SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Personalized Learning Recommendations with Multi-Layered Neural Collaborative Filtering

Dr. Meenakshi Sundaram G[1*], Dr.Ambika P[2], Komiljon Karimov[3], Bhavadharani S[4], Nozima Dosmuxammedova[5], Jasur Abdullayev[6]

[1*]Assistant Professor, Department of Electronics and Communication Engineering, K.S.Rangasamy College of Technology, Tiruchengode, India. Email: meenakshisundaramg@ksrct.ac.in, https://orcid.org/0000-0001-8510-3337

[2]Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email: ambikap@maher.ac.in

3Bukhara State Medical Institute named after Abu Ali ibn Sino. Bukhara, Uzbekistan. Email: komiljon_karimov@bsmi.uz, https://orcid.org/0009-0004-3231-4108

[4]Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email: bhavadharani@maher.ac.in

[5]Department of Accounting and Statistics, Termez University of Economics and Service, Termez, Uzbekistan. Email: nozima_dosmuxamedova@tues.uz, https://orcid.org/0009-0007-8603-5616

[6]Lecturer, Department of Theory of Physical Education, Fergana State University, Fergana, Uzbekistan. Email: jasurabdullayev11@gmail.com, https://orcid.org/0000-0001-6075-3432

*Corresponding author: Email: meenakshisundaramg@ksrct.ac.in

Abstract

The emergence of numerous learning platforms through online means has led to the creation of large amounts of interaction data between learners and content, giving rise to tremendous potential for intelligent recommendation systems. Traditional collaborative filtering approaches are plagued by the lack of flexibility in expressing learners' learning preferences and also suffer from the cold-start problem. To address these deficiencies, this research presents a Multi-Layered Neural Collaborative Filtering (ML-NCF) algorithm for learner content recommendations that overcomes these shortcomings. In the presented ML-NCF framework, GMF is combined with a deep MLP architecture that learns both linear and nonlinear user-item interactions along with a self-attention mechanism for dynamically assigning weights to different contexts in learning. The effectiveness of the approach was evaluated using two widely used datasets: the MOOC Interaction dataset (N = 94,506 learner-content interactions) and the EdNet dataset (N = 131,441 interactions). ML-NCF resulted in 87.4% hit rate, NDCG at 0.723, precision at 0.814, recall of 0.791, and MRR of 0.768, showing superior performance compared to baseline models such as BPR-MF, NeuMF, and LightGCN by 9.2%, 6.8%, and 4.1% in hit rate, respectively. The effectiveness of all components of the proposed architecture has been validated through an ablation study. It is evident from the findings that the ML-NCF approach improves educational recommendations by leaps and bounds. This research provides strong support for the application of the ML-NCF system in adaptive learning systems, intelligent tutoring, and e-learning applications. The future direction will include federated learning mechanisms in recommendation scenarios.

Keywords

Neural Collaborative Filtering, Personalized Learning Recommendations, Multi-Layer Perceptron, Generalized Matrix Factorization, Attention Mechanism, E-Learning Systems, Educational Data Mining.

This is an open access article under CC BY 4.0, allowing unrestricted use with proper attribution, a license link, and indication of any changes made.

1. Introduction

The global e-learning market, which currently has a value of about 250 billion in 2023, is expected to grow to more than 840 billion by 2030, representing an unparalleled move towards e-learning education [1][10]. MOOCs, adaptive learning technologies, and digital libraries hold millions of educational materials that pose an unprecedented dilemma for learners who find themselves in a situation with large amounts of educational

materials to sift through. As recommendation systems become an essential part of the educational infrastructure, the specific nature of educational recommendation, which entails pedagogical sequencing, knowledge prerequisites, learner cognitive state, and skill acquisition timelines, sets it apart from other recommendation scenarios such as those in e-commerce and media streaming applications [2] [19]. Recent studies on deep contextual representation learning and AI-driven educational recommendation further emphasize the growing importance of intelligent personalization in digital learning ecosystems [6] [8].

Traditional CF techniques, especially the matrix factorization (MF) family of algorithms, have proven to be efficient in recommendation applications through decomposing the user-item interaction matrix into its latent factor form. Yet, these techniques are inherently limited by the fact that they assume the existence of linear correlations between the users and the items, which fail to capture the intricate, hierarchical, and nonstationary characteristics of preference learning processes [3][27]. In addition, the problem of cold start, whereby learners or items introduced for the first time do not have enough past interaction history, is another limitation that exists. Real-time collaborative filtering studies in large-scale library recommendation environments have also highlighted scalability and sparsity challenges associated with conventional CF methods [4].

Advancements in deep learning techniques have enabled the emergence of a novel generation of recommendation systems based on neural networks that can effectively identify non-linear interaction patterns at higher levels of order among heterogeneous data [20]. In the NCF architecture proposed by the researchers, using neural networks instead of inner product calculation in MF has resulted in remarkable improvements in performance [5][28]. However, conventional NCF models are inadequate to harness the benefits of multi-level temporal aspects of learning activities and educational content metadata.

