

PAPER • OPEN ACCESS

Product recommender system using neural collaborative filtering for marketplace in indonesia

To cite this article: Arief Faizin and Isti Surjandari 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **909** 012072

View the [article online](#) for updates and enhancements.

You may also like

- [Becoming Heritage: A Place-Making Study of Old Neighbourhood Marketplace in Bandung](#)
A S Ekomadyo, A Nurfadillah, A Kartamihardja et al.
- [The seen and unseen: Exploring place identity of market place in Peninsular Malaysia](#)
Raziah Ahmad, Mohamed Ikhwan Nasir Mohamed Anuar, Zainuddin Ab Rahman et al.
- [eSciMart: Web Platform for Scientific Software Marketplace](#)
A P Kryukov and A P Demichev



244th Electrochemical Society Meeting

October 8 – 12, 2023 • Gothenburg, Sweden

50 symposia in electrochemistry & solid state science

Abstract submission deadline:
April 7, 2023

Read the call for
papers &
submit your abstract!

Product recommender system using neural collaborative filtering for marketplace in indonesia

Arief Faizin¹ and Isti Surjandari¹

¹Industrial Engineering Department, Faculty of Engineering, Universitas Indonesia, Kampus UI Depok, Indonesia

E-mail: arief.faizin@ui.ac.id

Abstract. Marketplace has the potential growth in Indonesia indicated by the continued increase in the number of customers. However, the marketplace has some limitations to deliver personalized purchasing experience. Recommender system can support marketplace to overcome that limitations so that customer can find items or services based on their preferences. This study propose to develop product recommender system based on Neural Collaborative Filtering (NCF) algorithm. The product recommender system to be built is using implicit feedback data in the form of customer purchase data. Implicit feedback is reliable data for building recommendation system. The results have shown that NCF achieve the best performance and outperforms over the other collaborative filtering methods.

1. Introduction

Marketplace is the one of business sectors that has the great potential growth in Indonesia. According to definition, marketplace is an online site that contains many retailers to sell their products or services that can be similar to other retailers [1]. Tokopedia, one of the top marketplaces in Indonesia, has had a monthly user of 6.790.000 at Q4 2019 and increased by 45.91% from Q1 2017 [2]. However, that number is still around 2.5% of the total population of Indonesia that is 271.066.400 [3]. In fact, Google predicts that 52% of Southeast Asian e-commerce activities will be dominated by Indonesia in 2025 [4] and we also know that competition among the marketplace in Indonesia is getting tougher. There are Top 10 marketplaces in Indonesia including Shopee Indonesia, Tokopedia, Bukalapak, Lazada Indonesia, Orami, Blibli, JD.id, Sociolla, Zalora Indonesia, and Bhinneka [5].

However, some Indonesian society is still often shop offline. Based on survey data on Indonesian people who shop for groceries, 63% still choose to shop offline and 36% state that convenience is the most important factor regardless of online and offline [6]. Marketplace also has limitations which do not have staff who can help customers like in offline stores. Customers can get complete information and get convenience at purchasing experience through assistance from staff. The marketplace also has difficulty managing large number of products in a product catalogue. These limits can be overcome using features such as search and directory, but that not efficient because requires extra effort from the customer. Currently, marketplace have innovative solutions to overcome that problem using the recommendation system [7].

The recommendation system can recognize customer preferences and give suggestions of products that customers are interested in and tend to buy through the utilization of customer behavior data and



product information data [8]. Research in the field of recommendation system is continuously improved. One of marketplace that has success in implementing a recommendation system is Amazon. Amazon utilize sales data that contains millions of products to identify potential products [9].

The recommender system algorithms are divide into collaborative filtering (CF), demographic filtering, content-based filtering, and hybrid filtering. In this study, we developed recommender system based on CF algorithm. CF is recommendation algorithm based on similarity of user past behaviour [10].

