



Movie Recommendation System for Educational Purposes Based on Field-Aware Factorization Machine

Fei Lang^{1,2} · Lili Liang^{2,3} · Kai Huang^{2,3} · Teng Chen^{2,3} · Suxia Zhu^{2,3} 

Accepted: 20 April 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

With rich resources, movies have been applied as instructional media in the domain of education, such as fields of Second/Foreign Language Learning, Communication, and Media Art. Factorization machine (FM) can effectively simulate common matrix factorization models by changing the form of real-value vector, which can be utilized in movies recommendation under the context of education. However, it is usually used to solve classification tasks. This paper applies the field-aware factorization machine (FFM) to solve movie rating prediction and help users select appropriate movies for learning purposes. In order to further enhance the availability of the model, clustering algorithm is also integrated in FFM for adding new fields. The experimental results demonstrate the effectiveness of the proposed methods in reducing the RMSE.

Keywords Movies recommendation · Education · Collaborative filtering · Field-aware factorization machine · Clustering

1 Introduction

In the era of big data, the amount of information is huge, the transmission speed is extremely fast, and the resources can be shared. In the context of education, the integration and optimal utilization of educational information resources are very significant in the information age. However, the explosion of information has brought convenience and confusion. It not only satisfies learners' requirement for knowledge, but also makes it impossible for learners to directly obtain the information they really need among massive data. Under the impact of a huge amount of information, learners have to spend a lot of time on information search, identification, and screening. From this perspective, recommendation systems play the role in the domain of education by retrieving and filtering the data through content and similar profiles [1]. Especially in the

context of smart education, as teaching and learning can be implemented anytime, anywhere with intelligent devices [2], recommendation systems are in support of this brand new learning process.

With E-learning being increasingly accepted among students from different levels, many studies indicate recommendation systems' role in various aspects of virtual education. Without assistance from teachers in the learning process, recommendation systems involve more in virtual educational context. It helps students make choices of online course and learning materials which are appropriate to them. With this regard, students' learning behaviors and users' function are analyzed based on hybrid and context-specific recommendation systems [3]. Thus, recommendation systems are proposed to assist students obtain personalized course materials in several studies [4, 5, 6]. Moreover, educational data are increasingly investigated automatically from various sources, including course assessments, students' navigation history, information systems, e-learning activities and so on [7]. Thus, recommend system utilizes those data to make recommendations in not only learning materials but also learning performance prediction. It provides information to predict students' success or failure and helps make intervention during learning process [8].

Collaborative filtering (CF) algorithm has been widely used in the recommendation systems. It can be roughly categorized into memory-based CF and model-based CF [9]. Memory-based CF algorithm depends on the

✉ Fei Lang
zhusuxia@hrbust.edu.cn

¹ School of Foreign Languages, Harbin University of Science and Technology, Harbin, Heilongjiang, China

² Research Center of Information Security and Intelligent Technology, Harbin University of Science and Technology, Harbin, Heilongjiang, China

³ School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang, China

user-item rating matrix, which can further be classified as the user-based algorithm and the item-based algorithm. Based on the similarities of the active user and the other users, the users with high similarities are selected as the neighbors of the active user in the user-based algorithm. Then the weighted average of the neighbors' rating on a specific item is calculated in the recommend system. The respective similarities are finally treated as weights. The memory-based CF has some limitations. Along with the increasing of users and commodities, a user can only generate a rating behavior for a few numbers of commodities, making the entire rating matrix very sparse. The similarity values are unreliable if data are sparse [10].

The model-based CF algorithm overcomes the shortcomings of the memory-based CF algorithm and achieves better prediction performance. It mainly uses data mining models or machine learning models to realize the prediction of the missing ratings, such as singular value decomposition (SVD), SVD++ [11, 12]. SVD is one of the basic models, which directly decomposes the user-item rating matrix. Later, SVD++ introduces time information based on SVD. However, the main drawback of matrix factorization algorithms is that it is built for specific input data and cannot be applied to different tasks. In order to effectively avoid the disadvantages of traditional matrix factorization models, a class of CF algorithms is proposed based on the hidden factor model, such as factorization machine [13] and field-aware factorization machine [14]. The FFM algorithm is a further enhancement to the FM algorithm, which introduces the concept of field into FM. The proposed CF algorithm in this paper is mainly based on FFM model.