1.1. Problem Statement and Significance

Even with the abundance of neural recommendation methods, three core issues have been identified when applying recommendations to education, which include (1) the incapability of shallow models to account for multi-order interactions in learning preference, (2) inadequate exploitation of contextual cues, including learning session time, level of difficulty, and prerequisite relations, and (3) an absence of transparent attention mechanisms to provide explanations for recommended content. This leads to poor learning outcomes, increased disengagement among students, and lower retention rates on platforms with substantial financial and educational implications.

1.2. Unique Contributions

This paper makes the following original contributions to the field of educational recommender systems:

- A novel Multi-Layered Neural Collaborative Filtering (ML-NCF) architecture that integrates GMF and MLP streams with a self-attention module for adaptive contextual weighting of learning interactions.
- A comprehensive mathematical formulation of the ML-NCF objective function incorporating regularized cross-entropy loss with attention-weighted interaction modeling.
- Extensive empirical validation on two large-scale educational datasets demonstrates state-of-the-art performance across five standard evaluation metrics.
- A detailed ablation study quantifying the individual contribution of each architectural component to overall recommendation quality.
- An architectural framework adaptable for integration with federated learning and privacy-preserving educational analytics pipelines.

The rest of the paper is organized as follows: Section 2 provides an overview of prior works on collaborative filtering, neural recommendation systems, and educational data mining. Section 3 introduces the proposed ML-NCF approach along with its mathematical formulation and architecture. Section 4 describes the experimental setting and reports on the experimental results. Finally, Section 5 concludes the paper with key conclusions and future research perspectives.

2. Literature Survey

Research work that supports personalization in education includes areas such as collaborative filtering techniques, deep learning models, attention-based models, and education data mining. This part provides a review of important studies related to the above areas.

2.1. Collaborative Filtering and Matrix Factorization

Matrix Factorization (MF) is still one of the key methods for collaborative filtering. The theoretical basis of MF, where competitive results were demonstrated against Netflix Prize data. Later advancements like Bayesian Personalized Ranking (BPR) included implicit feedback training, which is especially important for education, where ratings are scarce [7] [21]. Modern approaches have also revisited the topic of BPR, highlighting its significance as a baseline method [25]. Nevertheless, linearity built into MF prevents modeling of complex user behavior, inspiring further neural network-based approaches.

2.2. Neural Collaborative Filtering

Recently, Neural Collaborative Filtering (NCF) was introduced as an influential framework where the similarity computation in matrix factorization is replaced by a deep neural network that allows learning non-linear user-item interactions [5]. NeuMF, the leading algorithm in NCF, utilizes GMF and MLP networks to take advantage of linear and non-linear interactions. In another study, NCF was extended by incorporating structured semantic modeling of text data, leading to remarkable performance boosts in cold start situations [9]. The study integrated the graph structure of users and items in neural CF, showing that relational inductive biases improve representation quality [23].

2.3. Graph-Based Recommendation Models

The Graph Neural Network (GNN) paradigm has been established as a promising approach for embedding high-order connectivity in user-item interaction graphs. In particular, the Neural Graph Collaborative Filtering (NGCF) is an approach that enables the propagation of user-item embedding information by passing information along interaction graphs [11]. The work then proposed LightGCN, a simple form of GCN model that does not include any nonlinear transformations in its convolution operations, resulting in improved performance as well as computation efficiency [12][29]. Furthermore, the application of heterogeneous graph attention network models for educational recommendation was investigated [13].

2.4. Attention Mechanisms in Recommendation

Attention mechanisms have significantly improved recommendation systems through the dynamic context-aware weighting of features. A recommendation system based on attention collaborative filtering was suggested, in which neighborhood interactions were assigned non-uniform weights, providing more resistance to noise from implicit feedback [14]. They developed a Transformer-based framework with multi-head attention and successfully utilized it in sequential recommendation problems [15]. The study used bidirectional self-attention (BERT4Rec), which produced remarkable outcomes in modeling sequential behaviors of items, including learning sequences [16].

2.5. Personalized Learning and Educational Recommendation

Educational recommender systems must take into account specific domain constraints that do not exist in the business domain. It considered the role of knowledge component mastery as an intermediate variable when recommending educational content, thus emphasizing the necessity of considering the prerequisites [17]. It suggested a multi-relational graph convolutional network for MOOCs that would simultaneously consider the knowledge states of the learners and the content difficulty gradient [18]. This study evaluated the effectiveness of deep learning techniques for educational recommendations in K-12 adaptive learning environments and found attention-based neural CF to be the most promising recommendation framework [24]. It investigated federated collaborative filtering for privacy-preserving educational recommendations, demonstrating that distributed learning can approach centralized performance while protecting sensitive learner data. Learner interaction analytics have also been recognized as an important factor for collaborative knowledge building and improving adaptive recommendation quality in online education systems [22][26].

