Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

Course Learning Recommendation System Using Neural Collaborative Filtering

Hadist Laroibafi Mulyana[1*], **Fitrah Rumaisa**[2]
[1,2] Universitas Widyatama, Bandung, West Java 40125, Indonesia
[1]hadist.m@widyatama.ac.id, [2]fitrah.rumaisa@widyatama.ac.id

open access

***Corresponding Author**

Article History:
Submitted: 23-09-2024
Accepted: 26-09-2024
Published: 11-10-2024

Keywords:
Journal Paper; Online Learning; Recommendation System; Matrix Factorization; Machine Learning; Big Data; Neural Collaborative Filtering

ABSTRACT

The proliferation of e-learning platforms has created a need for sophisticated course recommendation systems. This paper presents an innovative online course recommendation system using Neural Collaborative Filtering (NCF), a deep learning technique designed to surpass traditional methods in accuracy and personalization. Our system employs a hybrid NCF architecture, integrating matrix factorization with multi-layer perceptron to capture complex user-course interactions. The proposed NCF-based recommendation system aims to address key challenges in the e-learning domain, such as diverse user preferences, varying course content, and evolving learning patterns. By leveraging the power of neural networks, our approach seeks to provide more relevant and personalized course suggestions to learners. Our research contributes to the intersection of deep learning and educational technology, offering new insights into how advanced machine learning techniques can be applied to improve online learning experiences. The proposed system has the potential to enhance the quality of course recommendations, leading to more effective learning pathways for users. This work has important implications for e-learning platforms, educational institutions, and lifelong learners navigating the vast landscape of online courses. By improving the match between learners and courses, we aim to increase engagement, completion rates, and overall satisfaction in online education. Future work will explore the long-term impact of such personalized recommendations on learning outcomes and skill development.

INTRODUCTION

The rapid growth of online education platforms has led to an unprecedented abundance of learning resources available to students worldwide (Ricci et al., 2015). However, this wealth of options often presents a challenge: how can learners efficiently navigate through thousands of courses to find those that best match their interests, goals, and learning styles? This problem has given rise to the need for sophisticated course recommendation systems that can effectively guide students towards relevant educational content (Jannach et al., 2010). Online course recommendation systems aim to analyze user behavior, preferences, and course characteristics to suggest personalized learning paths for individual students (Bobadilla et al., 2013). These systems not only enhance the learning experience by providing tailored suggestions but also contribute to increased engagement and completion rates on educational platforms (Ricci et al., 2015).

Traditional recommendation approaches, such as content-based filtering and collaborative filtering, have shown promise in various domains, including e-commerce and entertainment (Linden et al., 2003; Koren et al., 2009). However, the unique complexities of educational contexts—such as the long-term nature of learning goals, the interdependence of course topics, and the diversity of learning styles—call for more advanced techniques (Jannach et al., 2010; Bobadilla et al., 2013).

In recent years, deep learning approaches have demonstrated remarkable success in capturing intricate patterns and relationships within data (LeCun et al., 2015). Neural Collaborative Filtering (NCF), in particular, has emerged as a powerful framework for recommendation tasks, offering the ability to model complex user-item interactions more effectively than traditional matrix factorization techniques (He et al., 2017; Yang et al., 2020). Studies have shown that NCF consistently outperforms traditional methods, with improvements in recommendation accuracy of up to 7.7% in terms of Hit Ratio and 11.0% in terms of Normalized Discounted Cumulative Gain on certain datasets (He et al., 2017; Rendle et al., 2009). Furthermore, variants of NCF have demonstrated even higher performance gains, with some models achieving up to 12.9% improvement in Mean Average Precision compared to traditional collaborative filtering approaches (Rendle et al., 2009).

This paper presents an online course recommendation system that leverages Neural Collaborative Filtering to address the challenges specific to educational environments. Our approach combines the strengths of collaborative filtering with the representational power of neural networks to capture nuanced relationships between learners and courses (He et al., 2017). By incorporating both explicit feedback (such as ratings and reviews) and implicit feedback

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

(such as course enrollment and completion patterns), our system aims to provide more accurate and context-aware recommendations (Wang et al., 2019).

