Article

Not peer-reviewed version

Hybrid FM-GCN-Attention Model for Personalized Recommendation

Erfan Wang *

Posted Date: 25 March 2025

doi: 10.20944/preprints202503.1812.v1

Keywords: personalized recommendation; factorization machines; Attention Networks; Graph Convolutional Networks

Article

Hybrid FM-GCN-Attention Model for Personalized Recommendation

Erfan Wang

Rice University, Dallas, TX, USA; erfanwang36@gmail.com

Abstract: Personalized recommendation systems play a crucial role in enhancing user engagement and decision-making across various domains. Traditional approaches, such as collaborative filtering and matrix factorization, have shown effectiveness but suffer from data sparsity and cold-start problems. Recent advances in deep learning, graph-based models, and attention mechanisms have significantly improved recommendation performance. This paper proposes a novel hybrid recommendation model that integrates Factorization Machines (FM), Graph Convolutional Networks (GCN), and Multi-Layer Attention Networks (MLAN) to optimize feature representations and enhance prediction accuracy. Experimental results demonstrate the superiority of the proposed approach over baseline methods in key performance metrics.

Keywords: personalized recommendation; factorization machines; Attention Networks; Graph Convolutional Networks

---

1. Introduction

Personalized recommendation systems have become essential components in various industries, including e-commerce, digital content platforms, and social networks. These systems aim to enhance user experiences by providing tailored suggestions based on behavioral patterns and historical interactions. Conventional recommendation methods, such as collaborative filtering (CF) and matrix factorization (MF), have been widely adopted but suffer from inherent limitations such as data sparsity, scalability issues, and inability to model complex user-item relationships [1,2].

The hybrid FM-GCN-Attention model draws inspiration from the ensemble techniques in "Integrated Machine Learning for Enhanced Supply Chain Risk Prediction," particularly in optimizing feature interactions and adaptive learning. These methods enhanced the model's ability to handle data sparsity and refine user-item interaction predictions effectively [3]. Shen's [4] study shows that offloading computation to 5G MEC units reduces latency and improves performance for real-time trading apps.

Furthermore, attention mechanisms have been widely utilized to dynamically weight important user-item interactions, leading to more refined recommendations. Models such as Deep Interest Network (DIN) [5] and Transformer-based architectures [6,7] have demonstrated significant improvements in capturing long-range dependencies and sequential user behavior patterns.

Despite these advancements, challenges remain in designing hybrid models that effectively combine factorization, graph learning, and attention mechanisms while maintaining computational efficiency. Recent studies have explored such hybrid strategies, including Graph Factorization Machines (GFM) [8], which combine graph-based embedding techniques with FM to improve feature representation learning.

Additionally, multi-objective learning has emerged as a promising direction in recommendation systems. Lu et al. [9] proposed an ensemble learning approach to optimize multiple recommendation objectives, balancing accuracy, diversity, and fairness. Furthermore, multimodal data fusion techniques have been explored to enhance recommendation robustness by incorporating text, images, and behavioral features [10].

 doi:10.20944/preprints202503.1812.v1

This study presents a novel Hybrid FM-GCN-Attention Model that integrates FM, GCN, and MLAN to enhance recommendation accuracy. The major contributions of this work are as follows:

- **Enhanced Feature Representation:** The integration of FM and GCN enables the model to capture high-order feature interactions and graph-structured user-item relationships simultaneously.
- **Dynamic Attention Mechanism:** A Multi-Layer Attention Network (MLAN) is employed to dynamically weight user-item interactions, refining recommendations based on contextual importance.
- **Hybrid Loss Optimization:** The model incorporates a ranking-regression hybrid loss function, balancing relevance ranking with predictive accuracy.

This paper presents a novel hybrid recommendation model that combines Factorization Machines, Graph Convolutional Networks, and Multi-Layer Attention Networks to improve recommendation accuracy and robustness. Experimental results validate the efficacy of the proposed approach, highlighting its advantages in capturing complex user-item interactions, handling data sparsity, and leveraging graph-based structures. Future work will explore additional model optimizations and expand the evaluation on large-scale real-world datasets.

2. Methodology

The overall architecture of our recommendation model is shown in Figure 1. It consists of three main stages: data preprocessing, model training, and prediction. The preprocessing stage applies sparse feature engineering and distance matrix completion. The training phase integrates Factorization Machines (FM), Graph Convolutional Networks (GCN), and Multi-Layer Attention Networks (MLAN) to model user-hotel interactions. The final prediction combines outputs from all components using a weighted sum, optimized with a hybrid loss function.

