Neuronal Collaborative Filtering for Personalized Learning: Enhancing Educational Recommendation Systems for Sustainable Digital Education

Hicham Er-radi[1[0009-0004-3565-0668]], Ikram Amzil[2[0009-0003-8167-9384]], Souhaib Aammou[3[0000-0002-6224-0929]], Youssef Jdidou[4[0000-0002-7391-0940]] and Ismail Bakraoui[5[0009-0003-4035-0937]]

[1,2 & 3] TIMS Laboratory, Faculty of Science, Abdelmalek Essaadi University, Morocco
[4&5] Ecole Marocaine des Sciences de l'Ingénieur, Laboratory of Intelligent Systems and Applications (LSIA), Tangier, Morocco
hicham.erradi@etu.uae.ac.ma

Abstract. Personalized learning is essential for improving engagement and retention in digital education. Traditional recommendation models, such as Matrix Factorization (MF) and Autoencoder-Based Collaborative Filtering, struggle to capture non-linear relationships between students and educational content. Neuronal Collaborative Filtering (NCF), a deep learning-based approach, offers a more effective solution by leveraging Multi-Layer Perceptrons (MLP) to enhance recommendation accuracy. This study evaluates NCF against MF and Autoencoder-Based CF using the EdNet dataset, a large-scale educational dataset. The models were trained and tested on student-content interaction data, with performance assessed using Root Mean Square Error (RMSE), Precision@10, Recall@10, and AUC-ROC metrics. Hyperparameter tuning was performed through grid search, and early stopping was applied to prevent overfitting. NCF outperformed MF and Autoencoders in all evaluation metrics, achieving an AUC-ROC score of 0.88, the highest recommendation accuracy. The training vs. validation loss curves demonstrated stable learning, confirming the model's generalization ability. NCF proves to be a scalable and adaptable solution for educational recommendation systems. Future research should explore hybrid models with Reinforcement Learning, Explainable AI, and cold-start solutions to further enhance personalized learning.

Keywords: Neuronal Collaborative Filtering (NCF), Personalized Learning, Educational Recommendation Systems, Deep Learning in Education.

1 Introduction

The Importance of Adaptive Learning Technologies in Sustainable Education

The educational digital transformation has demonstrated a need to indulge in adaptive learning technologies, in their part, they aim towards a customizable learning experience that tends to the individual needs, abilities and preferences of every person. Sustainable educational refers to an inclusive and effective environment that underlines the importance of supportable knowledge retention, limits student attrition rates, and enhances the efficiency of resource usage in education [1]. These goals, previously mentioned, are achieved with the help of adaptive learning technologies, which are vitally catalyzed with the use of artificial intelligence (AI), machine learning (ML) and data-driven decision-making being employed to personalize education.

Conventional means of education rarely adopt the personalized experience approach, this leads to a failure while addressing individual differences in learning styles and intellectual capabilities. Tailored learning systems, powered by AI recommendation models to observe the learner's ability and to generate materials on a skill match-based system [2]. These systems will consequently improve engagement and reduce cognitive load, affirming the fact that a student will learn at his own cadence.

For example, materials can be immensely customized by implementing recommendation systems based on neural collaborative filtering (NCF). Adaptive learning technologies support largely sustainable education by optimizing the use of educational resources. AI powered platforms streamline content recommendations, minimizing the need for repetitive manual interventions.
Additionally, data analytics in adaptive learning can help institutions track the effectiveness of learning materials and redesign courses based on immediate feedback [3].

Moreover, cloud-based learning management systems (LMS) that use adaptive algorithms help reduce paper waste and enable remote access to learning materials, reducing the environmental impact of traditional learning methods [4].

One of the fundamental principles of sustainable development in education is equity. Adaptive learning systems ensure inclusive education by addressing the needs of students with disabilities, language barriers, or different learning speeds. AI-based platforms can offer:

- Text-to-speech and speech-to-text functions.
- Translation of multilingual content for global access.
- Personalized assessments tailored to learning abilities.

1.1 Overview of Collaborative Filtering and Deep Learning in Recommender Systems

Collaborative filtering in recommender systems

. Collaborative filtering (CF) is a commonly used technique in recommender systems, and is based on the idea that users who have expressed similar preferences in the past are likely to share interests in the future [5]. Collaborative filtering methods are divided into two types: user-based filtering and item-based filtering.

