused framework.
Background	Although AI-driven personalized learning has advanced rapidly, existing systems lack gender sensitivity, transparency, and contextual adaptability, limiting their effectiveness in promoting inclusive education.

Accepting Editor Faisal Syafar │ Received: February 7, 2026 │ Revised: April 29, May 9, May 16, May 30, 2026 │ Accepted: June 13, 2026.

Cite as: Nawaz, A., Hamid, M., Shabbir, A., Blfgeh, A., & Saleem, M. (2026). SHELF: A gender-sensitive hybrid AI framework for advancing educational equity in underserved regions. *Journal of Information Technology Education: Innovations in Practice, 25,* Article 32. https://doi.org/10.28945/5817

Methodology	A Systematic Literature Review (SLR) of 50 peer-reviewed studies (2015-2025) was conducted to analyze research trends in AI-powered recommender systems, gender-focused education, emerging technologies, and explainable AI in educational contexts.
Contribution	The key addition of the study is a conceptual ensemble hybrid AI framework, namely Sensitive Hybrid Ensemble for Learning Fairness (SHELF), which provides a comprehensive framework for simultaneously incorporating gender-sensitive learner profiling, bias-conscious recommendation, and low-bandwidth adaptability. Based on a systematic review of 50 studies, the framework provides an evidence-based solution to a critical gap in gender-blind AI-driven educational systems in underserved areas.
Findings	Results show that 48% of studies focus on AI-based recommender systems, 19% on adaptive and fair conceptual models, 11% on gender-specific education, and 12% on emerging technologies such as IoT and edge computing. Systematic reviews and meta-analyses of educational recommendation systems account for 7% of studies, while only 3% address transparency and explainability, highlighting a major research gap.
Recommendations for Practitioners	Educational stakeholders should adopt gender-aware, fair, and explainable AI systems to provide personalized learning and career guidance for women learners.
Recommendations for Researchers	Future research should emphasize inclusive AI design, bias mitigation, contextual adaptability, and explainability in educational systems.
Impact on Society	Gender-aware AI in education can enhance women's access to learning, reduce digital inequality, and support inclusive societal development.
Future Research	Further work should validate hybrid AI frameworks across diverse educational contexts and assess long-term impacts on women learners.

Keywords	artificial intelligence, AI, AI in education, gender-aware learning, personalized learning systems, learning equity, women's educational empowerment

INTRODUCTION

Recent advances in artificial intelligence (AI) have transformed digital learning environments through scalable, personalized education that supports data-driven decision making, adaptive content delivery, and learner analytics (Kuleto et al., 2021; McGrath et al., 2023; Qadir et al., 2025; Urdaneta-Ponte et al., 2021). The use of machine learning and deep learning has enabled automated identification of learning behavior, which benefits early detection of student failure and the optimization of learning pathways (Hamid et al., 2026). Different models have been applied to forecast student performance, prevent dropout, and dynamically adjust curricula (Kong et al., 2021; Saqr & López-Pernas, 2024; Tzirides et al., 2024). As institutional strategies are often tailored to the unique needs and preferences of learners, AI technologies contribute to providing a larger view of the personalized learning environment (Han et al., 2024; Raj & Renumol, 2022; Su & Yang, 2022).

Recommendation systems in education often use hybrid models that leverage collaborative filtering, content-based filtering, deep learning, and knowledge-based reasoning, achieving higher accuracy than single-method approaches (Adeleke & McSharry, 2022; Bagunaid et al., 2022; Holmes et al., 2022). These techniques have demonstrated improvements in learner satisfaction, provide personalized course suggestions and learning resources, reduce cognitive overload, and enhance learner motivation (Machado et al., 2025; Mahmud et al., 2024; Strzelecki, 2024). Additionally, the integration of

IoT, knowledge graph embedding, and neural graph networks has expanded the potential of recommendation systems, providing transparent and scalable solutions for complex learning content (Aladi, 2024; Huang et al., 2023; Q. Zhang et al., 2021).

However, several critical gaps and limitations persist in the present recommendation systems that include data sparsity, cold-start issues, lack of fairness and transparency, and insufficient scalability, which reduce the accuracy of recommendations for learners with limited digital access (Pimentel et al., 2021; Silva et al., 2023; Yang, 2024). Black Box models often favor students with more data-rich profiles and access to technologies, deepening the educational inequalities (Quan & Pu, 2023; Weber et al., 2022; P. Zhang & Zhang, 2024). The applicability of resource-intensive AI models is limited in low-resource areas, as these models require powerful infrastructure and large data sets, whereas connectivity and availability are restricted in diverse areas (Assegie et al., 2024; Benhamdi et al., 2017; Mohd Talib et al., 2023; Mohd Zaki et al., 2023).

The most unprecedented challenge in recommendation system design is the lack of gender-sensitive personalization for women of rural communities. These women often face barriers such as household responsibilities, income dependency, limited access to digital technology, and social mobility restrictions that are often neglected in the existing system (H. S. Lee & Lee, 2021; Y. Lee et al., 2023; Su & Yang, 2022). In this way, one-size-fits-all AI systems create the risk of educational disparities rather than mitigating them. This paper addresses three important gaps that current systems do not cover. Most AI-based recommendation systems are built with data-heavy learners with reliable access to devices, inherently excluding women in low-resource environments where connectivity, digital literacy, and availability to devices are limited (Assegie et al., 2024; Benhamdi et al., 2017; Mohd Talib et al., 2023; Mohd Zaki et al., 2023). There is no available framework that incorporates gender-sensitive personalization variables such as household responsibilities, time constraints, socio-cultural barriers, and income dependence (H. S. Lee & Lee, 2021; Y. Lee et al., 2023; Su & Yang, 2022). Only 4% of the studies discuss transparency and explainability, which makes algorithmic decisions unclear and diminishes the trust that isolated learners place in AI-supported educational directions. This paper directly addresses these three gaps by suggesting a hybrid AI framework.

To address these gaps, this paper is guided by the following research questions:

1. How can AI-based recommendation systems be designed to support personalized learning while ensuring ethical transparency and minimizing algorithmic bias?
2. Which methodologies improve the adaptability and scalability of systems across diverse educational environments?
3. How can recommendation systems be customized to address gender-specific challenges while supporting women's access to education in underserved regions?

The following research objectives are made based on the above research questions:

1. To critically analyze the existing studies on AI-based recommendation systems in education with a specific focus on personalization, fairness, and gender inclusion.
2. To identify key challenges, methodological gaps, and limitations of current systems in addressing the educational needs of women in low-resource settings.
3. To provide strategic and gender-sensitive recommendations for future research and development of gender-based educational recommendation platforms.

This paper contributes to this field by:

1. Reviewing the existing literature on AI-based educational recommendation systems to highlight the limitations of current models for the specific needs of women learners in diverse educational environments.
2. Identifying gaps, algorithmic bias, and scalability issues, this paper emphasizes the need for ethical and equitable AI educational systems.
3. Proposing a novel ensemble AI framework to address the identified challenges.

The rest of the paper is organized as follows. The next sections present the methodology and the systematic literature review, followed by an analysis of studies. The proposed conceptual framework is then presented, followed by answers to the three research questions. The recommendations for researchers and practitioners are then presented, followed by the conclusion, along with future directions and final considerations.

METHODOLOGY

This section describes the systematic methodology used to select and analyze the 50 studies. It presents the research methodology adopted to review the research topic. The methodology is divided into the following subsections: research strategy, study selection, inclusion and exclusion criteria, data extraction, and quality assessment of studies. The process is outlined below.

PHASE 1: PLANNING THE REVIEW

The initial phase of this study involved establishing a clear roadmap for a systematic review. The **research objectives** that guide this process and the subsequent framework development are detailed in the Introduction section. Following a structured **review protocol**, this study examines the role of AI-based recommendation systems in education, focusing on how they promote inclusive and fair education, particularly for women's access to education. A structured search strategy was applied across five major academic databases: IEEE Xplore, ACM Digital Library, Scopus, Web of Science, and Science Direct.

To ensure the search was fully reproducible, the following integrated Boolean search string was formulated and applied to the titles, abstracts, and keywords of the records:

> ("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Recommendation System") AND ("E-learning" OR "Personalized Learning" OR "Adaptive Learning" OR "Student Performance Prediction") AND ("Women Education" OR "Gender Inclusion" OR "Fairness" OR "Equitable Learning")

The keywords were formulated around three main categories:

- *Technology:* "Artificial Intelligence", "Machine Learning", "Deep Learning", "Recommendation System."
- *Education:* "E-learning", "Personalized Learning", "Adaptive Learning", "Student Performance Prediction."
- *Equity & Gender:* "Women Education", "Gender Inclusion", "Fairness", "Equitable Learning."

The selection of articles was performed in three steps, as shown in Figure 1. To maintain methodological rigor and minimize selection bias, each stage, from initial identification to full-text eligibility, was conducted independently by two researchers. To ensure the reliability of findings, each study was scored using a granular rubric based on five quality criteria (QA1–QA5). The scoring was applied as follows:

- *Yes:* The study provides explicit empirical evidence and a clear methodology for the criterion.
- *0.5 (partially):* The study addresses the criterion but lacks depth, specific metrics, or complete validation.
- *No:* The criterion is not addressed or is irrelevant to the study's scope.

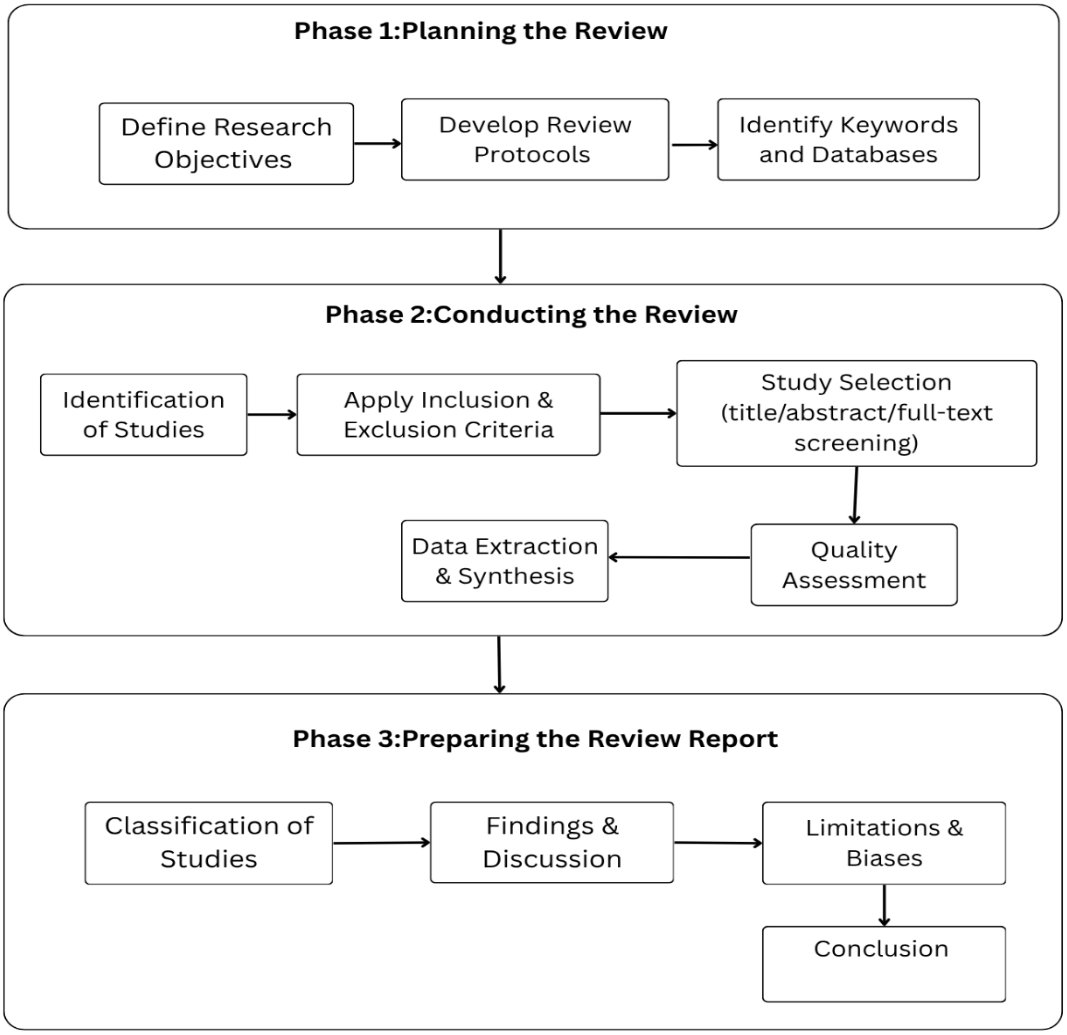


Figure 1. Phases of the review process

The search was limited to publications between January 2015 and March 2025, a ten-year range chosen to allow publications to reflect the development and maturation of deep learning in educational technology and to remain relevant to contemporary issues in system design. Papers published before 2015 were not considered, as AI-based recommendation systems in education were not common at that time.

