Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 47 (2026), No. 7, pp. 2607–2617
DOI : 10.47974/JIOS-2154

A machine learning approach for personalized course recommendation systems for learners

Anu Thomas *
Department of Computer Engineering
Lok Jagruti Kendra University
Ahmedabad 382210
Gujarat
India

Gayatri Pandi †
Department of Computer Science and Engineering
New L J Institute of Engineering and Technology
Sindhu Bhavan Road
Ahmedabad 380059
Gujarat
India

Abstract

This paper presents a novel approach to enhancing collaborative filtering for course recommendations by integrating Hybrid Fruit-fly Optimized Density-Based K-Means Clustering (HFO-KMC) with Atten-BERT-RU and a Coordinate Attention Module (CAM). The methodology improves user grouping and feature representation, resulting in more accurate and personalized recommendations. HFO-KMC optimizes clustering by reducing errors, while Atten-BERT-RU with CAM enhances feature extraction from course reviews. The system was validated using Coursera course reviews, showing notable improvements in recommendation accuracy and diversity over traditional models. The results highlight the

* *ORCID-ID:* **0009-0000-1915-8359**
† *ORCID-ID:* **0000-0002-3096-1836**

* *E-mail:* **anuthomas.phd20@ljku.edu.in (Corresponding Author)**
† *E-mail:* **gayatri.jain@ljinstitutes.edu.in**

effectiveness of combining optimized clustering with deep learning techniques, making this approach a significant advancement in recommendation systems.

Subject Classification: *68T05, 68T20.*

Keywords: *Recommendation systems, Machine learning, Collaborative filtering, Content-based filtering, Deep learning, Fruit-fly optimization, K-means clustering, Personalized learning, Educational data mining, Coursera, BERT, Attention mechanism, Hybrid models, Course recommendation, User modeling.*

1. Introduction

Personalized learning has become central to modern education, aiming to meet the varied needs of individual learners. Recommendation systems support this by using intelligent algorithms to suggest tailored content, improving learner engagement and academic outcomes. The shift to online education, accelerated by the COVID-19 pandemic, has further highlighted the need for adaptive digital learning tools, Singh et al. [1]. To predict learner preferences, recommendation systems utilize classification methods such as collaborative filtering, content-based filtering, and hybrid approaches, Burke [2]. Advances in deep learning, especially CNNs and RNNs, have improved the ability to model complex learner behaviors, Weamie et al. [3]. Meanwhile, optimization-enhanced clustering and data mining techniques are driving the development of more accurate and diverse recommendation models, Mali, Mishra, and Vijayalaxmi [4]. Platforms like Coursera increasingly rely on these methods, yet traditional systems still struggle with challenges like data sparsity. This paper proposes a novel model combining attention mechanisms, feature extraction, and optimized clustering to improve accuracy and diversity in course recommendations.