We used CF based on deep learning approach namely Neural Collaborative Filtering (NCF). NCF has been developed and showed the best performance compared to other collaborative filtering methods [11]. Some research has been done using NCF algorithm and shows the best performance such as commercial site recommendation [12] and e-learning recommendation [13]. Therefore, this paper proposed the implementation of NCF for product recommendation using implicit feedback data. We compare NCF with the conventional collaborative filtering method such as weighted matrix factorization (WMF) and bayesian personalized ranking (BPR) to measure the performance of the algorithm, and use some commonly adopted rank evaluation metrics consists of normalized discounted cumulative gain (NDCG), precision@k, and mean average precision (MAP).

2. Related Work

Recommender system generated by user feedback data. The user feedback can be collected either explicitly or implicitly. There are many research of product recommendation system. Research has been done in [14], [15], [16], [17], and [18]. However, That research mostly used explicit feedback based on users preference patterns from customer rating.

User interaction indirectly becomes input of implicit feedback data, so that the user is not disturbed or burdened, in contrast to explicit feedback data. Based on previous research, prediction algorithm that uses implicit feedback generate better prediction quality of user preferences than prediction algorithm that uses explicit ratings [19]. This paper uses implicit data in form of customer purchase data to generate recommender system model.

Next, we compare it with some recommender system algorithms such as WMF and BPR. In the study of Corso and Romani, there are some metrics that uses to evaluate the recommender system algorithm. In this study, the recommender system that will be built is using implicit feedback data, so the appropriate metrics are ranking metrics [20]. In this study we propose to use NDCG, precision@k, and MAP to compare the performance between algorithms.

3. Methods

3.1. Recommender System

The rapid growth of data causes users difficult to find products or services according to their needs and preferences. Recommender system can overcome that problems by suggesting products based on past behaviour of user [21]. There are 3 phases required to build a recommendation model [22] that depicted in Figure 1.

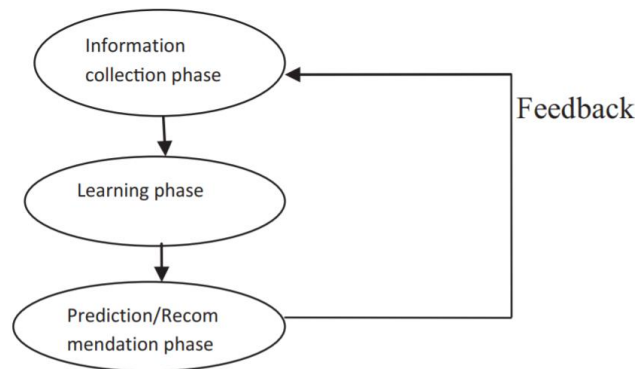


Figure 1. Recommendation phases.

1. Information collection phase is part where personal users information in form of user's attribute, behaviours or content is collected. There are two different input of recommender system, implicit feedback data and explicit feedback data. Explicit feedback data is user feedback that requires effort from user. Implicit feedback data is user feedback that automatically generated by system.
2. Learning phase is when the algorithm apply to learn the user feedback obtained in the previous phase.
3. Prediction/recommendation phase which is where the algorithm give recommendation or prediction result about user preferences.

3.2. Neural Collaborative Filtering

NCF is collaborative filtering algorithm that ensemble Matrix Factorization (MF) and Multi-Layer Perceptron (MLP) that collaborate non-linearity of MLP and linearity of MF and for modelling interaction of user and item. NCF can overcome the non-linearity problem by removing the inner product and replacing it with neural architecture. Figure 2 shows an overview of the NCF architecture [11].