PHASE 2: CONDUCTING THE REVIEW

Identification of studies

During the identification phase, 670 records were retrieved from the specified databases using the specified keywords. The search was limited to publications between January 2015 and March 2025, a ten-year range chosen to allow the publications to reflect the development and maturation of deep learning in educational technology.

Inclusion and exclusion criteria

The following inclusion and exclusion criteria were applied to refine the analysis.

Inclusion criteria

- The studies should be published in peer-reviewed journal articles, conference papers, or book chapters.
- The research explicitly addresses AI-based recommendation systems in education.
- The primary focus should be on personalization, fairness, scalability, or equity learning.
- The research that addresses gender-specific challenges, women's education, or underserved learners.
- The article should be published in English.
- The publication date should be between 2015 and 2025.

Exclusion criteria

- Studies not related to education and AI.
- Articles that do not have empirical evidence or evaluation.
- Editorials, books, dissertations, or non-peer-reviewed articles.
- Non-English publication.
- Publications before 2015.

Study selection

The selection of articles for the study followed a two-phase systematic process. The referees were from 2015 to 2025, as the selected period included new developments in AI-based recommendation systems and was also relevant to current educational technology practices. During the identification phase, 670 records were first retrieved from the specified databases using the specified keywords. The multi-stage screening process eliminated duplications and irrelevant studies, resulting in 550 records. These were then filtered in turn by title, abstract, and full-text evaluation, resulting in 35 studies added to the list of those that satisfied all predetermined criteria.

Two independent researchers took part in the screening and validation of the study to increase methodological rigor and minimize selection bias. Each stage of the research was assessed by both researchers independently in accordance with the predetermined inclusion and exclusion criteria. A consensus-based method was employed in cases of disagreement through discussion to maintain uniformity in selection decisions. Before full-scale screening, a sample of 20 records underwent a calibration exercise to ensure that inclusion rules were consistently applied. The snowballing method was applied to the reference lists of the first 35 studies in the second phase, yielding an additional 63 relevant records. These studies were also rigorously screened, and an additional 15 were included.

This thorough methodology led to a final sample of 50 studies for this review, as shown in Figure 2. Moreover, to improve the reliability and transparency of the synthesis process, the two researchers independently assessed the selected studies. Synthesis was conducted on extracted data to develop the frameworks, using thematic coding of the limitations and future work sections of each study. This strategy enabled the identification of recurrent design gaps that directly influenced the creation and architecture of the proposed hybrid AI framework.

Table 1 summarizes the selection process, and for each study, we extracted key information, including its primary contribution (e.g., system proposal, empirical analysis, literature review, datasets, and results limitations), the problem it addressed, the techniques used, and its publication year. This data was then synthesized to create the tables and analyses presented in the following sections. Table 2 provides a classification of studies by primary contribution, and Table 3 presents a year-wise classification of these studies.

Table 1. Study selection

Phase	Records (n)	Description
Phase 1	670	Records retrieved through database search using defined keywords & databases
Screening (multi-stage)	550	Removal of duplicates & irrelevant studies (title, abstract, methodology)
Eligibility	35	Studies assessed against inclusion criteria
Phase 2: Snowballing	63	Additional records identified through reference lists of 35 studies
Screening (snowballing)	15	Screening of 63 records with the same inclusion/exclusion process
Final inclusion	50	Total studies included (35 + 15)

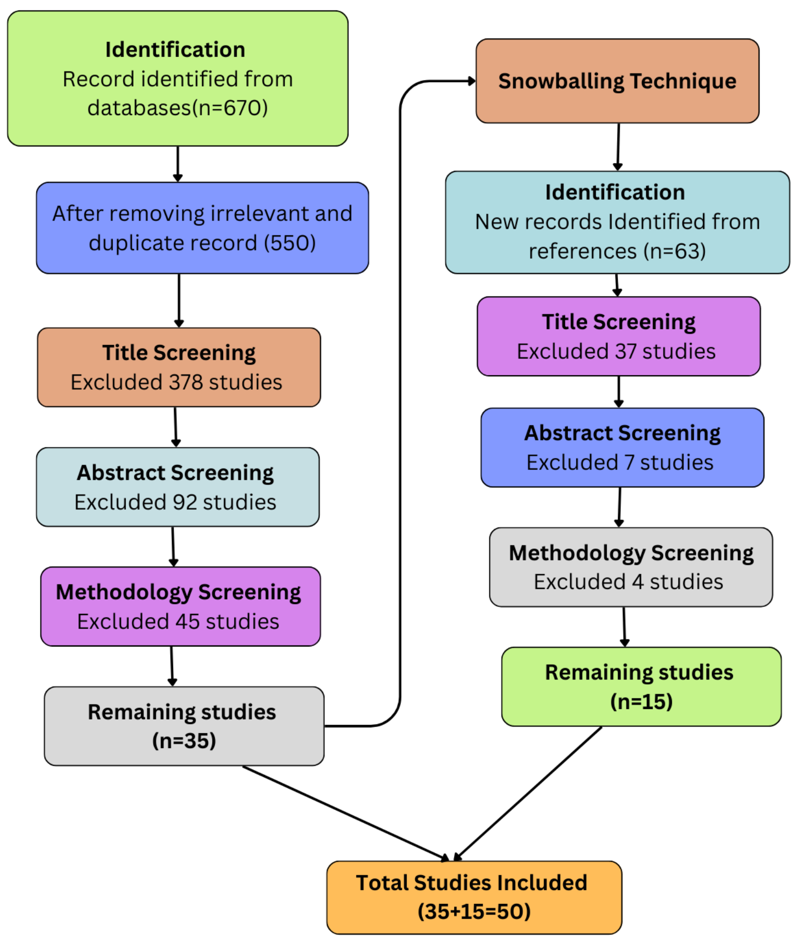


Figure 2. Flow of the study selection process applied in this SLR

Figure 4 shows that 48% of studies in the literature focus on AI-powered recommender systems for personalized and adaptive learning using hybrid recommendation approaches. The second major category, 19%, addresses student performance prediction and academic analysis, forecasting academic outcomes by machine learning and deep learning models. Gender-specific and women's education studies account for 11% of the literature. Emerging technologies such as the Internet of Things (IoT), edge computing, and visualization, which enhance students' learning and engagement, account for 12% of the research. Systematic reviews and meta-analyses of educational recommendation systems contribute 7% of studies. Finally, a smaller portion, 4%, highlighted the transparency and explainability of AI-driven educational tools.

Table 2. Primary classification of studies description

Primary Category	Description	No. of Studies	Percentage (%)
AI-Based Recommender Systems	Systems focusing on personalized recommendations in education	24	48%
Machine Learning in Education	Use of ML models (SVM, RF, LSTM, CNN, etc.) for prediction and analytics	10	20%
Explainable AI (XAI) in Education	Studies focusing on transparency and interpretability	2	4%
Adaptive Learning Systems	Systems adapting content dynamically to learners	5	10%
Gender & Educational Equity	Studies addressing women's education and gender disparities	4	8%
AI Ethics & Fairness	Studies focusing on bias, fairness, and ethical implications	3	6%
AI Integration & Educational Practices	Adoption, perception, and implementation of AI in education	2	4%
Total		**50**	**100%**

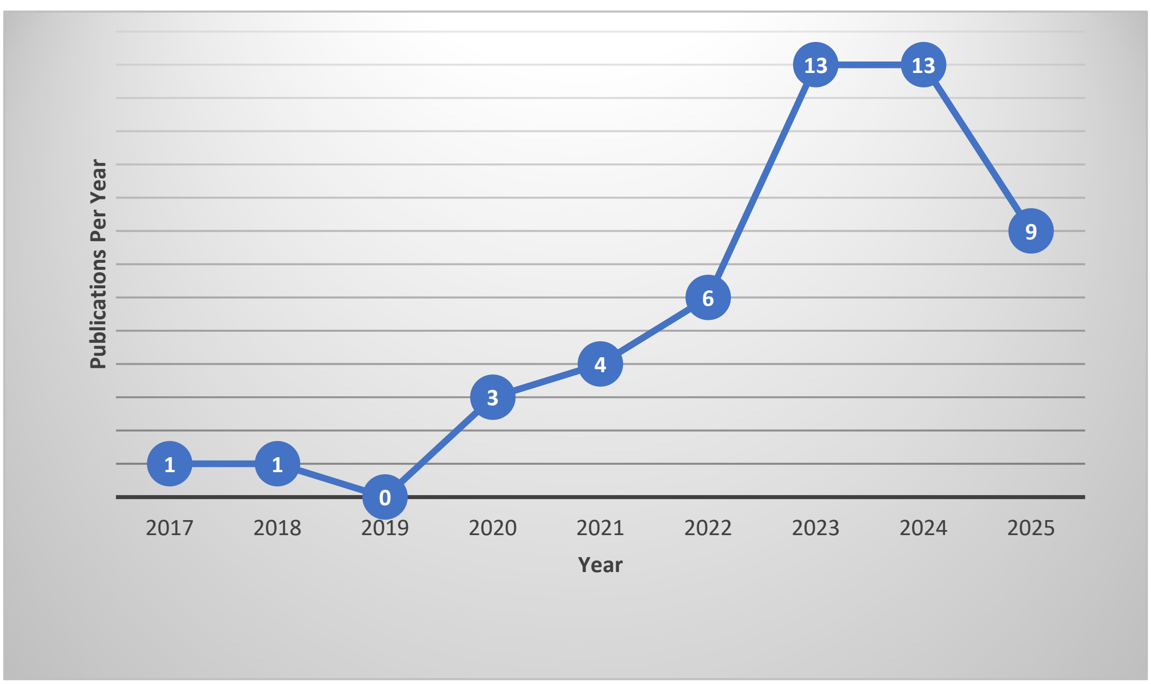


Figure 3. Demonstrates the publication trend of selected papers from 2017 to 2025

Table 3. Distribution of primary research contributions by publication year

Year	Model/ algorithm development	Framework/ conceptual architecture	Application to education contexts	Systematic reviews/ surveys	Gender & equity focused studies	Total unique papers per year
2017	1	0	0	0	0	**1**
2018	1	0	0	0	0	**1**
2020	0	2	1	0	0	**3**
2021	2	0	1	1	0	**4**
2022	1	2	2	1	0	**6**
2023	7	3	1	1	1	**13**
2024	8	2	1	0	2	**13**
2025	2	4	3	0	0	**9**
Total						**50**

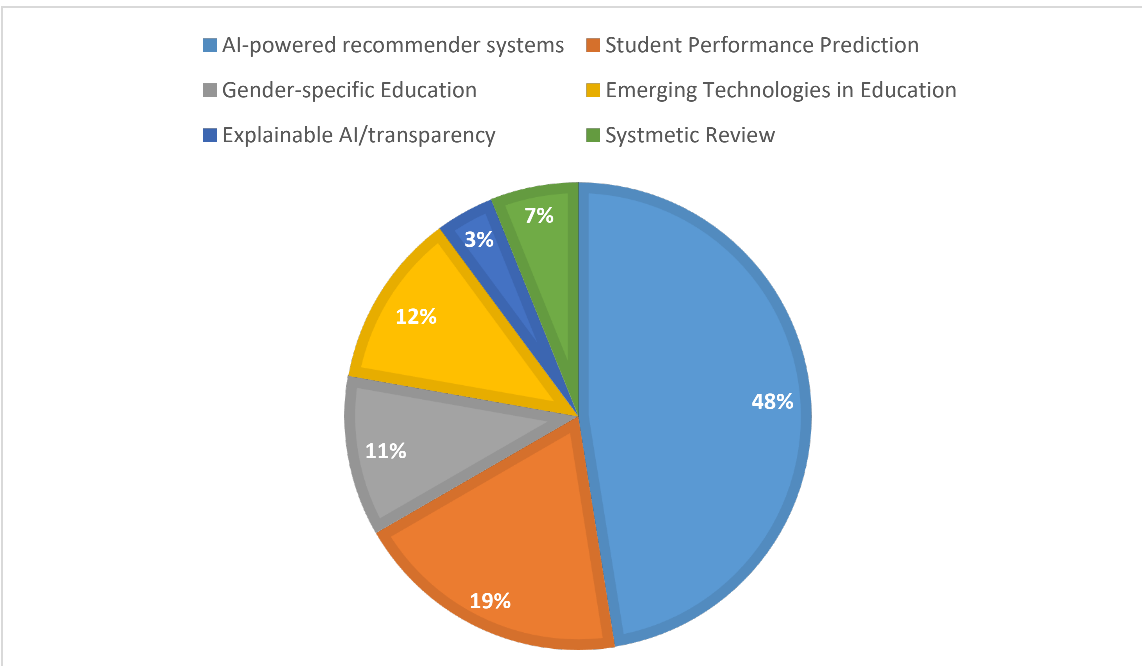


Figure 4. Research focus by problem area

Quality assessment of studies

In addition to the inclusion and exclusion criteria, to ensure the reliability of included studies, a quality assessment was performed using predefined criteria. Each study was assessed against five criteria. A quality assessment criterion (QAC) was established that aligned with the primary objectives of this study. Table 4 presents quality assessment criteria along with a three-point scoring scale – Yes (1 point), Partially (0.5 points), and No (0 points) – to evaluate the clarity of each paper. Table 5 presents the quality of the studies included in our final synthesis. The conceptual framework was inductively developed from trends observed during data extraction and quality measurement. During the first phase, the extracted data of each study were coded into five dimensions: AI technique, educational setting, gender-specificity, scalability concerns, and transparency. The second stage involved thematic coding of the literature to identify recurring limitations, namely the lack of gender-adaptive profiling, a focus on the constraints of low-resource deployment, and a focus on the mechanisms of explainability. The third step involved mapping these clustered gaps to framework design requirements: the gaps found in the literature are directly related to the modules of the proposed ensemble AI engine. This ensures the framework is based on synthesized evidence rather than a speculative creation.

