A Neural Network Based Explainable Recommender System

Jionghao Lin
Monash University
Clayton, VIC, Australia
jlin0028@student.monash.edu

Yiren Liu
University of Electronic
Science and Technology of China
Chengdu, Sichuan, China
yirenliu@std.uestc.edu.cn

Abstract

Recommendation system could help the companies to persuade users to visit or consume at a particular place, which was based on many traditional methods such as the set of collaborative filtering algorithms. Most research discusses the model design or feature engineering methods to minimize the root mean square error (RMSE) of rating prediction, but lacks exploring the ways to generate the reasons for recommendations. This paper proposed an integrated neural network based model which integrates rating scores prediction and explainable words generation. Based on the experimental results, this model presented lower RMSE compared with traditional methods, and generate the explanation of recommendation to convince customers to visit the recommended place.

1 Introduction

User-generated content such as rating and reviews about a place is gradually becoming a key factor for other people making a decision. Some platforms such as Yelp utilize recommender system to analyze the user preferences based on the reviews and rating to recommend a new place for other customers. According to (Liu et al., 2017), they introduce a phrase point-of-interests (POIs) to represent a set of places such as restaurant and tourist attractions where users are interested. In this pa-

per, we will continue to use this phrase and focus on restaurant rating prediction and explanation generation. Currently, much empirical research focuses on model design or feature engineering methods to improve model performance rather than discussing the ways to generate recommended reasons. This project proposed an integrated neural network based explainable recommender system, instead of using traditional methods such as user-based models, and Matrix Factorization, to generate explainable opinion-aspect pairs for a user, helping them explore their unfamiliar places, and predict rating score from that user for each POI and recommend POI with high predicted rating to each user. Our model could generate user latent preference from embedding layer in neural network after training rating predictions for POIs Instead of training models on user-POI rating matrix directly. The main contributions of this paper are summarized as follows:

- Predict rating stars for POIs based on user-POI rating matrix using neural-based model.

- Generate opinion-aspect pairs for users to explain reasons for recommended POI.

- Compare the model performance with the classical recommender system methods.

- Propose a way to evaluate the prediction of explainable pairs.

In the rest of the paper, we will discuss the related work in Section 2. Provide the de-

tails of data preprocessing, feature extraction, and data exploration in section 3. We will introduce the proposed model and other traditional models in Section 4. In Section 5 and Section 6, We discuss the method to evaluate the RMSE of predicting rating stars and the f-score of predicting explainable tuples. Then, we present the results including comparison with other baseline models. Finally, we conclude findings, limitation, and future work of this paper in section 7.

2 Related Work

Much empirical research focuses on either rating stars treating the scores as numerical data or explainable text generation offering the features of the POIs, and those studies adopt different models, features and evaluation methods (Liu et al., 2017; Bao et al., 2012; Zhang et al., 2015; Liu et al., 2013; Duan et al., 2017; Rendle et al., 2009). This project is motivated by these empirical studies on rating stars prediction, and we proposed an integrated model which can predict rating stars and generate explainable opinion-aspect pairs for users.

2.1 Collaborative Filtering for Rating Prediction

Collaborative Filtering is commonly used in recommender system(Chen et al., 2014; Konstas et al., 2009; Su and Khoshgoftaar, 2009; Wang and Blei, 2011) — exploiting similarity among the preference of users to generate recommendations. Research is done using users historical data to predict ratings, and deliver the results of recommendation to users (Balabanovi and Shoham, 1997; Rennie and Srebro, 2005). The traditional methods in rating prediction are the user-based model, and matrix factorization(MF). According to (Ricci et al., 2015), they introduced the user-based method which predicts the rating scores based on the other users who have similar preferences. Regarding matrix factorization, Ricci et al. mentioned the applications of two commonly used methods: Singular Value Decomposition(SVD) and Non-negative Matrix Factorization(NMF) (Ricci et al., 2015). Additionally, many researchers utilized neural-based collaborative filtering model for rating prediction (van den Oord et al., 2013; Wang et al., 2015). However, there are many limitations for the traditional algorithms such as scalability problems, and lacking bias terms. As for neural collaborative filtering, it performs better on rating prediction, but there is no description of generating explainable text in empirical works.

