Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

Friends Recommendation on Social Networks using the Bayesian Personalized Ranking-Matrix Factorization

Muhammad Haidir Ali, Z K A Baizal*

School of Computing, Informatics, Telkom University, Bandung, Indoensia
Email: [1]mhaidirali@student.telkomuniversity.ac.id, [2,*]baizal@telkomuniversity.ac.id
Correspondence Author Email: baizal@telkomuniversity.ac.id
Submitted: 10/01/2024; Accepted: 18/02/2024; Published: 20/02/2024

Abstract−In the digital landscape of social networking, the challenge of improving friend recommendation systems is pivotal for enhancing user interaction and fostering social connections. Addressing this challenge, the current study innovates by fusing Bayesian Personalized Ranking (BPR) with Matrix Factorization (MF), culminating in a novel BPR-MF model designed for the intricacies of social network relationships. The study harnesses a rich dataset from LastFM, comprising 27,806 interactions among 7,624 users, to analyze mutual follower patterns and augment the precision of friend recommendations. Through rigorous preprocessing and systematic evaluation of the BPR-MF model against different numbers of latent factors, the research uncovers that a configuration of 20 latent factors is most effective, achieving an RMSE of 0.156 and an AUC ROC of 0.800. This discovery addresses the critical problem of balancing computational complexity with prediction accuracy in recommendation models. It also demonstrates the necessity for a nuanced, data-driven approach to generate relevant social connections. The research sets a new direction for future studies aiming to capitalize on user interaction data to offer precise friend suggestions, all while upholding user privacy and avoiding reliance on personal data.

Keywords: Recommendation System; Social Network; Bayesian Personalized Ranking; Matrix Factorization; Alternating Least Squares

1. INTRODUCTION

In the current digital era, social networks have become an integral part of daily life, serving as the primary platform for communication and social interaction. These networks provide not only a means of communication but also a space for forming and expanding social relationships. In the context of social networking, effective friend recommendations play a vital role in enriching users social experiences by providing relevant and valuable friendship suggestions with shared interests or connections[1]. Therefore, the development of more accurate friend recommendation systems has become an increasingly important and interesting research topic. According to Mark S. Granovetter[2], social interaction is not limited to close relationships but also involves weaker ties, such as acquaintances, which can open access to new networks and information. This theory underscores the importance of recognizing and understanding various types of relationships in social networks to formulate effective friend recommendations. In an era where the number of social media users continues to grow, opportunities for interaction and forming new connections are expanding.

Amid the growth of social networks, the emerging challenge is how recommendation systems can effectively suggest relevant and meaningful friendships, especially when relying on limited data from existing user interactions. This limitation raises questions about how to maximize the understanding and utilization of existing interaction patterns to generate accurate recommendations. Given the importance of social interaction in forming relationships[3], this research focuses on understanding and analyzing relationships to develop more intuitive and useful recommendations. In this context, the importance of technology capable of interpreting the complexity of social relationships and interaction patterns without relying on personal user data becomes clear for greater flexibility. Therefore, this study proposes an approach that intelligently processes and analyzes available interaction data to provide more accurate and relevant friend recommendations.

Previous research has explored various aspects of recommendation systems in social networks, providing valuable insights for this study. One such study by Lamia Berkani[4]proposed a semantic and collaborative filtering (CF) approach that takes into account users' implicit profiles based on their interactions in the Yelp social network, such as similarities between active users and their friends. This study demonstrated an increase in recommendation accuracy by incorporating semantic aspects. Although this approach is effective, it has drawbacks, particularly in the complexity of integrating various types of data and metrics and the efficiency of processing dynamic data. Another study by Xu Chonghuan[5]proposed a recommendation method in social networks using matrix factorization (MF) and the application of K-Harmonic Means (KHM) and particle swarm optimization (PSO) techniques for clustering users in modeling using the MovieLens and Book-Crossing datasets. This study showed improved performance compared to traditional methods through comprehensive evaluation metrics but faced challenges in handling larger volumes and sparse data. Additionally, research by Jing Chen et al. [6]developed the implementation of Alternating Least Squares (ALS) in the context of recommendation systems, as it offers ease of integration, parallelization, and scalability on various platforms, especially in memory efficiency. This research achieved a significant improvement in performance speed, being 5.5 times faster in solving factorization problems.