In the proposed CF algorithm, the attribute data of the user and the item is expressed as a real-valued feature vector, and the user's rating value of the item is used as a class label. The recommendation question is then thereby formulated into a regression prediction problem [15]. There are plenty of categorical features in the input data, which will be encoded into multiple binary features by one-hot encoding technology. It also leads to the sparseness of the training data. Fortunately, the FFM model itself has a certain capacity for processing the sparse data. In addition, the clustering methods are combined in FFM model to enhance the performance of the rating prediction. The clusters are utilized as special field in FFM model to alleviate the sparseness of the input data [16].

2 Related work

According to the educational requirements, recommendation systems play the role in many fields such as application for university, course selection, e-learning, career decisions and so on. Being an essential part of higher

education, course selection plays a significant role in not only academic study but also other related issues like curriculum decision and materials selection suitable for students. A host of studies show that it gives students support for course selection through different methods in recommendation systems [2, 17].

However, as the amount of data increases, it becomes difficult for people to make correct decisions in a large amount of information. This phenomenon is called information overload [18]. The recommendation systems can skillfully alleviate the obstacle of information overload by effectively discovering the potential needs of users and picking out the information that users are interested in from a substantial quantity of candidate data. The recommendation systems can be split off into two categories in general, content-based recommendation and collaborative filtering [9].

Content-based recommendation is mainly to use the content information of the item to mine similar items. Further, specific clustering algorithm will be used to obtain the center of clustering of these vectors which characterize the user preferences to the article [19].

Collaborative filtering is another important branch of the recommendation systems, which can be categorized as memory-based CF and model-based CF [20]. Different from the content-based recommendation, collaborative filtering depends on the customers' ranking of the items. Because users who have similar ratings for the same items have similar preferences, collaborative filtering performs well, when the content information is scarce. In addition, collaborative filtering increases the diversity of recommendation results [21].

The memory-based CF algorithm directly utilizes all the historical rating to predict the user-item rating matrix. To effectively alleviate the data sparse problem, clustering algorithms are introduced in the memory-based CF, in order to divide the entire rating matrix into multiple segments. Then the CF algorithm can be applied on each small rating matrix. Based on the clustering results, the accuracy and scalability will be improved in condition of the missing ratings. Besides data sparse, the memory-based CF has high time and space complexity [19].

The model-based CF algorithm which effectively solves the problems of the memory-based CF is based on matrix factorization. Matrix factorization (MF) algorithm [22] has lots of attention due to their higher prediction accuracy and scalability, which copes with the data sparsity. It completes the original rating matrix by mapping both the users and items into the same latent feature space, extracting hidden semantic information and multiplying low rank hidden vector [18]. SVD is a typical method in early MF-based algorithms [23]. Then several MF-based algorithms are put forward in recommendation systems. Hofmann established a CF model by applying probabilistic latent semantic

analysis [24]. Srebro et al. proposed the MF-based model based on maximum-margin [25]. Kurucz et al. proposed a MF model with EM optimization [26]. In the Netfit Price Competition, the CF model that introduced the regular term achieved good results [27]. Several improved MF-based models are presented in recent years, such as biased SVD model [28], SVD++, and PMF model [29]. The main drawback of MF-based CF algorithm is that it is effective to specific input data and cannot be directly migrated to different tasks.

In 2010, the factorization machine proposed by Rendle simulated the common matrix factorization model by changing the form of the input. It is considered as a general model, which effectively solves the drawback of MF-based CF algorithm. FM model was firstly used in context recommendation in 2014 [30]. In FM, the attributes are represented as a real-valued feature vector, the ratings are treated as class labels. Then, CF can be solved as a regression problem. Field-aware factorization machine is a modified version of FM. Compared to FM, FFM introduces more structured control, which is effective in building cross-features. In FFM, features are organized according to the fields. A field can be considered as a class of features. FFM learns a different set of latent information, because each feature uses a different k -vector to interact with other features from different field. In addition, most FFM methods discard the global bias and linear terms [31].

Clustering methods can effectively alleviate the sparsity of input data, which reduces the consumption of computing time in the recommendation systems. In [32], the clustering for CF algorithm is proposed by Ungar L. et al., which clustered the users and items independently. But it does not improve the prediction accuracy of missing ratings. O'Connor et al. proposed a CF algorithm based on item clustering, whose scalability is improved [33]. Xue et al. proposed a cluster-based smoothing method which has a good performance on sparse datasets [34]. It not only improved the scalability, but also the prediction accuracy. However, the above methods mainly focus on user or item based CF. In FM and FFM, the sparsity problem is not solved. How to alleviate the sparseness in FFM model and improve the performance of recommendation are main problems in this paper.