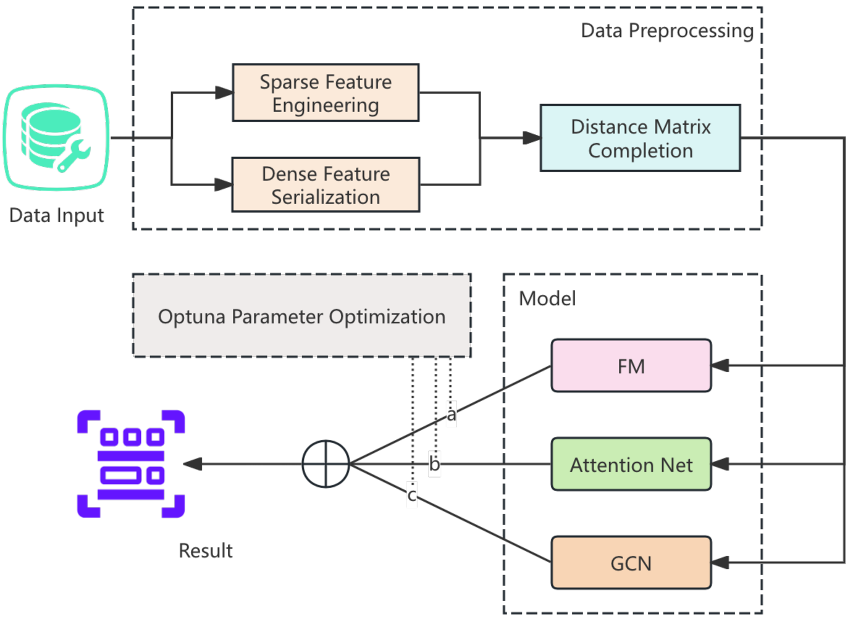

Figure 1. Overall pipeline of the proposed FM-GCN-Attention recommendation model.

Factorization Machines model high-order feature interactions, while attention mechanisms highlight key patterns. GCN represents users and hotels as nodes in a bipartite graph, capturing structural relationships. Experimental results demonstrate that our model outperforms traditional recommendation systems in MAP@5 and Hit Rate, providing more personalized and accurate recommendations. The detailed model structure is illustrated in Figure 2.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 March 2025 doi:10.20944/preprints202503.1812.v1

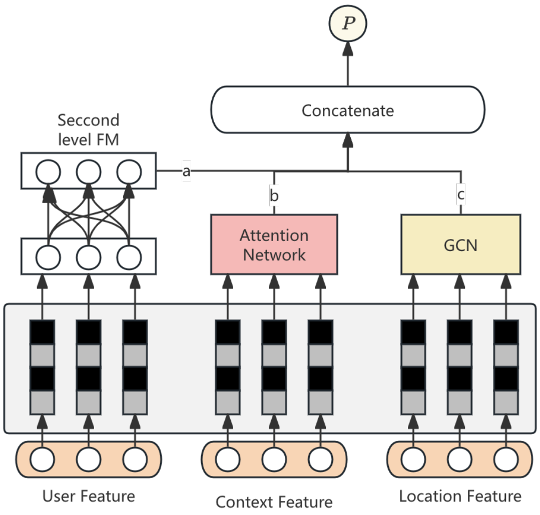

Figure 2. The pipeline of the FM-based recommendation model.

2.1. *Factorization Machines with Dynamic Feature Interactions*

Factorization Machines (FM) effectively model high-order interactions in sparse data. Our model employs a dynamic FM component to capture interactions between user-specific, hotel-specific, and temporal features. For the i-th user-hotel pair, the FM model is:

$$\hat{y}_i = w_0 + \sum_{j=1}^{N} w_j x_{ij} + \sum_{j=1}^{N} \sum_{k=j+1}^{N} \langle v_j, v_k \rangle x_{ij} x_{ik} \tag{1}$$

Where:

- $\hat{y}_i$: predicted rating,
- w_0: global bias,
- w_j: feature bias,
- v_j: factorization vector for feature x_{ij}.