2.2 Explainable Recommendation System

Explainable recommendation system (ERS) gradually attracts more researchers to explore the ways to generate convincing explanations for users. According to (Hou et al., 2018), they state two forms of ERS to output explanations. One is to extract noun words or phrases to represent item features, and there are many works utilizes this form for explainable recommendation. However, Hou et al. claim that it lacks consideration of sentiment for each aspect summarized from other visited customers which is less convincing for future users. Another method is to generate a set of words containing opinion and aspect phrases from reviews to explain fine-grained aspect information for future users (Hou et al., 2018). In this project, we investigated the second form of ERS, and summarized many empirical works about this type as follows:

Baccianella et al. introduced a method using SentiWordNet in extracting opinion words and quantifying the sentiment polarity and strength of each word to the POI (Baccianella et al., 2010). Pero et al. proposed a recommendation system by integrating opinion information with the rating score for the POI (Pero and Horvth, 2013). Chen et al. demonstrated an algorithm about tensor decomposition to predict rating based on opinion-aspect pairs in review over multiple domains (Chen et al., 2014). However, these method lacks of fine-grained sentiment differences based on some specific aspect of POIs (Hou et al., 2018).

Though given plenty of research focusing on POI rating prediction in recommender systems, less research has investigated the recommender system offering the fine-grained explanation for specific aspects of POIs. Former Collaborative Filtering research focuses on improving the precision of recommendation but lacks information to help target users better understand the proposed recommendation. Our proposed model could resolve the mentioned limitations from empirical works, which integrate rating prediction and explainable pairs generation, and provide better performance compared with other traditional models.

3 Data

3.1 Data Preprocessing and Feature Extraction

The data is composed of business ID, user ID, rating score (range from 1 to 5), reviews, and date in the city Pittsburgh from Yelp (Developers, 2018) challenge dataset. Considering the impact of time,we set 2017 (full year) as the training dataset and 2018 (until June) as test dataset which contains 26918 reviews (from 13225 users to 450 restaurants). After tokenizing and POS tagging, review texts are transformed into syntax relations (shown in Figure 1) by utilizing the spaCy CNN dependency parsing model (Kiperwasser and Goldberg, 2016; Goldberg and Nivre, 2012), and these relation pairs containing users opinions towards aspects are extracted in the form of tuples liked<opinion, aspect>.

Additionally, SentiWordNet (Baccianella et al., 2010) is incorporated to quantify the opinion sentiment polarity and strength, which will be described in detail in the evaluation section.

Figure 1: Aspect-opinion Pairs Examples

Finally, after lemmatizing and lowercasing the words, we use 100-dimension Glove word embedding pretrained by twitter corpus (Pennington et al., 2014) , and then horizontally concatenate the word embedding to tuple embedding.

3.2 Data Exploration Analysis

After preprocessing the data, the stars distribution is visualized in Figure 2 which indicates that users rarely give 1 or 2 stars for POIs.
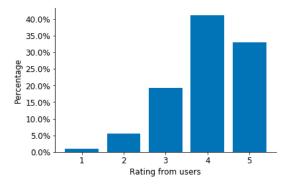

Figure 2: Rating Distribution Bar Chart

Then, we explore the opinion-aspect pairs distribution in the space of frequency and average rating. In Figure 3, these pairs are visualized based on the frequency greater than 700 and average stars that they appeared.
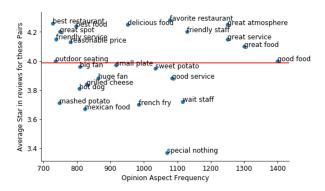

Figure 3: Word Pairs

Additionally, there are many examples having comparable high average stars which accord with their opinion words such as 'best restaurant', and 'best food'. On the other hand, many tuples such as 'good service' have positive opinion but present the position lower than the average stars line. This case will be discussed in our limitation part.

4 Method

There are many traditional model framework used for rating prediction. In this section, we

experimented with four algorithms: neural-based model, K nearest neighbor, SVD matrix factorization and NMF matrix factorization to predict the rating stars that a user would give for a POI based on the ratings of reviews posted in the past. Then, we demonstrate the ways of our models to generate opinion-aspect pairs for target users after completing the model training for RMSE.