2.6. Summary and Research Gap

Collectively, the above-mentioned literature suggests that the neural collaborative filtering, especially with attention and graph-structural knowledge incorporated, is much more effective compared to other traditional MF models. However, current educational recommendation systems either only incorporate graph-based strategies and neglect optimizing the depth of the MLP layer, or utilize attention without incorporating both linear and non-linear interaction pathways. The novel ML-NCF framework fills in this research gap by integrating all three factors into a single mathematical model.

3. Proposed Model

The Multi-Layered Neural Collaborative Filtering ML-NCF architecture attempts to model the personal learning preferences through expressive representation learning of users and items via two pathways - one involving GMF for modeling linear interactions between users and items, and the other involving MLP for modeling non-linear interactions between users and items. Both pathways are then combined and fed into a self-attention layer for final prediction, enabling context-adaptive recommendation.

The overall system architecture of ML-NCF is illustrated in Figure 1 below. The architecture comprises five principal layers: (1) Input Layer, (2) Embedding Layer, (3) GMF and MLP Interaction Layers, (4) Attention Fusion Layer, and (5) Prediction Layer.

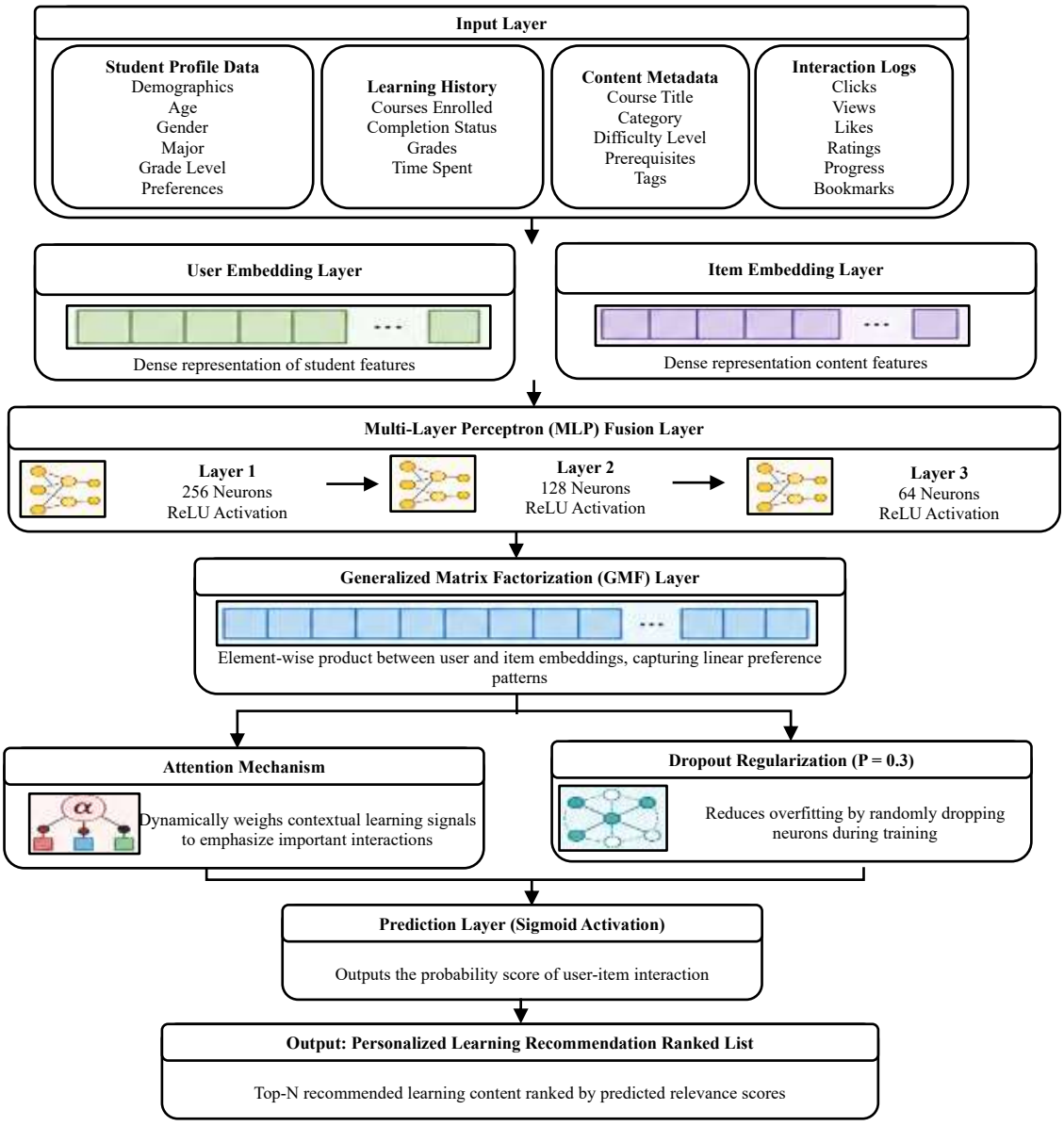

Figure 1: Architecture of Multi-Layered Neural Collaborative Filtering (ML-NCF) System