3 Method

The concept of field in FFM is used to add a clustering field. The clustering result of users or items is taken as a feature under this field. Firstly, users and items were clustering by Min-Batch k -Means, and then the clustering features were taken as two clustering fields of FFM model into features interaction. By introducing clustering information, the prediction accuracy of FFM model is further improved.

3.1 Field-aware factorization machine

Factorization Machine performs well in modeling the interactions between different kinds of features. In the highly sparse learning tasks, Factorization Machine is effective to model the interactions between features. It can simulate most of the MF-based CF algorithm through feature engineering, which has been widely applied to many important tasks, such as the recommendation systems, click-through rate estimation.

FM can capture the d -order interactions among n features of the input vector x . The 2-order FM model is defined as follows:

$$\hat{y}(x) = w_0 + \underbrace{\sum_{i=1}^n w_i x_i}_{\text{linear}} + \underbrace{\sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j}_{\text{polynomial}} \quad (1)$$

Where the parameters $\theta = \{w_0, w_1, \dots, w_n, v_1, \dots, v_n, k\}$, $w_0 \in R$, $V \in R^{n \times k}$, and $\langle \cdot, \cdot \rangle$ represents the inner product between two hidden vectors of dimension k :

$$\tau_{i,j} \approx \langle v_i, v_j \rangle = \sum_{f=1}^k v_{i,f} v_{j,f} \quad (2)$$

A row vector v_i of V , which contains k factors, is the i -th variable. The FM model's linear term is equivalent to a linear regression model. Instead of an independent parameter $\tau_{i,j}$, the polynomial term uses a factorized parameterization $\langle v_i, v_j \rangle$ to model the interaction between the i -th and j -th variables. In [35], it has been demonstrated that FM can be computed in linear time $O(kn)$. Eq. (1) can be reformulated as:

$$\hat{y}(x) = w_0 + \sum_{i=1}^n w_i x_i + \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{i,f} x_i \right)^2 - \sum_{i=1}^n v_{i,f}^2 x_i^2 \right) \quad (3)$$

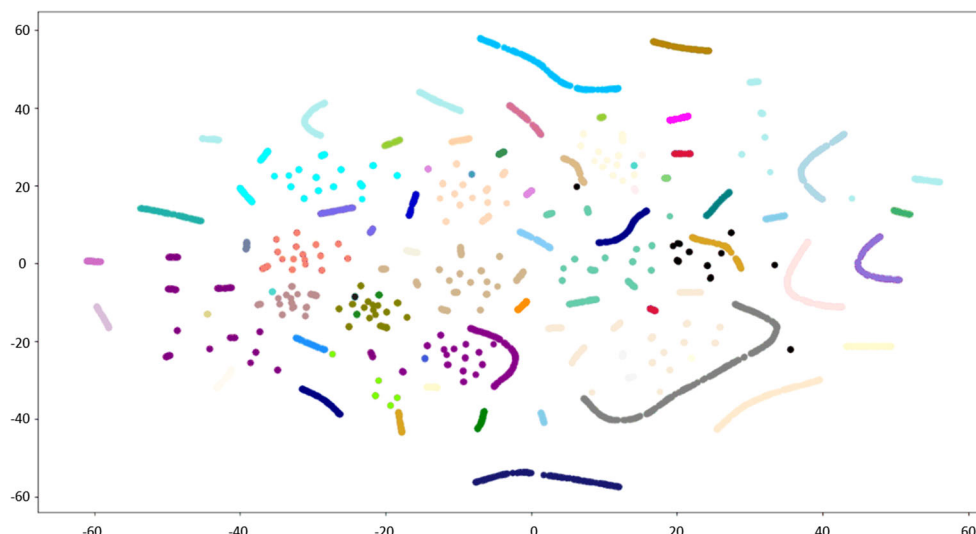
As a modified version based on FM, the features in FFM are organized according to the fields, that can be considered as a class of features. Comparing to one feature associated with a vector of latent factors in FM, each feature uses a corresponding k -vector, when it interacts with other features in different fields in FFM. The global bias and the linear terms used in FM are discarded in FFM. The detailed can be seen in [14]. The 2-order FFM is defined as follows:

$$\hat{y}(x) = \sum_{i=1}^n \sum_{j=i+1}^n \langle v_{i,f(j)}, v_{j,f(i)} \rangle x_i x_j \quad (4)$$

Where the field feature i belongs to $f(i)$.