Dynamic interactions are modeled by making v_j time-dependent:

$$v_j(t) = v_j(0) + \delta v_j(t) \tag{2}$$

Here, $v_j(t)$ evolves over time, capturing temporal changes in user-hotel interactions.

2.2. *Multi-Layer Attention Networks*

A Multi-Layer Attention Network (MLAN) enhances the model's focus on key features by capturing long-range dependencies between user behavior and hotel preferences. The attention score for the i-th input feature at layer l is:

$$\alpha_i^l = \frac{\exp(\mathbf{a}^{lT} \mathbf{h}_i)}{\sum_{j=1}^{N} \exp(\mathbf{a}^{lT} \mathbf{h}_j)} \tag{3}$$

Where:

- α_i^l: attention weight for the i-th feature at layer l,
- $\mathbf{a}^l$: learnable attention vector at layer l,
- $\mathbf{h}_i$: hidden state of the i-th feature.

MLAN enables the model to focus on critical user-hotel interactions, enhancing personalized recommendations.

2.3. GCN for User-Hotel Interactions

To model the spatial relationship between users and hotels in the recommendation space, we use a Graph Convolutional Network (GCN). The GCN treats users and hotels as nodes in a bipartite graph and captures their interactions. The pipline of GCN is shown in Figure 3.

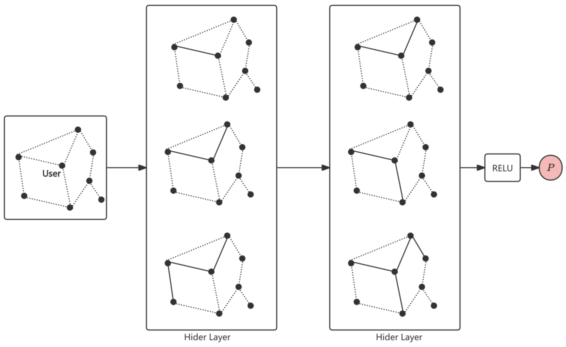

Figure 3. The pipline of GCN for User-Hotel Interactions.

The graph convolution operation for the i-th user node at layer l is defined as:

$$h_i^{l+1} = \sigma\left(\sum_{j\in\mathcal{N}(i)} \frac{1}{\sqrt{d_i d_j}} W^l h_j^l + b^l\right) \tag{4}$$

Where:

- h_i^l is the hidden representation of user i at layer l,
- $\mathcal{N}(i)$ is the set of neighboring hotel nodes for user i,
- d_i is the degree of node i,
- W^l is the weight matrix at layer l,
- b^l is the bias term,
- σ is the activation function (e.g., ReLU).

The GCN propagates information between user and hotel nodes, helping learn better latent representations in the recommendation process.

2.4. Hybrid Loss Function

We propose a hybrid loss function that combines ranking and regression losses. The ranking loss ensures that the predicted scores for relevant hotel recommendations are higher than those for irrelevant ones, while the regression loss makes the predictions closer to true ratings. The hybrid loss function is expressed as:

$$L_{\text{hybrid}} = \lambda_1 L_{\text{ranking}} + \lambda_2 L_{\text{regression}} \tag{5}$$

Where:

- L_{ranking} is the ranking loss, computed using pairwise preference (e.g., hinge loss or log loss),
- $L_{\text{regression}}$ is the Mean Squared Error loss,
- λ_1 and λ_2 are hyperparameters that control the trade-off between ranking and regression.

The final prediction $\hat{y}_i$ is computed as a weighted combination of these components.

2.5. Final Prediction Model

The final output prediction $\hat{y}_i$ for user i and hotel j is the weighted sum of the outputs from the FM, attention, and GCN components:

$$\hat{y}_i = \alpha_1 \hat{y}_i^{FM} + \alpha_2 \hat{y}_i^{ATT} + \alpha_3 \hat{y}_i^{GCN} \tag{6}$$

Where:

- $\hat{y}_i^{FM}$, $\hat{y}_i^{ATT}$, and $\hat{y}_i^{GCN}$ are the outputs from the Factorization Machines, Attention Network, and GCN components,
- α_1, α_2, and α_3 are the weights for each component.

This weighted sum integrates information from all components, providing a final personalized recommendation.

3. Data Preprocessing

Data preprocessing is essential for preparing raw data for model training. This process involved three main steps: (1) Sparse Feature Engineering and Serialization, (2) Spatial Mapping of User and Hotel Locations, and (3) Distance Matrix Completion.