2. Related Work

Modern learning environments rely heavily on recommendation systems to deliver personalized educational experiences by suggesting relevant learning materials, activities, and pathways based on individual student needs. Various classification techniques have been applied to these systems to effectively model and predict learner preferences. Collaborative filtering, which utilizes user-item interaction data for generating recommendations, remains one of the most widely used approaches. Content-based filtering, on the other hand, recommends items based on similarity metrics derived from item attributes. Hybrid recommendation systems, which merge collaborative and content-based approaches, have proven effective in increasing coverage and accuracy. He et al. [5] introduced neural collaborative filtering, which learns user-item interactions directly from data. Similarly, Wang et al. [6] proposed neural graph collaborative filtering, leveraging graph neural networks to capture relationships between users and items. Beyond traditional machine learning and deep learning models, clustering techniques have been explored to refine recommendations. Xia et al. [7] adapted density-based clustering methods, such as DBSCAN, to identify high-density user-item interaction regions. Optimization algorithms, including fruit-fly optimization, have further enhanced clustering accuracy and recommendation precision. Advancements in recommendation models also include sequential models that track temporal user behavior, session-based approaches that analyze short-term preferences, and reinforcement learning-based models that refine recommendation policies through continuous user interaction, Adomavicius and Tuzhilin [8], Ricci, Rokach, and Shapira [9]. Explainable AI has also gained traction, providing interpretable recommendations to improve user trust and engagement. Overall, the literature highlights the diversity of classification techniques in recommendation systems, ranging from conventional filtering methods to state-of-the-art deep learning and clustering-based techniques. This paper synthesizes these approaches, analyzing their benefits, limitations, and implications for personalized learning. Further, research has introduced innovative models such as attention-based, context-aware networks, which dynamically adjust recommendations based on user interests, Yuan et al. [10]. Deepak, Gerard, and Trivedi [11], proposed a deep learning strategy for course recommendation, demonstrating that hybrid attention mechanisms can significantly improve both accuracy and contextual relevance of recommendations in educational platforms. Deep learning has shown strong applicability in educational and cultural domains, Pandi and Aggarwal [12], who developed a 3D content model using neural networks to digitally preserve and enhance learning around heritage sites. Zuo et al. [13] provides a tag-aware comprehensive survey on deep learning-based recommendation systems, covering CNNs, RNNs, and GNNs, discussing recent advancements and emerging challenges. Graph-based methods, such as those introduced by Zhang, Li, and Zhao [14], utilize positional prompts to enhance structural user-item interactions. Lal et al. [15] demonstrated how clustering techniques using unstructured data, such as textual reviews, can enhance recommendation quality in sparse datasets. Zheng [16] explored reinforcement learning for news recommendations, showing how adaptive policies improve long-term user engagement. Context-aware recommendation techniques have also evolved, such as Kim [17] convolutional matrix factorization approach, which incorporates textual data to enhance document-based recommendations. Zhang et al. [18] conducted a comprehensive review of deep learning architectures used in recommendation systems, analyzing their training methodologies and evaluation techniques. Additional studies have focused on density-based clustering for personalized recommendations, collaborative clustering that incorporates implicit feedback, Renaud-Deputter, Xiong, and Wang [19], and enhanced K-means clustering for personalized course suggestions, Huang, Jia, and Zuo [20]. Graph-based clustering techniques, as explored by Wang [21] further improve accuracy, scalability, and interpretability. The role of big data in personalized learning has also been widely studied. Gandomi and Haider [22] discuss how data analytics can reshape educational methodologies, including recommendation systems.

Research by Alanya-Beltran [23] demonstrates how personalized recommendation algorithms can enhance student engagement and performance in e-learning environments. Educational data mining techniques, such as those proposed by Lei [24], optimize course selection and learning pathways by analyzing student behavior. Recent research has explored various enhancements in recommendation systems, including the use of online discussion forums for predicting student performance by means of genetic programming Ulloa-Cazarez [25], hybrid recommendation models in adaptive learning platforms Chen et al. [26], and affect-aware systems that personalize suggestions based on user attitudes, emotions and perception, Banihashem [27]. Collaborative filtering and clustering approaches have evolved to address modern challenges. While K-Means clustering remains a popular choice for user segmentation, its limitations in initialization and convergence have been mitigated using optimization techniques like the Fruit-Fly Optimization Algorithm (FOA), which refines cluster centroids to improve accuracy. Fu [28] successfully applied deep reinforcement learning in educational recommendations, suggesting its potential to dynamically adapt course suggestions based on evolving learner behavior. Deep learning models, particularly BERT, have advanced feature extraction from textual data, enhancing the quality of recommendations. However, traditional BERT models face difficulties in handling long-term dependencies, necessitating improvements such as recurrent units and attention mechanisms to maintain contextual integrity. Latifi, Jannach, and Ferraro [29] revisited BERT4Rec to explore bidirectional self-attention in sequential recommendation, which is a study on transformers and nearest neighbors. Building on these advancements, this study integrates optimized clustering with deep learning-based feature extraction to enhance recommendation precision. The Coordinate Attention Module (CAM) further refines feature selection and strengthens dependency management, ensuring a more adaptive, efficient, and robust recommendation framework. The importance of applying optimization and classification frameworks to enhance the efficiency of adaptive learning environments has been emphasized by Behl [30]. It studies optimized agile adoption in virtual learning environments.

3. Methodology

In this section, the methodology of the proposed model is discussed in detail. It first introduces the model, and then later, the entire training process is discussed.