From the above description, we can see that more structured control is introduced in FFM than in FM in modeling the interaction between features. The interactions are well represented.

Fig. 1 The result of users clustering



3.2 Clustering method

Clustering is a typical unsupervised algorithm whose core idea is to gather similar objects into a group. The typical clustering algorithm is k -Means. Although the idea is simple, it is still the commonly used clustering algorithm in the industry because of its efficiency and scalability.

Mini-batch k -Means is a variant of the k -Means algorithm, which optimizes the computational efficiency of the original

k -Means algorithm by using mini-batches in order to reduce the time complexity of the clustering algorithm. In Mini-batch k -Means, when the center point of each cluster is updated, a small subset named mini-batch data is randomly selected from all the data. According to this subset, the center points are then updated [36]. The above process is repeated in each iteration until the algorithm convergence.

The description of Mini-batch k -Means algorithm is as follows.

Algorithm 1 Mini-batch k -Means.

```

1: Given:  $k$ , mini-batch size  $b$ , iterations  $t$ , data set  $X$ 
2: Initialize each  $c \in C$  with an  $x$  picked randomly from  $X$ 
3:  $v \leftarrow 0$ 
4: for  $i = 1$  to  $t$  do
5:    $M \leftarrow b$  examples picked randomly from  $X$ 
6:   for  $x \in M$  do
7:      $d[x] \leftarrow f(C, x)$  // Cache the center nearest to  $x$ 
8:   end for
9:   for  $x \in M$  do
10:     $c \leftarrow d[x]$  // Get cached center for this  $x$ 
11:     $v[c] \leftarrow v[c] + 1$  // Update per-center counts
12:     $\eta \leftarrow \frac{1}{v[c]}$  // Get per-center learning rate
13:     $c \leftarrow (1 - \eta)c + \eta x$  // Take gradient step
14:   end for
15: end for

```

4 Experiments

4.1 Comparative models

This paper sets up two groups of comparative models.

In Memory-based CF algorithm, user-based CF algorithm calculates similarity between users based on user-item rating matrix. Item-based CF algorithm calculates similarity between items based on user-item rating matrix.

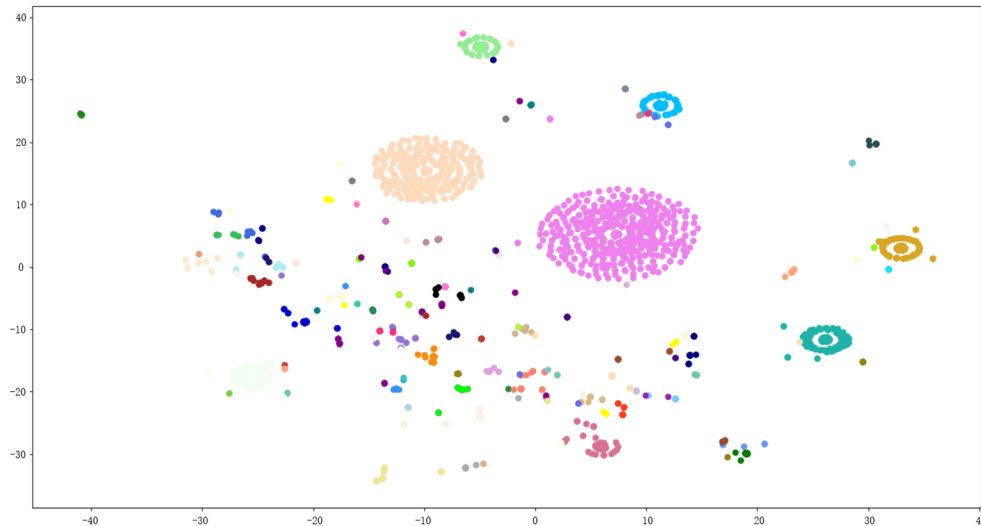


Fig. 2 The result of movies clustering

In Model-based CF algorithm, regularized singular value decomposition model (RSVD) [37] introduces regular terms based on traditional singular value decomposition. Probabilistic matrix factorization model (PMF) [38] introduces probability concept based on RSVD. Bayesian probabilistic matrix factorization model (BPMF) [29] uses the Markov chain Monte Carlo method efficiently. Non-negative matrix factorization model (NMF) [39] restricts the latent feature during the learning process. The factorization machine model can simulate common MF models only through feature engineering.