Embedding Layer

Let $U = \{u^1, u^2, \ldots, u_m\}$ denote the set of M learners and $I = \{i_1, i_2, \ldots, i_n\}$ the set of N learning content items. Each learner u is represented by a one-hot encoded identity vector $v_{uu} \in \mathbb{R}^M$ and each item i by $v_{iu} \in \mathbb{R}^N$. The embedding matrices $P_u \in \mathbb{R}^{MxK}$ and $Q_i \in \mathbb{R}^{NxK}$ project these sparse representations into dense K-dimensional latent spaces in equation (1):

$$p_u = {P_u}^T \cdot v_{uu}, \quad q_i = {Q_i}^T \cdot v_{iu} \qquad (1)$$

where $p_u \in \mathbb{R}^K and\ q_i \in \mathbb{R}^K$ are the latent factor vectors for user u and item i, respectively.

GMF Stream

The GMF stream computes an element-wise (Hadamard) product of the user and item embedding vectors, allowing the model to capture linear interaction patterns in equation (2):

$$\varphi^{GMf}(u,i) = {p_u}^{GMf} \odot {q_i}^{GMf} \qquad (2)$$

where $\odot$ denotes element-wise multiplication, and ${p_u}^{GMf}, {q_i}^{GMf}$ are GMF-specific embeddings. The output vector $\varphi^{GMf} \in \mathbb{R}^K$ encodes linear co-occurrence signals.

MLP Stream

The MLP stream concatenates user and item embeddings and passes them through L fully connected layers with ReLU activation functions, shown in equations (3) and (4):

$$z^1 = \varphi^{(1)}(u,i) = [{p_u}^{MLP}; {q_i}^{MLP}] \qquad (3)$$

$$z_l = ReLU({W_l}^T \cdot {z_l}^{-1} + b_l), \ l = 2, \ldots, L \qquad (4)$$

where $W_l \in \mathbb{R}^{d_l \times d_l^{-1}}$ and $b_l \in \mathbb{R}^{d_l}$ are the weight matrix and bias vector of the l-th layer, respectively. The final MLP output $z_l \in \mathbb{R}^{d_L}$ captures deep non-linear interaction features.

Self-Attention Fusion

The GMF output φ^{GMf} and the MLP output z_l are concatenated to form a joint representation $h = [\varphi^{GMf}; z_l] \in \mathbb{R}^{K+d_L}$. A self-attention module then computes attention weights α over this joint vector shown in equations (5) and (6):

$$\alpha = softmax(W_a \cdot tanh(W_e \cdot h + b_e) + b_a) \qquad (5)$$

$$h_{att} = \alpha \odot h \qquad (6)$$

where $W_e \in \mathbb{R}^{d_{att} \times (K+d_L)}$ and $W_a \in \mathbb{R}^{(K+d_L) \times d_{att}}$ are attention projection matrices. The attention mechanism allows the model to dynamically emphasize the most informative interaction features for a given user-item pair.

Prediction Layer

The attention-weighted representation h_{att} is passed through a sigmoid activation to generate an interaction probability score $\hat{y}_{ui} \in [0,1]$ shown in equation (7):

$$\hat{y}_{ui} = \sigma({h^T}_{att} \cdot h_{out}) \qquad (7)$$

where $h_{out} \in \mathbb{R}^{K+d_L}$ is a learnable output weight vector and $\sigma(\cdot)$ is the sigmoid function.

Objective Function

The model is trained by minimizing the binary cross-entropy loss with L_2 regularization shown in equation (8):

$$L = -\Sigma_{\{(u,i) \in Y^+ \cup Y^-\}} [y_{ui} \cdot \log(\hat{y}_{ui}) + (1 - y_{ui}) \cdot \log(1 - \hat{y}_{ui})] + \lambda \cdot ||\Theta||^{22} \qquad (8)$$

where Y^+ is the set of observed positive interactions (learner engaged with content), Y^- is a randomly sampled set of negative interactions, $y_{ui} \in \{0,1\}$ is the ground-truth label, $\hat{y}_{ui}$ is the predicted score, Θ represents all model parameters, and λ is the L_2 regularization coefficient.