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

In research [7], Yeon-Chang Lee et al. introduced a new approach to enhance Bayesian Personalized Ranking (BPR) in recommendation systems with Multi-Type Pair-Wise (M-BPR) to address the issue of invalid user-item assumptions in standard BPR. M-BPR successfully reduced uncertainty in user preferences, showing an 8.89% improvement in recommendation accuracy on the Normalized Discounted Cumulative Gain (NDCG) compared to the One Class Collaborative Filtering (OCCF) method. However, managing preferences with M-BPR adds complexity to the model, requiring more intensive processing of user interactions. Further, research [8] by Daizong Ding et al. introduced the concept of BayDNN, a fusion of BPR with deep neural networks (DNN) and the use of convolutional neural networks (CNN). This approach is unique as it extracts deeper features to capture personal user preferences using specialized pre-training. Results on the Epinions and Slashdot datasets showed a 5% improvement in NDCG. However, this approach faces specific challenges as BayDNN entails higher model complexity in analyzing different types of social data and characteristics, requiring substantially more computing resources.

While various studies have been explored, a significant gap remains in the effectiveness of friend recommendation systems in addressing the complexities of social relationships and the dynamics of user interactions. Previous studies have marked substantial advancements in recommendation accuracy through semantic, collaborative approaches, and optimization techniques. Nonetheless, challenges persist, particularly in integrating diverse types of data and metrics, overcoming limitations of large and sparse data, and processing dynamic data efficiently. Moreover, although approaches like BayDNN offer enhancements in capturing user preferences, the complexity of the model and the substantial need for computational resources introduce new limitations. This research endeavors to bridge this gap by proposing a more intuitive and efficient approach in user preference modeling in friend recommendation systems. By employing the Bayesian Personalized Ranking and Matrix Factorization method (BPR-MF) [9], this study aims to address the shortcomings of previous approaches and provide a solution that not only enhances recommendation accuracy but also manages the complexity of social data more effectively [10]. Distinctively, this research concentrates on optimizing latent factors to reflect user preferences more accurately, relying on patterns of social interaction rather than personal user data [11]. The integration of BPR with MF, particularly through Alternating Least Squares (ALS) techniques, is anticipated to offer improvements in efficiency, tackle the challenges of sparse data, and expedite data processing [12]. It represents a stride beyond conventional methodologies, which frequently sideline the nuanced tapestry of social relationships, thereby marking a novel chapter in the landscape of friend recommendation systems.

2. RESEARCH METHODOLOGY

2.1 Research Stages

This research was conducted based on the stages shown in the flowchart in Figure 1.

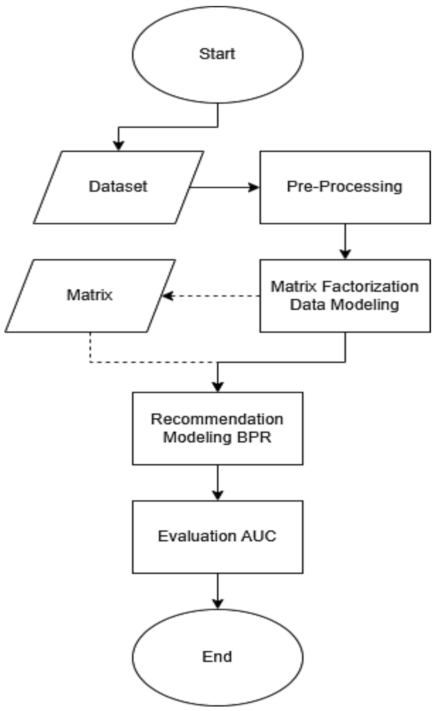

Figure 1. Research Flow

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

Figure 1 shows designing a recommendation system BPR-MF, commencing with the initial Start phase, the process transitions into the Dataset stage, where a collection of user-item interaction data is compiled, forming the bedrock of subsequent analysis. The Pre-Processing phase follows, wherein the dataset is rigorously cleansed to remove any anomalies or irrelevant information. This stage is critical for enhancing the data quality, which is instrumental in ensuring that the subsequent modeling is relevant and reliable. In the Matrix stage, the study meticulously implements Matrix Factorization (MF) techniques, with the Alternating Least Squares (ALS) method being particularly central to this phase. This technique is instrumental in extracting latent structures embedded within the user-item interaction data, revealing intricate patterns that may not be overtly discernible. After the matrix factorization data modeling, the model's predictive accuracy is rigorously evaluated using the Root Mean Square Error (RMSE). This statistical measure is employed to quantify the differences between the values predicted by the model and the actual values, thus providing a clear indicator of the model's precision and guiding further refinements to enhance the recommendation system's performance [13].