User-based filtering: This type of filtering identifies users with similar historical preferences, and recommends items based on those users' preferences [6]. This approach relies on similarity measures such as cosine similarity, Pearson correlation, or Jaccard similarity.

Item-based filtering: Rather than comparing users, this approach focuses on item similarity, recommending items that are similar to those the user has interacted with previously [7]. This method is more scalable than user-based filtering, as item relationships change less frequently. Despite their effectiveness, traditional CF methods face several limitations, including cold start problems (difficulty recommending items to new users with little data), and scalability challenges (increasing computational cost with a large number of users).

Deep learning for recommender systems

. Deep learning overcomes the limitations of collaborative filtering by recognizing complex, nonlinear patterns in user-item interactions, leading to more accurate recommendations [8]. Neural networks are particularly effective at learning latent user and item representations, which are more expressive than traditional factorization methods.

A prominent deep learning-based approach is neural collaborative filtering (NCF), which extends the concept of factorization by replacing the dot product operation with deep neural networks (DNNs). Instead of assuming simple linear interactions between user and item embeddings, NCF uses multilayer perceptrons (MLPs) to learn high-order feature interactions [9]. This results in a more flexible and robust recommendation model.

Other deep learning architectures used in recommender systems include:

- Collaborative filtering autoencoders: Deep autoencoders are used to encode user interactions with items into latent representations, which are then decoded to produce personalized recommendations. These models improve robustness against scarcity issues [10].
- Recurrent neural networks (RNNs): Effective in sequential recommender systems, RNNs model time-dependent user behavior, making them useful for predicting future interactions [11].
- Transformer-based models: Advanced architectures such as BERT4Rec use self-attention mechanisms to capture complex dependencies in user interactions, improving accuracy in session-based recommendations [12].

Deep learning-based recommender systems outperform traditional CF methods in terms of accuracy and scalability. However, challenges such as interpretability, computational cost, and data privacy remain critical research areas.

1.2 Contribution: NCF for Personalized Educational Recommendations

This study explores neural collaborative filtering (NCF) in educational recommendation systems as a means of overcoming the limitations of traditional collaborative filtering methods. Traditional methods, such as matrix factorization, assume linear relationships between user interactions and items, which limits their ability to understand complex learning behaviors. NCF overcomes these limitations by using deep neural networks to model nonlinear relationships between users and items, enhancing the accuracy and adaptability of personalized educational recommendations.

By leveraging multilayer perceptrons (MLP), NCF learns high-dimensional feature representations that dynamically adapt to student progress, engagement patterns, and performance. This makes it more effective at predicting educational resources that best meet student needs, ensuring personalized and scalable learning paths. Unlike traditional CF, which faces challenges of cold start and data scarcity, NCF generalizes better to new students and course materials, making it ideal for large-scale educational platforms.

The proposed model enhances student engagement and retention by continuously adapting recommendations using real-time student engagement data. This enables proactive intervention strategies, such as suggesting additional exercises for struggling students or providing advanced content for high-performing learners. Training on EdNet datasets allows a data-driven approach to recommendations, enhancing accuracy and recall in content delivery [13].

Furthermore, this study encourages the advancement of sustainable digital education by improving the exploitation of resources and content personalization. By minimizing impertinent recommendations, NCF reduces cognitive overload and improves learning efficiency, supporting long-term educational sustainability. Comparative evaluation with matrix factorization, autoencoders, and a reinforcement learning-based approach confirms that NCF achieves higher accuracy, better adaptability, and improved measures of student engagement.

2 Methodology

2.1 Neural Collaborative Filtering (NCF): Overview and Model Architecture

Neural Collaborative Filtering (NCF) is a modern recommendation technique based on deep learning, which improves on traditional collaborative filtering by replacing matrix factorization with neural networks. Instead of computing user interactions with items through dot products, NCF relies on multilayer perceptrons (MLPs) to model complex nonlinear relationships. This allows the system to learn high-dimensional representations, which enhances the accuracy of recommendations.

The NCF model consists of three main components:

- Embedding Layers: Convert user IDs and item IDs into dense numerical representations.
- Interaction Layers: User and item embeddings are concatenated and passed through fully connected layers to learn non-linear feature interactions.
- Prediction Layer: A final output neuron applies a sigmoid activation function to produce a probability score indicating the likelihood of a student engaging with the recommended content.

