

Research Article

Research on PMF Model Based on BP Neural Network Ensemble Learning Bagging and Fuzzy Clustering

Zhengjin Zhang ¹, Guilin Huang,¹ Yong Zhang,¹ Siwei Wei,¹ Baojin Shi,² Jiabao Jiang,¹ and Baohua Liang ³

¹Chaohu University, Hefei 238000, Anhui, China

²Bengbu University, Bengbu 233030, Anhui, China

³School of Information Engineering, Chaohu University, Hefei, China

Correspondence should be addressed to Baohua Liang; liangbh426@126.com

Received 21 May 2021; Accepted 2 July 2021; Published 21 July 2021

Academic Editor: Muhammad Javaid

Copyright © 2021 Zhengjin Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Probability matrix factorization model can be used to solve the problem of high-dimensional sparsity of user and rating data in the recommender systems. However, most of the existing methods use the user to model the item rating, ignoring the relationship between the user and the item, so the accuracy of user-item rating prediction is still low. Therefore, this paper proposes a probabilistic matrix factorization model based on BP neural network ensemble learning, bagging, and fuzzy clustering. Firstly, the membership function of fuzzy clustering and the selection of cluster center are used to calculate the user-item rating matrix; secondly, BP neural network trains the user-item scoring matrix after clustering, further improving the accuracy of scoring prediction; finally, the bagging method in ensemble learning is introduced, which takes the number of user-item scores as the base learner, trains the base learner through BP neural network, and finally obtains the score prediction through the voting results, which improves the stability of the model. Compared with the existing PMF models, the root mean square error of the PMF model after fuzzy clustering is increased by 9.27% and 3.95%, and the average absolute error is increased by 21.14% and 1.11%, respectively; then, the performance of the first mock exam is introduced. The root mean square error of the ensemble method is increased by 4.02% and 0.42%, respectively, compared with the existing single model. Finally, the weights of BP neural network training based learner are introduced to improve the accuracy of the model, which also verifies the universality of the model.

1. Introduction

In recent years, matrix factorization technology, with good scalability and high recommendation accuracy, has developed rapidly [1]. After the famous Netflix recommendation contest, matrix factorization has received more attention. The basic idea of matrix factorization technology is to assume that users' preferences and project characteristics can be described by potential factors and find the minimum sum of squares of the distance between the original scoring matrixes. The representative ones are probability matrix factorization, Bayesian probability matrix factorization, and fast parallel matrix factorization.

Koren put forward a new SVD++ model by combining the matrix factorization model with the domain-based recommendation method [2]. Salakbuttinov and Mnih analyzed the principle of matrix factorization from the angle of probability and put forward the probabilistic matrix factorization (PMF) model [3], which extended matrix factorization to any maximum likelihood solution. Later, the Bayesian probabilistic matrix factorization (BPMF) was put forward [4].

The idea of ensemble learning was also adopted to improve the accuracy of the recommendation system. Fang et al. [5] integrated the recommendation methods based on user similarity, used different similarity measures to generate different recommendation models, and weighted sum to get

the final prediction score, which improved the prediction accuracy of the model. Cui et al. [6] constructed a new dataset by combining user-based and product-based prediction score difference with real scores and then trained and predicted with the XG-boost model. All the above integration methods are based on a content-based recommendation algorithm, which has the defects of high time complexity and relatively low prediction accuracy. When applied to high-dimensional sparse data, users or commodities with 0 similarities may appear, which can reduce the prediction accuracy of the algorithm.

Based on the above analysis, we can conclude that the probability decomposition matrix has inherent defects in the face of high-dimensional sparsity. In this paper, a probability matrix factorization model by fusing the ensemble learning bagging method based on BP neural network and fuzzy clustering is proposed. The main work is as follows:

- (1) The scoring matrix of users and items is calculated by using the membership function of fuzzy matrix and the selection of cluster center, which is more accurate than the traditional probability matrix method and can better construct the scoring matrix of users and items.
- (2) The bagging method in ensemble learning is proposed to generate different training sets by self-sampling, and ensemble learning is introduced into this model, thus increasing the parallelism and improving the accuracy and stability of scoring prediction.

2. System Model

In this section, we review the literature related to our work and discuss their differences with our contributions.