Upon the completion of the data modeling, the Recommendation Modeling BPR phase is initiated. Here, the Bayesian Personalized Ranking (BPR) technique is integrated into the model. BPR leverages individual user preferences and interaction patterns, facilitating the generation of personalized and precise recommendations. The terminal stage of the research, Evaluation AUC, involves a quantitative assessment of the model's performance. The Area Under the Curve (AUC) from the Receiver Operating Characteristic (ROC) curve is computed, offering a robust metric for evaluating the model's discriminative capacity to accurately distinguish between positive and negative user-item interactions. The culmination of these methodical stages results in the End Phase, This process was undertaken to affirm the reliability and validity of the resultant recommendation model..

2.2 Dataset

The dataset used in this study utilizes the dataset from the UCI Machine Learning Repository [14], which consists of a total of 27,806 interactions between 7,624 users. This dataset is specifically employed to analyze mutual follower patterns among users of the LastFM music streaming service, with the primary aim of developing a recommendation system that exploits the social network structure to provide an accurate depiction of the outcomes achievable by the BPR-MF model. The study focuses on inter-user relationships and the identification of recommendations based on mutual followers, with the data format presented in Table 1.

Table 1. Dataset Example

Nodes (User)	Edges (User Interaction)
0	747
1	4257
1	2194
1	580
1	6478

Table 1 presents the used dataset segments, highlighting nodes as users and edges as interactions between them in the LastFM. In Table 1, there are node 0 users who do not have mutual connections, while node 1 users have mutual followers. Edge provides a numerical measure of engagement, indicating the level of user engagement and their connectivity within this social network. This part of the data set is critical for examining the intricacies of social networking patterns and serves as a foundation for refining the BPR-MF recommendation model.

2.3 Preprocessing

Preprocessing is crucial in preparing data before modeling to structure the data more effectively and ensure that the developed recommendation model is based on valid, representative, and analysis-ready data. The preprocessing stages include:

a. Data Cleansing
   Enhancing the quality of the dataset by removing isolated nodes and edges that do not contribute to interactions, thereby improving the relevance and effectiveness of the recommendation model.
b. Data Transformation
   Transforming data into an interaction matrix is a preliminary step in modeling interactions from relationships between subjects and objects, providing a consistent data structure for machine learning techniques.
c. Data Splitting
   Dividing the data into training and testing sets ensures that the model can be trained and validated effectively, testing its predictive capacity on different sample data to support robust performance evaluation.
d. Checking Sparsity
   Checking for sparsity provides insights into the interaction density within the dataset, offering crucial information for adjusting the recommendation model to be more efficient in handling sparse data.

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

Table 2 shows the results of the preprocessing process before and after preprocessing.

Table 2. Preprocessing Result Example

	Process			
Metrics	Before Cleansing	After Cleansing	Transformation	Sparsity
Nodes	7624	5870	N/A	N/A
Edges	27806	26052	N/A	N/A
Training	N/A	N/A	4772x4766	99.919%
Testing	N/A	N/A	4772x4766	99.965%

Based on Table 2, there are identical dimensions of the training and testing matrices, namely 4772 x 4766, which are crucial elements in the development of a recommendation system using BPR-MF. This uniformity in dimensions is essential for matrix factorization to be effective in decomposing the interaction matrix into latent factors, reflecting the hidden attributes of subjects and objects [15]. Meanwhile, BPR, focusing on optimizing the personal ranking of items, requires similar dimensions between the training and testing sets for consistent learning and prediction validation. Considering the high sparsity value, which indicates that most potential interactions in the network have not yet occurred, the recommendation system faces a common challenge in learning this sparse data structure to generate accurate predictions [16].