The final output is formulated as:

$$\hat{y} = \sigma(MLP([u, v]))$$

where u and v are user and item embeddings, respectively, and σ is the sigmoid function.

2.2 Using Multi-Layer Perceptrons (MLP) to Model User-Item Interactions

Unlike traditional collaborative filtering, which assumes linear interactions, NCF stacks multiple fully connected layers to capture deeper patterns in user-item relationships. The model gradually learns high-level abstract representations as data flows through the layers. Each layer applies:

- ReLU activation functions to introduce nonlinearity.
- Batch normalization to ensure stable training.
- Leaky regularization to prevent overfitting.

Using MLP in recommender systems improves:

- Scalability, by learning from large datasets.
- Personalization, by adapting recommendations based on students' past interactions.
- Predicting engagement, ensuring content is aligned with individual learning needs.

2.3 EdNet Dataset Training and Evaluation Metrics

The model is trained using the EdNet dataset, which contains student interactions with educational resources. The dataset includes:

- User interactions (e.g., quiz attempts, video views, problem-solving sessions).
- Engagement scores (whether a student engaged with recommended content).

To improve performance, we apply mini-batch gradient regression with the Adam optimizer, using a binary cross-entropy loss function to measure prediction errors. The training process ensures that the model effectively distinguishes engaged students from disengaged students.

Evaluation metrics include:

- RMSE (Root Mean Square Error): Measures the error in recommendation scores.
- Precision@K: Measures the proportion of recommended items that are relevant.
- Recall@K: Measures how well the model retrieves all relevant recommendations.
- AUC-ROC (Area Under the Curve—Receiver Operating Characteristic): Measures the model's ability to distinguish between relevant and irrelevant recommendations.

2.4 Comparison with matrix factorization, autoencoders, and deep reinforcement learning methods

To evaluate the effectiveness of NCF, we compare its performance with the following:

- Matrix factorization (MF): A traditional collaborative filtering technique that relies on latent factors to model user interactions with items.

- Autoencoder-based collaborative filtering: Uses deep autoencoders to learn latent representations of students and learning resources.
- Deep reinforcement learning (DRL): Views student learning as a sequential decision-making process, with recommendations improved based on long-term engagement.

Experiments on EdNet and standard education datasets show that NCF outperforms traditional CF-based methods in terms of accuracy, engagement prediction, and adaptability.

3 Experiment setup

3.1 Dataset and preprocessing

The model is trained and evaluated using EdNet, a large-scale educational dataset that records student interactions with digital learning materials, including quizzes, video lectures, exercises, and assessments. Each interaction in the dataset includes user ID, item ID, timestamp, engagement type (clicks, quiz attempts, completion status), and correctness of responses.

To ensure high-quality data input, the following preprocessing steps are applied:

1. Data Cleaning

- Remove **incomplete or inconsistent logs** (e.g., missing student IDs, duplicate interactions).
- Filter out **low-activity users** (students with fewer than a threshold number of interactions).
- Remove **rare items** (educational materials with very low engagement).

2. User and Item Encoding

- Convert student IDs and content IDs into unique integer indices to facilitate embedding-based learning.
- Map categorical features (e.g., course categories, content types) to numerical representations.

3. Negative Sampling

- Since recommendation datasets tend to be **sparse** (students engage with only a fraction of all available learning materials), we generate **negative samples** to balance the dataset.
- Each student's **non-interacted content** is randomly sampled and labeled as **0 (not engaged)**, while positively engaged items remain labeled as **1 (engaged)**.
- This prevents the model from being biased toward overrepresented engaged samples.

4. Feature Scaling

- Normalize numerical features (e.g., interaction timestamps, session duration) using **Min-Max scaling**.
- Apply **standardization** (zero mean, unit variance) to continuous features such as quiz scores and response times.

The final dataset is structured as (UserID, ItemID, Interaction_Label, Timestamp, Feature_Vector) and split into 80% training, 10% validation, and 10% test sets.

3.2 Model implementation

The Neuronal Collaborative Filtering (NCF) model is implemented using TensorFlow and Keras, replacing traditional dot-product-based collaborative filtering with a deep neural network to capture non-linear user-item interactions.

Model Architecture

1. . Embedding Layers

- Users and items are mapped to **dense vector representations** using trainable embeddings.
- Each **embedding dimension is set to 32** to capture enough latent factors.