The above models use the default parameters except of FM. In FM and FFM, the learning rate is set to 0.01, the hidden factor is set to 100, and the iteration number is set to 1000. The early-stop mechanism is adopted.

4.2 Data sets

Movies are not only the ways of entertainment, but also an interesting learning tool for students. Take Teaching English as a foreign language as an example, English movies offer learners authentic English input, which may improve their motivations of language learning [40]. However, the representation of real-life English has different impacts on learners from different levels. Thus, it is necessary to help users select movies better appropriate for learning purposes. The actual datasets are used in the experiments. MovieLens-1M is from a research website MovieLens at the University of Minnesota [41]. Thousands of users visit the site and rates the movies in every week. MovieLens-1M collected data for nearly 3 years, which contains over one million ratings from 6040 users on 3952 movies. The rating is from 1 to 5 stars. The larger the rating is, the more satisfied the user is with the movie. The experiments are based on 100 K and 1 M samples,

respectively. The input data is split into training set and test set according to the ratio of 80% and 20%.

4.3 Evaluation index

As the rating prediction is a regression problem, we use Mean Absolutely Error (MAE) and Root Mean Square Error (RMSE) as metrics to evaluate the prediction result.

MAE is defined as follows:

$$MAE = \frac{\sum_{i \in D} |p_i - r_i|}{|D|} \quad (5)$$

RMSE is defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i \in D} (p_i - r_i)^2}{|D|}} \quad (6)$$

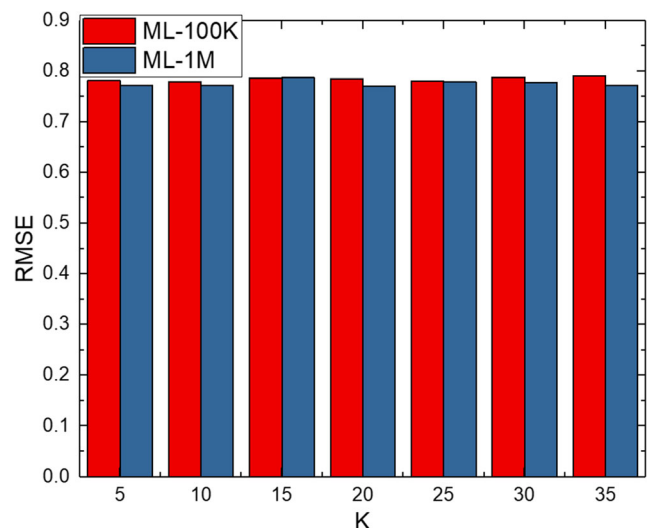


Fig. 3 The prediction performance with different K in FFM (RMSE)

Table 1 The prediction performance based on different CF algorithms (RMSE)

CF algorithm	ML-100 K	ML-1 M
Item-based CF	0.9354	0.9057
User-based CF	0.9536	0.9253
Biased SVD	0.9473	0.8770
PFM	0.9667	0.9234
FM	0.8078	0.7928
FFM	0.7785	0.7714
FFM+ Mini-Batch k-Means	0.7745	0.7648

Where p_i is the predicted rating, r_i is the real rating, and D is the test data collection. The lower the MAE and RMSE are, the better recommendation accuracy is.

4.4 Experimental results

Firstly, visual analysis is conducted on the clustering results of users and movies in 1 M dataset, respectively. The user's age, occupation, gender, postcode and other information are used for clustering in the Mini-Batch k-Means. The users are clustered to 95 groups. It is obvious in different groups of users. Figure 1 shows the result of users clustering.

The category of movies is utilized in movies clustering. There are 18 categories in the dataset, in which a movie can belongs to many categories. The movies are clustered to 80 groups. Most movies are clustered to some main groups. However, the distribution shows a long tail. Figure 2 shows the result of movies clustering.

Secondly, the influence of hidden vector dimension parameter on the prediction error of final score are analyzed. From Fig. 3, FFM is sensitive to the parameter K . But the fluctuation range is not large.