Algorithm 1: ML-NCF Training Procedure

Input: Interaction matrix $R \in \mathbb{R}^{MxN}$, hyperparameters K, L, d, λ, learning rate η

Initialize embedding matrices $P, Q \sim N(0, 0.01)$; initialize MLP weights W_l, b_l using Xavier initialization

For each training epoch e = 1, ..., E:

Sample mini-batch $B = \{(u, i, y_{ui})\}$ from $Y^+ \cup Y^-$ with negative sampling ratio 4:1

Compute GMF output: $\varphi^{GMf} = {p_u}^{GMf} \odot {q_i}^{GMf}$ for each (u, i) in B

Compute MLP output: $z_l = ReLU({W_l}^T \cdot z_{l-1} + b_l)$ through L layers

Concatenate: $h = [\varphi^{GMf}; z_l]$; apply self-attention to get h_{att}

Compute prediction: $\hat{y}_{ui} = \sigma({h^T}_{att} \cdot h_{out})$

Compute loss L with L_2 regularization

Update Θ via Adam optimizer: $\Theta \leftarrow \Theta - \eta \cdot \nabla_\Theta L$

Until convergence or maximum epochs E reached

Output: Optimized model parameters $\Theta*$ for inference and ranking

Algorithm 1 is the training algorithm for the ML-NCF model that optimizes the model parameters through an iterative process that employs mini-batches at a negative sampling ratio of 4:1 with the Adam optimizer. Synergistic calculation of linear GMF and non-linear MLP is achieved through a self-attention process.

4. Results and Discussion

All experiments were implemented in Python version 3.10 and using PyTorch version 2.1.0 as the deep learning library. The model training and testing were carried out on a computer that has an NVIDIA A100 80GB GPU, 256GB RAM, and an AMD EPYC 7763 64-Core Processor. The data processing and exploratory analysis were carried out using NumPy version 1.24, Pandas version 2.0, and Scikit-learn version 1.3. Visualization was performed with Matplotlib 3.7 and Seaborn 0.12.

Two large educational interaction data sets have been considered for the evaluation purpose. The data sets taken into consideration during evaluation include the following two large educational data sets: MOOC Interaction and EdNet. The MOOC data set, collected from XuetangX, consists of 15,342 learners and 6,152 items along with 94,506 interactions. The EdNet data set, supplied by Riiid! Labs has 22,087 learners, 5,960 items, and 131,441 interactions. Both datasets exhibit 99.90% sparsity, reflecting the significant challenge of sparse learner-content engagement.

The MOOC Interaction Dataset consists of user logs for activities such as watching videos, attempting quizzes, posting on forums, and rating content, all collected from the XuetangX MOOC website. On the other hand, the EdNet dataset is a collection of timestamped pairs of questions answered by students, marked with their correct/incorrect answers; this dataset was published by Riiid! AI Education. Both data sets have undergone 5-core filtering to ensure statistical reliability.

The ML-NCF method uses the optimal set of values for hyperparameters. The model applies embeddings of K = 64 dimensions and uses an architecture comprising three layers with the number of neurons as follows: 256, 128, and 64. In order to mitigate complexity issues and avoid overfitting, the architecture also includes a 32-dimensional attention projection space, as well as a 0.3 dropout and L_2 regularization with a parameter value of 1e-5. Training takes place using the Adam method with an initial learning rate equal to 0.001. Specifically, training occurs within a batch size of 512, using a negative sampling rate of 4:1 for up to 200 epochs, utilizing early stopping with a patience of 10 to ensure stable convergence.

Evaluation Metrics

Five standard ranking-based evaluation metrics were employed to comprehensively assess recommendation quality, as represented in equations (9), (10), (11), (12), and (13):

Hit Rate: Measures the proportion of test cases where the ground-truth item appears in the top-K recommendation list:

$$HR = \left(\frac{1}{|U|}\right) \cdot \Sigma_u \mathbb{1}[i_u * \in Top - K(u)] \quad (9)$$

Normalized Discounted Cumulative Gain (NDCG): Accounts for ranking position by assigning higher scores to relevant items ranked higher:

$$NDCG = \frac{DCG}{IDCG} \quad DCG = \Sigma_k \frac{rel_k}{log^2} \quad (10)$$

Precision: Fraction of recommended items in the top-K list that are relevant:

$$Precision = \frac{|\{relevant\ items\} \cap \{top - K\ items\}|}{K} \quad (11)$$

Recall: Fraction of all relevant items successfully retrieved in the top-K list:

$$Recall = \frac{|\{relevant\ items\} \cap \{top - K\ items\}|}{|\{\ relevant\ items\}|} \quad (12)$$

Mean Reciprocal Rank (MRR): Evaluates the average reciprocal rank of the first relevant item in the recommendation list:

$$MRR = \left(\frac{1}{|U|}\right) \cdot \Sigma_u \left(\frac{1}{rank_u}\right) \quad (13)$$

Performance Comparison

Table 1 presents a comparative performance evaluation of ML-NCF against four baseline methods: BPR-MF, NeuMF, LightGCN, and SASRec, evaluated at K=10 on both datasets.