2.1. Probabilistic Matrix Factorization (PMF). Salakhutdinov et al. proposed the PMF, which is a well-known approach for recommendation systems. Table 1 summarizes the notations of PMF, and Figures 1 and 2 show the overview of the graphical model of PMF. We suppose that M users, N items, a rating matrix $R \in R^{k \times N}$, and item latent matrix $V \in R^{k \times N}$ to reconstruct the rating matrix R . The goal of the PMF is to determine the optimal matrix U , V and minimize the loss function ε as follows:

$$\min \varepsilon(U, V) = \sum_i^N \sum_j^N \frac{I_{ij}}{2} (r_{ij} - u_i^T v_j)^2 + \frac{\lambda_U}{2} \sum_i^N u_i^2 + \frac{\lambda_V}{2} \sum_i^M u_i^2. \quad (1)$$

After the objective function is determined, the stochastic gradient descent method is used to update u and v iteratively to minimize the objective function:

$$\begin{aligned} u_i &\leftarrow u_i - \alpha \cdot ((r_{ij} - u_i^T v_j) v_j + \lambda u_i), \\ v_j &\leftarrow v_j - \alpha \cdot ((r_{ij} - u_i^T v_j) u_i + \lambda v_j), \end{aligned} \quad (2)$$

TABLE 1: Notation.

Notation	Description
R	Rating matrix
N	Number of users
u_i	Latent factors of user i
M	Number of MAE
v_j	Latent factors of item j
r_{ij}	Rating of item j given by user i
\hat{r}_{ij}	Predicted rating of item j given by user i
U	User latent factor
V	Item latent factor
k	Size of latent factor
I	Indicator, $I_{ij} = 1$ if $r_{ij} \neq 0$; otherwise $I_{ij} = 0$
$\delta^2, \delta_U^2, \delta_V^2$	Variance

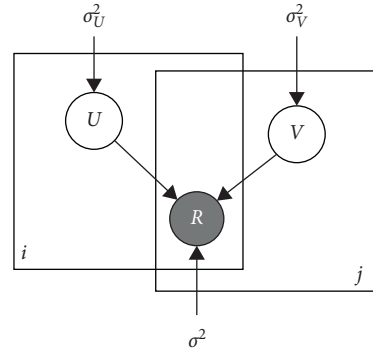


FIGURE 1: Graphical model of PMF [2].

where α is the learning rate. When a certain number of iterations or the change of the objective function is less than a certain threshold, the iteration stops. Finally, the U , V characteristic matrix is trained to predict the score.

2.2. Fuzzy C-Means. Fuzzy c-means is an unsupervised clustering algorithm in which each point has a certain strength of association between the nodes and the particular community [7].

The FCM minimizes an objective function J_f :

$$J_f(\vec{U}, \vec{C}) = \sum_{i=1}^n \sum_{j=1}^k u_{ij}^f x_i - c_j^2, \quad (3)$$

where u_{ij} is the membership degree of the i -th node to the j -th cluster and $d_{ij} = x_i - c_j$ is the distance between the i -th node and the center of the j -th cluster. During optimizing J_f , the constraint $\sum_{j=1}^k u_{ij} = 1$ must be satisfied. The parameters of controlling the fuzziness of the algorithm. As f turns out to be larger, the process is fuzzier. c_j can be calculated by the following equation:

$$c_j = \frac{\sum_{i=1}^n u_{ij}^f x_i}{\sum_{i=1}^n u_{ij}^f}. \quad (4)$$

u_{ij} can be calculated via the following equation:

$$u_{ij} = \frac{1}{\sum_{l=1}^k (d_{ij}/d_{lj})^{2/(f-1)}}. \quad (5)$$

J_f can be minimized by iterative optimization with the update of membership degree u_{ij} and the cluster center c_j .

2.3. Ensemble Learning. Ensemble learning is to use a series of basic learners to learn [8] and then integrate the learning results based on certain rules, to get a better learning method than a single learner. Usually, there are some differences between base learners, either different algorithms or the same algorithms (with different parameters or super parameters). Generally speaking, the greater the difference between base learners, the better the final learning results. Ensemble learning has great advantages in performance improvement, so it is widely used in theoretical research and practical application. The classical ensemble learning methods mainly include bagging and boosting. In this paper, the bagging method is used, so the principle of this method is introduced in detail.

Bagging (bootstrap aggregation) is a classic parallel ensemble learning algorithm. It is a bootstrap sampling ensemble learning algorithm. It can obtain a lower prediction error and improve the accuracy of the recommendation algorithm. The general idea of the algorithm is as follows: given a dataset D containing K samples, a sample is randomly taken out and put into the sampling set, and then the sample is put back into the original dataset, so that the next sampling of the sample may be selected. Because of the use of the sample put back, a sample may appear many times in the sample set, or it may not appear once. In theory, after K times of random sampling, the sampling set D' containing K samples can be obtained. It is worth noting that if the probability of each sampling in the initial training set is $1/K$, the probability that the sample will not be collected in K sampling is $(1 - 1/k)^k$, and the limit is

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{k}\right)^k \rightarrow \frac{1}{e} \approx 0.368. \quad (6)$$

From the above formula, the probability of being sampled is

$$1 - \lim_{n \rightarrow \infty} \left(1 - \frac{1}{k}\right)^k = 1 - \frac{1}{e} \approx 0.632. \quad (7)$$

In other words, the probability of each sample in the sample set obtained by autonomous sampling is 63.2%. Using the above method, G sampling sets D containing K samples can be sampling sets $\{D_1, \dots, D_G\}$, a base learner is trained based on each sampling set, and then the base learner is integrated to generate the model prediction. Figure 2 shows the structure of the bagging model.