Table. 1 shows the prediction performance based on different CF algorithms. From the experimental results, we can see that the performance of FM and FFM are generally better than the traditional memory-based and part of the MF-based CF algorithm. When Mini-Batch k-Means is introduced, the prediction performance of the FFM model has been further improved, which proves the effectiveness of the proposed method.

5 Conclusion

It is difficult to obtain valuable educational information due to Internet information overload. The collaborative filtering algorithm is utilized to solve movies recommendation for educational purpose. The field-aware factorization machine effectively avoids the drawbacks of traditional model-based CF

algorithms. The paper utilized the field-aware factorization machine combined with clustering algorithm to predict the ratings of the movies, which contributes to movies recommendation. Experiments on MovieLens dataset shows that the proposed method gets better performance.

Funding This research is partly supported by The Heilongjiang Higher Education Teaching Reform Project (SJGY20200320), The National Natural Science Foundation of China (No.60903083), The Scientific Research Foundation for The Overseas Returning Person of Heilongjiang Province of China (LC2018030), The Fundamental Research Foundation for Universities of Heilongjiang Province (JMRH2018XM04).

References

1. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions [J]. *IEEE Trans Knowl Data Eng* 17(6):734–749
2. Lin J, Pu H, Li Y, Lian J (2018) Intelligent recommendation system for course selection in smart education[J]. *Procedia Comput Sci* 129:449–453
3. Ansari M H, Moradi M, NikRah O, et al. (2016) CodERS: a hybrid recommender system for an E-learning system[C]. 2016 2nd international conference of signal processing and intelligent systems (ICSPIS). IEEE, pp 1–5
4. Zhuhadar L, Nasraoui O, Wyatt R, et al. (2009) Multi-model ontology-based hybrid recommender system in e-learning domain[C]. 2009 IEEE/WIC/ACM international joint conference on web intelligence and intelligent agent technology. IEEE 3:91–95
5. Santos OC, Boticario JG (2011) Requirements for semantic educational recommender systems in formal e-learning scenarios [J]. *Algorithms* 4(2):131–154
6. Ghauth K I B, Abdullah N A (2009). Building an e-learning recommender system using vector space model and good learners average rating[C]. 2009 ninth IEEE international conference on advanced learning technologies. IEEE, pp 194–196
7. Prieto L P, Rodríguez-Triana M J, Martínez-Maldonado R, et al. (2019) Orchestrating learning analytics (OrLA): supporting inter-stakeholder communication about adoption of learning analytics at the classroom level [J]. *Australas J Educ Technol*, pp 35(4)
8. Baker RS, Inventado PS (2014) Educational data mining and learning analytics [M]. *Learning analytics*. Springer, New York, pp 61–75
9. Yao L, Sheng QZ, Ngu AHH et al (2014) Unified collaborative and content-based web service recommendation [J]. *IEEE Trans Serv Comput* 8(3):453–466
10. Luo X, Zhou M, Xia Y et al (2014) An efficient non-negative matrix-factorization-based approach to collaborative filtering for recommender systems [J]. *IEEE Trans Ind Inform* 10(2):1273–1284
11. Ricci F, Rokach L, Shapira B (2011) Introduction to recommender systems handbook [M]. *Recommender systems handbook*. Springer, Boston, pp 1–35
12. Koren Y, Bell R, Volinsky C (2009) Matrix factorization techniques for recommender systems [J]. *Computer* 42(8):30–37
13. Rendle S (2012) Factorization machines with libfm [J]. *ACM Trans Intell Syst Technol (TIST)* 3(3):1–22
14. Chislenko A, Lashkari Y, Tiu D D, et al. (n.d.) Method and apparatus for efficiently recommending items using automated