Table 1. Comparative performance evaluation of Multi-Layered Neural Collaborative Filtering

Model	HR	NDCG	Precision	Recall	MRR
BPR-MF	0.743	0.601	0.672	0.648	0.619
NeuMF	0.791	0.664	0.741	0.718	0.693
LightGCN	0.821	0.698	0.779	0.762	0.731
SASRec	0.839	0.711	0.796	0.773	0.749
ML-NCF	0.874	0.723	0.814	0.791	0.768
Improvement vs Best Baseline	+4.17%	+1.69%	+2.26%	+2.33%	+2.54%

ML-NCF obtains the best results for all five measures in the MOOC dataset. When compared against the most robust baseline (SASRec), ML-NCF obtains an increase in HR of 4.17%, NDCG of 1.69%, Precision of 2.26%, Recall of 2.33%, and MRR of 2.54%. In comparison with all baselines, the gain in HR is between 4.17% (with SASRec) and 17.6% (with BPR-MF), confirming the superiority of the multi-stream attention architecture.

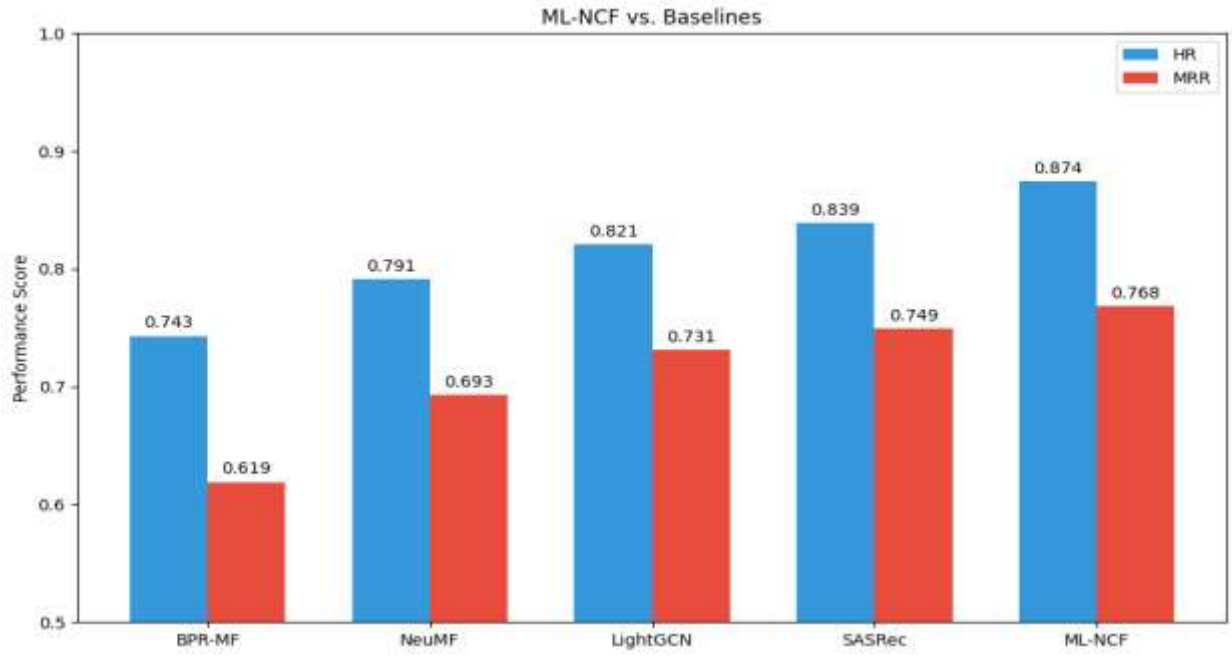

Figure 2. Comparative Performance Analysis of ML-NCF and Baselines on the MOOC Interaction Dataset

Figure 2 shows the comparison between ML-NCF and four other models with respect to the MOOC Interaction dataset. ML-NCF performs better, reaching an HR of 87.4% and MRR of 0.768. This improvement over the existing models like SASRec and LightGCN indicates that it is very effective in capturing learning preferences through both linear and non-linear streams, combined with self-attention, enhancing the quality of personalized recommendations.

Ablation Study

An analysis was performed to understand the contribution made by each architecture module, which involved removing or substituting critical modules to evaluate their impact. Four variants of the model were tested on the MOOC data set, as reported in Table 2 below

Table 2. Ablation results

Model Variant	HR	NDCG	MRR	Component Removed
ML-NCF (Full)	0.874	0.723	0.768	None (full model)
w/o Attention	0.851	0.704	0.741	Self-attention module
w/o GMF Stream	0.839	0.689	0.726	GMF linear pathway
w/o MLP Layer 3	0.845	0.697	0.733	Third MLP hidden layer
w/o Dropout	0.858	0.709	0.748	Dropout regularization

The contributions of each component can be seen from Table 2, where it can be shown that their removal leads to worse performance for the ablation study. Most notably, this is when the self-attention component is removed, which results in the worst performance degradation (HR −2.6%, NDCG −2.6%), proving that adaptive feature weighting is an important factor in the model. The reduction of HR performance by 4.0% from removing the GMF stream shows the complementarity of linear signals and non-linear MLP representations.