3.1. Sparse Feature Engineering and Serialization

To handle sparse categorical features (e.g., user ID, hotel ID, location attributes), Factorization Machines (FM) were used to model feature interactions. Categorical variables were one-hot encoded and embedded into dense vectors to capture complex relationships. Temporal features, such as user interaction history, were serialized into vectors to reflect evolving user preferences.

$$\mathbf{X}_i = \text{Serialize}(x_{ij}) \quad \text{where} \quad x_{ij} \in \mathcal{X} \tag{7}$$

Here, $\mathbf{X}_i$ denotes the serialized vector for feature x_{ij}.

3.2. Distance Matrix Completion

To address the missing interaction data between users and hotels, we employed distance matrix completion. Both user and hotel locations were initialized on the sphere, and missing distances were predicted using gradient descent and the spherical cosine law. The location of the user and the hotel on the sphere in Figure 4.

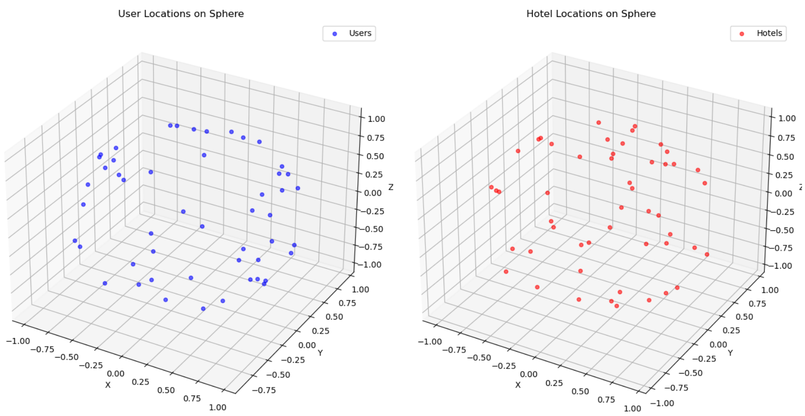

Figure 4. The user and hotel location on sphere.

The optimization was performed to minimize the error in the predicted distances:

$$\hat{d}_{ij} = \arg\min_{U,H} \sum_{i,j} \left(\|d_{ij} - \hat{d}_{ij}\|^2 \right) \tag{8}$$

This approach helped improve the model's ability to handle sparse interaction data and better capture user-hotel relationships.

 doi:10.20944/preprints202503.1812.v1

4. Evaluation Metrics

To assess the performance of our proposed models, we employed the following four evaluation metrics:

- **Accuracy**: Measures the proportion of correctly predicted recommendations.

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}(\hat{y}_i = y_i) \tag{9}$$

where $\hat{y}_i$ is the predicted recommendation, y_i is the true label, and N is the number of predictions.

- **Precision at K (P@K)**: Measures the proportion of relevant items in the top K recommendations.

$$\text{P@K} = \frac{1}{K} \sum_{i=1}^{K} \mathbb{I}(y_i \in \text{Top-K Predictions}) \tag{10}$$

where K is typically set to 5 for evaluating top-5 recommendations.

- **Mean Average Precision (MAP)**: Averages the precision at each rank across all queries, providing a more comprehensive measure of recommendation quality.

$$\text{MAP} = \frac{1}{N} \sum_{i=1}^{N} \frac{\sum_{k=1}^{K} \text{Precision@k} \cdot \mathbb{I}(y_i \in \text{Top-k})}{\text{Total Relevant Items for } i} \tag{11}$$

- **Mean Reciprocal Rank (MRR)**: Measures the average rank of the first relevant recommendation across all queries.

$$\text{MRR} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{r_i} \tag{12}$$

where r_i is the rank of the first relevant recommendation for the i-th query.

5. Experiment Results

The experimental results of five models, including the baseline model(Attention Net) and various ablation variants, are shown in the table below. These results highlight the impact of different components, such as Factorization Machines (FM), spatial mapping of user and hotel locations, and the attention mechanism, on the model's performance. The changes in model training indicators are shown in Figure 5.