3.1 *Proposed Model*

The proposed system begins with a comprehensive data analysis and preprocessing phase to ensure consistency and reliability. The dataset, comprising Coursera course reviews, undergoes normalization, missing value handling, and format standardization. Courses are represented as $C = \{c_1, c_2 \ldots, c_n\}$, where c_i is a specific course and users as $U = \{u_1, u_2, \ldots, u_m\}$, where u_j is a specific user. Let R_{ij} represent the rating given by user u_j to course c_i, where $R_{ij} \in \{1,2,3,4,5\}$. Following preprocessing, the Hybrid Fruit-Fly Optimized Density-Based K-Means Clustering (HFO-KMC) method refines user segmentation by integrating the Fruit-Fly Optimization Algorithm (FOA) with K-Means clustering. FOA dynamically optimizes cluster centroids, mitigating K-Means' challenges such as poor initialization and convergence to local minima. The system partitions users into k clusters $S=\{S_1,S_2,\ldots,S_k\}$. Minimize intra-cluster variance:

$$J = \sum_{I=1}^{k} \sum_{uj\epsilon Si} ||uj - ui||^2 \quad (1)$$

where μ_i is the centroid of cluster S_i and || || represents the Euclidean distance. FOA optimizes clustering by iteratively adjusting centroids using the formula:

$$\text{Position}_{new}=\text{Position}_{old}+\alpha\cdot(\text{BestPosition}-\text{Position}_{old}) \quad (2)$$

where α is the learning rate. For feature extraction, the Atten-BERT-RU Model processes course reviews by leveraging BERT embeddings with adaptive recurrent units. Each review r_i is transformed into a tokenized vector using BERT. Let r_i be represented as: $T_i=\text{BERT}(r_i) = \{t_1,t_2,\ldots,t_n\}$ where t_i are the tokenized vectors of words.

To enhance feature extraction, the Coordinate Attention Module (CAM) is incorporated, applying an attention mechanism that assigns weight scores to tokens: $V_r = \sum_{i=1}^{n} \alpha_i t_i$ (3) where α_i represents the attention score for token $t_{i.}$ ensuring the model effectively captures crucial information in user reviews. Transitioning from feature extraction to prediction, the Hybrid Deep Convolutional Dilated Residual Network with Stacked LSTM (HDC-RSL) model is deployed. Dilated convolutions expand the receptive field without increasing computational costs, with outputs computed as:

$$y_i^l = \sum_{k=1}^{K} W_k x_{i+r(k-1)}^l \quad (4)$$

where r is the dilation rate, K is the kernel size, and W_k are the convolution weights. To handle sequence data in user interactions: $h_t=\text{LSTM}(x_t,h_{t-1})$ (5) where h_t is the hidden state at time t, and x_t is the input at time t. For recommendation synthesis, the Weighted Soft Attention with Multilayer Perceptron (WSA-MLP) model is utilized. This model refines recommendations by computing the final score for for course c_i for user u_j is: $S_{ij}=\text{MLP}(\alpha\cdot R_{ij}+\beta\cdot V_{ri})$ (6) where α and β are learned weights, R_{ij} is the rating prediction, and V_{ri} is the feature vector from the reviews. To further improve recommendation diversity and accuracy, the Greedy Re-Ranking Approach is applied. This iterative strategy prioritizes the most suitable courses by maximizing diversity between recommendations. If $D(c_i,c_j)$ represents the distance between courses c_i and c_j, the objective is to maximize:

$$\text{Max} \sum_{i=1}^{k} D(c_i + c_{i+1}) \quad (7)$$

By integrating these advanced clustering, feature extraction, and ranking techniques, the system delivers precise, diverse, and learner-specific course recommendations.

4. Experimentation

4.1. *Dataset Description*

The dataset used for this study comprises Coursera course reviews obtained from Kaggle. It includes user interactions, course metadata, and ratings.

4.2. *Experimental Setup*

The model was implemented using Python with TensorFlow and PyTorch. Training was conducted in a GPU-accelerated environment to optimize efficiency.

4.3. *Parameter Tuning and Training*

Content-based and collaborative filtering techniques were employed to extract meaningful insights from course descriptions and user behaviors.

5. Results and Discussion

The effectiveness of the proposed model was evaluated using multiple performance metrics, including accuracy, specificity, F1-score, precision, testing loss, Mean Absolute Error (MAE), and Mean Squared Error (MSE). These metrics collectively assess the model's capability in delivering accurate and reliable course recommendations.

5.1. *Performance Metrics*

The proposed system consistently outperformed baseline deep learning models such as ResNet, RNN, Bi-LSTM and LSTM, demonstrating superior accuracy and feature extraction quality. The integration of Hybrid Fruit-Fly Optimized Density-Based K-Means Clustering (HFO-KMC) with Atten-BERT-RU and Coordinate Attention Module (CAM) contributed to this improvement by enhancing user segmentation and refining feature extraction. Specificity, which measures the model's ability to correctly identify relevant recommendations, showed a marked improvement, reducing false positives. Precision was also notably higher, the F1-score, precision, and recall, confirmed the model's ability to maintain accuracy while minimizing errors in course recommendations.