- collaborative filtering and feature-guided automated collaborative filtering: U.S. Patent 6,092,049[P]. 2000-7-18
15. Loni B, Shi Y, Larson M et al (2014) Cross-domain collaborative filtering with factorization machines[C]. European conference on information retrieval. Springer, Cham, pp 656–661
 16. Shuai L, Wang S, Xinyu L et al (2021) Fuzzy detection aided real-time and robust visual tracking under complex environments. *IEEE Trans Fuzzy Syst* 29(1):90–102
 17. Chang PC, Lin CH, Chen MH (2016) A hybrid course recommendation system by integrating collaborative filtering and artificial immune systems [J]. *Algorithms* 9(3):47
 18. Gao S, Luo H, Chen D, Li S, Gallinari P, Ma Z, Guo J (2013) A cross-domain recommendation model for cyber-physical systems [J]. *IEEE Trans Emerg Top Comput* 1(2):384–393
 19. Yang Z, Wu B, Zheng K, Wang X, Lei L (2016) A survey of collaborative filtering-based recommender systems for mobile internet applications [J]. *IEEE Access* 4:3273–3287
 20. Yang H, Ling G, Su Y, Lyu MR, King I (2015) Boosting response aware model-based collaborative filtering [J]. *IEEE Trans Knowl Data Eng* 27(8):2064–2077
 21. Liu S, Guo C, Al-Turjman F et al (2020) Reliability of response region: a novel mechanism in visual tracking by edge computing for IIoT environments. *Mech Syst Sig Process* 138:106537
 22. Adomavicius G, Tuzhilin A (2005) Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions [J]. *IEEE Trans Knowl Data Eng* 17(6):734–749
 23. Sarwar B, Karypis G, Konstan J, et al. (2000) Application of dimensionality reduction in recommender system-a case study[R]. Minnesota Univ Minneapolis Dept of Computer Science
 24. Hofmann T (2004) Latent semantic models for collaborative filtering [J]. *ACM Trans Inform Syst (TOIS)* 22(1):89–115
 25. Srebro N, Rennie J, Jaakkola T S (2005) Maximum-margin matrix factorization[C]. *Advances in neural information processing systems*, pp 1329–1336
 26. Kurucz M, Benczúr A A, Csalogány K (2007) Methods for large scale SVD with missing values[C]. *Proceedings of KDD cup and workshop*. San José, USA: Citeseer 12:31–38
 27. Gorrell G. (2006) Generalized Hebbian algorithm for incremental singular value decomposition in natural language processing[C]. 11th conference of the European chapter of the association for computational linguistics
 28. Yao-Ning F, Yun-Fei G, Xue-Tao D (2013) An improved regularized singular value decomposition recommender algorithm based on tag transfer learning [J]. *J Electron Inf Technol* 35(12):3046–3050
 29. Mnih A, Salakhutdinov RR (2007) Probabilistic matrix factorization [J]. *Adv Neural Inf Proces Syst* 20:1257–1264
 30. Rendle S, Gantner Z, Freudenthaler C, et al. (2011) Fast context-aware recommendations with factorization machines [C]. *Proceedings of the 34th international ACM SIGIR conference on research and development in Information Retrieval*, pp 635–644
 31. Yan P, Zhou X, Duan Y (2015) E-commerce item recommendation based on field-aware factorization machine [M]. *Proceedings of the 2015 international ACM recommender systems challenge*, pp 1–4
 32. Ungar LH, Foster DP (1998) Clustering methods for collaborative filtering [C]. *AAAI Work Recomm Syst* 1:114–129
 33. O'Connor M, Herlocker J (1999) Clustering items for collaborative filtering[C]. *Proceedings of the ACM SIGIR workshop on recommender systems*. UC Berkeley, pp 128
 34. Xue G R, Lin C, Yang Q, et al. (2005) Scalable collaborative filtering using cluster-based smoothing[C]. *Proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval*, pp 114–121
 35. Rendle S (2010) Factorization machines [C]. *2010 IEEE international conference on data mining*. IEEE, pp 995–1000
 36. Shuai Liu, Shuai Wang, Xinyu Liu, et al (2021) Human memory update strategy: a multi-layer template update mechanism for remote visual monitoring. *IEEE Trans Multimedia*, pp 99:1–1
 37. Nguyen J, Zhu M (2013) Content-boosted matrix factorization techniques for recommender systems [J]. *Stat Anal Data Min: ASA Data Sci J* 6(4):286–301
 38. Sarwar B M, Karypis G, Konstan J, et al. (2002) Recommender systems for large-scale e-commerce: scalable neighborhood formation using clustering[C]. *Proceedings of the fifth international conference on computer and information technology*, 1:291–324
 39. Lee D D, Seung H S (2001) Algorithm for non-negative matrix factorization[C]. *Advances in neural information processing systems*, pp 556–562
 40. Sari A, Sugandi B (2015) Teaching english through english movie: advantages and disadvantages [J]. *J Engl Lit Educ: Teach Learn Engl Foreign Lang* 2(2):10–15
 41. Harper FM, Konstan JA (2015) The movielens datasets: history and context [J]. *Acm Trans Interact Intell Syst (tiis)* 5(4):1–19

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.