Table 4. Quality assessment description

Criterion	Description
QA1: Clarity of Objectives	Does the study clearly state its research goals/questions?
QA2: Methodological Rigor	Is the methodology (AI model, dataset, evaluation) clearly described and replicable?
QA3: Validity of Results	Are the results supported with empirical data and performance metrics (accuracy, precision, recall, F1, RMSE, etc.)?
QA4: Contribution Relevance	Does the study contribute directly to AI-based recommendations in education (personalization, fairness, scalability, or equity)?
QA5: Consideration of Limitations/Future Work	Does the study acknowledge its limitations and propose directions for improvement?

Table 5. Quality assessment interception

Score range	Interpretation	Number of studies
1	High Quality	28
0.5	Moderate Quality	15
0	Low Quality (limited rigor, missing validation, or weak education focus)	7
Total		50

PHASE 3: PREPARING THE REVIEW REPORT

Analysis of studies

With the growth of digital technology, the integration of AI and recommendation systems in the educational landscape has improved personalized learning, enabled the prediction of student performance, and supported self-education. These systems are supported by AI, using machine learning, deep learning, and knowledge-based frameworks to align educational content with learners' needs. This section is a synthesis of the results from 50 studies across four thematic domains: explainability and transparency in educational AI, AI integration into educational frameworks, machine learning in educational assessment, and recommender systems development.

Explainability and transparency in educational AI

This analysis examines key developments and methodologies, especially focused on enhancing educational access for women. Khosravi et al. (2022) explored the application of Explainable Artificial Intelligence (XAI) in education (XAI-ED), highlighting the growing need for transparency and trust in AI-driven educational tools. The study aimed to develop a framework to guide explainable AI design in education by merging insights from AI, human-computer interaction, and cognitive science. The case study demonstrated success in student-teacher collaboration with the RiPPLE and AcaWriter systems; however, deploying XAI in an educational context remained complex and risky due to the potential for incomplete explanations. Future work can extend the XAI-ED framework to diverse cultural contexts while reducing implementation complexity. Additionally, Saqr and López-Pernas (2024) presented the application of XAI in education to predict student performance and provide personalized feedback. They aimed to explore the utility of instance-level using an explainable AI model for individual performance prediction and its potential to inform decision-making and feedback processes.

Multiple Machine Learning (ML) algorithms, such as Random Forest, SVM, and other Shapley Additive explanations, were applied for local interpretability, achieving 74% accuracy with Random Forest. However, sometimes misprediction indicates human oversight in interpreting AI output. They further suggested developing hybrid AI-driven protocols that can reduce misprediction in performance analytics. Moreover, Gloerfeld et al. (2020) highlighted the potentials and limitations of recommender systems in higher education as AI technologies increasingly support self-study programs. The study aimed to analyze how AI-powered recommender systems enhance self-regulated learning and how constraints affected their adoption. The proposed model combined intelligent feedback, automatic assessment, and a learning strategy within an inquiry-based learning framework to support students through personalized content recommendations. The results suggested that these systems enhance autonomous learning and positively impact student competencies. However, further exploration is needed in education to adopt AI in different disciplines. Their suggestion includes expanding transparent AI models to an interdisciplinary self-regulated learning environment to promote their adoption. Table 6 summarizes key studies that examine explainability and transparency in AI-driven educational systems, with a focus on fairness, trust, and interpretability.

Table 6. Summary of explainability and transparency in educational AI

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(Khosravi et al., 2022)	XAI-ED framework, decision tree, rule-based model, and clustering	3000+ students across three German universities	35% improvement in students' engagement	Complexity in deploying XAI in diverse contexts	Expand personalized explanation and ethical guidelines
(Saqr & López-Pernas, 2024)	Random Forest, XGBoost, Neural network, SVM, Linear Regression	126 medical students	R2 0.30 for Random Forest	Misprediction based on passive engagement requires human validation	Develop hybrid human-AI collaboration to contextualize prediction

All the reviewed studies discussed in this subsection support the idea that explainability in educational AI is technically achievable and pedagogically meaningful. A major shortcoming is, however, evident: current XAI frameworks were constructed and tested in highly resourced, Western institutionalized settings. All the reviewed studies focus on how explainability mechanisms can be implemented to support learners with low levels of digital literacy, or learners in situations where cultural distrust of automated systems is not technical but has a cultural basis. Transparency is rather a precondition than a technical characteristic of an algorithm for women learners in marginalized settings, where transparency can serve to strengthen existing power dynamics. This is the gap that will directly inform the transparency layer that is embedded in the proposed framework. Collectively, these studies demonstrate that explainability (XAI) is technically achievable and essential for building learner trust. However, they remain unresolved regarding how these mechanisms function for low-literacy users or in non-Western cultural contexts where distrust is rooted in socio-power dynamics rather than technical flaws. This unresolved gap justifies the **Embedded XAI Transparency Layer** in the SHELF framework, which focuses on providing culturally adapted and accessible explanations.

Integration of AI in educational frameworks

Personalized learning and adaptive systems have influenced AI to tailor the educational experience, yet they face challenges in mitigating bias, scaling, and validating in the real world. Wang et al. (2024) explored the field of artificial intelligence in education (AIED) and highlighted a limited understanding of its conceptual structure and limited research. Identifying key topics, methods, and theories in AIED through a mixed-method approach and content analysis of recent literature was a key objective. The study map research trends highlighted that adaptive learning and personalized tutoring were the most used methods, and major areas additionally call for greater use of emerging AI technologies. They also highlighted the need to incorporate preschool education and emerging AI technologies into adaptive learning research. Furthermore, Bagunaid et al. (2022) presented the integration of AI into e-learning systems to enhance student assessment and personalized recommendations.

For distance education, e-learning offers flexibility and accessibility to learners. By developing an AI-based Student Assessment and Recommendation system (AISAR), the study aimed to address issues like predicting student performance and providing tailored recommendations. Their proposed model used an RNN to estimate student scores based on engagement and exam results, achieving a high accuracy of 97.21% in performance prediction and outperforming traditional machine learning models. Despite this, the study acknowledges limitations, including reliance on specific datasets and the need for further validation across various educational contexts. AISAR can validate on multi-institutional datasets to improve generalizability. Moreover, Huang et al. (2023) investigated the impact of AI-enabled personalized video recommendations on students' motivation, outcomes, and engagement in a flipped classroom. The study aimed to replace traditional methods with interactive activities, but motivation remains a challenge. A hybrid recommendation system (logistic regression + Bayes theorem) increases motivation among students with moderate motivation, indicating notable improvement. Although the study was limited to short-term effects, a single course may lack its generality. Their work suggested that future studies should explore long-term motivational effects of hybrid recommenders in flipped classrooms.

Additionally, Ahmadian Yazdi et al. (2024) explore educational recommender systems that personalize the selection of learning resources based on individuals' interests. The objective of their study was to develop a dynamic recommender model that adapts to behavioral changes by integrating deep learning techniques, especially the Long Short-Term Memory (LSTM) network. Their approach involves a two-module architecture: one module captures long-term user interest, while the other captures short-term, session-based interest, with weighted attention to recent activities. The model achieved an accuracy of 0.9978, indicating high-quality recommendations. The study was limited to a specific dataset. So future work could explore scalability for larger datasets and real-world validation.

In the same way, Q. Zhang et al. (2021) explored the role of recommender systems in e-learning, focusing on their ability to provide personalized learning experiences as online resources are increasing day by day. The objective was to highlight how content-based, collaborative filtering-based, and knowledge-based approaches enhance personalized learning experiences by adapting to learner preferences and pedagogical requirements. Their proposed framework integrated these methods with deep learning and fuzzy logic to handle uncertainties in learner profiles and item attributes. The study found that hybrid systems outperformed a single-method approach, yet reliance on predefined ontologies in knowledge-based systems compromises their privacy and scalability. They further suggested developing ontology-free knowledge-based recommenders to improve system scalability. However, Wei et al. (2021) explored the integration of AI and educational psychology to develop a personalized online learning recommendation system. They aimed to address the gap by aligning learning resources with students' abilities and applying AI and educational psychology principles through a proposed recommendation scheme. A Linear Upper Confidence Bound-based algorithm was employed to explore personalized coefficients to balance known preferences with new resource discovery. Experimental results demonstrated effectiveness, achieving 70% precision for active students, and showed significant improvement over traditional methods such as user-based and item-based collaborative filtering. However, the study examined static learning abilities over a short period, which does not account for long-term changes in student behavior.

Similarly, Y. Lee et al. (2023) focused on developing a personalized recommender system for education using big data and machine learning techniques to address the limitations of traditional filtering methods based on behavior. They aimed to enhance the accuracy of class recommendations by introducing recommendation indicators extracted from extensive learning activity data. The proposed approach analyzed factors such as student participation, quiz scores, and class engagement. The model showed improvement in F1 score from 0.616 to 0.844 as the volume of data increased, achieving high accuracy. However, the study is limited by its dependence on the quality and completeness of collected data. In the future, real-time bias detection can be integrated into collaborative filtering for fair recommendations.

Following these studies, Oubalahcen and Tamym (2023) explored the application of AI in e-learning recommender systems, focusing on how AI enhances personalized learning experiences. The study aimed to compare AI-based recommender systems in e-learning, contrasting machine learning and deep learning approaches through collaborative filtering, content-based filtering, knowledge-based filtering, and hybrid recommendation systems. Their results showed that AI-driven recommenders enhance personal education; however, they lack data privacy, incur high implementation costs, and lack real-time adaptability. However, limitations include dependence on quality data and computational complexity. Future work can optimize deep learning recommenders for real-time adaptability. Focusing on this, Gao et al. (2021) explored the integration of AI and big data into e-learning and e-education, highlighting their impact on modern education. However, challenges such as data complexity, resource management, and real-time adaptability persist in distance education. They examined fuzzy association rules for public opinion mining and machine cooperation for resource sharing, adaptive index updating, and edge computing for mobile education. Findings demonstrated significant improvements, such as 94-96% accuracy in semantic classification and a reduced 0.22% index redundancy. The paper has limitations, including computational costs and scalability issues in real-world deployment. Future work can focus on developing fuzzy systems that are less costly and require fewer resources to operate.

While generative AI could improve teaching and learning processes, such as feedback and assessment (Giannakos et al., 2025), experts cautioned against being overly optimistic, as there are risks of disruption, ethical concerns, and misuse. They warn against uninformed adoption without evidence and pedagogical soundness, and without ecosystem considerations. Moreover, Zhu et al. (2024) presented an augmented machine learning-based course enrollment recommender system focused on academic advising by manipulating contextual data to enhance course recommendations. The purpose of their study was to develop a hybrid recommendation model integrated with student demographics and course descriptions to improve accuracy. Their proposed model combined matrix factorization and collaborative filtering, augmented with contextual features like a student's job status and course similarity. Also, heatmaps were used to measure weight similarity. The results demonstrated that the model outperforms basic collaborative filtering across diverse university settings. Future work can incorporate gender-specific features to enhance course recommendations for underrepresented groups, such as women in rural areas. Furthermore, Nadimpalli et al. (2025) introduced Nestor, a hybrid AI algorithm that enhanced the learning management system (LMS) by integrating learners' characteristics to provide a personalized learning path. Their study employed a Bayesian Network to integrate psychological models and learning element preferences, enabling data-driven recommendations. The model was evaluated using real and synthetic datasets; its results showed an 8.6-11.6% difference in log-likelihood scores and a reduced hamming distance of up to 36.36% with synthetic data augmentation. Still, the study found that limitations include biases in psychological questionnaires and the challenge of collecting large-scale data. They also highlighted that automating psychological data collection via LMS can reduce questionnaire bias.