4.1 Baseline Model for Rating Prediction

4.1.1 User-based collaborative filtering

The first baseline model is a user-based collaborative filtering algorithm based on the K nearest neighbor (KNN) algorithm. According to (Ricci et al., 2015), they claim that this method utilizes a ratings matrix to compute centered cosine similarity (Pearson relationship) between each user. Then, predict the rating score for the POI where the target user has never been there before based on the weight of similarity and scores from other top K users who are most similar to target user (Ricci et al., 2015). This method could provide decent performance in rating prediction. However, this method is based on the KNN algorithm which will suffer serious scalability problems when the number of users and the number of items increase (Shani and Gunawardana, 2011; Ricci et al., 2015).

4.1.2 Matrix Factorization

Matrix factorization is a commonly used collaborative filtering method (Koren, 2010) which decomposes the user-restaurant rating matrix into two latent feature matrices of user and restaurant and can be used to predict the rating stars of restaurants the users who have not visited. Singular value decomposition (SVD) and non-negative matrix factorization (NMF) are two commonly used matrix factorization methods and will be used as baseline models in this paper.

Matrix Factorization algorithms are known to have high computational efficiency and generate user and item embedding containing information about users preferences and item features, which can be extracted for further

analysis. However, traditional Matrix Factorization methods have a hindered performance compared to neural based models due to their ignorance of users and items bias terms.

4.2 Integrated Neural Collaborative Filtering

Our proposed neural-based model is motivated by (He et al., 2017), which is comprised of two parts — rating prediction and explainable pair generation (shown in Figure 4). As for rating prediction, the model is trained using the user-POI rating matrix and predicts the rating stars of each user on each POI. Concerning explainable pairs generation, we use the user embedding learned from embedding layer to find the top-K users who have the most similar preferences to the target user and then generate explanations based on their reviews over each specific POI. Our model resolves the issues mentioned in the baseline models and considers the bias terms when calculating users and POIs embedding. Additionally, the most similar users could be detected from user embedding instead of using other formulas to compute the correlation.
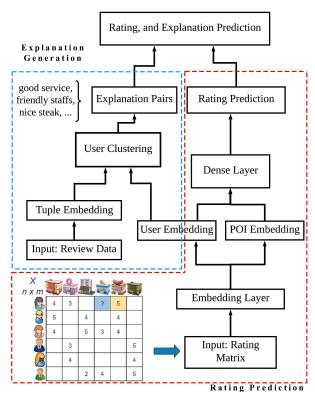

Figure 4: Integrated Neural Collaborative Filtering

4.2.1 Rating Prediction

We predict ratings for POIs that a user has not visited before. In rating prediction part (shown in Figure 4. right block), our model takes user-POI matrix as the input and predicts rating stars for the user over the POI. After training, the model vectorizes the latent preferences of users about POIs in the embedding layers (He et al., 2017), which will be used as the input for the explanation generation. (shown in Figure 4. left block).

After the score of each POI is predicted, the top-K POIs with highest ratings that the user has not visited in the past are offered as the recommended POIs for the user. Our neural-based model is established using Keras in Python. The model is trained using the review data of restaurants in Pittsburgh in 2017 and validated using data in 2018 and sets RMSE as the loss. We trained our model in 200 epochs, converging at around 0.17 (shown in Figure 5).
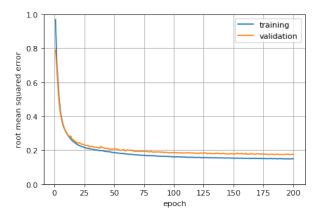

Figure 5: Learning curve for the NCF method

4.2.2 Explainable Pairs Generation

As the neural model completed the training phase, the user embedding layer is extracted from user-POI matrix which contains the information about the latent preferences of each user. After determining the POIs which will be recommended to each target user, we use the extracted user embedding to find the top-K most similar users of our target user who have been to the same POIs. The similarity between users is measured by calculating the cosine similarity using the user embedding. Then, we extract the pairs from the reviews of

the top-K users generated for the POIs using dependency parsing and sample from these pairs as the explanation provided to the target user.