2.4 Matrix Factorization Data Modeling

Matrix factorization with the Alternating Least Squares (ALS) approach is applied to construct the recommendation system in data modeling. Matrix factorization is effective in identifying the latent structure of user-item interaction data. In this study, an item can be interpreted as a mutual follower occurring among users, allowing for the discovery of implicit patterns. ALS, as an optimization method, alternately updates user and item factors, efficiently handling sparse data and reducing computational complexity. This approach lies in its ability to combat overfitting through regularization, with the optimization process carried out through repeated iterations. Each iteration involves adjusting the user and item factor matrices to minimize prediction errors. This process involves converting interaction data into a matrix, followed by matrix factorization to decompose interactions into latent factors, followed by iterative adjustments of these factors to uncover hidden patterns in the data [17]. The formulas for matrix factorization and ALS are as follows:

$$R_{i,j} \approx \sum_{k=1}^{K} P_{i,k} Q_{k,j} \tag{1}$$

$$E = \sum_{i,j} (R_{i,j} - P_i Q_j^T)^2 + \lambda \left(\sum_i \|P_i\|^2 + \sum_j \|Q_j\|^2 \right) \tag{2}$$

Based on formula (1) the user-item interaction matrix R can be approximated by the product of two lower-dimensionality latent factor matrices, P and Q, where P represents user-specific latent factors, Q denotes item-specific latent factors, and K is the number of latent factors. Formula (2) defines the loss function that ALS attempts to minimize, which comprises the squared error between the actual interaction matrix R and the estimated interactions from the product of P and Q, regularized by the sum of the squared norms of the latent factors weighted by a regularization parameter λ to control overfitting. For evaluating the accuracy of the matrix factorization model, the Root Mean Square Error (RMSE) is frequently used and can be formulated as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i,j} \left(R_{i,j} - \hat{R}_{i,j} \right)^2} \tag{3}$$

$\hat{R}_{i,j}$: The predicted rating from the matrix factorization model for the interaction between user i and item j
$R_{i,j}$: The actual observed rating
N : The total number of observed user-item interactions

On the formula (3) The RMSE computes the square root of the average squared differences between the predicted and actual ratings, providing a measure of the model's prediction accuracy. Lower values of RMSE indicate better model performance, reflecting a closer fit to the observed data. The use of RMSE is particularly effective in highlighting the impact of larger errors due to the squaring of the error terms [18].

2.5 Recommendation Modeling Bayesian Personalized Ranking

In this modeling, the Bayesian Personalized Ranking (BPR) technique is adopted as an advanced method in recommendation modeling, building on the foundation established through the Matrix Factorization (MF) process with the Alternating Least Squares (ALS) approach. The implementation of BPR aims to enhance the effectiveness of the recommendation system by accommodating individual user preferences. These preferences are revealed through their interactions with various items in the network by the latent factors applied in MF modeling using ALS techniques. BPR specifically focuses on comparing liked (positive interactions) and disliked (negative interactions) items by users. This process involves identifying samples of positive and

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

negative interactions from the interaction matrix. Positive interactions are identified as items that have interacted with the user, while negative interactions are defined as items that have not interacted with them. The process involves differentiating prediction scores between positive and negative items. The BPR algorithm iteratively adjusts these factors based on user preferences. Hyperparameters such as learning rate and lambda are adjusted to achieve optimal model performance [19]. The formula for Bayesian personalized ranking is as follows:

$$BPR - Opt := \sum_{(u,i,j)\in D_s} ln\sigma\left(\hat{x}_{u,i,j}\right) - \lambda_\theta \|\theta\|^2 \quad (4)$$

Formula (4) represents the optimization objective, D_s is the set of user-item interaction triplets, where each triplet *(u,i,j)* signifies that user u has shown a preference for item i over item j. The function σ represents the sigmoid function applied to the predicted difference in preferences, $\hat{x}_{u,i,j}$, for a pair of items i and j for a user u. The sigmoid function is chosen because it maps the real-valued input into the (0, 1) range, representing the probability of user u preferring item i over item j. The term $\lambda_\theta\|\theta\|^2$ is a regularization term where λ_θ is the regularization parameter that controls the extent of regularization applied to the model parameters θ, which includes user and item latent factors. This regularization is essential to prevent overfitting by penalizing the magnitude of the parameters. Optimization of this objective function seeks to increase the likelihood of observed user preferences (positive interactions over negative ones) while simultaneously constraining the complexity of the model, as captured by the magnitude of the latent factors. The BPR model thus iteratively updates the user and item latent factors to maximize this objective function, effectively ranking items for each user in a personalized manner.