5.2. *Loss and Error Metrics*

A key advantage of the proposed system is its significantly lower testing loss, which highlights its ability to generalize effectively from training data to unseen data. Furthermore, the model's MAE and MSE values were substantially lower compared to those of ResNet, RNN, Bi-LSTM, and LSTM. These improvements validate the effectiveness of HFO-KMC in optimizing user clustering and Atten-BERT-RU with CAM in refining feature extraction.

5.3. *Comparative Analysis*

The proposed model demonstrated superior performance across various metrics. A comparison of performance metrics is provided in the table below.

Table 1

Performance Metrics-Comparative Analysis

Metric	Proposed Model	ResNet	RNN	Bi_LSTM	LSTM
Accuracy	0.9792	0.94	0.93	0.90	0.88
Precision	0.9545	0.92	0.91	0.88	0.86
Specificity	0.987	0.94	0.93	0.90	0.88
F1 Score	0.9642	0.92	0.91	0.88	0.86

Table 2

Error Metrics

Metric	Proposed Model	ResNet	RNN	Bi_LSTM	LSTM
MAE	0.0312	0.05	0.06	0.10	0.11
MSE	0.0312	0.05	0.06	0.10	0.11

Table 1 indicates that the proposed model outperforms traditional deep learning models across all evaluation metrics. Figure 1 demonstrates that the model achieves the highest accuracy of 97.92%, demonstrating its ability to classify data more effectively. Figure 2 shows the precision of 95.45% suggests a reduced false positive rate, while Figure 3 shows the specificity of 98.7%, and Figure 4 shows the F1 score. It highlights its capability to correctly identify negative cases. In contrast, Bi_LSTM and LSTM models exhibit relatively lower accuracy and precision, indicating their limitations in handling the dataset. Table 2, error metrics further validate the superiority of the proposed model. Figure 5 shows lower MAE (0.0312), and Figure 6 shows MSE values (0.0312) compared to other models.