LITERATURE REVIEW

Deep Learning

Deep learning has emerged as a transformative force in artificial intelligence, driven by advances in computational power and data availability (LeCun et al., 2015). Key architectures like Convolutional Neural Networks (CNNs) have revolutionized computer vision (Krizhevsky et al., 2012), while Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) networks, have advanced sequential data processing (Hochreiter & Schmidhuber, 1997). The introduction of the Transformer architecture has led to significant progress in natural language processing, forming the basis for powerful language models (Vaswani et al., 2017).

Recent developments in deep learning include Generative Adversarial Networks (GANs), which have opened new avenues in generative modeling and unsupervised learning (Goodfellow et al., 2014). Despite its successes, deep learning continues to face challenges in areas such as interpretability, robustness against adversarial attacks, and ethical considerations. Ongoing research aims to address these issues while further advancing the field's capabilities in transfer learning, self-supervised learning, and application to diverse domains.

Recommendation System

Recommendation systems have become integral to many online platforms, helping users navigate vast amounts of information and products. Recent developments in recommendation systems have focused on incorporating deep learning techniques to capture complex user-item interactions and context information. Neural network-based approaches, such as neural collaborative filtering and deep factorization machines, have shown promising results in various domains (He et al., 2017). Despite their success, recommendation systems continue to face challenges in areas such as cold-start problems, scalability, and privacy concerns. Ongoing research aims to address these issues while improving recommendation accuracy and user satisfaction.

Neural Collaborative Filtering

Neural Collaborative Filtering (NCF) has emerged as a powerful approach in recommendation systems, combining the strengths of neural networks with collaborative filtering techniques. Traditional collaborative filtering methods, such as matrix factorization, have been widely used but often struggle to capture complex user-item interactions. NCF addresses this limitation by leveraging deep neural networks to learn non-linear relationships from user-item interaction data (He et al., 2017). The key innovation of NCF lies in its ability to replace the inner product used in traditional matrix factorization with a neural architecture, allowing for more expressive models of user-item interactions.

The fundamental work on NCF by He et al. (2017) introduced a general framework that unifies matrix factorization and multilayer perceptron models. This approach has since inspired various extensions and improvements. For instance, Wang et al. (2019) proposed the Neural Graph Collaborative Filtering (NGCF) model, which explicitly encodes the collaborative signal in the form of high-order connectivity in user-item bipartite graphs. More recent advancements have focused on incorporating additional contextual information and improving model efficiency. Despite its success, NCF still faces challenges in areas such as model interpretability and the cold-start problem. Ongoing research aims to address these issues while further improving recommendation accuracy and generalization ability.

RESEARCH METHODOLOGY

Research Design

This research proposes a system design that can recommend data using Neural Collaborative Filtering (NCF). There are several stages in achieving the final result, which form a common architecture used in almost all types of algorithms. Below is the implementation:

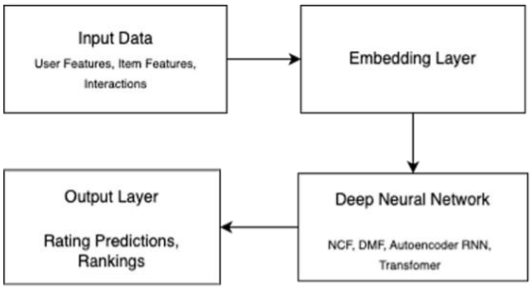

Figure 1. Research Design

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

In the illustration above, input data such as user features, item features, and interactions are passed through an embedding layer to obtain latent representations. These latent representations are then processed by deep learning architectures such as NCF, DMF, autoencoder, RNN, or Transformer to learn complex patterns. Finally, the output layer produces a prediction of ratings or a ranking of recommended items

Dataset

The data used consists of 8,093 online course listings available online, obtained from several online learning platforms. Additionally, there are about 1,000 personalized user data entries, each with specified interests. The data will be combined and is expected to serve as the foundational material for the model training to be conducted this time

Neural Collaborative Filtering

Neural Collaborative Filtering (NCF) first emerged around 2017, building on traditional neural network-based recommendation techniques. NCF captures user-item interactions through a multi-layer representation, typically expressed as $\hat{y}_{ui}$. The data used in this context is implicit feedback derived from interactions between news articles and readers, resulting in a user-item interaction matrix with values as expressed in the following equation,

$$\hat{y}_{u,i} = f(\boldsymbol{u}, \boldsymbol{v})$$

The output value can be represented as follows: Let $\boldsymbol{u}$ be the user representation vector, $\boldsymbol{v}$ be the item representation vector, and $\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})$ be a function that describes the interaction between the user and the item. In many implementations, the function $\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})$ is implemented as a neural network that produces a score or probability, indicating how likely it is that the user will like a particular item.