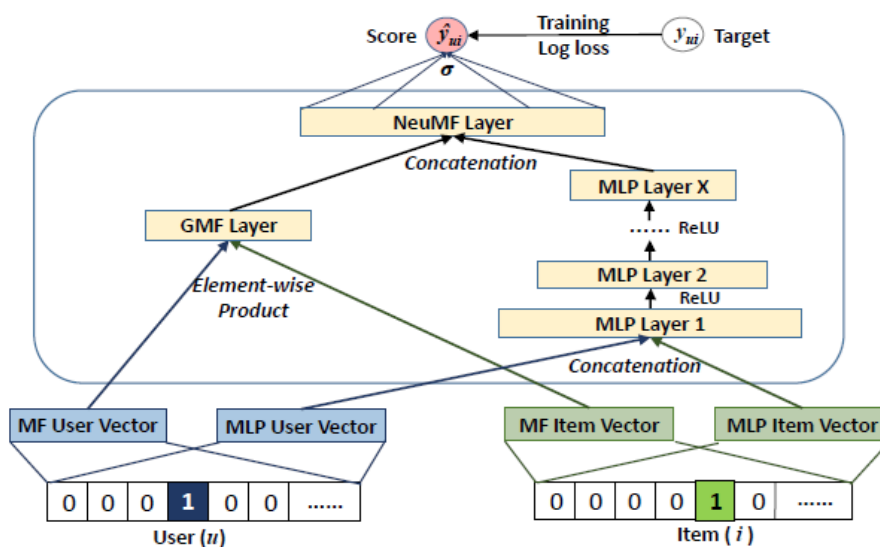


Figure 2. NCF architecture

The input layer consist of item feature vectors v_i^i and user feature vectors v_u^u which contains one hot encoded of items i and users u data. The next step is to get the latent vectors of item i and user u by

transform the sparse vectors v_u^u and v_i^i into dense vector p_u and q_i by the embedding layer. User-item interactions are modeled separately between linear and nonlinear perspectives at the Neural CF layer, then the two previous models are combined and used as the final result. The calculations can be defined as Equation 1-3 [11,12]:

$$\phi^{GMF} = p_u^G \odot q_i^G \quad (1)$$

$$\phi^{MLP} = \alpha_x \left(w_x^T \left(\alpha_{x-1} \left(\dots \alpha_2 \left(w_2^T \begin{bmatrix} p_u^M \\ p_u^M \end{bmatrix} + b_2 \right) \dots \right) \right) + b_x \right) \quad (2)$$

$$\hat{y}_{ui} = \sigma \left(h^T \begin{bmatrix} \phi^{GMF} \\ \phi^{MLP} \end{bmatrix} \right) \quad (3)$$

Item embedding of the MLP and GMF parts are indicated in q_i^G and q_i^M . User embedding of the MLP and GMF are indicated in p_u^G and p_u^M . The output and hidden activation functions described as σ and α_x . The x-th layer's perceptron weight matrix and bias vector described as W_x and b_x . Cross entropy loss for implicit feedback is the NCF model optimization function, which is defined as Equation 4 [11,12].

$$L = - \sum_{(u,i) \in y \cup y^-} y_{ui} \log \hat{y}_{ui} + (1 - y_{ui}) \log (1 - \hat{y}_{ui}) \quad (4)$$

3.3. Evaluation Metrics

There are some popular ranking-based evaluation metrics to measure the performance for each recommender algorithm consists of precision@k, MAP, and NDCG [23], [24].

Precision@K is extension of a conventional precision metric. Precision@K is the top K list prediction items that actually consumed by the user divided by the K list generated by the model for each user in the data. The higher precision@k value indicates that the model generate more accurate predictions. The average of all user precision values indicates the final of precision value. The Precision@K calculation is calculated as shown as follows [23].

$$Precision@K = \frac{|C(K, u)|}{|K|} \quad (5)$$

MAP is very suitable for measuring performance for a ranking recommendation engine. The MAP value calculates how accurate the recommendations are in ranking items. The system will provide relevant advice among the first few recommended products if they have a high MAP score. The system will offer appropriate suggestions among the first few recommended products if they have a high MAP score. The first step in determining the MAP value is to calculate 'precision at the cutoff k' (P(k)) where precision is calculated by considering the list of recommendation subsets in order from first rank to k-th rank. The Average Precision (AP@k) can then be obtained using Equation 6. The MAP@k is calculated as follows [24].