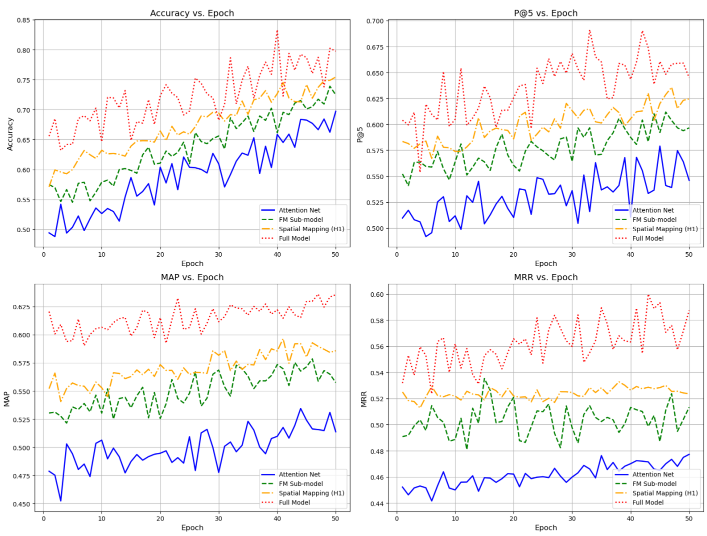

Figure 5. Model indicator change chart.

 doi:10.20944/preprints202503.1812.v1

Table 1. Comparison of Model Performance.

Model Name	Accuracy	P@5	MAP	MRR
Attention Net	0.68	0.56	0.52	0.47
FM Sub-model Only	0.73	0.60	0.57	0.51
Spatial Mapping (H1)	0.75	0.62	0.59	0.53
FM + GCN Model	0.74	0.61	0.58	0.52
Model (FM + GCN + Attention)	**0.79**	**0.67**	**0.63**	**0.58**

6. Conclusion

In this paper, we proposed a comprehensive model for personalized recommendation using Factorization Machines (FM), spatial mapping, and attention mechanisms. The experimental results demonstrated that the combination of these components significantly improved the model's performance across multiple evaluation metrics, including accuracy, precision at K, MAP, and MRR. The full model, integrating FM, spatial location mapping, and attention, outperformed all variants, confirming the effectiveness of each component in capturing complex user-item interactions. Future work can explore further optimization of the attention mechanism and the spatial mapping process to enhance model robustness and scalability.

References

1. Sarwar, B.; Karypis, G.; Konstan, J.; Riedl, J. Item-based collaborative filtering recommendation algorithms. In Proceedings of the Proceedings of the 10th international conference on World Wide Web, 2001, pp. 285–295.
2. Koren, Y.; Bell, R.; Volinsky, C. Matrix factorization techniques for recommender systems. *Computer* **2009**, *42*, 30–37.
3. Jin, T. Integrated Machine Learning for Enhanced Supply Chain Risk Prediction **2025**.
4. Shen, G. Computation Offloading for Better Real-Time Technical Market Analysis on Mobile Devices. In Proceedings of the Proceedings of the 2021 3rd International Conference on Image Processing and Machine Vision, 2021, pp. 72–76.
5. Zhou, G.; Zhu, X.; Song, C.; Fan, Y.; Zhu, H.; Ma, X.; Yan, Y.; Jin, J.; Li, H.; Gai, K. Deep interest network for click-through rate prediction. In Proceedings of the Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, 2018, pp. 1059–1068.
6. Kang, W.C.; McAuley, J. Self-attentive sequential recommendation. In Proceedings of the 2018 IEEE international conference on data mining (ICDM). IEEE, 2018, pp. 197–206.
7. Sun, F.; Liu, J.; Wu, J.; Pei, C.; Lin, X.; Ou, W.; Jiang, P. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In Proceedings of the Proceedings of the 28th ACM international conference on information and knowledge management, 2019, pp. 1441–1450.
8. Zheng, Y.; Wei, P.; Chen, Z.; Cao, Y.; Lin, L. Graph-convolved factorization machines for personalized recommendation. *IEEE Transactions on Knowledge and Data Engineering* **2021**, *35*, 1567–1580.
9. Lu, J. Optimizing e-commerce with multi-objective recommendations using ensemble learning. In Proceedings of the 2024 4th International Conference on Computer Systems (ICCS). IEEE, 2024, pp. 167–171.
10. Li, S. Harnessing multimodal data and mult-recall strategies for enhanced product recommendation in e-commerce. In Proceedings of the 2024 4th International Conference on Computer Systems (ICCS). IEEE, 2024, pp. 181–185.