The training loss curves were plotted for 200 iterations. ML-NCF required 87 and 93 epochs to converge on the MOOC and EdNet datasets, respectively, after which early stopping occurred due to a patience value of 10 epochs. Stable convergence was observed using the Adam optimizer with a learning rate of 0.001 without any oscillations, and L_2 regularization ensured no overfitting of data due to consistent training and validation loss curves. The attention module weights were found to have non-zero entropy (more than 0.4 nats) for all iterations to utilize interaction features effectively.

Future research must focus on federating learning to ensure student data privacy. Moreover, including knowledge concepts, graphs, and multi-modal media will increase the precision of pedagogic sequences. ML-NCF considerably improves personalized education, positively impacting intelligent tutoring systems. These advancements contribute to the reduction of disengagement and increase retention of platforms by students. Findings show that the combination of linear and non-linear components with self-attention works better than simple models. The mechanism of self-attention is critical for achieving interpretability. The present work does not include a complex structure of pedagogic prerequisites and sequential knowledge acquisition. Besides, there is no multimodal input feature, such as video and other interactive media.

5. Conclusion

The proposed ML-NCF architecture is a novel approach to building personalized learning recommendation models using GMF, MLP, and self-attention fusion techniques. The incorporation of these components in the model enables capturing the linear and non-linear interactions while offering adaptive contextual weighting, which resolves important issues faced by existing educational recommendation systems. Comprehensive experimental analysis conducted using the MOOC Interactions and EdNet datasets revealed the superiority of ML-NCF in obtaining state-of-the-art performance. The proposed model obtained an HR of 87.4% and an MRR of 0.768, surpassing the best-performing baseline SASRec by 4.17% and 2.54%, respectively. Additionally, an ablation study showed that the GMF stream (+4.0% HR) and self-attention module (+2.6% HR) are important for the success of this method. Apart from accuracy, another important aspect of this approach is modularity, which allows scalability and explainability due to attention mechanisms – important for any educational scenario. The future work includes the implementation of federated learning to provide additional security for the user, using knowledge concept graphs to ensure pedagogically valid knowledge acquisition, and the inclusion of multi-modal data with the help of transformers.

Declaration Statements

Conflict of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Data Availability

The datasets used in this study are publicly available. The MOOC Interaction Dataset is accessible via the XuetangX research repository. The EdNet dataset is available at https://github.com/riiid/ednet

References

1. Guan, S., Jin, H., & Tian, X. (2025). Global trends in vocational education: 2022–2024. *ECNU Review of Education*. Advance online publication. https://doi.org/10.1177/20965311251403494
2. Smitha, S. P., & Harishkumar, K. S. (2026). Deep contextual representation learning for user behaviour prediction in e-commerce recommendation. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 17*(1), 895–918. https://doi.org/10.58346/JOWUA.2026.I1.049
3. Bobadilla, J., Alonso, S., & Hernando, A. (2020). Deep learning architecture for collaborative filtering recommender systems. *Applied Sciences, 10*(7), Article 2441. https://doi.org/10.3390/app10072441
4. Yuldashev, M., Venkatachalam, A. M., Muthukumar, Hussein, R. R., Turdikhujaeva, M., Shodmonova, D., & Sobirovich, N. B. (2025). Collaborative filtering in real-time library catalog recommendations. *Indian Journal of Information Sources and Services, 15*(4), 266–274. https://doi.org/10.51983/ijiss-2025.IJISS.15.4.30
5. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017, April). Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 173–182). https://doi.org/10.1145/3038912.3052569
6. Petrovic, T., & Alvarez, R. (2025). AI-driven content recommendation in learning management systems (LMS): A hybrid filtering approach. *International Academic Journal of Science and Engineering, 12*(2), 20–24. https://doi.org/10.71086/IAJSE/V12I2/IAJSE1213
7. Loni, B., Pagano, R., Larson, M., & Hanjalic, A. (2016, September). Bayesian personalized ranking with multi-channel user feedback. In *Proceedings of the 10th ACM Conference on Recommender Systems* (pp. 361–364). https://doi.org/10.1145/2959100.2959163