2.4. BP Neural Network. BP neural network is composed of an input layer, hidden layer, and output layer, which can realize continuous nonlinear mapping [9]. BP neural network is a kind of multilayer feed-forward neural network,

which is characterized by signal forward propagation and error backpropagation. In the process of forward propagation, the signal is processed layer by layer from the input layer through the hidden layer and finally reaches the output layer. Figure 3 shows the topological structure of the BP neural network.

BP neural network is a supervised learning algorithm, which completes the mapping from input to output by minimizing the objective function. The main idea of the algorithm flow of bagging algorithm-integrated BP neural network is shown in Algorithm 1.

The basic processing framework of the BP neural network is shown in Figure 3, where $X = (X_1, X_2, \dots, X_n)$ is the set of n values that are input from the outside or other neurons output; $W = (W_1, W_2, \dots, W_n)$ is called the weight, representing the connections strength between the neuron and other neurons; $\sum WX$ is called the activation value that is equal to the total input of the artificial neuron; O refers to the output of the neuron; b refers to the threshold of this neuron, and if the weighted sum of the input signal is greater than b , the artificial neuron is activated. In this way, the output of the artificial neurons can be described as follows:

$$O = f(\sum WX - b). \quad (8)$$

In equation (8), $f(\cdot)$ is called the activation function. The activation function used in this paper is a nonlinear transformation function and bipolar sigmoid function ($\tanh(x)$ function). In the process of error backpropagation, the problem of derivation with respect to the activation function is involved, and the $\tanh(x)$ function solves the problem of derivative discontinuity and the output problem of zero-centered effectively, so it is used as the activation function of this paper. so it is used as the activation function of this paper. It is defined as follows:

$$f(x) = \frac{(1 - e^{-x})}{(1 + e^x)}. \quad (9)$$

The basic processing framework of the BP neural network is shown in Figure 4.

This paper uses a three-layer BP neural network with a single hidden layer structure to simulate the change of the outburst.

3. Probability Matrix Factorization Model with Fuzzy Clustering

To further improve the prediction accuracy of probability matrix factorization for high-dimensional and sparse matrices, this paper uses the FCM method to process the score matrix by fuzzy clustering. On the one hand, the FCM algorithm is suitable for solving the problem of high-dimensional [10] and sparse data and has strong scalability; on the other hand, it can solve the shortcomings of hard clustering, that is, it does not force a certain score to be classified into a certain category but expresses the degree of a certain category's score belonging to a certain category in the

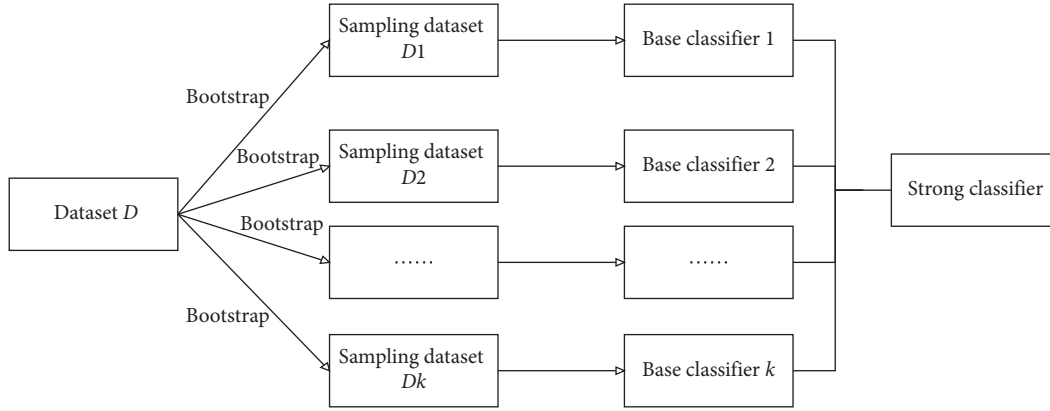


FIGURE 2: Bagging model.