2.6 Evaluation AUC

Performance evaluation in recommendation systems is a critical element essential for determining the effectiveness of the model in generating relevant recommendations. This study adopts the Area Under the Curve (AUC) method from the Receiver Operating Characteristic (ROC) curve to measure the performance of the BPR-MF recommendation model. This evaluation function utilizes factors developed for users and items, as well as data from testing interactions, to generate a score reflecting the extent to which the model can accurately predict user preferences. For each test interaction case, the model predicts scores for positive and negative interactions based on the factorization results performed by BPR-MF. The ROC curve is then plotted using false positive rate (FPR) and true positive rate (TPR) values [20]. The AUC metric provides an effective quantitative measure to assess how well the model can differentiate between relevant and irrelevant recommendations, and a high AUC score indicates that the model is effective in performing its task. The formula for AUC is as follows:

$$TPR = \frac{TP}{TP+FN} \quad (5)$$

$$FPR = \frac{FP}{FP+TN} \quad (6)$$

$$AUC = \sum_{i=1}^{n-1} \frac{(FPR_{i+1} - FPR_i)x(TPR_{+i} + TPR_i)}{2} \quad (7)$$

Formula (5) represents the true positive rate (TPR), also known as recall or sensitivity. It is the ratio of the number of true positive predictions (TP) to the sum of the number of true positive predictions and the number of false negatives (FN), where a true positive is an interaction correctly predicted as positive and a false negative is a positive interaction that was incorrectly predicted as negative. Formula (6) denotes the false positive rate (FPR), which is the ratio of the number of false positive predictions (FP) and the number of true negatives (TN). A false positive is a negative interaction incorrectly predicted as positive, and a true negative is an interaction correctly predicted as negative. Formula (7) provides the method for calculating the AUC. This is achieved by summing the area of trapezoids under the ROC curve, which is plotted with FPR on the x-axis and TPR on the y-axis across various threshold settings. The index i runs over all sorted observed values of the decision function, which are thresholds at which the TPR and FPR are recalculated. The AUC value is a measure of the model's ability to discriminate between positive and negative classes. A value of 1 indicates perfect discrimination, while a value of 0.5 suggests no discrimination, equivalent to random guessing.

3. RESULT AND DISCUSSION

To gain a deep understanding of the performance of the recommendation model that combines Bayesian Personalized Ranking (BPR) with Matrix Factorization (MF), this study conducts a series of structured experiments. The main dataset is divided into two main segments, namely training data and testing data, with a 70:30 distribution ratio. This division is intended to ensure that the model can learn from most of the data (70%) and be tested on previously unseen data (30%) to validate its prediction accuracy. The learning process using the MF model is carried out using the training dataset, while the model evaluation is conducted with the testing dataset. The main goal is to measure the effectiveness of the model in predicting unknown user interactions, which is an important indicator of the reliability of the recommendation system. To achieve optimal results from

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

the BPR-MF model, a series of experiments are conducted by varying the number of latent factors between 20 and 100. Each latent factor configuration is tested five times to identify the configuration that provides the best performance. The MF model, enhanced with the Alternating Least Squares (ALS) approach, is assessed based on the Root Mean Square Error (RMSE) metric to analyze its performance before being implemented in the BPR process. RMSE provides insights into how accurately the model predicts the actual values in the testing dataset. Meanwhile, AUC ROC is used to measure the BPR-MF model's ability to differentiate between positive and negative interactions. Comparing the performance of the model with various latent factor configurations using these metrics provides a basis for analysis to determine the most effective approach to achieving the best accuracy.

3.1 Result

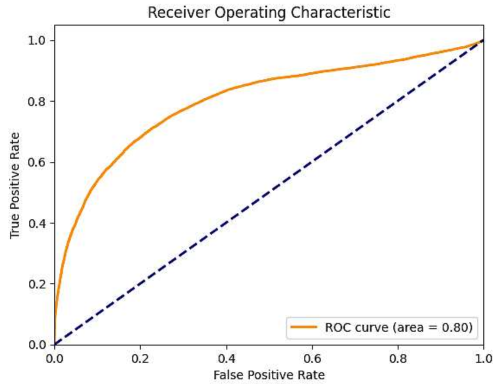

Figure 2. AUC ROC Accuracy with 20 Latent Factor (K)

Figure 2 illustrates the first research scenario in the application of the BPR-MF model with 20 factors used. The results show an RMSE of 0.156 and an AUC ROC accuracy of 0.800.