Obeid et al. (2018) explored the application of ontology-based recommender systems, as a significant problem in this domain was students' difficulty in selecting major courses and universities. The purpose of this study was to develop a hybrid model combining an ontology-based method with machine learning to provide diverse recommendations for students. The proposed system collected explicit and implicit data from students and clustered the profiles using k-means and hierarchical clustering to identify a suitable major and university for each student. The system improves recommendation accuracy but requires extensive ontology development. Future work could explore semi-automated, domain-adaptive ontology methods to reduce development time and improve system adaptability across different educational disciplines. In the same way, X. Zhou et al. (2025) focused on improving personalized course recommendation, as existing recommendation systems struggled with data sparsity and noise, limiting their accuracy in dynamic educational environments. Their proposed EdGCL, a LightGCN-based model, integrated learnable noise modules and contrastive loss to optimize feature embedding. Results achieved 8.53% higher Recall@10 on MovieLens-1M and a 1.2% improvement in NDCG@10 on MoocCube, compared to a baseline such as SimGCL. Yet its experimental validation was limited to only two datasets. Future work can scale EdGCL across diverse educational datasets while maintaining low computational overhead.

Additionally, Cai and Liu (2024) analyzed the role of adaptive learning systems in personalized education as traditional education struggled with uniform teaching methods. Their purpose was to analyze how adaptive learning systems can enhance personalized education by dynamically adjusting content, difficulty, and feedback. Using Squirrel AI, the system analyzed student data and teaching strategies. Results showed improvement in learning efficiency and engagement. However, technological dependency, privacy concerns, and occasional gaps limited it. They both highlighted that integrating emotional and motivational factors into Squirrel AI's adaptive algorithms can help reduce the occasional gap. Table 7 summarizes key studies on integrating AI into educational frameworks, highlighting the reported system architectures, learning models, and implementation strategies.

Table 7. Summary of integration of artificial intelligence in educational frameworks

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(Wang et al., 2024)	Mixed method approach (bibliometric and content analysis)	AIED scholarly articles	Heavily skewed towards applications	Less inclusion of latest AI technologies and unexplored initial educational stages	Expand research to preschool education
(Bagunaid et al., 2022)	AISAR system (RNN+DBSCAN Clustering+ TMR + reinforcement learning)	32,593 students from OULAD	97.21% accuracy, 0.97 true-positive rate, 0.91 precision	Potential variability in engagement data quality	Integrate real-time adaptive learning features in educational settings
(Huang et al., 2023)	Hybrid recommendation system (Logistic Regression +Bayes' theorem)	102 college students enrolled in a systems programming course	Significant improvement in motivation, engagement, and learning outcomes for moderate level students	Limited to a single course, short-term effects	Explore the long-term impact of AI recommendations and extend the model to diverse disciplines

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(Wei et al., 2021)	LinUCB-based recommendation algorithm	26934 students-video pairs from Coursera	Achieved 70% precision for active students	Considered static abilities, long-term behavior changes are ignored	Incorporate dynamic ability tracking
(Y. Lee et al., 2023)	Collaborative Filtering	Integrated learning activities data from online platforms	F1-score improved from 0.616 to 0.844	Dependence on data quality	Incorporate more indicators and adapt to changing behaviors
(Gao et al., 2021)	Fuzzy association rules, man-machine cooperation	Small-scale implementation	94-96% accuracy in semantic classification	High computational cost, Scalability issues	Integrate real-time adaptive learning
(Nadimpalli et al., 2025)	Hybrid Bayesian Network (Nestor)	Real(N=54) + synthetic (50-50k samples)	LL:8.6-11.6% difference; HD improved by 36.36%	Biased data	Dynamic adoption using log data

Collectively, this body of research demonstrates significant advances in hybrid AI systems, showing they consistently outperform single-method approaches in terms of accuracy and learner engagement. However, a critical unresolved trend persists. High-performing models remain heavily dependent on large datasets and robust digital infrastructure. This dependency creates a 'structural exclusion' effect, in which learners in low-resource, low-connectivity settings who stand to benefit most from personalized education are systematically underserved. Furthermore, current frameworks fail to incorporate vital socio-cultural constraints, such as household responsibilities, time fragmentation, or social permission structures, into their profiling logic. This dual gap directly justifies the architectural logic of the SHELF framework, which is designed to operate in low-bandwidth environments while integrating contextual profiling variables that capture the specific lived realities of women learners.

Machine learning model for educational assessment

Student performance prediction using AI offers more powerful insights for early intervention. Qadir et al. (2025) applied hybrid models to educational data mining for predicting student performance. They aimed to evaluate and compare ensemble methods, including CatBoost, XGBoost, and LightGBM, with deep neural networks (CNN, LSTM, and DNN) to develop more reliable prediction systems. Additionally, used synthetic data generation via the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance and improve model robustness. CatBoost achieved the highest predictive accuracy of 0.7143, while CNN also performed strongly on sequential data. However, overfitting and high computational cost remained concerning, and they suggested enhancing CatBoost's robustness to noisy, imbalanced datasets through synthetic data augmentation.

Additionally, Jin (2025) explored various machine learning techniques, such as Support Vector Regression (SVR), k-Nearest Neighbor (k-NN), and Artificial Neural Network (ANN) with Principal Component Analysis (PCA) to evaluate moral education programs at the university, indicating that SVR is more accurate with an accuracy of 98.2%. The study recommended that an AI-driven moral assessment system lead to fair evolution. The limitation includes reliance on specific machine learning models that may not evaluate all facets of moral education. Future work could expand moral education assessment to include non-Western frameworks. Moreover, Christou et al. (2023) focused on predicting student performance and study duration in higher education using ML techniques. This study aimed to develop a predictive model to improve the accuracy of academic planning and student retention efforts. A hybrid feature construction method combined feature engineering via genetic programming with neural network modeling, achieving the lowest mean squared error in 9 out of 12 regression datasets and the highest accuracy across all datasets; however, the generalizability of the results may be limited by its reliance on a single university. He further suggested applying the FSC4RPF algorithm to multi-institutional datasets for broader validation. In the same way, Silva et al. (2023) explored Convolutional Neural Networks in education, focusing on addressing challenges in teaching and learning. Lack of comprehensive studies on CNN applications tailored to various educational contexts was a problem in the Global South. The SLR aimed to analyze current CNN applications in education by examining publication trends, datasets, model performance, and research gaps, based on a review of 133 papers from IEEE Xplore, ACM Digital Library, and Scopus. Results showed that CNN models achieved high accuracy in various tasks, including emotion recognition and student performance prediction. However, a limitation of the study was the lack of randomized controlled trials and insufficient attention to generalizability.

Focusing on performance prediction, Mohd Talib et al. (2023) discussed that predicting students' academic performance was becoming critical as the institutions promoted holistic development among learners. The study aimed to enhance the identification of students' behavior patterns to better predict and inform academic performance in a timely manner. The authors applied a combination of K-means clustering and SVM classification to cluster students into three classes. The results showed higher accuracy with cluster labels than with traditional methods, leading to improved assessment metrics. A key recommendation from this study is to add psychological attributes to SVM clustering to focus on gender-specific analysis. For instance, Mohd Zaki et al. (2023) focused on classifying students' performance using an SVM with a polynomial kernel based on students' online activities. The main objective of this paper was to develop a model to predict students' performance in identifying at-risk students and understanding factors affecting their success. The author proposed a dataset containing both academic and behavioral features, followed by preprocessing and SVM with polynomial kernels. Results indicated that the SVM model performed well with gamma = 0.005, regularization parameter = 0.1, and degree = 1. They recommended developing gender-balanced SVM kernels to reduce bias in online classification.

Further, Narimani et al. (2024) focused on academic advising in online higher education, as traditional methods relied on generic prerequisites, failing to account for individual performance and suboptimal course selections. The purpose of this study was to integrate association rule mining with grade-based metrics to predict enrollment outcomes. They applied the Apriori algorithm to mine course sequences and introduced grade change to measure performance shifts between courses. Their results showed the model predicted grade drops with 88.44% accuracy and raises with 81.02% for informatics students. However, the limitations include reliance on a fixed threshold for rule mining and low applicability for diverse programs. Future efforts, they argue, should extend association rule mining across diverse academic programs. Assegie et al. (2024) applied machine learning techniques to predict students' performance across various educational domains. Their objective was to compare RF and SVM models for classifying students based on their academic performance, as measured by grade scores. For this purpose, the authors utilized Portuguese SGS datasets, including various demographic and student assessment details. Their results showed that the SVM model outperformed the RF model, particularly in the linear kernel, achieving an accuracy rate of 75.72%. Future work could expand the comparative model to analyze gender-specific barriers and social behavioral factors. Additionally, S. Liu et al. (2024) addressed the inefficiency of existing cognitive diagnosis methods in web-based intelligent education systems. Their study proposed an Inductive Cognitive Diagnosis Model (ICDM) to infer students' masses without retraining, using student-centered graphs and the CAGT process. The model has achieved 96.87% accuracy with an inference time of 26 ms for 355 new students. However, limitations include handling new exercises and reliance on static knowledge concepts. They suggested enhancing ICDM to handle dynamic knowledge concepts in cognitive diagnosis. Table 8 summarizes key studies on machine learning models for educational assessment, including approaches for predicting learner performance, adaptive evaluation, and personalized feedback.

Table 8. Summary for a machine learning model for educational assessment

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(Qadir et al., 2025)	CatBoost, CNN, XGBoost, DNN, RNN, Gradient Boosting, Light	Higher Education Datasets	0.7143 for CatBoost	Performance degrades for noisy and imbalanced data	Expand the realism of dataset
(Jin, 2025)	SVR, ANN, and KNN for moral assessment education	University student's data	SVR has superior predictive accuracy with the lowest error rate	Datasets may not encompass all facets of moral education	Explore various datasets and ML models to enhance reliability
(Christou et al., 2023)	FSC4RPF algorithm	Six departments at the University of Ioannina	Lowest mean square error in 9/12 datasets, highest accuracy in all datasets	Small size of the dataset	Broader dataset includes data from various universities
(Silva et al., 2023)	Various CNN architectures	FER-2013, CK+, MNIST, custom datasets	90.30% accuracy (emotions),92.86% (performance)	Lack of randomized trials	Explore few-shot/transfer learning
(Mohd Talib et al., 2023)	k-means clustering and SVM	200 first-term students	enhancement in accuracy from 76 to 90	Limited size of the dataset	Add psychological and behavioral attributes for gender specification
(Mohd Zaki et al., 2023)	Polynomial SVM	Online behavior of 101 students	Accuracy improved from 90.5 to 95.3%	Imbalanced data	Combine SVM and behavior for females' access in educational technology

Collectively, the reviewed studies demonstrate that performance-prediction models have achieved impressive accuracy in identifying at-risk students. However, these technical gains often conceal a significant equity blind spot; with the notable exception of Mohd Talib et al. (2023), existing models systematically fail to treat gender as a dynamic design variable. By treating all learners as statistically similar, current systems risk furthering educational disparities rather than mitigating them. This remains a critical unresolved issue for women learners whose educational paths are frequently non-linear, often interrupted by household responsibilities, limited mobility, or economic constraints. Traditional measures of performance prediction fail to account for these interruptions, providing a direct justification for the Gender-Sensitive User Layer in the SHELF framework. This component is designed to move beyond population standards by incorporating socio-cultural constraints as primary variables, ensuring that performance feedback is tailored to the personal objectives and lived realities of women learners.

Advancement of recommender systems in education

AI-powered recommender systems hold transformative potential for personalized learning in education. Machado et al. (2025) explored the integration of AI into educational resource recommendation systems within gamified intelligent tutoring systems (ITS). They examined teachers' high cognitive workload by comparing workload perceptions across manual, automated, and semi-automated recommendation scenarios in an experiment with 151 teachers who customized a gamified intelligent tutoring system. Results showed significantly less workload, having p=0.024 with female teachers and higher education instructors, but a limited simulation environment and non-representative sampling may limit the findings.

Similarly, Rahayu et al. (2022) review the use of ontologies in e-learning recommender systems, analyzing existing research on ontology-based recommendation methods, development practices, and evaluation strategies. The study reviews 28 articles (2010-2020) that use OWL and AI techniques to examine how ontologies model learner profiles and content. Results showed positive outcomes from personalized learning, mostly from prototype systems, though a small sample size and a lack of real-world validation remain limitations. However, standardized ontological frameworks for cross-platform e-learning can be focused on in the future. Following Urdaneta-Ponte et al. (2021), a systematic review of recommendation systems in education was presented, addressing information overload and enhanced learning by relevant educational resources. They aimed to analyze existing education recommendation systems, developmental approaches, recommended elements, and gaps to guide future research. The review analyzed 98 articles from major databases (IEEE, ACM, and WOS) published between 2015 and 2020, categorizing recommendation systems by education type and techniques .59% of RSs targeted formal education,33% filtering and hybrid approaches 21% while machine learning techniques gained prominence in recent years, but insufficient details on platform implementation were included in their limitations.