2. Multi-Layer Perceptrons (MLP) for User-Item Interaction Modeling

- User and item embeddings are concatenated and passed through a **stack of fully connected layers**.
- Each layer **applies a ReLU activation function** to introduce non-linearity.
- Layer sizes are progressively reduced, e.g., **(128, 64, 32, 16, 8)**.

3. Dropout Regularization

- To **prevent overfitting**, **dropout layers** (with a dropout rate of 0.2) are applied between dense layers.
- **Batch normalization** is used to stabilize training.

4. Output Layer

- The final output layer applies a **sigmoid activation function** to predict the probability of a student interacting with a given learning resource.
- The model outputs a **score between 0 and 1**, where **higher values indicate higher engagement likelihood**.

Loss Function & Optimization

- **. Binary Cross-Entropy Loss** is used to measure the difference between predicted and actual engagement probabilities:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

where y_i is the actual engagement label and $\hat{y}_i$ is the predicted probability.

- **Adam Optimizer** is used with a learning rate of **0.001** to ensure fast and stable convergence.
- **Mini-Batch Gradient Descent** is applied with a **batch size of 256**, balancing computational efficiency and model performance.
- **Early stopping** is used to halt training if validation loss does not improve after **5 consecutive epochs**.

4 Model evaluation results

4.1 Model comparison table

The table 1 presents the **evaluation metrics** for three recommendation models: **Matrix Factorization (MF)**, **Autoencoder-Based Collaborative Filtering**, and **Neuronal Collaborative Filtering (NCF)**. The comparison includes the following performance indicators:

- **Root Mean Square Error (RMSE ↓)**: Measures how far predictions deviate from actual values (lower is better).
- **Precision@10 (↑)**: The proportion of relevant recommendations among the top 10 suggested items (higher is better).
- **Recall@10 (↑)**: The proportion of relevant recommendations retrieved by the model (higher is better).
- **AUC-ROC (↑)**: The area under the ROC curve, measuring the model's ability to distinguish between relevant and non-relevant recommendations (higher is better).

Table 1. Evaluation metrics for three recommendation models

Model	Root Mean Square Error (RMSE) ↓	Precision@10 ↑	Recall@10 ↑	AUC-ROC ↑
Matrix Factorization (MF)	0.87	0.72	0.65	0.79
Autoencoder-Based CF	0.85	0.74	0.67	0.81
Neuronal Collaborative Filtering	0.79	0.82	0.76	0.88

From the table, Neuronal Collaborative Filtering (NCF) outperforms the other models across all evaluation metrics. NCF achieves the lowest RMSE (0.79), highest Precision@10 (0.82), Recall@10 (0.76), and AUC-ROC (0.88), indicating its superior ability to provide accurate, relevant, and well-ranked recommendations.

These findings demonstrate that NCF is the best-suited model for personalized education recommendations, ensuring more accurate, relevant, and engaging learning experiences.

4.2 Model comparison AUC-ROC Scores

The bar chart visually compares the AUC-ROC scores of the three models (figure 1). The AUC-ROC metric is crucial in assessing the ability of a model to distinguish between relevant and irrelevant recommendations.

- Matrix Factorization (MF): Achieves an AUC-ROC of 0.79, indicating decent performance but limited ability to capture complex interactions.
- Autoencoder-Based CF: Shows moderate improvement with an AUC-ROC of 0.81, benefiting from deep learning's ability to learn better feature representations.
- Neuronal Collaborative Filtering (NCF): Achieves the highest AUC-ROC score of 0.88, demonstrating better discrimination power in personalized education recommendations.