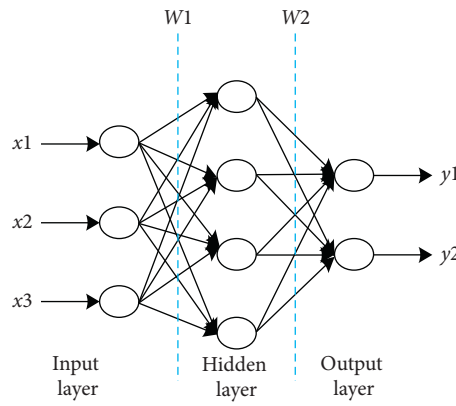


FIGURE 3: BP neural network mode.

Input: the normalized rating matrix from training dataset $D = \{(x_i, y_i)_{i=1}^m\}$

Output: the rating prediction result of this sample x of this test set

- (1) for $t = 1, \dots, k$ do (k is the number of base models),
 - 1.1: randomly select cluster center with FCM and calculate the fuzzy membership matrix F with membership function. F matrix represents an association between the clusters for ratings of users. Select $k-1$ samples from the training set.
 - 1.2: training the BP neural network with this sample to obtain the base model.
- (3) Averaging all the base models to get a strong learner.
- (4) The strong learner is used to test the dataset.

ALGORITHM 1: The algorithm flow of bagging algorithm integrated BP neural network.

form of membership function, to better divide scoring users without clear boundaries.

3.1. Algorithm Thought. FCM is introduced into the scoring matrix [11], where n users score m items. Every element in the matrix x_{ik} represents the score of user i on item k , and the row of the matrix $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$, where $i \in [1, n]$ represents the users; the column of the matrix $x_j = (x_{1j}, x_{2j}, \dots, x_{mj})$, where $j \in [1, M]$ denotes the project. Users are clustered according to the scores of each user, and the whole user is divided into the number of c clusters so that the similarity of user scores in the same cluster is the highest,

and the clustering results are expressed by the membership matrix U . The objective function J_f of fuzzy clustering based on user-item scoring matrix is as follows.

The FCM minimizes an objective function J_f :

$$J_f(\vec{U}, \vec{C}) = \sum_{i=1}^n \sum_{j=1}^k u_{ij}^f x_i - c_j^2. \quad (10)$$

The membership matrix needs to be generated by the fuzzy clustering Algorithm 2, and the fuzzy similarity matrix needs to be constructed by the data similarity in the matrix. The construction methods of fuzzy similarity matrix include maximum and minimum calculation method, cosine angle

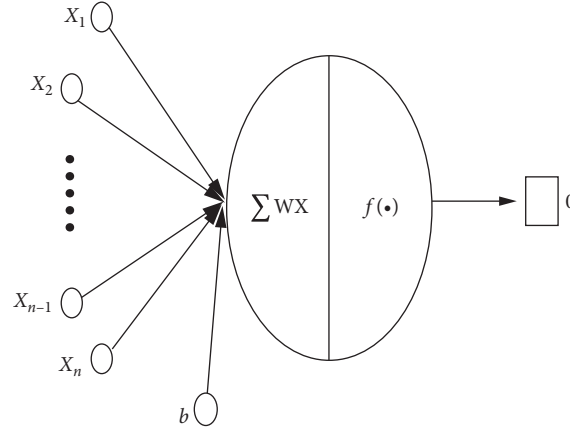


FIGURE 4: Artificial neuron model.

Input: constructing the rating matrix from training dataset $D = \{(x_i, y_i)_{i=1}^m\}$.

Output: the rating prediction result of this sample x of this test set.

- (1) Initialize related parameters.
- (2) Randomly select cluster center with FCM and calculate the fuzzy membership matrix F with membership function. F matrix represents an association between the clusters for ratings of users.
- (3) Apply PMF model on clustered data, initialize P and Q with gauss distribution.
- (4) Rating prediction.

ALGORITHM 2: The algorithm flow of rating prediction of PMF model based on FCM.

method, and correlation coefficient method. This paper mainly adopts the correlation coefficient method.

3.2. Algorithm Description. In Figure 5, we demonstrate the workflow of our work in which, first, the training dataset is engaged and then FCM is used to classify the users in the training dataset by applying the similarity of the user rating matrix. The useful rating predictions are delivered to the users who received the effects of FCM and PMF models.

4. Probability Matrix Factorization Model Ensemble Learning Bagging with BP Neural Network

The probabilistic matrix factorization model and similarity-based recommendation algorithms have greatly improved the efficiency and prediction accuracy. However, due to the characteristics of the data itself, the high-dimensional sparsity and the setting of random initial value lead to the instability of the model, resulting in the large variance of the prediction score, which affects the accuracy of the recommendation.

Considering that the accuracy of the single weak learner algorithm is not high, we choose the bagging ensemble learning method. At the same time, to further improve the generalization ability of the learner, we choose the probability matrix factorization model of bagging ensemble BP

neural network to effectively improve the accuracy of scoring prediction.