$$AP@k = \frac{1}{\min(m, k)} * \sum_{i=1}^k (P(i) * rel(i)) \quad (6)$$

$$MAP@k = \frac{1}{|U|} * \sum_{u=1}^{|U|} (\frac{1}{m} * \sum_{i=1}^k (P_u(i) * rel_u(i))) \quad (7)$$

A larger MAP indicates better recommendation performance. The weighted rank of the recommended item is considered by NDCG metrics. Items that have a high ranking have a greater weight on the final result. NDCG is calculated by adding up the weighted rankings of items for users. The ranking list on the proposed model shows many items that are suitable if they have high NDCG values. NDCG can be obtained by calculating DCG(u) and Z(u) first. The value Z(u) indicates the user ideal value of DCG(u) which is calculated as the best item rating for the user. Furthermore NDCG calculations can be obtained using Equations 8. and 9.

$$DCG(u) = \sum_{j=1}^n \frac{2^{r(j)} - 1}{\log_2(1 + j)} \quad (8)$$

$$NDCG(u) = \frac{DCG(u)}{Z(u)} \quad (9)$$

4. Experiment

4.1. Data preparation

The data was collected at one of the marketplaces in Indonesia in 2019. The data consists of 11.918 transaction data of implicit feedback in form of customer purchase data. The data has been used consists of customer data, product data, and customer purchase data.

4.2. Implementation

At the data processing stage, the data is divided into 2 parts, 0.70 is used as training data and 0.30 is used as testing data. Then the training data is trained using a neural collaborative filtering algorithm. First, we should set the parameter model on the NCF algorithm including learning rate, batch size, and the number of predictive factors with a value of $k = 10$ (recommended number of top items). The last hidden layer of NCF is as a predictive factors because it determines the model capabilities. We experimented the batch size of 256, learning rate of 0.001 and 4 trials of predictive factors of [8, 16, 32, and 64]. For WMF and BPR methods, the number of predictive factors is obtain from the number of latent factors. The model that is formed from the previous data processing is measured using the NDCG, Precision@10, and MAP metrics.

5. Result and Discussion

The result of evaluation shown in Table 1 below. The results show that the greater the predictive factor, the better the performance of the model.

Table 1. Evaluation Results

Metric	NCF				BPR				WMF			
	Predictive Factors				Predictive Factors				Predictive Factors			
	8	16	32	64	8	16	32	64	8	16	32	64
<i>Prec@</i>	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.00	0.00	0.01	0.01
10	76	63	95	03	01	01	01	01	41	68	15	37
MAP	0.07	0.06	0.08	0.09	0.03	0.03	0.03	0.03	0.00	0.01	0.04	0.05
	79	51	98	96	97	96	97	97	95	79	13	79
NDCG	0.09	0.08	0.10	0.11	0.06	0.06	0.06	0.06	0.01	0.02	0.05	0.06
	69	40	93	94	60	60	60	60	50	66	30	97

Based on the evaluation results of the comparison in Table 1, NCF algorithm has shown the best performance on predictive factor of 64, this is indicated by the highest values of precision@10, MAP, and NDCG.

6. Conclusion

This paper studies the product recommender system for marketplace based on neural collaborative filtering. First, we generate the prediction model using NCF algorithm and then evaluate the performance of NCF method using ranking metrics consist of precision@k, MAP, and NDCG. Finally we compare NCF with BPR and WMF methods based on previous metrics. The results indicate that NCF shows best performance over the other methods.