The future can expand recommendation systems research to informal education. In this way, Raj and Renumol (2022) focused on personalized learning and adaptive content recommenders in e-learning environments. The system aimed to provide learning resources tailored to individuals' learning styles, preferences, and knowledge. This study aimed to systematically review research conducted from 2015 to 2020 on the development, implementation, and evaluation of recommender systems. The data consisted of 52 relevant publications focused on recommendation strategies and evaluation metrics, which evaluated prediction accuracy and considered learner satisfaction, achieving a success rate of 80-90%. A limitation of this study was that it focused on non-cognitive learner attributes. A key recommendation for this study is to incorporate cognitive attributes into non-cognitive recommender systems. Moreover, Weber et al. (2022) focused on SIDDATA, a web-based recommendation system for higher education aimed at developing a scalable platform that enables the implementation of recommendations and evaluation of students' learning goals. This utilized a modular Django framework integrated with Stud.IP learning management systems have been effectively tested on over 3000 students and have demonstrated their practical usability and adaptability. However, deploying across different LMS is challenging due to variations in systems and data formats. Future work can develop universal API adapters or SIDDATA to support heterogeneous LMS platforms.

Y. Zhou et al. (2025) addressed the challenges of accurately evaluating teaching quality in a blended classroom. They aimed to develop a data-driven, intelligent framework that integrates multisource and multi-model data for comprehensive assessment. The model captures complex relationships and applies attention mechanisms using Graph Attention Networks (GAT) for dynamic evaluation. The model achieved 90% accuracy in western China; however, large-scale deployment and real-time scalability remained challenging. Optimize GAT models for edge devices to enable real-time feedback on teaching quality in the future. This investigation by McGrath et al. (2023) with 194 university teachers showed that they are more inclined to use AI to achieve equitable outcomes for first-generation students and those with learning disabilities. It also uncovered significant concerns about a lack of knowledge, fairness, and the use of AI in Higher Education. Only in the archetypical and disability scenarios, but not in the first-generation scenario, did gender, age, faculty, and academic position affect responses. Furthermore, Benhamdi et al. (2017) explored personalized recommender systems in e-learning environments, as traditional static systems fail to adapt to individual learner needs. The study aimed to enhance the quality of recommendations by incorporating learners' preferences, background knowledge, and memory capacity. The proposed NPR-eL system employed a custom taxonomy and multi-dimensional similarity metric to enable effective clustering of learners and accurate prediction of learning materials. Results showed high accuracy in recommendations, demonstrating significant improvements in precision, recall, and F1 metrics when memory capacity was included. However, automating taxonomy generation via Natural Language Processing (NLP) can scale NPR-El to massive OER repositories.

Cha et al. (2024) investigated an AI-based course recommender system focused on students' course selection, a problem often considered in course evaluation. Their study aimed to understand students' perceptions and expectations regarding how CRS can support them at different phases of the decision-making process. A semi-structured interview and speed dating with storyboards were employed to explore students' views and analyze the role of AI in the research and evaluation phases. Results showed that students value AI for streamlining information retrieval and that it served as a benchmark in the evaluation phase. Future work can involve real-world interaction to expand the research. Following the above perspective, J. Zhang et al. (2023) explored an AI-based course recommendation system for higher education platforms, focusing on enhancing personalized learning through advanced algorithms. This study aimed to overcome the inefficiency of traditional recommendation methods by developing a system that integrated multiple AI techniques. The proposed model evaluated four algorithms: content-based, collaborative filtering, hybrid, and deep learning-based, and analyzed their accuracy, coverage, and data sparsity. Their results showed that collaborative filtering achieved an accuracy of 96%. However, its limitations include sensitivity to data sparsity and high computational demands for deep learning. However, federated learning techniques can be used to combat data sparsity in hybrid recommendation systems.

Dhar and Jodder (2020) focused on developing an effective recommendation system using machine learning classification algorithms to estimate the best education program for students after passing class 10. By analyzing students' performance across 10 classes, the study aimed to address the difficulty students face in selecting the most suitable higher education path based on their academic history, leading to potential mismatches in career choice. The proposed approach involved data collection, preprocessing, feature selection, and multiple machine learning models, achieving 85% accuracy in forecasting students' performance metrics, recommending the most suitable program, and demonstrating high effectiveness. However, the study is limited by its reliance on data from a single geographic region and a specific educational institution. They suggested expanding ML models to include socio-economic factors beyond the grades.

Similarly, P. Zhang & Zhang (2024) focused on improving student performance and learning evolution by integrating advanced machine learning techniques in the education sector. The authors aimed to analyze the effectiveness of combined Support Vector Machine (SVM) and Teaching-Learning-Based Optimization (TLBO) to enhance predictive accuracy in the educational sector. The proposed model, SVM + TLBO, achieved superior predictive performance for student learning compared to traditional methods such as decision trees and neural networks. By extending SVM + TLBO to analyze gender barriers in education sectors, we can improve women's access to education, particularly in STEM fields. The above studies (Zulfiqar & Kuskoff, 2024) explored the relationship between education and women's empowerment in Pakistan. They reviewed the international gender-parity approach, which assumed that education empowered women by enabling them to challenge traditional norms. The problem lies in oversimplified gender-parity approaches that ignore contextual barriers, such as societal and familial pressures. An ecological approach was employed, conducting semi-structured interviews with 35 highly educated Pakistani women. Results showed that education provided women with employment opportunities, but it also burdened them with work and domestic duties. Future work can develop systems that consider domestic factors alongside education to provide educational opportunities for women.

Kinol et al. (2023) evaluated the predictive accuracy of the AI method for women's education post-independence. The objective of this study was to compare the performance of AlexNet with linear regression to identify which method predicts better. A semi-supervised technique was employed, indicating that AlexNet outperformed linear regression, achieving an accuracy of approximately 84.22% compared to 80.07%. Future work could expand the dataset and refine the model for better generalization. The study by Kuleto et al. (2021) aimed at addressing some of the major issues, such as a common understanding of AI and ML by the research bases, best practices in the use of AI and

ML in HEI, knowledge of AI and ML by the students, and their attitudes towards the opportunities and challenges of AI and ML in HEI.

Following this, Pasha (2024) examined Pakistan's preserved gender difference in education, with girls facing lower access to education and reduced income as compared to boys. The purpose of this study was to analyze how household income and socioeconomic characteristics impact educational disparities. Using a two-model approach, an ordered logit model for educational attainment and a logit model for current enrollment were applied to cross-sectional data from the PSLM survey from 2005 to 2019. Findings indicated that household income significantly increased the likelihood of a higher educational level, with boys having more access to education. Future work should design income-sensitive policies for girls to access education. Moreover, Askari et al. (2022) examined women's education in Pakistan, focusing on gender disparities in access to education and literacy, highlighting the urban-rural divide and systemic barriers. Their objective was to analyze the challenges and identify the opportunities for improvement through technology and government. The study demonstrated the key issue of the scarcity of female teachers and cultural biases using a qualitative literature review of Pakistan. Results showed that urban female literacy was 80%, compared with 20% in rural areas. Future work should explore localized intervention and longitudinal impact assessment for females in education.

Emerging technologies such as AI, IoT, and edge computing are significantly impacting the field of education. Steele (2023) explored the integration of AI in education, particularly tools like ChatGPT, to enhance learning while addressing ethical concerns. The study found that 80% of students had an effective understanding of complex topics and improved their information literacy through the use of AI. The study lacks empirical data on how learners use ChatGPT, which would enable future work to investigate ChatGPT's longitudinal impact on information literacy. Moreover, Veeramanickam et al. (2023) focused on enhancing personalized learning by developing a smart education system that used IoT and Case-Based Reasoning (CBR). Their study aimed to design a model that dynamically recommends relevant learning materials by tracking student engagement and performance. The data was collected via IoT sensors, processed multimedia resource access, and then applied CBR to map past cases and generate personalized recommendations. The study indicated correct responses ranging from 43.67% to 75.32%, demonstrating the model's effectiveness. They recommended expanding IoT-CBR systems to support multilingual learning resources. Renţea et al. (2025) examined the effectiveness of visualization in machine learning education, focused on enhancing students' learning outcomes. Their study addressed the challenge of teaching complex mathematical concepts and split students into groups to expose them to both interactive and static visualizations of gradient descent and PCA. PCA understanding improved ($p = 0.0032$), but did not affect the motivation level. Limitations include a small sample size and short-term assessments, which can be overcome through large-scale implementation to validate the findings.

Chang et al. (2022) focused on evaluating the challenge of prioritizing key factors of e-learning systems. Their main objective was to develop a structured framework for assessing the key factors of e-learning systems, employing the Analytic Hierarchy Process (AHP) and expert input to hierarchically weight factors such as IoT integration, teaching monitoring, and smart classroom functions, enabling systematic prioritization. The system identified educational resource optimization (37%) and personalized teaching (24%) as the most vital factors. However, the study primarily focused on technological factors; pedagogical considerations are still overlooked. Validating the AHP framework with pedagogical experts in the future can balance teaching priorities. In the same vein, Afreen et al. (2025) explored the application of path-based recommendation methods using knowledge graphs to address data sparsity and the need for transparent reasoning in recommender systems in education. They aimed to evaluate the generalizability of KG-based reasoning methods for course recommendations by comparing their utility, beyond-utility, and explainability factors. The authors constructed KGs from three educational datasets and found that PEARLM achieved the highest utility, while PGRP outperformed in diversity. Future work can improve KG completeness via automated ontology refinement for PEARLM. Following this, Y. Liu et al. (2024) investigated the role of machine learning in digitized physical education learning (DPEL) focused on student engagement and performance assessment. The study aimed to develop an intelligent feedback system to enhance DPEL outcomes. They employed a mixed-method approach to the proposed Student Engagement Learning and Feedback System (SELFS), integrating machine learning and natural language processing to predict low engagement. The results indicated a strong correlation between dwell time and visit frequency (ϱ = 0.835) and 81.25% accuracy. Table 9 summarizes key studies on the advancement of recommender systems in education, highlighting personalized learning, adaptive content delivery, and learner engagement strategies.

Collectively, the literature indicates that the field of educational recommender systems is maturing, with hybrid methods, knowledge graphs, and neural networks significantly increasing accuracy and learner satisfaction. However, three convergent gaps remain unresolved that compromise the equity potential of this technical progress: systematic gender blindness in learner profiling, evaluation metrics that prioritize mathematical accuracy over long-term empowerment, and an 'infrastructure-first' design that marginalizes learners in low-bandwidth regions (Askari et al., 2022; Pasha, 2024; Zulfiqar & Kuskoff, 2024). Most critically, current systems fail to account for 'time fragmentation' – the non-linear study patterns necessitated by women's domestic responsibilities. This major unresolved gap provides a direct justification for the **Awareness and Training Modules** in the SHELF framework, which use 10-15-minute micro-learning sessions rather than presupposing long, uninterrupted blocks of study time. Furthermore, the systematic failure of existing systems to address socio-cultural barriers justifies SHELF's integration of multimedia-based content and gender-sensitive variables, ensuring the framework serves not just as a technical tool, but as a context-aware intervention for women's educational empowerment.

Table 9. Summary of advancements in recommender systems in education

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(Urdaneta-Ponte et al., 2021)	Collaborative, hybrid, content-based, machine learning	98 articles from IEEE, ACM, Scopus, and WOS	59% targeted formal education: ML adoption increased	76% studies lack platform implementation details	Broder's educational levels and real-world validations
(Y. Zhou et al., 2025)	Graph Attention Network (GAT)	70000 primary and secondary students and 8000 teachers in Western China	90% accuracy in teaching quality, HR personalized recommendation 0.90 ± 0.015	Issue in ensuring real-time scalability	Integrate lightweight AI model with edge computing for a recommendation mechanism
(J. Zhang et al., 2023)	Hybrid model (Content-based, CF, Hybrid model, deep learning)	University course data	CF:98% accuracy, Hybrid:82% coverage	High computational cost and sparsity sensitivity	Enhance deep learning efficiency and address sparsity issues
(Veeramanickam et al., 2023)	Case-Based Reasoning STS with IOT	500-1600 students across different subjects	Quiz improvement from 43.67% to 75.32%	Higher RMSE for larger group; limited to specific learning resources	Expand it to diverse resources

Reference	Model/ methodology	Dataset	Accuracy/ results	Limitations	Future work
(P. Zhang & Zhang, 2024)	SVM+TLBO	297 students' survey	Outperformed predictive performance	Limited size of the dataset	Extend SVM +TLBO to analyze gender barriers in the education sectors
(Y. Liu et al., 2024)	Random forest prediction, EM clustering	64 undergraduate students	81.25% prediction accuracy, RMSE = 2.06	Short-term study, small dataset	Expand the sample into a large dataset
(Renţea et al., 2025)	Controlled experiment	40 CS students (grouped into 20)	PCA: p=0.0032(M=0.52 VS 0.32)	Self-reported qualitative data	Refine real-time feedback
(Afreen et al., 2025)	PEARLM (Generative)	COCO, MOOPer, MOOCube	NDCG: 0.41 (COCO), 0.60 (MOOPer), 0.39 (MOOCube) 310% higher NDCG vs. best baseline (BPRFM)	Susceptible to hallucination (though less than PLM); relies on KG completeness	Improve adaptability to diverse data
(Assegie et al., 2024)	SVM+RF	Portuguese SGS datasets	An accuracy rate of 75.72% achieved	Performance dependency on the dataset	Explore social behavioral factors that affect women learners
(Cai & Liu, 2024)	Squirrel AI (Adaptive Learning System)	K12 student data	Improved learning efficiency	Reliance on AI/data, privacy risks	Enhance AI precision, integrate emotional/motivational factors
(Kinol et al., 2023)	AlexNet+Linear Regression	Women's education data for post-independence	84.22% accuracy (AlexNet)	External validations limited	Incorporate non-linear models
(Askari et al., 2022)	Qualitative Literature Review	Pakistan demographics and academic reports	Urban female literacy 80%, rural:20%	Reliance on secondary data	Localized interventions