Matrix Factorization Algorithm

NCF is a neural matrix factorization model that combines Generalized Matrix Factorization (GMF) and Multi-Layer Perceptron (MLP) to capture both linear and non-linear relationships between users and items. GMF effectively models linear interactions, while MLP learns complex patterns. The following figure illustrates how this hybrid approach enables NCF to deliver accurate and personalized recommendations:

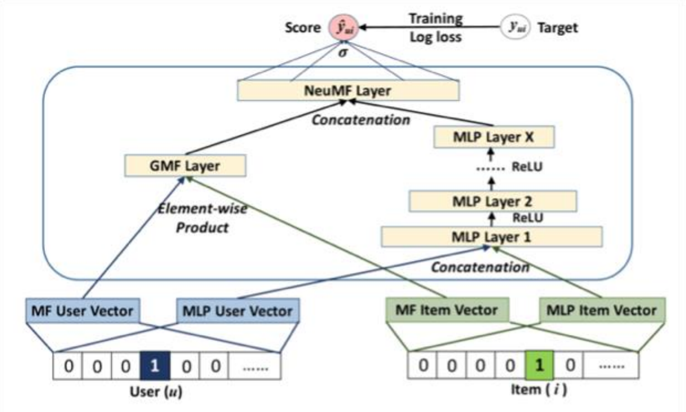

Figure 2. Neural Collaborative Filtering Framework

This figure illustrates the use of latent vectors for items and users, followed by the fusion of outputs from the GMF layer (left) and the MLP layer (right). In the following sections, we will introduce this framework and explain how to learn the model parameters.

The Generalized Matrix Factorization Model

GMF extends standard MF by introducing a neural CF layer as the output layer. This generalization allows for greater flexibility in modeling user-item interactions. For instance, by learning the edge weights without constraints, GMF can assign varying importance to latent dimensions. Additionally, employing a non-linear activation function allows the model to capture more complex relationships compared to linear matrix factorization (MF). The formula depicted as follows:

$$\hat{r}_{ui} = a_{out}(h^T(q_i \odot p_u))$$

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

The formula above predicts user-item interactions by combining latent representations of users (p_u) and items (q_i). It first multiplies these representations element-wise, then calculates a weighted sum using a hidden representation (h^T), and finally applies an activation function (a_{out}) to obtain the predicted rating. This approach allows GMF to learn complex patterns and relationships between users and items, providing more accurate and personalized recommendations.

The Multi-Layer Perception Model

CF employs two pathways to model user-item interactions: 1) element-wise multiplication of latent vectors and 2) concatenation of latent vectors. To learn complex interactions between concatenated user and item features, a standard MLP model is applied. This approach provides the model with a high degree of flexibility and non-linearity, enabling it to effectively capture intricate relationships between users and items:

$$z_1 = \phi_1(p_u, q_i) = \begin{bmatrix} p_u \\ q_i \end{bmatrix}$$

And the details of the hidden layers and output layer for the MLP are as follows:

$$\phi_l(z_l) = a_{out}(W_l^T z_l + b_l), (l = 2,3, \dots, L-1)$$

and:

$$\hat{r}_{u,i} = \sigma(h^T \phi(z_{L-1}))$$

Above, W_l , b_l, and a_{out} represent the weight matrix, bias vector, and activation function for the perceptron in the l-th layer, respectively. For the activation functions in the MLP layers, commonly used options include sigmoid, hyperbolic tangent (tanh), and Rectified Linear Unit (ReLU), among others. Given that this is a binary classification task, the activation function for the output layer is set to the sigmoid function, $\sigma(x) = 1 + e - x1$ ensuring the predicted score is constrained within the range (0,1)