References

- [1] Nisafani A S, Wibisono A, and Revaldo M H T 2017 Analyzing the effectiveness of public e-marketplaces for selling apparel products in indonesia *Procedia Computer Science* **124** 274–279
- [2] iPrice 2020 Online: <https://iprice.co.id/insights/mapofecommerce/en/>
- [3] BPS 2020 Online: <https://www.bps.go.id/statistictable/2014/02/18/1274/proyeksi-penduduk-menurut-provinsi-2010---2035.html>
- [4] Freischlad N 2016 Online: <https://www.techinasia.com/google-temasek-ecommerce-data-indonesia>
- [5] ECommerceIQ 2017 Online: <https://ecommerceiq.asia/top-ecommerce-sites-indonesia/#1476946678150-d101b526-dcf7>
- [6] Snapcart 2020 Online: <https://snapcart.global/indonesian-consumer-insights-customer-loyalty>
- [7] Hwangbo H, Kim Y S, and Cha K Y 2018 Recommendation system development for fashion retail e-commerce *Electronic Commerce Research and Applications* **28** 94–101
- [8] Choi K, Yoo D, Kim G, and Suh Y 2012 A hybrid online-product recommendation system: combining implicit rating-based collaborative filtering and sequential pattern analysis *Electronic Commerce Research and Application* **11** 309–317
- [9] Jiang B, Jerath K, and Srinivasan K 2011 Firm strategies in the “mid-tail” of platform-based retailing *Marketing Science* **30** 757–775
- [10] Bobadilla J, Ortega F, Hernando A, and Gutiérrez A 2013 Recommender systems survey *Knowledge-Based Systems* **46** 109–132
- [11] He X, Liao L, Zhang H, Nie L, Hu X, and Chua, T-S 2017 Neural collaborative filtering *In WWW* 173–182
- [12] Li N, Guo B, Liu Y, Jing Y, Ouyang Y, Yu Z 2018 Commercial site recommendation based on neural collaborative filtering *UbiComp* 2018 138–141
- [13] Deng X, Li H, and Huangfu F 2018 A trust-aware neural collaborative filtering for e-learning recommendation *Educational Sciences: Theory & Practice* **18**
- [14] Wan M, Ni J, Misra R, and McAuley J, 2020 Addressing marketing bias in product recommendations *Proceedings of the 13th International Conference on Web Search and Data Mining* 618–626
- [15] Dixit V S, and Gupta S 2020 Personalized recommender agent for e-commerce products based on data mining techniques *Advances in Intelligent Systems and Computing* 77–90
- [16] Zhang X, Liu H, Chen X, Zhong J, and Wang D 2020 A novel hybrid deep recommendation system to differentiate user’s preference and item’s attractiveness *Information Sciences* 306–316

- [17] Ding N, Lv J, and Hu L, 2019 Application of improved collaborative filtering algorithm in recommendation of batik products of miao nationality *IOP Conference Series: Materials Science and Engineering* **677**
- [18] Bandyopadhyay S, and Thakur S S 2019 Product prediction and recommendation in e-commerce using collaborative filtering and artificial neural networks: a hybrid approach *Intelligent Computing Paradigm: Recent Trends* 59-67
- [19] Palanivel K, and Sivakumar R 2010 A study on implicit feedback in multicriteria e-commerce recommender system *Journal of Electronic Commerce Research* **11** 140-156
- [20] Del Corso, and G M Romani F 2019 Adaptive nonnegative matrix factorization and measure comparisons for recommender systems *Applied Mathematics Computation* **354** 164–179
- [21] Acilar AM, and Arslan A 2009 A collaborative filtering method based on artificial immune network *Expert System Application* 2009 **36** 8324–8332
- [22] Isinkaye O, Folajimi Y O, and Ojokoh B A 2015 Recommendation systems: principles, methods, and evaluation *Egyptian Informatics Journal* **16** 261–273
- [23] Zhao F, Shen Y, Gui X, and Jin H 2019 SDBPR: social distance-aware bayesian personalized ranking for recommendation *Future Generation Computer Systems* **95** 372–381
- [24] Manning C D, Raghavan P, and Schütze H 2008 *An Introduction to Information Retrieval* (New York: Cambridge University Press)