4.1. Algorithm Thought. Firstly, the FCM algorithm is used to initialize the sample dataset D , and the number of clusters is $\{D_1, D_2, \dots, D_G\}$. The difference is that, in order to ensure that each user and product has training samples in each sampling set, each sampling first randomly selects one of the scoring data participated by each user and product as a sample, with a total of $(m+n)$ samples ($(m+n \ll k)$), and then conducts self-help sampling on the overall training set to obtain a sampling set containing K samples. Then, for each sample set D_g , the BP neural network algorithm is used to train the optimal weights, and then the PMF model is used to predict the score.

For a regression task, let (x, y) be a piece of data in dataset D , where x is the eigenvector and Y is the true value. Multiple regression models are trained through the dataset, and then the features are put into the regression model to produce the corresponding predictive values $\Phi(x, D)$. The integrated prediction value is the average value predicted by multiple models on dataset D :

$$\Phi_A(x) = E_D \Phi(x, D), \quad (11)$$

where x is the fixed input value and y is the output value; then,

$$E_D (y - \Phi(x, D))^2 = y^2 - 2yE_D \Phi(x, D) + E_D \Phi^2(x, D). \quad (12)$$

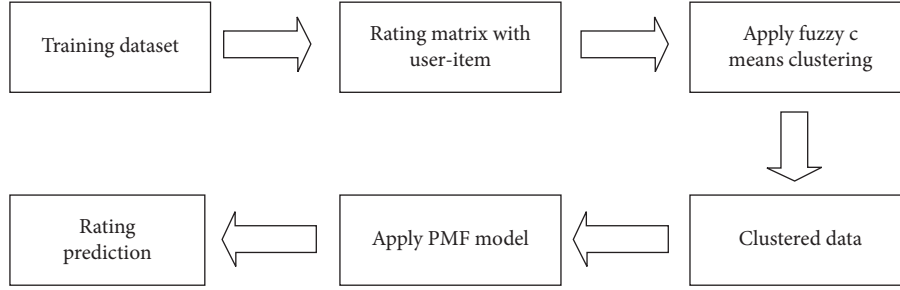


FIGURE 5: Rating prediction of the PMF model based on FCM.

Applying equation (10) and inequality $EZ^2 \geq (EZ)^2$ and then applying the change of equation (12), we can get

$$E_D(y - \Phi(x, D))^2 \geq (y - \Phi_A(x))^2. \quad (13)$$

It can be seen from equation (12) that the root mean square error (RMSE) of the predicted value $\Phi(x)$ generated by the ensemble methods is less than the average value $\Phi(x, D)$ of RMSE, and the more unstable $\Phi(x, D)$ is, the greater the ensemble methods' improvement of model performance.

Figure 6 shows the PMF model based on FCM and bagging-BP.

4.2. Algorithm Description. The algorithm flow of bagging algorithm with BP neural network and PMF model is given in Algorithm 3.

5. Experiments

In this part, we mainly test our hypothesis through several groups of experiments: FCM clustering methods are applied to the PMF model from different aspects to achieve the purpose of prediction accuracy. At the same time, the prediction accuracy of this method is verified, and the mean error (MAE) and root mean square error (RMSE) of the prediction are reduced:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{r_{ij} \in C} |\hat{r}_{ij} - r_{ij}|}{N}, \\ \text{RMSE} &= \sqrt{\frac{\sum_{r_{ij} \in C} |\hat{r}_{ij} - r_{ij}|^2}{N}}, \end{aligned} \quad (14)$$

where \hat{r}_{ij} is the prediction score, r_{ij} is the actual score of the test set, and N is the number of data pieces contained in the test set. From the definition of MAE and RMSE, MAE can well reflect the prediction error, while RMSE is more sensitive to outliers with a larger error. By calculating the root mean square of the sum of the square error between the predicted user score and the actual user score to predict the accuracy, the smaller the RMSE value, the better the recommendation quality. The smaller the MAE and RMSE, the higher the accuracy of recommendation.

For the models under the same evaluation method, this paper will choose the evaluation index used in the comparison model to evaluate the accuracy of scoring prediction.

5.1. Relevant Parameter Settings. Without losing generality, we take 80% of the data as the training data according to the clustering results and then predict the remaining 20% of the recommended accuracy, and set the regularization factor of this experiment $\lambda_U = \lambda_V = \lambda_{bu} = \lambda_{bv} = 0.01$; the learning rate of SGD $\alpha = 0.03$. The number of hidden layers of BP is 100. The datasets selected in this paper are MovieLens and FilmTrust, which are, respectively, applied to PMF, FCM-PMF, and FCM-bagging-BP-PMF models for comparison and conclusion.