PROPOSED CONCEPTUAL FRAMEWORK

The proposed framework, SHELF (Figure 5), is an ensemble framework designed to directly address the limitations and gaps identified across the 50 reviewed studies. The core objective of this framework is to promote personalized, equitable, and scalable learning with a special focus on empowering women in education. Each component of the framework is grounded in the evidence synthesized in the literature review. The gender-sensitive input layer responds to the finding that no existing system collects household responsibilities, time availability, or socio-economic constraints as profiling variables (Askari et al., 2022; Pasha, 2024; Zulfiqar & Kuskoff, 2024). The lightweight, low-bandwidth architecture responds to evidence that device and connectivity constraints disproportionately exclude rural women (Cai & Liu, 2024; Renţea et al., 2025). The embedded XAI transparency layer directly addresses the critical finding that only 4% of reviewed studies address explainability directions (Ahmadian Yazdi et al., 2024; Benhamdi et al., 2017). The ensemble AI engine, which combines collaborative filtering, content-based filtering, and reinforcement learning, responds to evidence that hybrid systems consistently outperform single-method approaches while reducing data sparsity (Benhamdi et al., 2017; Dhar & Jodder, 2020; X. Zhou et al., 2025). Finally, the real-time feedback and assessment module addresses the recurring limitation of static models that cannot adapt to women's changing learning contexts (Ahmadian Yazdi et al., 2024; Cai & Liu, 2024). The hybrid AI model guides women through structured learning pathways, providing continuous performance feedback to promote inclusiveness in recommendations and embed transparency and fairness throughout, counter algorithmic bias, and build learner trust.

INPUT LAYER

This layer gathers all data comprehensively and serves as a foundation. It consists of two components: the user and data foundation layers.

User layer

In addition to standard demographic variables, the user profile captures time availability per day (accounting for domestic responsibilities); preferred learning modality (audio/visual/text based on literacy level); device type and connectivity tier; language preference, including regional dialect, household income bracket (for content cost filtering); and social permission level (degree of family/community support for education). These variables are absent from all reviewed recommendation systems and are designed specifically to represent the lived constraints of women learners in marginalized settings.

Data foundation layer

This component serves as the system's knowledge repository, storing learner profile data, educational trends, and external data feeds.

ENSEMBLE AI ENGINE

Awareness module

For new learners with limited prior exposure to digital education, this module provides culturally adapted onboarding content in local languages, uses female role models and locally relevant success narratives, and sets initial expectations that are achievable within time-fragmented learning schedules, like micro-sessions of 10–15 minutes rather than hour-long blocks.

Training module

This module, called the instructional component, aims to deliver personalized, adaptive lessons through virtual memory support, with his explanation and motivation. The goal of this module is to produce knowledge gain and confidence in the path.

Skill development

This module uses training progress and learner goals to translate theoretical knowledge into hands-on ability. It provides hands-on coding tasks and evaluates learners' performance to unlock advanced challenges. This module results in verified coding skills and a project portfolio.

Recommendation system

This module acts as a long-term strategic guide for learners by analyzing performance and engagement to recommend more effective learning. Pathways created a roadmap for microlearning to maintain student momentum and advanced topics to accelerate growth in specific fields of learning.

Feedback & assessment module

This component formalizes the evaluation process. It logs mission evaluation and interaction data to perform auto-assessment and detect gaps. After evaluation, it provides more directions to failure and, if mastery is achieved, it validates the skill.

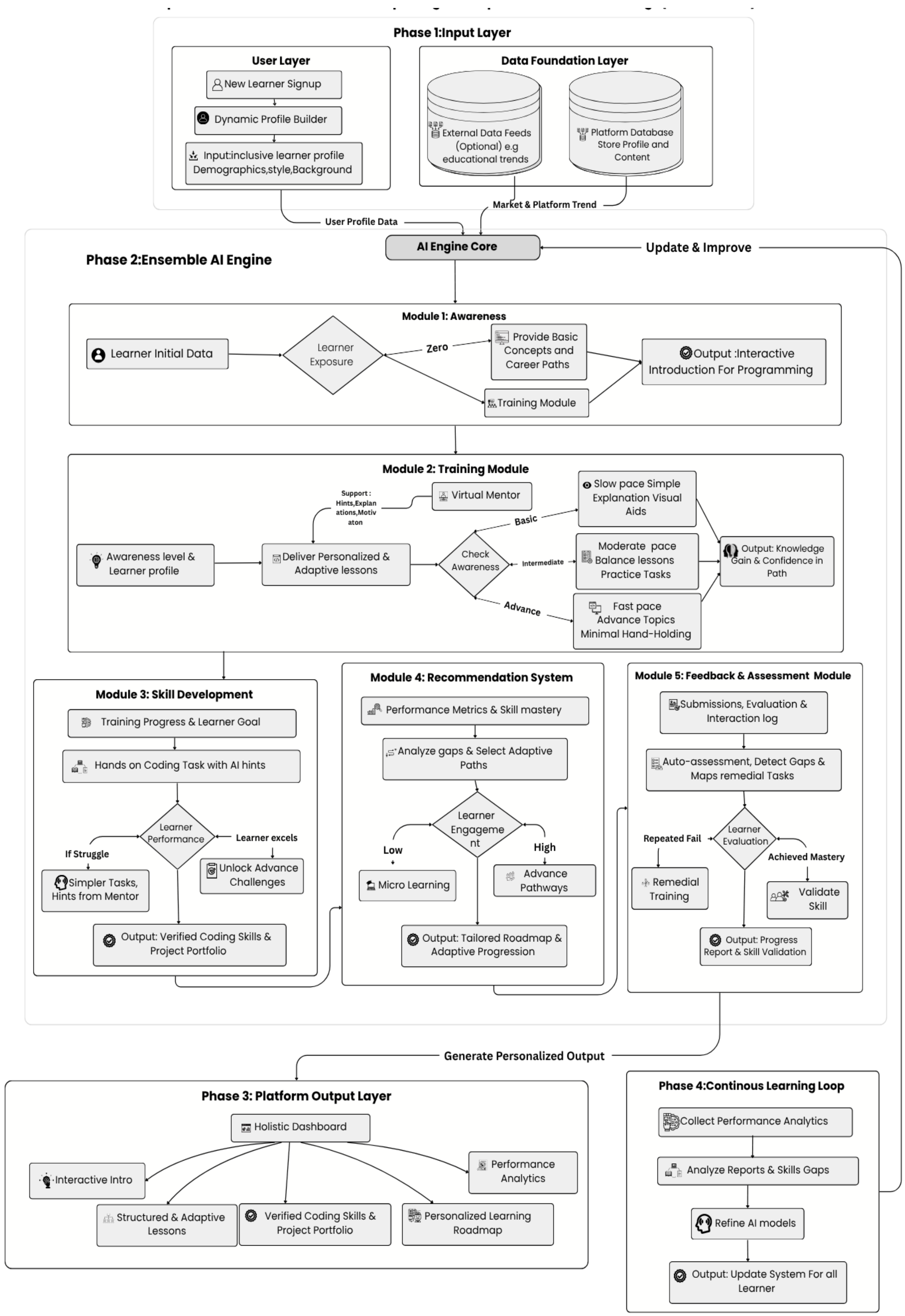


Figure 5. Proposed conceptual framework for SHELF

Platform Output Layer

The Holistic Dashboard is a central hub for learners. It displays all the personalized elements generated by the AI engine, having an interactive introduction, adaptive lessons, verified coding skills, a personalized learning roadmap, and performance analytics.

Continuous Learning Loop

This loop continuously refines the AI model by collecting performance analytics from the platform's output. It analyzes reports and skills gaps across the entire user base and refines the model. The final output is an updated system for all learners, making the platform more accurate and supportive with every cycle.

This model will benefit female learners through context-sensitive data, local-language support, and privacy, and will also provide data-driven decision support for curriculum design in institutions. Additionally, it provides a blueprint for AI deployment in education to promote educational equity and support women's empowerment through learning.

To provide a clear link between our empirical findings and the proposed architecture, Table 10 summarizes how the SHELF framework addresses the specific gaps identified in the analysis of studies. This mapping ensures that every design component is grounded in evidence rather than speculative reasoning, directly addressing the technical and socio-cultural barriers faced by women learners. As illustrated in the mapping, the SHELF framework transitions from a one-size-fits-all approach to a context-aware model. By operationalizing these design imperatives, the framework serves not only as a technical tool but also as a socially responsive intervention for educational equity.

Table 10. Mapping of empirical evidence to SHELF design components

Empirical gap identified in analysis	Supporting evidence (citations)	Corresponding SHELF component
Structural Exclusion: High bandwidth and modern device requirements.	Cai and Liu (2024); Rențea et al. (2025)	Architectural Logic: Low-bandwidth, offline-first adaptive engine.
Time Fragmentation: Ignoring domestic duties and non-linear paths.	Pasha (2024); Zulfiqar and Kuskoff (2024)	Awareness & Training Modules: 10–15-minute micro-learning sessions.
Gender-Blind Profiling: Lack of socio-cultural variables in AI models.	Mohd Talib et al. (2023); Askari et al. (2022)	Input Layer (User Layer): Gender-sensitive profiling variables.
Transparency Gap: Lack of XAI for low-literacy users.	Khosravi et al. (2022); Saqr and López-Pernas (2024)	Embedded XAI Layer: Contextual and local language explanations.
Static Feedback: Failure to adapt to changing learner contexts.	Ahmadian Yazdi et al. (2024); Bagunaid et al. (2022)	Continuous Learning Loop: Real-time performance analytics & model refinement.

Answers to Research Questions

This study examined the role of AI-based recommendation systems in education by analyzing 50 research papers. The study explored how AI technologies are currently utilized, their limitations, and future directions to support equitable education, especially for women in marginalized contexts. This section directly addresses the research questions within the broader context of existing literature.

For RQ1, the literature highlighted the growing demand for personalized educational content. Studies have shown that recommendation systems improve learning outcomes by providing learners with content based on their individual preferences (Urdaneta-Ponte et al., 2021), cognitive style (Weber et al., 2022), and prior performance. Despite these benefits, the costs of transparency, fairness (Gloerfeld et al., 2020), and bias in algorithmic decisions create ethical concerns when learners do not fully understand or trust systems' choices. To overcome these shortcomings, AI-based recommendation systems need to incorporate explainability frameworks that enable users to understand how recommendations are made. Moreover, algorithms that are sensitive to fairness must be implemented at the architectural level to identify and counter bias across user groups. Thus, to design an effective system, a transition towards ethically based architectures that reconcile personalization with transparency and fairness must be made. This understanding directly guides the addition of explainability and fairness features to the framework proposed.

RQ2 focuses on methodologies to improve the adaptability and scalability of education systems for women. As the learning context varied, scalability and adaptability remained technical bottlenecks in the adoption of recommendation systems. AI recommendation systems enhance learning outcomes by tailoring content to personal preferences, cognitive styles, and prior results (Urdaneta-Ponte et al., 2021; Weber et al., 2022). But ethical issues arise when decisions are not made transparently, and learners are unable to comprehend or have confidence in system selections (Gloerfeld et al., 2020). Existing systems prioritize personalization at the expense of transparency. Architectural-level fairness-conscious algorithms and explainable AI (XAI) mechanisms remain underrepresented, leaving current methodologies ill-suited for real-world implementation across a variety of educational settings. To be adaptable, systems should include context-aware modeling that accounts for environmental constraints, such as connectivity and device limitations, as well as user behavior patterns. Effective systems require a transition toward ethically grounded architectures that reconcile accuracy with interpretability, as reflected in the proposed framework (SHELF).

RQ3 is central to this review contribution. The nature of recommendation systems has to shift to fundamentally alter their strategy, as gender ceases to be a demographic filter and becomes a design variable, directly integrating socio-cultural constraints into each level of the system. Evidence in this review records five distinct challenges that are part of the current systems being systematically overlooked: household roles that divide available educational time (Pasha, 2024); income dependency in which household income predicts access to education better than ability; socio-cultural constraints such as the need to seek family permission and limited mobility (Askari et al., 2022); a digital literacy gap with rural female literacy rates of five key features of design are required to overcome these barriers and, therefore, recommendation systems need to have them.