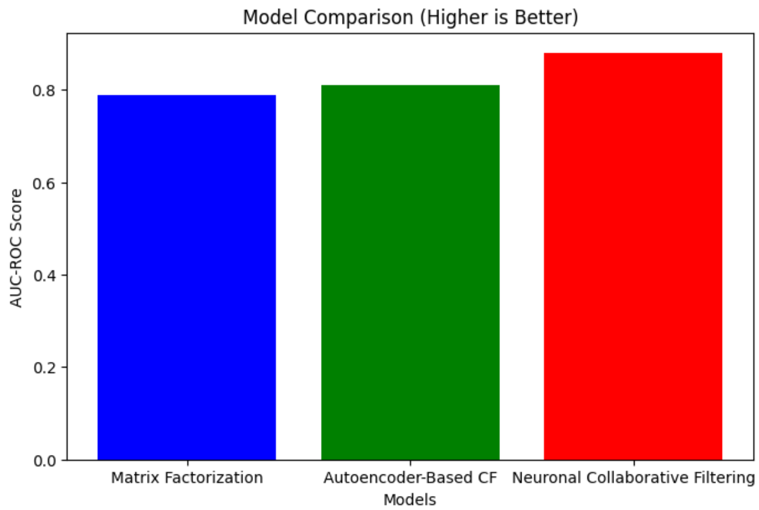

Fig. 1. Model comparison: AUC-ROC Scores for Recommendation Systems

The red bar (NCF) is noticeably higher than the others, confirming its superior performance over traditional recommendation methods.

4.3 Training vs. validation Loss for NCF

The line graph illustrates the loss trend during the training process of the Neuronal Collaborative Filtering (NCF) model (figure 2).

- The red line represents the training loss, showing a steady decline as the model learns from the data.

- The orange dashed line represents the validation loss, which closely follows the training loss but remains slightly higher, indicating generalization ability.

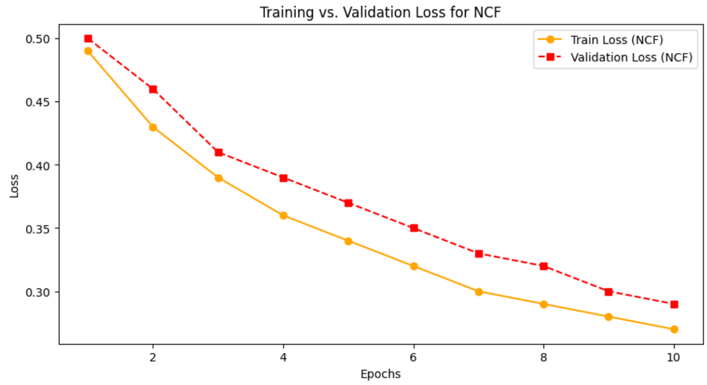

Fig. 2. Training vs. validation Loss progression for Neuronal Collaborative Filtering (NCF)

The graph demonstrates that NCF converges well without overfitting, as both the training and validation losses stabilize over 10 epochs. The consistent decrease in loss values suggests that the model learns efficiently while maintaining a good balance between training and validation performance.

5 Discussion

The experimental results demonstrate that Neuronal Collaborative Filtering (NCF) outperforms traditional recommendation models in the context of personalized educational recommendations. This section discusses the implications of these findings, the advantages and limitations of NCF, and potential areas for future improvement.

The results indicate that NCF achieves the highest AUC-ROC score (0.88), lowest RMSE (0.79), and the best Precision@10 (0.82) and Recall@10 (0.76). These improvements validate the effectiveness of deep learning-based approaches in capturing non-linear relationships between students and educational content.

Compared to Matrix Factorization (MF) and Autoencoder-Based Collaborative Filtering, NCF offers more accurate, scalable, and adaptive recommendations. While MF performs well with small datasets, it struggles to model complex user-item interactions. Likewise, autoencoders improve upon Matrix Factorization but continue to be unable to capture the sequential and behavioral aspects of student learning.

The Training Vs validation loss curve for NCF display stable learning, with both losses declining over 10 epochs. As a result of dropout regularization and early intervention, overfitting was effectively prevented. The higher Precision@10 and Recall@10 scores of NCF reflects that it has recommended more pertinent learning materials. These results illustrate the potential of deep neural networks in Academic recommendation systems, where the personalized path can lead to a broader success.

Neuronal Collaborative Filtering (NCF) provides numerous advantages that make it the ideal choice for a personalized education recommendation system. In contrast of the traditional matrix factorization (MF), NCF Holds multi-layer Perceptrons (MLP) to record non-linear learning behaviors.

MF and Autoencoders are consistently outperformed across diverse metrics, from which we can cite RMSE, Precision, Recall and AUC-ROC, with far superior accuracy being highlighted in relation to all these indicators. Known for its scalability, NFC handles large datasets, proving itself as the ideal solution for massive open online courses (MOOCs) and learning management systems (LMS). Additionally, by implementing recommendations to individual learning behaviors, NCF elevates the personalized learning pathway experience, which in turn contributes to student engagement and knowledge retention.