5.1.1. Datasets Information. This experiment is carried out on MovieLens and FilmTrust datasets, both of which contain the user's rating information of the project. The rating values are 1–5 discrete values, and the sparsity is 4.47% and 1.04%, respectively, which belong to a high-dimensional and sparse matrix. The specific information of the dataset is shown in Table 2.

This paper studies the clustering number of the model. The experiment shows that different clustering numbers have different effects on the performance of the model. In the experiment, we set the clustering number as 10, 20, 30, 40, and 50. The experimental results on the MovieLens dataset are shown in Figure 7.

5.2. Comparison of Recommendation Accuracy. To verify the accuracy of the proposed model, the PMF model based on FCM and bagging BP is evaluated by experiments, and the results are compared with the existing models MF and PMF in two datasets. The comparison results of RMSE and MAE of different models in different datasets are shown in Table 3.

It can be seen from Table 3 that the performance of the model with the fuzzy clustering method is better than that without clustering. The RMSE and MAE of the PMF model with fuzzy clustering method on MovieLens (1 M) dataset are about 0.9305 and 0.95268, respectively. The RMSE and MAE of the FCMPMF model with fuzzy clustering method are about 0.83781 and 0.74131, respectively, which are improved by 9.27% and 21.14%. Finally, the RMSE and MAE of the PMF model based on bagging BP and fuzzy clustering are about 0.79765 and 0.73074, respectively. The effect of fuzzy clustering is improved by 4.02% and 1.06%. The results of the three models in RMSE and MAE of the MovieLens (1 M) dataset are shown in Figures 8 and 9.

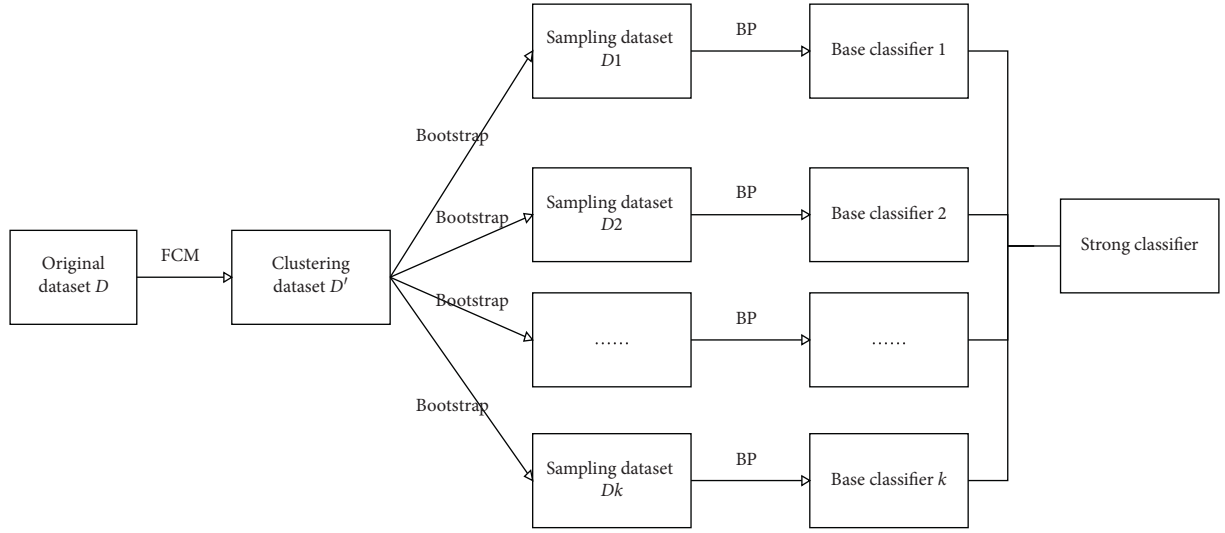


FIGURE 6: PMF model based on FCM and bagging-BP.

Input: constructing the rating matrix from training dataset $D = \{(x_i, y_i)_{i=1}^m\}$. Initialize related parameters.

Output: the rating prediction result of this sample x of this test set.

- (1) for $t = 1, \dots, k$ do (k is the number of base model)
 - 1.1: randomly select $k-1$ samples from the training set DT (sampling with replacement)
 - 1.2: training the BP neural network with this sample to obtain the base model E_i
- (2) Averaging all the base models $\{E_{i=1}^k\}$ to get a strong learner $F(x) \leftarrow 1/b \sum_{i=1}^k E_i(x)$
- (3) The strong learner is used to test the dataset
- (4) Apply PMF model with clustered data; initialize P and Q with gauss distribution
- (5) Rating prediction

ALGORITHM 3: The algorithm flow of bagging algorithm with BP neural network and PMF model.