To begin with, time-conscious scheduling enables students to define how much time they have each day (between 15 minutes and several hours), and the system will suggest micro-learning units that can be accommodated within those limits, rather than presupposing continuous study periods. Second, income-sensitive filtering marks data-cost resources and considers offline or low-bandwidth options to learners in lower-income households. Third, permission-aware design adds optional private mode settings with no notifications and stealth learning features to situations where family/community permission to learn is conditional or limited. Fourth, multimodal content delivery offers audio-first and visual-first options in addition to more traditional text, and supports regional dialects rather than assuming standard language knowledge or text literacy. Fifth, the flexible learning pathways provide pause and resume without penalty and permit learners to adjust the pathway's intensity based on life situations, since women can have more or less time to learn depending on household, harvest, or family needs. Importantly, these properties cannot be optional extensions to the system architecture; they have to be core. These needs are operationalized in the proposed SHELF framework through its gender-sensitive input layer, which directly measures variables unaddressed by all the analyzed systems: time availability per day (measured by a sliding scale during onboarding), household responsibility load (modeled by the number of dependents and self-reported busyness), type of device and connection level, income bracket to content cost filter, social permission level (self-reported These variables can be modeled as first-class inputs instead of optional metadata, allowing recommendation systems to dynamically adapt to realities of women learners in underserved areas, turning them into technically functioning but contextually unfitting instruments into actually helpful educational collaborators.

RECOMMENDATIONS

Based on the findings and discussion above, the following recommendations are proposed for researchers and practitioners to inform future research and system engineering. These suggestions will promote equity, scalability, and inclusivity, especially among women learners in low-resource settings.

INTEGRATE ALGORITHMIC FAIRNESS AND GENDER-AWARE PERSONALIZATION

AI-based recommendation systems must also include fairness-aware algorithms and explicitly address gender-specific factors, such as time, device availability, and socio-economic factors. The literature has already shown that deep learning models can effectively predict women's educational outcomes; however, economic and cultural factors are not adequately addressed in existing AI systems (Askari et al., 2022; Kinol et al., 2023). To achieve fair results, systems must go beyond general customization and implement context-sensitive recommendation practices that are responsive to the realities of women learners. Moreover, it is necessary to incorporate explainable AI (XAI) mechanisms to enhance transparency, build user trust, and enable informed decision-making.

DESIGN FOR LOW-RESOURCE AND DIVERSE EDUCATIONAL CONTEXT

Recommendation systems need to be low-bandwidth-friendly and engineered to work under poor connectivity and infrastructure. This involves creating lightweight architectures, providing offline access, and adapting localized content through linguistic and socio-cultural customization (Assegie et al., 2024). These design considerations play a significant role in enhancing accessibility and scalability in underserved and rural areas.

HYBRID AI MODELS WITH REAL-TIME FEEDBACK MECHANISM

LightGCN and knowledge graphs are used to overcome data noise and sparsity (Afreen et al., 2025; X. Zhou et al., 2025). Next-generation systems must leverage hybrid AI techniques, including collaborative filtering, content-based filtering, and knowledge-based approaches, to improve recommendation quality and reduce data sparsity. Moreover, the provision of real-time feedback and adaptive learning mechanisms can facilitate consistent system enhancements and individualized learning adaptation. This proves advantageous, especially for women learners who need flexible learning directions due to competing tasks.

CAREER ALIGNMENT AND SOCIAL LEARNING OPPORTUNITIES

Recommendation systems must be able to go beyond academic content to provide career-related advice and skill-building. The motivation and engagement of learners can be improved with the integration of AI-driven mentorship, career suggestions, and peer learning networks. These characteristics may empower women by aligning educational opportunities with work and future objectives. All these suggestions point to the necessity of AI-based educational systems that are not merely technically efficient but also fair, contextual, and socially-responsive. Future systems can be designed to better serve women learners by incorporating equity, flexibility, and user-friendliness to help alleviate educational disparities.

CONCLUSION

This paper makes an evidence-based contribution to the field of AI in education by systematically reviewing and critically analyzing 50 studies on AI-based recommendation systems, with a particular focus on personalization, fairness, and gender inclusion. It identifies key challenges, methodological gaps, and limitations in existing systems, especially in addressing the educational needs of women in low-resource and marginalized contexts. The results of a methodical review of literature published in 2015-2025 indicate that the field has achieved significant technical progress in personalized learning systems, but it still falls short on three key equity aspects: algorithmic transparency, gender-sensitive personalization, and the potential for scaling in low-resource contexts. These are not marginal gaps but constitute the core of inclusion or exclusion of AI-based educational systems. The unique value of the proposed hybrid AI framework SHELF is not the innovation of the specific techniques used, e.g., hybrid recommenders, collaborative filtering, or ensemble methods, but the logic behind it.

All parts of the framework are clearly grounded in the empirically identified gaps, and its gender-sensitive adaptations are grounded in the recorded socio-cultural limitations. Moreover, considerations of fairness are integrated into the architecture rather than treated as an optional improvement. In this regard, the framework transcends by applying a gender lens to current systems; it transforms the functional aim of AI-powered educational platforms for learners exposed to structural imbalances. Practically, the study has several implications. Equity-oriented evaluation criteria must be included with traditional performance metrics by policymakers and educational institutions in low-resource settings, and gender-disaggregated reporting must be required in AI-enabled educational systems. Educational designers must abandon the concept of generic personalization in favor of context-sensitive solutions that overcome time limitations, connectivity issues, and socio-cultural impediments. The small sample of explainability (4% of the studies reviewed) is a serious gap in research, as researchers in AI believe they must provide clear and open AI models that can be easily comprehended and trusted by various user populations, and even those with low literacy levels.

Although the suggested framework has a good conceptual background, it needs empirical testing. Further studies must focus on: (1) pilot applications in different low-resource contexts to evaluate the extent to which the research methodology is applicable; (2) constructing equity-grounded measures of performance, which can be used alongside the conventional measures of accuracy; and (3) longitudinal studies on the educational and socio-economic outcomes in the long-term regarding women learners using gender-adaptive systems. To sum up, AI-based education technologies are at a crossroads. Even though they have shown promise to revolutionize learning outcomes, their effect will be minimal -and even have the opposite effect as they increase inequalities- unless equity is designed as an essential design element. The framework proposed in this paper is one such step in this direction, and it calls for the use of AI systems that not only work within the current inequalities but also help eliminate them.

REFERENCES

Adeleke, O., & McSharry, P. E. (2022). Female enrollment, child mortality and corruption are good predictors of a country's UN Education Index. *International Journal of Educational Development*, *90*, 102561. https://doi.org/10.1016/j.ijedudev.2022.102561

Afreen, N., Balloccu, G., Boratto, L., Fenu, G., Malloci, F. M., Marras, M., & Martis, A. G. (2025). Can path-based explainable recommendation methods based on knowledge graphs generalize for personalized education? *Proceedings of the 33rd ACM Conference on User Modeling, Adaptation and Personalization* (pp. 273–278). Association for Computing Machinery. https://doi.org/10.1145/3699682.3728323

Ahmadian Yazdi, H., Seyyed Mahdavi, S. J., & Ahmadian Yazdi, H. (2024). Dynamic educational recommender system based on Improved LSTM neural network. *Scientific Reports*, *14*, 4381. https://doi.org/10.1038/s41598-024-54729-y

Aladi, C. C. (2024). IT Higher Education Teachers and Trust in AI-Enabled Ed-Tech: Implications for Adoption of AI in Higher Education. *Proceedings of the 2024 Computers and People Research Conference* (Article 19). Association for Computing Machinery. https://doi.org/10.1145/3632634.3655852

Askari, A., Jawed, A., & Askari, S. (2022). Women education in Pakistan: Challenges and opportunities. *International Journal of Womens Empowerment*, *8*, 27-32. https://doi.org/10.29052/2413-4252.v8.i1.2022.27-32

Assegie, T. A., Salau, A. O., Chhabra, G., Kaushik, K., & Braide, S. L. (2024). Evaluation of random forest and support vector machine models in educational data mining. *Proceedings of the 2nd International Conference on Advancement in Computation & Computer Technologies, Gharuan, India*, 131-135. https://doi.org/10.1109/InCACCT61598.2024.10551110

Bagunaid, W., Chilamkurti, N., & Veeraraghavan, P. (2022). AISAR: Artificial intelligence-based student assessment and recommendation system for e-learning in big data. *Sustainability*, *14*(17), 10551. https://doi.org/10.3390/su141710551

Benhamdi, S., Babouri, A., & Chiky, R. (2017). Personalized recommender system for e-Learning environment. *Education and Information Technologies*, *22*, 1455-1477. https://doi.org/10.1007/s10639-016-9504-y

Cai, Q., & Liu, Q. (2024). Exploration and practice of personalized education based on adaptive learning systems. *Proceedings of the 2024 International Symposium on Artificial Intelligence for Education* (pp. 54-58). Association for Computing Machinery. https://doi.org/10.1145/3700297.3700307

Cha, S., Loeser, M., & Seo, K. (2024). The impact of AI-based course-recommender system on students' course-selection decision-making process. *Applied Sciences*, *14*(9), 3672. https://doi.org/10.3390/app14093672

Chang, S.-H., Chang, C.-W., Chen, H.-H., & Wu, M.-C. (2022). Measuring the importance of smart E-learning education system. *Proceedings of the 6th International Conference on Digital Technology in Education* (pp. 421-428). Association for Computing Machinery. https://doi.org/10.1145/3568739.3568810

Christou, V., Tsoulos, I., Loupas, V., Tzallas, A. T., Gogos, C., Karvelis, P. S., Antoniadis, N., Glavas, E., & Giannakeas, N. (2023). Performance and early drop prediction for higher education students using machine learning. *Expert Systems with Applications*, *225*, 120079. https://doi.org/10.1016/j.eswa.2023.120079

Dhar, J., & Jodder, A. K. (2020). An effective recommendation system to forecast the best educational program using machine learning classification algorithms. *Ingénierie des Systèmes d Information*, *25*(5), 559-568. https://doi.org/10.18280/isi.250502

Gao, P., Li, J., & Liu, S. (2021). An introduction to key technology in artificial intelligence and big data driven e-learning and e-education. *Mobile Networks and Applications*, *26*(5), 2123-2126. https://doi.org/10.1007/s11036-021-01777-7

Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2025). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, *44*(11), 2518-2544. https://doi.org/10.1080/0144929X.2024.2394886

Gloerfeld, C., Wrede, S., de Witt, C., & Wang, X. (2020). Recommender-potentials and limitations for self-study in higher education from an educational science perspective. *International Journal of Learning Analytics and Artificial Intelligence for Education*, *2*(2), 34. https://doi.org/10.3991/ijai.v2i2.14763

Hamid, M., Malik, S., Saleem, M., Zahary, A. T., & Jaghdam, I. H. (2026). Enhancing educational assessment through automated question classification using a RoBERTa-based ensemble model. *Scientific Reports, 16*, Article 13754. https://doi.org/10.1038/s41598-026-45486-1

Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., & Peppler, K. A. (2024). Teachers, parents, and students' perspectives on integrating generative AI into elementary literacy education. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Article 678). Association for Computing Machinery. https://doi.org/10.1145/3613904.3642438

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., & Bittencourt, I. I. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, *32*(3), 504-526. https://doi.org/10.1007/s40593-021-00239-1

Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of artificial intelligence-enabled personalised recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, *194*, 104684. https://doi.org/10.1016/j.compedu.2022.104684

Jin, T. (2025). Methods and reliability study of moral education assessment in universities: A machine learning-based approach. *Alexandria Engineering Journal*, *125*, 20-28. https://doi.org/10.1016/j.aej.2025.03.095

Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, 100074. https://doi.org/10.1016/j.caeai.2022.100074

Kinol, A. M. J., Bommi, R. M., Nalini, M., & Uganya, G. (2023). Evaluating the predictive accuracy and precision of AlexNet and linear regression for women's education post-independence. *Proceedings of the 2nd International Conference on Automation, Computing and Renewable Systems, Pudukkottai, India*, 1110-1114. https://doi.org/10.1109/ICACRS58579.2023.10405081

Kong, S.-C., Cheung, W. M.-Y., & Zhang, G. (2021). Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. *Computers and Education: Artificial Intelligence*, *2*, 100026. https://doi.org/10.1016/j.caeai.2021.100026

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13*(18), 10424. https://doi.org/10.3390/su131810424

Lee, H. S., & Lee, J. (2021). Applying artificial intelligence in physical education and future perspectives. *Sustainability*, *13*(1), 351. https://doi.org/10.3390/su13010351

Lee, Y., Han, E., Lee, A., & Hong, M. (2023). Personalized recommendation system based on recommendation indicator in education platform. *Proceedings of the 2023 International Conference on Research in Adaptive and Convergent Systems* (Article 37). Association for Computing Machinery. https://doi.org/10.1145/3599957.3606224