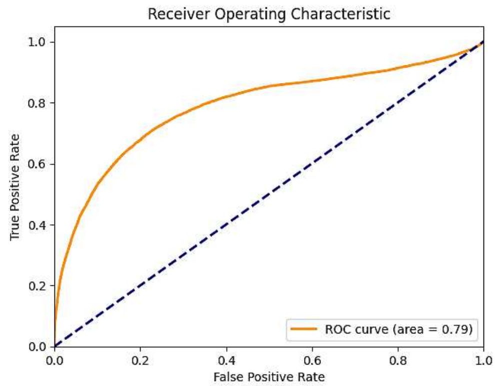

Figure 3. AUC ROC Accuracy with 40 Latent Factor (K)

Figure 3 depicts the second research scenario in the implementation of the BPR-MF model, with an increased number of factors to 40. The findings indicate a decrease in RMSE to 0.144 but a slight decrease in AUC ROC accuracy by 1% to 0.787.

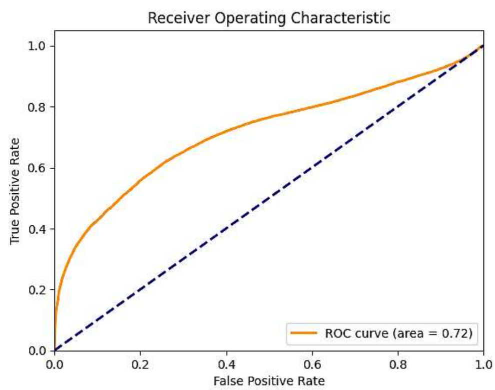

Figure 4. AUC ROC Accuracy with 60 Latent Factor (K)

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

Figure 4 presents the third research scenario with an increased number of factors to 60. The results obtained show an RMSE of 0.137, but there is a significant decrease in accuracy, about 8%, to 0.716, which is far from the highest accuracy achieved in the first scenario.

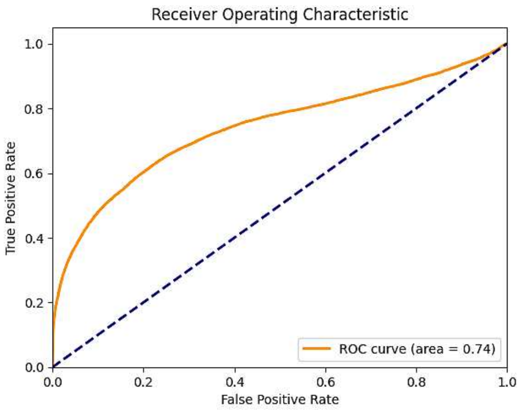

Figure 5. AUC ROC Accuracy with 80 Latent Factor (K)

Figure 5 displays the fourth research scenario, with the number of factors increasing to 80. The results yield an RMSE of 0.141, with a slight increase from the third scenario by 2% to 0.740.

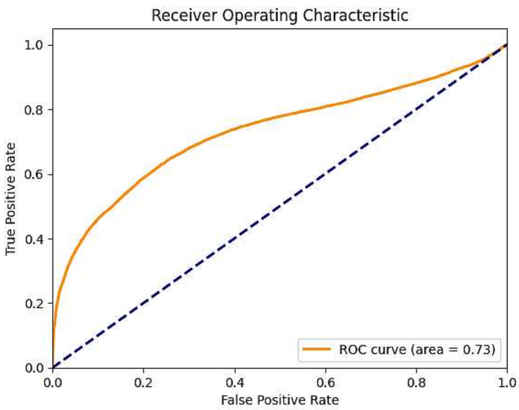

Figure 6. AUC ROC Accuracy with 100 Latent Factor (K)

Figure 6 shows the final research scenario, with the number of factors at 100. The outcomes indicate an RMSE of 0.134 and an accuracy level of 0.730.