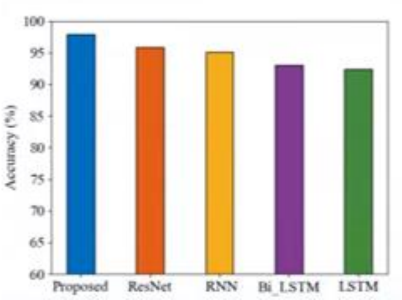

Figure 1
Performance Metric-Accuracy

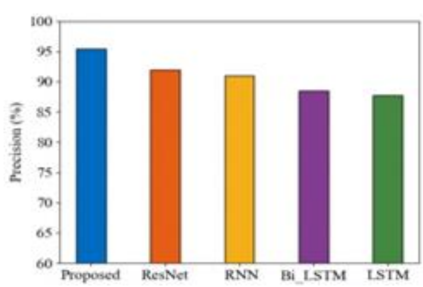

Figure 2
Performance Metric-Precision

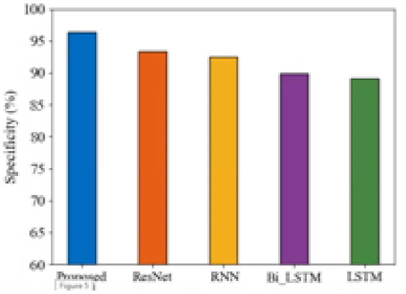

Figure 3
Performance Metric- Specificity

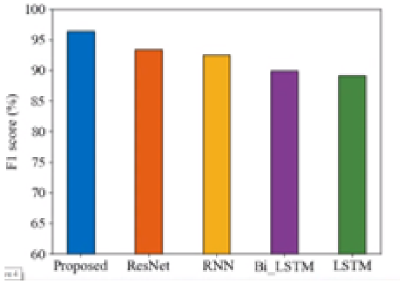

Figure 4
Performance Metric-F1 Score

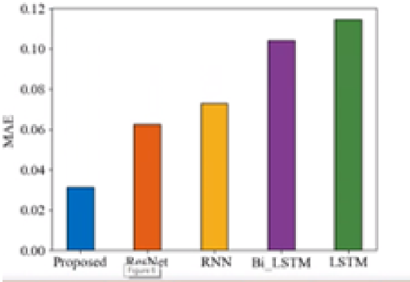

Figure 5
Loss and Error Metric-MAE

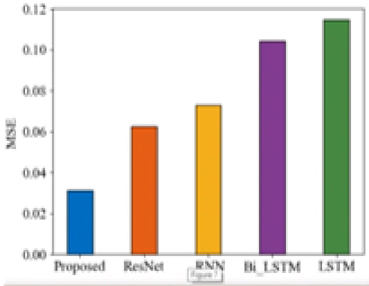

Figure 6
Loss and Error Metric-MSE

6. Conclusion

This study introduced an enhanced approach to collaborative filtering in course recommendation systems by integrating Hybrid Fruit-Fly Optimized Density-Based K-Means Clustering (HFO-KMC) with Atten-BERT-RU and the Coordinate Attention Module (CAM) for advanced feature extraction. The proposed model demonstrated outstanding results across comprehensive evaluation metrics, including accuracy, precision, and F1-score, outperforming traditional models such as ResNet, RNN, Bi-LSTM, and LSTM. HFO-KMC improved user clustering, while Atten-BERT-RU with CAM refined feature representation, resulting in more precise and diverse recommendations.

7. Future Work

Future research can focus on expanding this methodology to other domains, such as e-commerce and healthcare, where personalized recommendations play a crucial role. Incorporating dynamic user profiling could enable the system to continuously adapt to evolving preferences, making recommendations more relevant over time. Additionally, integrating reinforcement learning and graph neural networks could further enhance accuracy and diversity in predictions. Finally, deploying the model in real-world applications and incorporating user feedback would provide valuable insights for refining performance and improving adaptability in live environments.

Acknowledgement

The authors acknowledge the support provided by L. J University for conducting this research.

References

[1] S. Singh, R. Aggarwal, R. Bali, G. Bapat, and S. Mohite, "Transformative impact: Online learning evolution in higher education post-COVID-19 pandemic," *J. Inf. Optim. Sci.*, vol. 44, no. 8, pp. 1627–1647 (2023), doi: 10.47974/JIOS-1481.

[2] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331–370 (2002).

[3] S. J. Y. Weamie, V. Kolluru, A. B. B. Jallah, and Y. Challagundla, "Hybrid deep learning-based IoT intrusion detection: A comparative study of CNN, GRU, LSTM, and hybrid architectures," *Journal of Information and Optimization Sciences*, vol. 46, no. 6, pp. 1983–1994 (2025), doi: 10.47974/JIOS-2027.

[4] M. Mali, D. Mishra, and M. Vijayalaxmi, "Bi-clustering based recommendation system," *Journal of Information and Optimization Sciences*, vol. 45, no. 4, pp. 1029–1039 (2024), doi: 10.47974/JIOS-1625.

[5] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T. S. Chua, "Neural collaborative filtering," in *Proc. 26th Int. Conf. on World Wide Web (WWW)*, pp. 173–182 (2017).

[6] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, "Neural graph collaborative filtering," in *Proc. 42nd Int. ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR)*, Paris, France, pp. 165–174 (2019), doi: 10.1145/3331184.3331267.

[7] F. Xia, X. Xu, L. Chen, and S. Zhang, "Density-based spatial clustering of applications with noise for personalized recommendation," *Information Sciences*, vol. 512, pp. 185–197 (2020).

[8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749 (2005).

[9] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems," in *Recommender Systems Handbook*, Boston, MA, USA: Springer, pp. 1–35 (2015), doi: 10.1007/978-1-4899-7637-6_1.

[10] W. Yuan, H. Wang, X. Yu, N. Liu, and Z. Li, "Attention-based context-aware sequential recommendation model," *Information Sciences*, vol. 510, pp. 122–134 (2020), doi: 10.1016/j.ins.2019.09.007.

[11] D. Deepak, G. Gerard, and I. Trivedi, "A hybridized deep learning strategy for course recommendation," *International Journal of Adult Education and Technology*, vol. 14, no. 1, pp. 1–16 (2023), doi: 10.4018/IJAET.321752.

[12] G. Pandi and K. P. Aggarwal, "Deep learning-based 3-D model for the cultural heritage sites in the state of Gujarat, India," in *Artificial Intelligence and Sustainable Computing (ICSISCET 2022)*, Singapore: Springer, (2023), doi: 10.1007/978-981-99-1431-9_59.