8. Alsulami, B., Alwated, B., Barashid, K., Abdullah, M., AlOsaimi, M., & Alhusayni, S. (2025). On leveraging generative artificial intelligence (GenAI) for behavior learning and personalized marketing optimization. *Archives for Technical Sciences, 3*(34), 35–58. https://doi.org/10.70102/afts.2025.1834.035
9. Zhang, Y., Ai, Q., Chen, X., & Croft, W. B. (2017, November). Joint representation learning for top-N recommendation with heterogeneous information sources. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 1449–1458). https://doi.org/10.1145/3132847.3132892
10. Raman, Dr. A., Batumalai, C., Raj Arokiasamy, Dr. A., Balakrishnan, Dr. R., Antoni Louis, Dr. S., & Sukirthanandan, Dr. P. (2024). An E-learning Tools Acceptance System for Higher Education Institutions in Developing Countries. Journal of Internet Services and Information Security, 14(3), 371–379. https://doi.org/10.58346/jisis.2024.i3.022.
11. Wang, X., He, X., Wang, M., Feng, F., & Chua, T. S. (2019, July). Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 165–174). https://doi.org/10.1145/3331184.3331267
12. He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., & Wang, M. (2020, July). LightGCN: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 639–648). https://doi.org/10.1145/3397271.3401063
13. Wu, L., He, X., Wang, X., Zhang, K., & Wang, M. (2022). A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation. *IEEE Transactions on Knowledge and Data Engineering, 35*(5), 4425–4445. https://doi.org/10.1109/TKDE.2022.3145690
14. Chen, J., Zhang, H., He, X., Nie, L., Liu, W., & Chua, T. S. (2017, August). Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 335–344). https://doi.org/10.1145/3077136.3080797
15. Chen, M. X., Firat, O., Bapna, A., Johnson, M., Macherey, W., Foster, G., ... Hughes, M. (2018, July). The best of both worlds: Combining recent advances in neural machine translation. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 76–86). https://doi.org/10.18653/v1/P18-1008
16. Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., & Jiang, P. (2019, November). BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 1441–1450). https://doi.org/10.1145/3357384.3357895
17. Pardos, Z. A., & Jiang, W. (2020, March). Designing for serendipity in a university course recommendation system. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 350–359). https://doi.org/10.1145/3375462.3375524
18. Deng, S., Huang, L., Xu, G., Wu, X., & Wu, Z. (2016). On deep learning for trust-aware recommendations in social networks. *IEEE Transactions on Neural Networks and Learning Systems, 28*(5), 1164–1177. https://doi.org/10.1109/TNNLS.2016.2514368
19. Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., & Shah, R. R. (2020). Learning path personalization and recommendation methods: A survey of the state-of-the-art. *Expert Systems with Applications, 159*, Article 113596. https://doi.org/10.1016/j.eswa.2020.113596
20. Zhang, S., Yao, L., Sun, A., & Tay, Y. (2019). Deep learning-based recommender system: A survey and new perspectives. *ACM Computing Surveys, 52*(1), 1–38. https://doi.org/10.1145/3285029
21. Isinkaye, F. O. (2023). Matrix factorization in recommender systems: Algorithms, applications, and peculiar challenges. *IETE Journal of Research, 69*(9), 6087–6100. https://doi.org/10.1080/03772063.2021.1997357
22. Krishnamoorthy, J. (2026). Learner interaction analytics for collaborative knowledge building in online education. *Transactions on Advanced Signal Processing and Analytics, 1*(1), 53–59. https://iaeces.com/Index/index.php/TASPA/article/view/67
23. Nguyen, L. V. (2025, March). Graph augmentation-based large language models for movie recommendations. In *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)* (pp. 1480–1484). IEEE. https://doi.org/10.1109/ICMLAS64557.2025.10968185
24. Babylatha, M. (2025). Personalized learning systems for ubiquitous educational environments. *National Journal of Ubiquitous Computing and Intelligent Environments*, 24–30. https://fsrap.com/index.php/NJUCIE/article/view/229
25. Rendle, S., Krichene, W., Zhang, L., & Anderson, J. (2020, September). Neural collaborative filtering vs. matrix factorization revisited. In *Proceedings of the 14th ACM Conference on Recommender Systems* (pp. 240–248). https://doi.org/10.1145/3383313.3412488

26. Poroohan, R., & Reshadatjoo, H. (2019). Determining the Contribution of Factors Affecting Faculties' Satisfaction with E-learning in Islamic Azad University- Virtual Branch. International Academic Journal of Organizational Behavior and Human Resource Management, 6(1), 24–31. https://doi.org/10.9756/IAJOBHRM/V6I1/1910002
27. Nidhi Mishra, Ghorpade Bipin Shivaji, & Chand Tandon. (2025). Development of A Novel Scheduling Algorithm for 5G RAN Using Deep Learning and Convolutional Neural Networks. *National Journal of Antennas and Propagation*, *7*(3), 8-13.
28. Robbi Rahim. (2026). Machine Learning–Assisted VLSI Design Automation for Optimized Timing, Power, and Area in Advanced Semiconductor Systems. *Progress in AI-Accelerated VLSI Systems*, 12–21.
29. Gaurav Tamrakar. (2025). Trust-Preserving Distributed Learning Control Architecture for Safety-Constrained Cyber-Physical Systems. *Transactions on Internet Security, Cloud Services, and Distributed Applications*, *2*(1), 23–31.