TABLE 2: Datasets information.

Dataset	Number of users	Number of MAE	Number of scoring records	Sparsity (%)
MovieLens	6450	3706	1000209	4.47
FilmTrust	1642	2071	35497	1.04

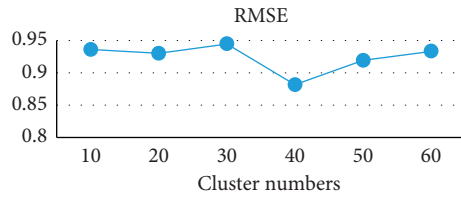


FIGURE 7: Influence of cluster numbers on RMSE.

TABLE 3: Comparison of RMSE and MAE of test data in MovieLens dataset.

Model	RMSE	MAE
PMF	0.93052	0.95268
FCMPMF	0.83781	0.74131
FCM-bagging-BP-PMF	0.79765	0.73074

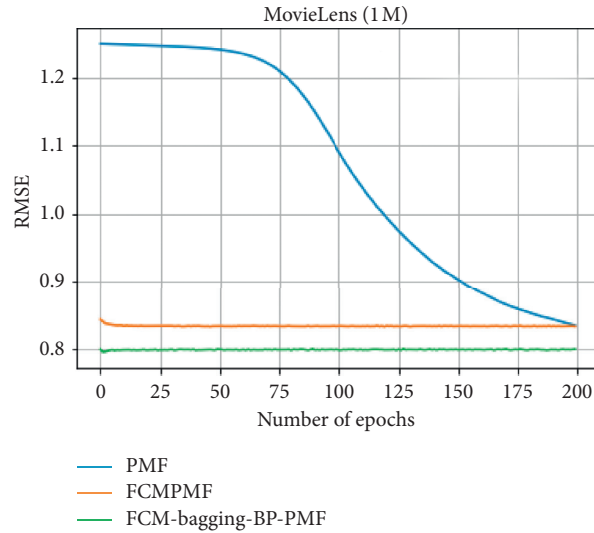


FIGURE 8: MovieLens (1 M) RMSE comparison.

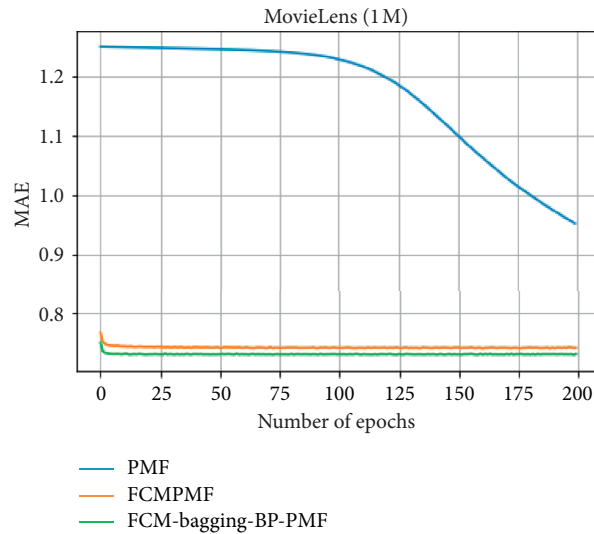


FIGURE 9: MovieLens (1 M) MAE comparison.

TABLE 4: Comparison of RMSE and MAE of test data in FilmTrust dataset.

Model	RMSE	MAE
PMF	1.44094	1.83424
FCMPMF	1.40143	1.82315
FCM-bagging-BP-PMF	1.39723	1.79593

It can be seen from Table 4 that the overall performance of the model with the fuzzy clustering method is better than that without clustering. The RMSE and MAE of the PMF model with fuzzy clustering method on the FilmTrust dataset are about 1.440940 and 1.83424, respectively. The RMSE and MAE of the FCMPMF model with fuzzy clustering method are about 1.401439 and 1.82315, respectively, and the effect is

improved by about 3.95% and 1.11%. Finally, the RMSE and MAE of the PMF model based on fuzzy clustering and bagging BP are about 1.397237 and 1.79593, respectively, and the effect of fuzzy clustering is improved by about 0.42% and 3.66%. The results of the three models in RMSE and MAE of the FilmTrust dataset are shown in Figures 10 and 11.