Fusion of Generalized Matrix Factorization and Multi Layer Perception

To enhance flexibility, we allow GMF and MLP to learn separate embeddings and combine their outputs by concatenating their final hidden layers. Then we get ϕ^{GMF} from GMF:

$$\phi_{u,i}^{GMF} = p_u^{GMF} \odot q_i^{GMF}$$

and obtain ϕ^{MLP} from MLP:

$$\phi_{u,i}^{MLP} = a_{out}(W_L^T(a_{out}(\dots a_{out}(W_2^T \begin{bmatrix} p_u^{MLP} \\ q_i^{MLP} \end{bmatrix} + b_2) \dots)) + b_L$$

Then, the results of the GMF and MLP models are combined to generate the final prediction:

$$\hat{r}_{u,i} = \sigma(h^T \begin{bmatrix} \phi^{GMF} \\ \phi^{MLP} \end{bmatrix})$$

This model effectively leverages the complementary strengths of Matrix Factorization (MF) and Deep Neural Networks (DNNs) to model user-item latent structures.

Binary Cross-Entropy

Binary Cross-Entropy (BCE) is a widely used loss function in binary classification tasks, including in our model where we predict whether a user will be interested in a course or not. The BCE formula is:

$$BCE(y, y\ \ ) = -[ylog(y\ \ ) + (1-y)log(1-y\ \ )]$$

Where y is the actual label (0 or 1), and $\hat{y}$ is the predicted probability. This function penalizes the model based on how far the predicted probability deviates from the true label. We use BCE because it handles probabilistic outputs effectively, offering smooth gradient updates, accurate probability estimates, and the ability to balance class imbalances when necessary. In this context, it helps ensure the model minimizes errors in predicting user preferences, improving recommendation accuracy.

Brilliance
Research of Artificial Intelligence
E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

DISCUSSION

The testing scenario discusses the test data and the testing scenario for accuracy using Neural Collaborative Filtering (NCF) on the course online dataset of 8,093 entries and 1,000 personalized user data entries with their interests linked with course dataset based on their categories. We evaluated the impact of different learning rates and batch sizes on the model's loss and number of positive predictions. The objective was to identify the optimal configuration that balances performance (as measured by loss) and the model's ability to classify positive cases. We explored two learning rates, 0.001 and 0.0001, across three batch sizes: 128, 256, and 512. Below, we discuss the findings based on these hyperparameters and their effects on model behavior.

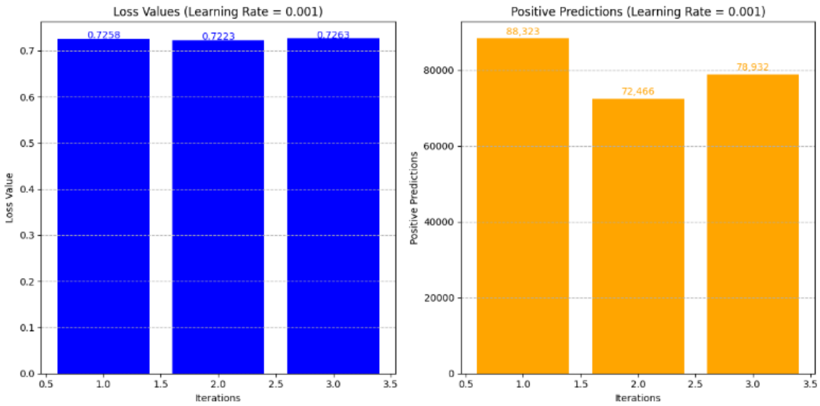

Figure 3. Performance of learning rate 0.001

Table 1. List of learning iteration based on different learning rate and batch sizes

Iteration	Learning Rate	Batch Size	Loss	Positive Prediction
1	0.001	128	0.7258	88323
2	0.001	256	0.7223	72466
3	0.001	512	0.7263	78932
4	0.0001	128	0.7273	82494
5	0.0001	256	0.7239	78808
6	0.0001	512	0.7334	80528

Performance Overview

The bar chart illustrates the performance of a Neural Collaborative Filtering model at a learning rate of 0.001 across three iterations. It compares loss values (in blue) and positive predictions (in orange) for each iteration. The chart shows that the loss decreases from 0.7258 to 0.7223 in the first two iterations, indicating effective learning, but slightly increases to 0.7263 in the third iteration, suggesting potential instability. Meanwhile, positive predictions fluctuate, starting at 88,323, dropping to 72,466, and then rising to 78,932. This visualization highlights the relationship between model performance metrics and underscores the importance of monitoring both loss and predictions during training.