3.2 Discussion

To see more comparison the best scenario test results from the five tested, it can be seen ini the following Table 3:

Table 3. Comparative Analysis of RMSE and AUC ROC Scores Across Different Test Scenarios

Iteration	Latent Factor (K)	RMSE	AUC ROC
1	20	0.156	0.800
2	40	0.144	0.787
3	60	0.137	0.716
4	80	0.141	0.740
5	100	0.134	0.730

Based on Table 3 the results of the experiments conducted, the optimal scenario was achieved through five iterations, focusing on the use of varying numbers of latent factors (K) in each scenario. The research demonstrated that the first scenario, which utilized 20 latent factors in the Bayesian Personalized Ranking-Matrix Factorization (BPR-MF) model for social network friend recommendations, was the most effective. This scenario, leveraging the user interaction dataset from LastFM, successfully generated accurate recommendations by maximizing the model's understanding of interaction patterns. The results obtained in this scenario indicated a root mean square error (RMSE) of 0.156 and an area under the curve (AUC) score from the receiver operating characteristic (ROC) curve of 80%. The RMSE value of 0.156, while being relatively low, signifies a high level of precision in the model's predictions, indicating that the matrix factorization process is effectively minimizing the prediction errors. This level of accuracy in the RMSE demonstrates the model's robustness in capturing and

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

replicating the complex patterns of user interactions within the social network. It's noteworthy that the lowest RMSE value, which was 0.137, was obtained in the third research scenario. However, this scenario also exhibited the lowest accuracy. This juxtaposition highlights the nuanced balance between minimizing RMSE and maintaining high accuracy, underscoring the complexity of optimizing recommendation models for social networking environments. After obtaining the optimal configuration, Table 4 presents the final recommendation outcomes from the BPR-MF model, delineating personalized item suggestions for individual users.

Table 4. Personalized Recommendations from the BPR-MF Model

User ID	Recommendation Item IDs
1	650, 6105, 6891, 6789, 6792
2	3530, 6360, 5127, 3450, 6955
3	3530, 3450, 5127, 6105, 7488

Based on Table 4, user 1 with a preference profile determined from LastFM interaction data, receives a unique set of item recommendations, exemplified by IDs such as 650 and 6789. This selection reflects a nuanced understanding of the user's preferences by the BPR-MF model. User 2 and User 3 also receive curated lists, with an interesting overlap for item IDs 3530 and 5127, indicating these items' universal relevance or popularity. This mapping underscores the BPR-MF model's practical application in generating tailored recommendations within a social network context, affirming its ability to discern and adapt to complex user interaction patterns. The findings also illustrate the model's proficiency in personalizing content, a critical aim in recommendation system design. The choice to implement 20 latent factors proved instrumental in enhancing the model's interpretive capability, striking a balance between predictive accuracy and computational efficiency. The favorable outcomes from this configuration offer insights into optimizing model parameters for meaningful social networking recommendations.

4. CONCLUSION

Based on results and discussion, the research undertaken provides a comprehensive analysis of the effectiveness of integrating Bayesian Personalized Ranking (BPR) with Matrix Factorization (MF) techniques in refining friend recommendation systems for social networks, with a particular focus on data derived from LastFM. The optimized selection of 20 latent factors emerged as a pivotal element in this study, yielding the most optimal Area Under the ROC Curve (AUC) score of 0.800 and a Root Mean Square Error (RMSE) of 0.156, indicative of a high degree of accuracy in the model's predictions. The intricate equilibrium between the number of latent factors and the model's accuracy underscores the nuanced intricacies involved in developing sophisticated recommendation systems. The mapping of specific item recommendations to individual user IDs further exemplifies the model's capability to generate personalized suggestions, illustrating its practical application and the tangible benefits it holds for enhancing the user experience on social platforms. These insights not only contribute substantively to the field of recommendation systems but also highlight the importance of meticulous parameter tuning to achieve a balance between precision and generalizability. The advancement evidenced by the BPR-MF model offers a significant step forward from traditional recommendation methods, moving towards more nuanced, privacy-conscious, and user-centric approaches that are well-suited to the dynamic nature of social networking sites. The implications of this research are twofold: it enriches the academic dialogue on predictive analytics and serves as an empirical foundation for future innovations aimed at cultivating more contextually relevant and engaging social networking environments.