5 Model Evaluation

5.1 Evaluation of Rating Prediction

The introduced recommendation strategy relies on the prediction of ratings over unvisited POIs. To evaluate the rating prediction model, we use root mean square error (RMSE) to measure its performance (Lu et al., 2015). As mentioned in the Section 3, we treat 2017 as training set and 2018 as testing set, so we predict the rating stars that a user will give for an unvisited POI in 2018, and use the ratings in test dataset to measure RMSE of the result.

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}(P_i - R_i)^2}{n}} \qquad (1)$$

5.2 Evaluation of Explanation Generation

To evaluate the quality of the generated explanations, we calculate the cosine similarity between the pairs that we predicted for POIs and pairs extracted from the user review of corresponding POIs in the test set. Since many opinion-aspect pairs present the same meaning but different word, this project sets the cosine similarity greater than 0.8 as true, otherwise false. Then, we calculate the median number of pairs of the target user in the training dataset and predict same amount of pairs according the median number.

The way that we propose for evaluation of explanation prediction will bring some problems including false positive and false negative. However, if we use exact match, it will ignore the same meaning between tuples with semantic similarity but different words such as 'great beef' and 'great steak'. Since we focus more on related pairs recommendation to users, we decide to propose this way to evaluate explainable pairs. After examining many tuples in the dataset, we found that setting the

cosine similarity threshold as 0.8 can generally distinguish relevant pairs from irrelevant ones (shown in Table 1).

Pair 1	Pair 2	Similarity	Label
('good', 'food')	('nice', 'food')	0.969 (+)	True
('great', 'sides')	('great','snacks')	0.821 (+)	True
('great', 'beef')	('nice', 'steak')	0.855 (+)	True
('bad', 'food')	('bad', 'pricing')	0.761 (+)	False
('good', 'service')	('bad', 'service')	0.952 (-)	False
('friendly', 'staff')	('cool', 'vibe')	0.579 (+)	False

Table 1: Examples for Evaluating Tuples

Furthermore, we found that many pairs have the same aspect but different opinion words like 'good' or 'bad' which cannot be effectively distinguished simply using GloVe vectors. Therefore, this paper introduces a cosine similarity based F-score with sentiment penalty and uses it as the evaluation method. In this method, if two words have similarity greater than 0.8 but with opposite sentiment polarity, we consider this as a false case (shown in equation 2). In Table 1, the cosine similarity with (+) means two words have the same sentiment polarity and (-) means different. Although 'good service' and 'bad service' have quite high similarity (0.952), this case will be labelled with False because the sentiment penalty presents (-). We will discuss more details about the impact without using sentiment penalty in Section 6.

To calculate the F-score during evaluation, we first determine the median number of pair counts of each user generated from reviews for each POI in training set. We set the median count number as the number of explanation pairs generated using our model individually for each user. According to (Shani and Gunawardana, 2011), they listed ways to calculate the recall rate (shown in equation 3), precision (shown in equation 4), and F-score (shown in equation 5). Then, we following these ways to calculate F-score for our case.

The evaluation method accommodates the error caused by using GloVe vectors during explanation evaluation. The modified F-score

consolidates sentiment penalty requiring a match with a cosine similarity under 0.8 and both opinion words having the same sentiment polarity, which can be written as follows:

$$n_i = \begin{cases} 1, & \text{if } s_i \cdot \overline{s_i} > 0 \text{ and } similarity > 0.8 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

$$recall_{n_i} = \frac{number\ of\ True\ predicted\ pairs}{total\ number\ of\ pairs\ in\ reviews} \quad (3)$$

$$precision_{n_i} = \frac{number\ of\ True\ predicted\ pairs}{total\ number\ of\ pairs\ predicted} \quad (4)$$

$$F_1 = \frac{2 \cdot recall_{n_i} \cdot precision_{n_i}}{recall_{n_i} + precision_{n_i}} \quad (5)$$

Where s_i is the sentiment score of opinion in the pair generated by our model predicted using SentiWordNet (Baccianella et al., 2010), $\overline{s_i}$ is the sentiment score of opinion in the pair extracted from users review.

6 Results Discussion and Analysis

We apply the same dataset to the three baseline methods: KNN, SVD matrix factorization, and NMF matrix factorization, which refer to the algorithms in (Ricci et al., 2015; Luo et al., 2014). The result presents that our proposed integrated neural collaborative filtering method (NCF) outputed the lowest RMSE compared with other methods (shown in Table 2).