The bar chart presents the performance of a Neural Collaborative Filtering model at a learning rate of 0.0001 across three iterations. It displays the loss values (in blue) and positive predictions (in orange) for each iteration. The loss initially decreases from 0.7273 in the fourth iteration to 0.7239 in the fifth iteration, indicating some improvement. However, it then increases to 0.7334 in the sixth iteration, suggesting instability in learning. The positive predictions fluctuate, starting at 82,494, decreasing to 78,808, and then rising to 80,528. This visualization emphasizes the need to monitor both loss and prediction metrics closely, as they provide valuable insights into the model's performance during training.

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

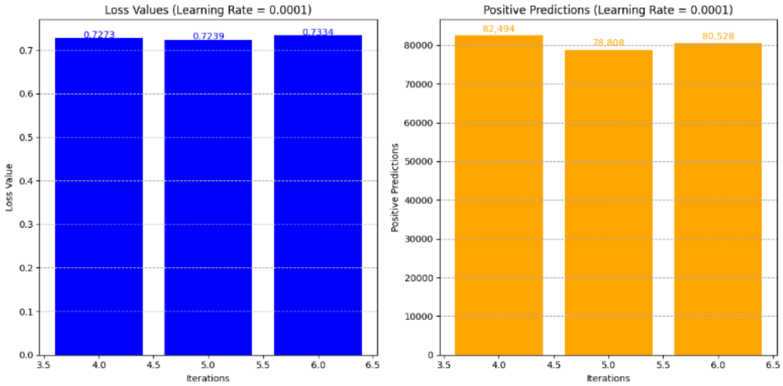

Figure 4. Performance of learning rate 0.0001

Analysis of Batch Size Impact

While the primary focus of the study was on learning rates, the impact of batch size on model performance was also examined. However, the results did not reveal a clear correlation between batch size and loss values within the iterations observed. It is essential to note that the variations in loss were predominantly influenced by the learning rates rather than the batch sizes used in this experiment.

The lack of significant impact from batch size suggests that other factors, such as the complexity of the model and the characteristics of the dataset, may play a more pivotal role in determining how effectively the model learns from the data. This finding prompts further investigation into the interplay between batch size and learning rate, as well as how they collectively influence the training dynamics.

RESULT

Overall best result

The second iteration with a learning rate of 0.001 combined with a batch size of 256 achieved the best overall result for the Neural Collaborative Filtering model, recording a loss value of 0.7223. This represents a significant improvement over the first iteration, which had a higher loss of 0.7258, despite generating the highest positive predictions at 88,323. The lower loss in the second iteration, coupled with a reasonable number of positive predictions (72,466), indicates that the model was more effectively capturing underlying data patterns while also maintaining accuracy in its predictions.

Table 2. Iteration of learning

Epoch	Accuracy	Loss
1	0.6663	0.9588
2	0.7060	0.7455
3	0.7150	0.7355
4	0.7099	0.7377
5	0.7108	0.7341
6	0.7134	0.7287
7	0. 7173	0. 7304

The training results across the first seven epochs highlight a positive trend in model performance, showcasing steady improvements in both accuracy and loss. Beginning with Epoch 1, the model achieved an accuracy of 66.63% and a loss of 0.9588, setting a solid foundation for learning. By Epoch 2, the model made significant strides, reaching an accuracy of 70.60% and reducing the loss to 0.7455—a clear indication of effective learning.

The momentum continued through subsequent epochs, with accuracy climbing to a peak of 71.73% by Epoch 7 and loss decreasing to 0.7304. This consistent improvement reflects the model's ability to adapt and refine its predictions over time. Overall, these results demonstrate a successful training process, showcasing the model's capacity for growth and its potential for further enhancements in future iterations.

The training progress over 7 epochs reveals interesting dynamics. Training accuracy improves steadily from 66.63% to 71.73%, while training loss decreases from 0.9588 to 0.7304, indicating consistent learning. However, validation metrics show high volatility: validation accuracy fluctuates between 72.66% and 86.92%, and validation loss varies from 0.2620 to 0.3486. This instability in validation performance suggests potential overfitting or dataset peculiarities.