REFERENCES

[1] L. A. Adamic and E. Adar, "Friends and neighbors on the Web," *Soc Networks*, vol. 25, no. 3, pp. 211–230, Jul. 2003, doi: 10.1016/S0378-8733(03)00009-1.
[2] M. S. Granovetter, "The Strength of Weak Ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, May 1973, doi: 10.1086/225469.
[3] M. Ameri, E. Honka, and Y. Xie, "From Strangers to Friends: Tie Formations and Online Activities in an Evolving Social Network," *Journal of Marketing Research*, vol. 60, no. 2, pp. 329–354, Apr. 2023, doi: 10.1177/00222437221107900.
[4] L. Berkani, "A semantic and social-based collaborative recommendation of friends in social networks," *Softw Pract Exp*, vol. 50, no. 8, pp. 1498–1519, Aug. 2020, doi: 10.1002/spe.2828.
[5] C. Xu, "A novel recommendation method based on social network using matrix factorization technique," *Inf Process Manag*, vol. 54, no. 3, pp. 463–474, May 2018, doi: 10.1016/j.ipm.2018.02.005.
[6] J. Chen, J. Fang, W. Liu, T. Tang, X. Chen, and C. Yang, "Efficient and Portable ALS Matrix Factorization for Recommender Systems," in *2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, IEEE, May 2017, pp. 409–418. doi: 10.1109/IPDPSW.2017.91.

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 5, No. 2, February 2024, Page 297-305
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v5i2.4804

[7] Y.-C. Lee, T. Kim, J. Choi, X. He, and S.-W. Kim, "M-BPR: A novel approach to improving BPR for recommendation with multi-type pair-wise preferences," *Inf Sci (N Y)*, vol. 547, pp. 255–270, Feb. 2021, doi: 10.1016/j.ins.2020.08.027.

[8] D. Ding, M. Zhang, S.-Y. Li, J. Tang, X. Chen, and Z.-H. Zhou, "BayDNN," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, New York, NY, USA: ACM, Nov. 2017, pp. 1479–1488. doi: 10.1145/3132847.3132941.

[9] Akhilesh Narapareddy, "Recommender System Using Bayesian Personalized Ranking," towardsdatascience.com.

[10] A. Ameri, A. Bose, and M. Soltanalian, "Comprehensive Personalized Ranking Using One-Bit Comparison Data," in *2019 IEEE Data Science Workshop (DSW)*, IEEE, Jun. 2019, pp. 338–342. doi: 10.1109/DSW.2019.8755595.

[11] Shah and Shalin, "A Survey of Latent Factor Models for Recommender Systems and Personalization," *Authorea Preprints*, 2023.

[12] X. He, Z. He, X. Du, and T.-S. Chua, "Adversarial Personalized Ranking for Recommendation," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, New York, NY, USA: ACM, Jun. 2018, pp. 355–364. doi: 10.1145/3209978.3209981.

[13] N. Taneja and H. K. Thakur, "Evaluating the Scalability of Matrix Factorization and Neighborhood Based Recommender Systems," *International Journal of Information Technology and Computer Science*, vol. 15, no. 1, pp. 21–29, Feb. 2023, doi: 10.5815/ijitcs.2023.01.03.

[14] UCI Machine Learning Repository, "LastFM Asia Social Network." DOI, 2020. doi: https://doi.org/10.24432/C5S904.

[15] H. Zhang, I. Ganchev, N. S. Nikolov, Z. Ji, and M. O'Droma, "Weighted matrix factorization with Bayesian personalized ranking," in *2017 Computing Conference*, IEEE, Jul. 2017, pp. 307–311. doi: 10.1109/SAI.2017.8252119.

[16] Dilip Yadav, Tushar Vaghasiya, Savita Ravate, and Prof. Vinit Raut, "Sparsity and Matrix Factorization in Recommender System," *IJARCCE*, vol. 10, no. 4, Apr. 2021.

[17] M. Nitti, L. Atzori, and I. P. Cvijikj, "Friendship Selection in the Social Internet of Things: Challenges and Possible Strategies," *IEEE Internet Things J*, vol. 2, no. 3, pp. 240–247, Jun. 2015, doi: 10.1109/JIOT.2014.2384734.

[18] A. Al-Nafjan, N. Alrashoudi, and H. Alrasheed, "Recommendation System Algorithms on Location-Based Social Networks: Comparative Study," *Information*, vol. 13, no. 4, p. 188, Apr. 2022, doi: 10.3390/info13040188.

[19] B. G. Galuzzi, I. Giordani, A. Candelieri, R. Perego, and F. Archetti, "Hyperparameter optimization for recommender systems through Bayesian optimization," *Computational Management Science*, vol. 17, no. 4, pp. 495–515, Dec. 2020, doi: 10.1007/s10287-020-00376-3.

[20] H. Ko, S. Lee, Y. Park, and A. Choi, "A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields," *Electronics (Basel)*, vol. 11, no. 1, p. 141, Jan. 2022, doi: 10.3390/electronics11010141.