Liu, S., Shen, J., Qian, H., & Zhou, A. (2024). Inductive cognitive diagnosis for fast student learning in web-based intelligent education systems. *Proceedings of the ACM Web Conference 2024* (pp. 4260-4271). Association for Computing Machinery. https://doi.org/10.1145/3589334.3645589

Liu, Y., Lu, D., Xu, D., Zheng, Y., Zhou, T., Li, N., & Feng, R. (2024). Exploring intelligent feedback systems for student engagement in digitized physical education learning through machine learning. *Proceedings of the 2024 International Conference on Artificial Intelligence and Future Education* (pp. 114-118). Association for Computing Machinery. https://doi.org/10.1145/3708394.3708415

Machado, A., Tenório, K., Santos, M. M., Barros, A. P., Rodrigues, L., Mello, R. F., Paiva, R., & Dermeval, D. (2025). Workload perception in educational resource recommendation supported by artificial intelligence: A controlled experiment with teachers. *Smart Learning Environments*, *12*, Article 20. https://doi.org/10.1186/s40561-025-00373-6

Mahmud, M. M., Wong, S. F., Mohd A'seri, M. S., Ahmad, R., Yaacob, Y., Qazi, A., Mustamam, N. I., & Nagasundram, U. (2024). Exploring academic perceptions and challenges in AI integration in higher education. *Proceedings of the 2024 16th International Conference on Education Technology and Computers* (pp. 167-172). https://doi.org/10.1145/3702163.3702410

McGrath, C., Pargman, T. C., Juth, N., & Palmgren, P. J. (2023). University teachers' perceptions of responsibility and artificial intelligence in higher education – An experimental philosophical study. *Computers and Education: Artificial Intelligence*, *4*, 100139. https://doi.org/10.1016/j.caeai.2023.100139

Mohd Talib, N. I., Abd Majid, N. A., & Sahran, S. (2023). Identification of student behavioral patterns in higher education using K-means clustering and support vector machine. *Applied Sciences*, *13*(5), 3267. https://doi.org/10.3390/app13053267

Mohd Zaki, M. H., Abdul Aziz, M. A., Sulaiman, S., & Hambali, N. (2023). Student performance classification using support vector machine (SVM) with polynomical kernel on online student activities. *Journal of Electrical and Electronic Systems Research*, *23*(1), 80-90. https://doi.org/10.24191/jeesr.v23i1.009

Nadimpalli, V. K., Maier, R., Ezer, T., Bugert, F., Staufer, S., Röhrl, S., Hauser, F., Grabinger, L., & Mottok, J. (2025). Nestor: A personalized learning path recommendation algorithm for adaptive learning environments. *Proceedings of the 6th European Conference on Software Engineering Education* (pp. 49-59). Association for Computing Machinery. https://doi.org/10.1145/3723010.3723016

Narimani, A., Barbera, E., & Lundqvist, K. Ø. (2024). Academic advising for online higher education enrolment based on course association rules. *Proceedings of the 2024 16th International Conference on Education Technology and Computers* (pp. 324-331). Association for Computing Machinery. https://doi.org/10.1145/3702163.3702433

Obeid, C., Lahoud, I., El Khoury, H., & Champin, P.-A. (2018). Ontology-based recommender system in higher education. Companion proceedings of the the web conference 2018, https://doi.org/10.1145/3184558.3191533

Oubalahcen, H., & Tamym, L. (2023). The use of AI in E-learning recommender systems: A comprehensive survey. *Procedia Computer Science*, *224*, 437-442. https://doi.org/10.1016/j.procs.2023.09.061

Pasha, H. K. (2024). Gender differences in education: Are girls neglected in Pakistani society? *Journal of the Knowledge Economy*, *15*, 3466-3511. https://doi.org/10.1007/s13132-023-01222-y

Pimentel, J. S., Ospina, R., & Ara, A. (2021). Learning time acceleration in support vector regression: a case study in educational data mining. *Stats*, *4*, 682-700. https://doi.org/10.3390/stats4030041

Qadir, H. M., Suleman, M. T., Khan, R. A., Sohaib, M., Hasan, M. J., & Hussain, S. A. (2025). Optimizing learning outcomes: A deep dive into hybrid AI models for adaptive educational feedback. *Journal of Big Data*, *12*, Article 144. https://doi.org/10.1186/s40537-025-01187-6

Quan, Z., & Pu, L. (2023). An improved accurate classification method for online education resources based on support vector machine (SVM): Algorithm and experiment. *Education and Information Technologies*, *28*, 8097-8111. https://doi.org/10.1007/s10639-022-11514-6

Rahayu, N. W., Ferdiana, R., & Kusumawardani, S. S. (2022). A systematic review of ontology use in e-learning recommender system. *Computers and Education: Artificial Intelligence*, *3*, 100047. https://doi.org/10.1016/j.caeai.2022.100047

Raj, N. S., & Renumol, V. (2022). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020. *Journal of Computers in Education*, *9*, 113-148. https://doi.org/10.1007/s40692-021-00199-4

Renţea, I., Migut, G., & Krijthe, J. (2025). Are Interactive Visualizations in Machine Learning Education Helping Students? *Proceedings of the 30th ACM Conference on Innovation and Technology in Computer Science Education V. 1* (pp. 2-8). Association for Computing Machinery. https://doi.org/10.1145/3724363.3729032

Saqr, M., & López-Pernas, S. (2024). Why explainable AI may not be enough: Predictions and mispredictions in decision making in education. *Smart Learning Environments*, *11*, Article 52. https://doi.org/10.1186/s40561-024-00343-4

Silva, L. C., Sobrinho, A. A. d. C. C., Cordeiro, T. D., Melo, R. F., Bittencourt, I. I., Marques, L. B., da Cunha Matos, D. D. M., da Silva, A. P., & Isotani, S. (2023). Applications of convolutional neural networks in education: A systematic literature review. *Expert Systems with Applications*, *231*, 120621. https://doi.org/10.1016/j.eswa.2023.120621

Steele, J. L. (2023). To GPT or not GPT? Empowering our students to learn with AI. *Computers and Education: Artificial Intelligence*, *5*, 100160. https://doi.org/10.1016/j.caeai.2023.100160

Strzelecki, A. (2024). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive learning environments*, *32*(9), 5142-5155. https://doi.org/10.1080/10494820.2023.2209881

Su, J., & Yang, W. (2022). Artificial intelligence in early childhood education: A scoping review. *Computers and Education: Artificial Intelligence*, *3*, 100049. https://doi.org/10.1016/j.caeai.2022.100049

Tzirides, A. O. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., You, Y.-l., dos Santos, T. A., Searsmith, D., & O'Brien, C. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open*, *6*, 100184. https://doi.org/10.1016/j.caeo.2024.100184

Urdaneta-Ponte, M. C., Mendez-Zorrilla, A., & Oleagordia-Ruiz, I. (2021). Recommendation systems for education: Systematic review. *Electronics*, *10*(14), 1611. https://doi.org/10.3390/electronics10141611

Veeramanickam, M. R. M., Dabade, M. S., Borhade, R. R., Barekar, S. S., Navarro, C., Roman-Concha, U., & Rodriguez, C. (2023). Smart education system to improve the learning system with CBR based recommendation system using IoT. *Heliyon*, *9*(7), e17863. https://doi.org/10.1016/j.heliyon.2023.e17863

Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, *252*, 124167. https://doi.org/10.1016/j.eswa.2024.124167

Weber, F., Schrumpf, J., Dettmer, N., & Thelen, T. (2022). A web-based recommendation system for higher education: SIDDATA: History, architecture and future of a digital data-driven study assistant. *International Journal of Emerging Technologies in Learning*, *17*(22), 246-254. https://doi.org/10.3991/ijet.v17i22.31887

Wei, X., Sun, S., Wu, D., & Zhou, L. (2021). Personalized online learning resource recommendation based on artificial intelligence and educational psychology. *Frontiers in Psychology*, *12*, 767837. https://doi.org/10.3389/fpsyg.2021.767837

Yang, Q. (2024). A review of the role and impact of generative artificial intelligence on education. *Proceedings of the 2024 7th International Conference on Educational Technology Management* (pp. 121-127). Association for Computing Machinery. https://doi.org/10.1145/3711403.3711435

Zhang, J., Wang, L., Chen, X., Jiang, Y., & Tian, Z. (2023). Artificial intelligence-based education platform course recommendation system. *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering* (pp. 835-841). Association for Computing Machinery. https://doi.org/10.1145/3652628.3652767

Zhang, P., & Zhang, F. (2024). Enhancing education and teaching management through data mining and support vector machine algorithm. *International Journal of e-Collaboration*, *20*(1), 1-16. https://doi.org/10.4018/IJeC.357998

Zhang, Q., Lu, J., & Zhang, G. (2021). Recommender systems in e-learning. *Journal of Smart Environments and Green Computing*, *1*, 76-89. https://doi.org/10.20517/jsegc.2020.06

Zhou, X., Zhang, S., Kong, R., & Wu, X. (2025). Online education recommendation model combining dynamic noise perturbation and contrastive learning. *Proceedings of the 2025 2nd International Conference on Informatics Education and Computer Technology Applications* (pp. 300-304). Association for Computing Machinery. https://doi.org/10.1145/3732801.3732855

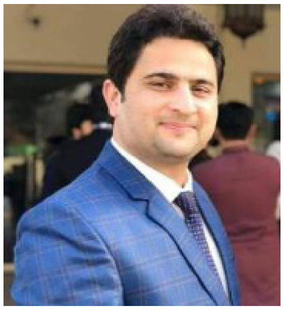
Zhou, Y., Zou, S., Liwang, M., Sun, Y., & Ni, W. (2025). A teaching quality evaluation framework for blended classroom modes with multi-domain heterogeneous data integration. *Expert Systems with Applications*, *289*, Article 127884. https://doi.org/10.1016/j.eswa.2025.127884

Zhu, L., Perchyk, O., & Wang, X. (2024). An augmented machine learning-based course enrollment recommender system. *Proceedings of the 2024 ACM Southeast Conference* (pp. 319-320). Association for Computing Machinery. https://doi.org/10.1145/3603287.3656163

Zulfiqar, A., & Kuskoff, E. (2024). Developing a contextual understanding of empowerment through education: Narratives from highly educated women in Pakistan. *Gender and Education*, *36*(6), 665-681. https://doi.org/10.1080/09540253.2024.2359519

AUTHORS

Ayesha Nawaz received her BSCS degree from the Department of Computer Science, GC Women University Sialkot, Pakistan. She is currently pursuing her MS degree in Computer Science at the Department of Computer Science, GC Women University Sialkot, Pakistan. Her research interests include Artificial Intelligence, Machine Learning, and Deep Learning.



Muhammad Hamid holds an M.Sc. in Information Technology, an MPhil, and a PhD in Computer Science. His doctoral research focused on enhancing Pakistan's software export potential by addressing recurring industry challenges through the application of Artificial Intelligence (AI). During his PhD studies, he conducted collaborative research with the Department of Computer Science and Operations Research at the University of Montreal, Canada. He brings over 15 years (6 years post-PhD) of combined experience in teaching, research, and academic administration. He is currently serving as an Associate Professor at University of Kamalia, Pakistan. His research interests include artificial intelligence, software engineering, and the integration of AI in educational systems. He actively contributes to the academic community as a reviewer, technical committee member, and editorial board member for several reputable national and international conferences and journals.

Aisha Shabbir is currently serving as an Assistant Professor at the Faculty of Computing and Information Technology, Sohar University, Oman. She has served as an Assistant Professor at the School of Civil & Environmental Engineering (SCEE), National University of Sciences and Technology (NUST), Islamabad. She has completed a PhD in Computer Science from University Technology Malaysia. She has completed her MS from Ghulam Ishaq Khan Institute (GIKI). Her research interests include data analytics and visualization, data mining, interdisciplinary research, machine learning, and big data processing techniques.

Aisha Blfgeh is an Assistant Professor in the Department of Computer Science and Artificial Intelligence at the University of Jeddah, Saudi Arabia. She received her PhD in Computer Science from Newcastle University. Her research interests lie at the intersection of applied artificial intelligence, knowledge engineering, and sustainable computing ecosystems. Having transitioned from foundational work in semantic web ontologies and structural AI, her recent research focuses on functional AI applications, including distributed security, federated intelligence, deep learning, and intelligent IoT frameworks for sustainable energy management.

Muhammad Saleem holds a Master's degree in Computer Science and Communications Engineering from the University of Duisburg-Essen, Germany, and a PhD in Engineering from the University of the Federal Armed Forces, Munich, Germany. Dr Saleem has more than 15 years of teaching, research, and administrative experience in the Department of Industrial Engineering at the University of Duisburg-Essen, Germany, and King Abdulaziz University, Kingdom of Saudi Arabia. Dr Saleem also worked as a project manager in Siemens Power Generation's energy sector in Munich, Germany, for 3 years. His work is focused mainly on Industrial Quality Control, Artificial Intelligence, and Renewable Energy. He is actively involved in the curriculum development and accreditation processes of engineering programs.