NCF	KNN	MF - SVD	MF - NMF
0.1733	0.2555	0.3032	0.3890

Table 2: Rating Prediction Results

Then, we evaluate our model in terms of explainable pairs over this dataset. The baseline model we introduced is to randomly sample pairs extracted from all users' reviews under the predicted POIs which correspondes to human random guess. The random sampling method provides information extracted

from historical reviews of the recommended POI but does not consider users' preferences. By using this as a baseline, we can validate how much improvement our model achieved through incorporating information of users' preferences. To collect unbiased result, we use ten folds cross-validation for each experiment, and the result shows (in Table 3) that NCF with sentiment penalty outcomes an F-score of 0.5088, which is higher than 0.0349 of random sampling. As we mentioned the sentiment penalty in the process of evaluating our explainable results before, it could help to modify the labeling results. Since 'bad service' and 'good service' have 0.952 similarity but they present different sentiment, we add sentiment penalty to label these cases as False. Therefore, after incorporating the sentiment penalty, the F-score drops from 0.5696 to 0.5088 (shown in Table 3), which indicates the deficiency without applying penalty.

NCF	NCF*	Baseline
0.5696	0.5088	0.0349

* represents using sentiment penalty

Table 3: Explanation Generation Results

7 Conclusion

This paper introduces a neural-based explainable recommender system comprised of 2 parts: providing recommendation to users based on rating prediction and integrating the recommendation with explanations that can convince users to visit the proposed recommendation. We deployed a neural collaborative filtering model to predict ratings for the target user. Further, we utilize the user embedding layer which contains information on users' latent preference to generate explanations based on users' historical review data. Then, the neural based rating prediction model is evaluated using Yelp dataset and compared with commonly used collaborative filtering methods. Finally, we introduce cosine similarity based F-score with sentiment penalty to evaluate our explanation generation method. Result shows that our model achieves a better performance in rating pre-

diction than classical collaborative filtering model and a significant improvement over the random sampling baseline in explanation generation.

Limitation and Future Work

The evaluation results of rating and explanation present that our model is useful to solve some problems and could be generalized to other types of POI recommendation such as hotels and tourist attractions. However, there are also some limitations which are discovered during our experiments. We listed these limitations and discussed the ways to improve our model performance in future work.

In opinion-aspects pair extraction, we only explore the aspect-opinion pairs in the form of adjectival modifier and noun using dependency parsing. Current method incorporated excludes part of POI features that are not presented in the form of opinion-aspect pairs, such as 'cafe near the river' and 'restaurant in the city center'. If we consider higher level n-gram (n>2) for extracting opinion-aspect pairs by using dependency parsing method, more noise pairs will also be extracted. Hence, we consider using deep learning model for explanation generation in the future.

The rating prediction phase only considers the rating scores between users and POIs. Since the customer reviews also relate to the rating stars, we will incorporate other information about users such as social links to improve the model performance in future work. Review information may be incorporated by adding another embedding layer which encodes users' reviews and import reviews as input of the layer.

Our model assumes that the users have visiting records before, and continue to use the website which contains user generated content about POIs. Therefore, we will explore the methods for the users who don't have any visited record in browser or website.

Contribution

The authors of this project are Jionghao Lin

and Yiren Liu. Both of us contribute to idea brainstorming, code implementation, and report writing. Jionghao was mainly responsible for preprocessing dataset, model evaluation and generate graphs/figures for the report. Yiren Liu mainly focused on model designing and implementation.

References

Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec*, volume 10, pages 2200–2204.

Marko Balabanovi and Yoav Shoham. 1997. Fab: Content-based, Collaborative Recommendation. *Commun. ACM*, 40(3):66–72.

Jie Bao, Yu Zheng, and Mohamed F. Mokbel. 2012. Location-based and Preference-aware Recommendation Using Sparse Geo-social Networking Data. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 199–208, New York, NY, USA. ACM.

Chaochao Chen, Xiaolin Zheng, Yan Wang, Fuxing Hong, Zhen Lin, and others. 2014. Context-Aware Collaborative Topic Regression with Social Matrix Factorization for Recommender Systems. In *AAAI*, pages 9–15.