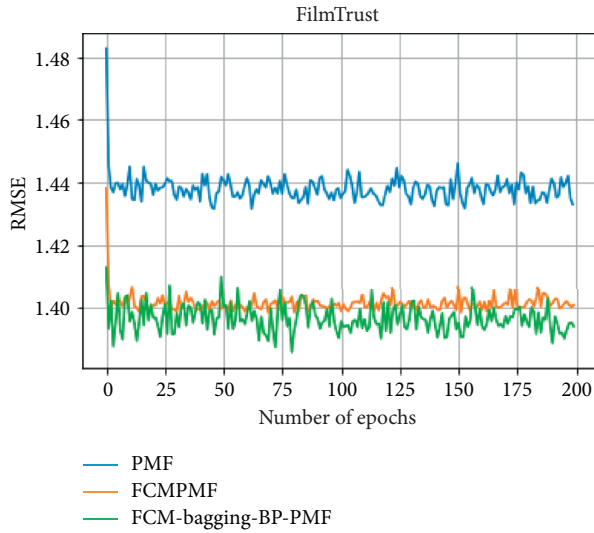


FIGURE 10: FilmTrust RMSE comparison.

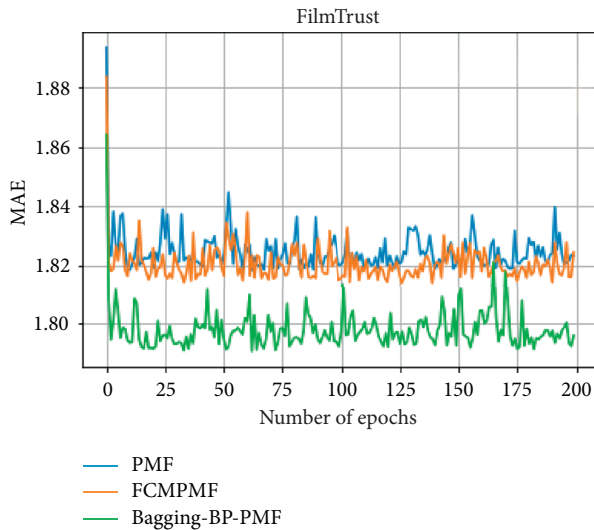


FIGURE 11: FilmTrust MAE comparison.

6. Conclusions

In this paper, a probability matrix factorization model based on BP neural network ensemble learning and fuzzy clustering is proposed. By using the similarity of the scoring matrix, the fuzzy clustering method is used to divide the users, which effectively solves the problem of scoring consistency; each base learner uses BP neural network to find the optimal weights and then carries out integrated processing to build a strong learner. The PMF model is built on the strong learner to improve the accuracy of the model prediction score.

Data Availability

We use the public datasets of MovieLens (1 M) and FilmTrust, and our model and related hyperparameters are provided in our paper.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by Key Projects of Natural Science Research in Universities of Anhui Province: Research and Implementation of Decoding Unit for Ternary Optical Processor (nos. KJ2020A0681 and KJ2019A0682) and Research on Key Technologies of Digital Survival of Ceramic Cultural Relics and Institute of Networks and Distributed System of Chaohu University.

References

- [1] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.
- [2] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proceedings of the International Conference on Neural Information Processing Systems*, Vancouver, Canada, December 2007.
- [3] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo," in *Proceedings of the International Conference on Machine Learning*, Helsinki, Finland, July 2008.
- [4] P. Akulwar and S. Pardeshi, "Bayesian probabilistic matrix factorization—a dive towards recommendation," in *Proceedings of the International Conference on Inventive Computation Technologies IEEE*, Coimbatore, India, August 2016.
- [5] Y. Fang, Y. Fu, and J. Zhou, "Personalized recommendation algorithm based on Ensemble Learning," *Computer Engineering and Application*, vol. 47, no. 10, pp. 1–4, 2011.
- [6] Y. Cui, W. Qi, and H. Pang, "Extreme gradient boosting recommendation algorithm with collaborative filtering," *Computer Application Research*, vol. 37, no. 1, pp. 62–65, 2020.
- [7] R. Katarya, O. P. Verma, Verma et al., "Recommender system with grey wolf optimizer and FCM," *Neural Computing and Applications*, vol. 30, no. 5, pp. 1679–1687, 2018.
- [8] T. Hastie, R. Tibshirani, and J. Friedman, *Ensemble Learning*, Springer, New York, NY, USA, 2009.
- [9] J. Liao, J. Lv, Y. Shao et al., "Application of BP neural network ensemble model based on bagging algorithm," *International Journal of Machine Learning and Computing*, vol. 9, no. 2, pp. 121–128, 2019.
- [10] H. Ichihashi, K. Honda, A. Notsu et al., "FCM classifier for high-dimensional data," in *Proceedings of the IEEE International Conference on Fuzzy Systems*, New Orleans, LA, USA, June 2008.
- [11] J. Wu and Z. Wu, "Improved fuzzy C-means clustering for personalized product recommendation," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 6, no. 3, pp. 393–399, 2013.