[13] Y. Zuo, S. Liu, Y. Zhou, and H. Liu, "TRAL: A tag-aware recommendation algorithm based on attention learning," *Applied Sciences*, vol. 13, no. 2, Art. no. 814 (2023), doi: 10.3390/app13020814.

[14] J. Zhang, C. Li, and Z. Zhao, "Lightweight yet efficient: An external attentive graph convolutional network with positional prompts for sequential recommendation," *ACM Transactions on Information Systems*, vol. 43, no. 3, Art. no. 79, pp. 1–25 (2025), doi: 10.1145/3719343.

[15] N. Lal, M. Singh, S. Pandey, and A. Solanki, "A proposed ranked clustering approach for unstructured data from dataspace using VSM," in *Proc. Int. Conf. on Computer Science and Applications (ICCSA)*, pp. 80–86 (2020), doi: 10.1109/ICCSA50381.2020.00024.

[16] G. Zheng, "DRN: A deep reinforcement learning framework for news recommendation," in *Proc. The Web Conference (WWW '18)*, Lyon, France, pp. 167–176 (2018), doi: 10.1145/3178876.3185994.

[17] D. Kim, "Convolutional matrix factorization for document context-aware recommendation," in *Proc. ACM Conference on Recommender Systems (RecSys)*, Boston, MA, USA, pp. 233–240 (2016), doi: 10.1145/2959100.2959165.

[18] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM Computing Surveys*, vol. 52, no. 1, pp. 1–38 (2019).

[19] S. Renaud-Deputter, T. Xiong, and S. Wang, "Combining collaborative filtering and clustering for implicit recommender system," in *Proc. IEEE Int. Conf. on Advanced Information Networking and Applications (AINA)*, Barcelona, Spain, pp. 748–755 (2013), doi: 10.1109/AINA.2013.65.

[20] J. Huang, Z. Jia, and P. Zuo, "Improved collaborative filtering personalized recommendation algorithm based on k-means clustering and weighted similarity on the reduced item space," *Mathematical Modelling and Control*, vol. 3, no. 1, pp. 39–49 (2023), doi: 10.3934/mmc.2023004.

[21] S. Wang, "Graph learning based recommender systems: A review," in *Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI)*, Montreal, Canada, pp. 4644–4652 (2021), doi: 10.24963/ijcai.2021/630.

[22] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *International Journal of Information Management*, vol. 35, no. 2, pp. 137–144 (2015).

[23] J. Alanya-Beltran, "Personalized learning recommendation system in e-learning platforms using collaborative filtering and machine learning," in *Proc. 2024 Int. Conf. on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, India, pp. 1–5 (2024), doi: 10.1109/ACCAI61061.2024.10602322.

[24] L. Lei, "Research on personalized education recommendation algorithm based on artificial intelligence," in *Proc. Int. Conf. on Applied Physics and Computing (ICAPC)*, Ottawa, ON, Canada, pp. 531–535 (2023), doi: 10.1109/ICAPC61546.2023.00104.

[25] R. L. Ulloa-Cazarez, "Prediction of online students' performance by means of genetic programming," *Applied Artificial Intelligence*, vol. 32, no. 9–10, pp. 858–881 (2018), doi: 10.1080/08839514.2018.1508839.

[26] Y. Chen, X. Li, J. Liu, and Z. Ying, "Recommendation system for adaptive learning," *Appl. Psychol. Meas.*, vol. 42, no. 1, pp. 24–41 (2018), doi: 10.1177/0146621617697959.

[27] S. K. Banihashem, "Modeling teachers' and students' attitudes, emotions, and perceptions in blended education: Towards post-pandemic education," *Int. J. Manage. Educ.*, vol. 21, no. 2, Art. no. 100803 (2023), doi: 10.1016/j.ijme.2023.100803.

[28] M. Fu, "A deep reinforcement learning recommender system with multiple policies for recommendations," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 2049–2061 (Feb. 2023), doi: 10.1109/TII.2022.3209290.

[29] S. Latifi, D. Jannach, and A. Ferraro, "Sequential recommendation: A study on transformers, nearest neighbors and sampled metrics," *Inf. Sci.*, vol. 609, pp. 660–678 (2022), doi: 10.1016/j.ins.2022.07.079.

[30] G. Behl, "Optimizing agile adoption in virtual learning environments: A framework for classification and prioritization," *J. Inf. Optim. Sci.*, vol. 45, no. 6, pp. 1705–1716 (2024), doi: 10.47974/JIOS-1706.

Received March, 2025