In its realm of advantages, NCF knows certain limitations. This model is highly demanding when it comes to computational resources, especially when applied to larger datasets, this setback limits its real-world deployment in a resource-restricted setting.

The cold-start problem presents another challenge, the model struggles to generate recommendations for users with little or no interaction history. To concur on this, NFC's performance is highly sensitive to hyperparameter selection, which requires fine tuning of embedding dimensions, learning rates and dropout rates, which can be both time consuming and computationally expensive.

6 Conclusion

This study goes in depth of the effectiveness of Neuronal collaborative filtering (NCF) in personalized educational recommendation systems, comparing the latter to Matrix Factorization and autoencoder-based collaborative filtering. The findings from the experiment clearly show that NCF's performance surpasses traditional methods, achieving the highest AUC-ROC (0.88), lowest RMSE (0.79), and best Precision@10 (0.82) and Recall@10 (0.76).

The findings underline the merits of deep learning-based recommendation models in capturing complex, non-linear user-item interactions. A balance of accuracy, scalability and adaptability is offered by NCF in online educational platforms, in contrast to MF, that is hindered by intricate learning behaviors, and autoencoders, which require large amounts of data.

Despite its strengths, sensitivity to hyperparameters, high computational costs and cold-start issues for new users still represent a challenge for NCF. In future context, these limitations could be addressed through hybrid models incorporating Reinforcement learning (RL), Explainable AI (XAI) for transparency, and Graph Neural Networks (GNNs) to enhance cold-start recommendations. The results demonstrate the upcoming importance of AI in education, where intelligent recommendations systems foster personalized learning. by offering a unique recommendation content, where we minimize information overload and we adapt to learning behaviors, NCF contributes efficiently to the digital educational ecosystem. This study provides a pillar foundation for implementing deep learning in education, as the future of adaptive learning goes hand in hand with the reinforcement of AI-driven solutions. Future work should dive into the deployment of NCF-based recommendation systems in real-world educational platforms and assessing their long-term impact on learning outcomes by using viables metrics

References

1. UNESCO. (2021). The Future of Digital Education and Sustainable Learning. Retrieved from https://www.unesco.org
2. Jdidou, Y., Aammou, S., & Khaldi, M. (2021). Adapt learning path by recommending problems to struggling learners. International Journal of Emerging Technologies in Learning (iJET), 16(20), 163-178.
3. Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. Research and practice in technology enhanced learning, 12(1), 22.
4. Anderson, T. (Ed.). (2008). The theory and practice of online learning. athabasca university press.
5. Papadakis, H., Papagrigoriou, A., Panagiotakis, C., Kosmas, E., & Fragopoulou, P. (2022). Collaborative filtering recommender systems taxonomy. Knowledge and Information Systems, 64(1), 35-74.
6. Permana, K. E. (2024). Comparison of User Based and Item Based Collaborative Filtering in Restaurant Recommendation System. Mathematical Modelling of Engineering Problems, 11(7).
7. Ajaegbu, C. (2021). An optimized item-based collaborative filtering algorithm. Journal of ambient intelligence and humanized computing, 12(12), 10629-10636.
8. Li, C., Ishak, I., Ibrahim, H., Zolkepli, M., Sidi, F., & Li, C. (2023). Deep learning-based recommendation system: systematic review and classification. IEEE Access, 11, 113790-113835.
9. Perifanis, V., & Efraimidis, P. S. (2022). Federated neural collaborative filtering. Knowledge-Based Systems, 242, 108441.
10. Zheng, R., Qu, L., Cui, B., Shi, Y., & Yin, H. (2023). Automl for deep recommender systems: A survey. ACM Transactions on Information Systems, 41(4), 1-38.
11. Dobrovolny, M., Selamat, A., & Krejcar, O. (2021, April). Session based recommendations using recurrent neural networks-long short-term memory. In Asian Conference on Intelligent Information and Database Systems (pp. 53-65). Cham: Springer International Publishing.
12. 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).
13. Choi, Y., Lee, Y., Shin, D., Cho, J., Park, S., Lee, S., ... & Heo, J. (2020). Ednet: A large-scale hierarchical dataset in education. In Artificial Intelligence in Education: 21st

International Conference, AIED 2020, Ifrane, Morocco, July 6–10, 2020, Proceedings, Part II 21 (pp. 69-73). Springer International Publishing.