E-ISSN : 2807-9035
Volume 4, Number 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

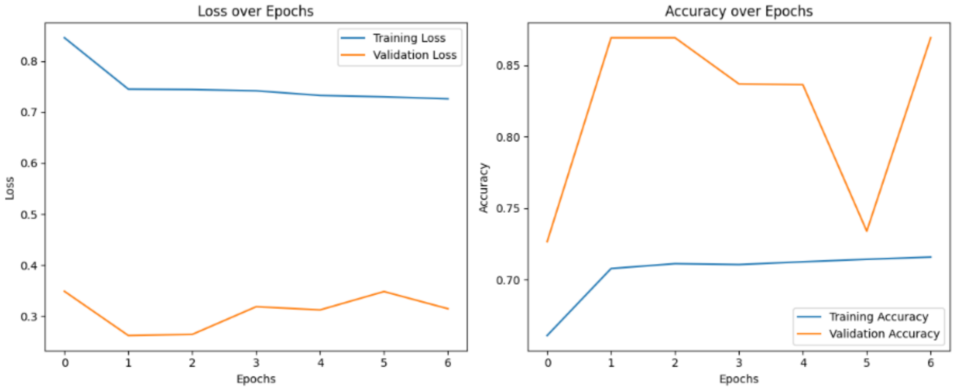

Figure 5. Performance overview

The learning rate is adjusted during training, starting at 0.001 and decreasing to 0.00025 by the 7th epoch, likely to fine-tune model performance. Final training metrics show a loss of 0.7260 and accuracy of 71.59%. However, the model's predictive performance reveals challenges, with very small predicted values (around 0.00013) and a tendency to predict mostly negative cases (class 0).

Performance metrics indicate a trade-off between precision (0.3093) and recall (0.7156), resulting in an F1-score of 0.4319.

CONCLUSION

The results from this experiment suggest that a learning rate of 0.001 combined with a batch size of 256 provides the best overall performance, minimizing loss while maintaining a reasonable number of positive predictions. For tasks where precision and low loss are prioritized, the batch size of 256 is the optimal choice. Conversely, for applications where recall is more important, the batch size of 128 can be considered, as it results in more positive predictions at the expense of slightly higher loss. A batch size of 512 represents a middle ground, balancing between minimizing loss and increasing positive predictions but does not outperform the smaller batch sizes.

Further experimentation could explore learning rate scheduling and other batch size combinations to fine-tune the model further. Additionally, it would be beneficial to assess these configurations on a validation set to determine generalization capabilities beyond the training set.

REFERENCES

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, *46*, 109-132.

Chen, J., Zhang, H., He, X., Nie, L., Liu, W., & Chua, T. S. (2017). 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).

Chen, J., Zhang, H., He, X., Nie, L., Liu, W., & Chua, T. S. (2017). Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 173-182).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (Vol. 27).

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T. S. (2017). Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 173-182).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780.

Huang, Z., Chen, H., & Zeng, D. (2004). Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems*, *22*(1), 116-142.

Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: An introduction*. Cambridge University Press.

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, *42*(8), 30-37.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (Vol. 25).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, *7*(1), 76-80.

Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2009). BPR: Bayesian personalized ranking from

Brilliance
Research of Artificial Intelligence

E-ISSN : 2807-9035
Volume 4, **Number** 2, November 2024
https://doi.org/10.47709/brilliance.v4i2.4699

implicit feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence* (pp. 452-461).

Ricci, F., Rokach, L., & Shapira, B. (2015). *Recommender systems handbook*. Springer.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30).

Wang, H., Wang, N., & Yeung, D. Y. (2015). Collaborative deep learning for recommender systems. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1235-1244).

Wang, X., He, X., Wang, M., Feng, F., & Chua, T. S. (2019). Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 165-174).

Yang, C., Bai, L., Zhang, C., Yuan, Q., & Han, J. (2020). Neural collaborative filtering vs. matrix factorization revisited. In *Proceedings of the 14th ACM Conference on Recommender Systems* (pp. 81-89).

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.