Yelp Developers. 2018. Yelp Dataset.

Yaqian Duan, Xinze Wang, Yang Yang, Zi Huang, Ning Xie, and Heng Tao Shen. 2017. POI Popularity Prediction via Hierarchical Fusion of Multiple Social Clues. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, pages 1001–1004, New York, NY, USA. ACM.

Yoav Goldberg and Joakim Nivre. 2012. A dynamic oracle for arc-eager dependency parsing. *Proceedings of COLING 2012*, pages 959–976.

Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, pages 173–182, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee.

Yunfeng Hou, Ning Yang, Yi Wu, and Philip S. Yu. 2018. Explainable recommendation with fusion of aspect information. *World Wide Web*.

Eliyahu Kiperwasser and Yoav Goldberg. 2016. Simple and Accurate Dependency Parsing Using Bidirectional LSTM Feature Representations. *arXiv:1603.04351 [cs]*. ArXiv: 1603.04351.

Ioannis Konstas, Vassilios Stathopoulos, and Joemon M. Jose. 2009. On Social Networks and Collaborative Recommendation. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 195–202, New York, NY, USA. ACM.

Yehuda Koren. 2010. Factor in the Neighbors: Scalable and Accurate Collaborative Filtering. *ACM Trans. Knowl. Discov. Data*, 4(1):1:1–1:24.

Xin Liu, Yong Liu, Karl Aberer, and Chunyan Miao. 2013. Personalized Point-of-interest Recommendation by Mining Users' Preference Transition. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management*, CIKM '13, pages 733–738, New York, NY, USA. ACM.

Yiding Liu, Tuan-Anh Nguyen Pham, Gao Cong, and Quan Yuan. 2017. An Experimental Evaluation of Point-of-interest Recommendation in Location-based Social Networks. *Proc. VLDB Endow.*, 10(10):1010–1021.

Jie Lu, Dianshuang Wu, Mingsong Mao, Wei Wang, and Guangquan Zhang. 2015. Recommender system application developments: A survey. *Decision Support Systems*, 74:12–32.

X. Luo, M. Zhou, Y. Xia, and Q. Zhu. 2014. An Efficient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems. *IEEE Transactions on Industrial Informatics*, 10(2):1273–1284.

Aaron van den Oord, Sander Dieleman, and Benjamin Schrauwen. 2013. Deep content-based music recommendation. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 2643–2651. Curran Associates, Inc.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.

tefan Pero and Tom Horvth. 2013. Opinion-Driven Matrix Factorization for Rating Prediction. In *User Modeling, Adaptation, and Personalization*, Lecture Notes in Computer Science, pages 1–13. Springer Berlin Heidelberg.

Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, UAI '09, pages 452–461, Arlington, Virginia, United States. AUAI Press.

Jasson D. M. Rennie and Nathan Srebro. 2005. Fast Maximum Margin Matrix Factorization for Collaborative Prediction. In *Proceedings of the 22Nd International Conference on Machine Learning*, ICML '05, pages 713–719, New York, NY, USA. ACM.

Francesco Ricci, Lior Rokach, and Bracha Shapira. 2015. Recommender Systems: Introduction and Challenges. In Francesco Ricci, Lior Rokach, and Bracha Shapira, editors, *Recommender Systems Handbook*, pages 1–34. Springer US, Boston, MA.

Guy Shani and Asela Gunawardana. 2011. Evaluating Recommendation Systems. In Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, editors, *Recommender Systems Handbook*, pages 257–297. Springer US, Boston, MA.

Xiaoyuan Su and Taghi M. Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques.

Chong Wang and David M. Blei. 2011. Collaborative Topic Modeling for Recommending Scientific Articles. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pages 448–456, New York, NY, USA. ACM.

Hao Wang, Naiyan Wang, and Dit-Yan Yeung. 2015. Collaborative Deep Learning for Recommender Systems. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, pages 1235–1244, New York, NY, USA. ACM.

Jia-Dong Zhang, Chi-Yin Chow, and Yu Zheng. 2015. ORec: An Opinion-Based Point-of-Interest Recommendation Framework. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, pages 1641–1650, New York, NY, USA. ACM.