2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM) | 978-1-6654-7650-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/CENIM56801.2022.10037471

Skincare Recommender System Using Neural Collaborative Filtering with Implicit Rating

Chaira Qalbyassalam, Reza Fuad Rachmadi, Arief Kurniawan
Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
chaira.18072@mhs.its.ac.id, fuad@its.ac.id, arifku@ee.its.ac.id

***Abstract*—Skincare products are essential cosmetics for women, especially in this modern era. Many e-commerce services provide a variety of skincare products in their catalogs. One problem with purchasing skincare products online is that users cannot try the product and depend on other customers' rating reviews. However, rating reviews on a scale of 1 to 5 are considered insufficient to represent product quality, and users need to read review texts written by other users to get more specific information about the quality of the product. This paper investigated NCF (Neural Collaborative Filtering) for skincare recommender systems. Instead of using explicit rating as usually used on standard recommender systems, we adapted the sentiment score as a rating which, in our experiments, proved can improve the classifier's performance. We collected 180,104 rows of data with 11 data attributes and 1,339 skincare products to evaluate our proposed method. Experiments on the dataset show that the proposed NCF with explicit ratings achieved an RMSE of 0.8033, and the NCF with implicit ratings achieved an RMSE of 0.4931.**

***Index Terms*—Neural Collaborative Filtering, Skincare Recommender System, sentiment analysis**

I. Introduction

Skincare is a beauty product useful for caring for the skin and is used to make the physical appearance more attractive. Skincare in Indonesia has grown quite rapidly, as evidenced by data on the market share of the cosmetic industry in Indonesia. Based on Statista data as of July 2020, the largest market of the Indonesian cosmetic industry is the skin care segment, with a market volume of USD 1,673 million in 2019 and will continue to increase. The high skincare market segment makes entrepreneurs in the cosmetic industry increasingly interested in issuing many types of skincare products in Indonesia. With the skincare industry's development and technology, users can buy skincare online anywhere and anytime.

In contrast to making purchases in stores, online purchases that are not made directly can increase the risk of uncertainty in shopping. Therefore, before purchasing products online, users can search for product information in various ways [1]. One way is to take the benefit of reviews from other users regarding the information on a product ranging from advantages to disadvantages [2] to be taken into account in decision making.

When identifying information from reviews, prospective users are not enough to look at the explicit rating, ranging from 1 to 5. Because different users who give the same rating number do not necessarily have the same level of satisfaction because the characteristics of users are different, therefore, potential users need to read the review text to get a more objective assessment. However, the many reviews written by users who have tried this skincare product have another impact. The amount of information that needs to be identified in these thousands of reviews can make it difficult for potential users to make the right decision before buying a skincare product [3]. As a result, the ease of buying skincare products online is reduced, and it can lead to wrong conclusions about a product. A personalized skincare product recommendation system that utilizes reviews written by other users and takes advantage of profiles of prospective users using Deep Learning is demanding to increase user online skincare shopping experience. Neural Collaborative Filtering (NCF) is a strong candidate to be used for the problems. NCF method was the deep learning version of the standard collaborative filtering method [4].

In this paper, we proposed a variant of the NCF method using implicit rating input. We utilize sentiment analysis scores for the implicit rating on user text reviews. Our contributions can be listed as follows.

- We build a skincare recommender dataset using web scrapping methods consisting of 180,104 data with 12 columns (11 data attributes and one user text review).
- We built NCF model with implicit rating input. Experiments show that the NCF model with implicit rating input is superior to the NCF model with explicit rating.
- We also perform experiments using Matrix Factorization (MF) methods for comparison. Experiments show that NCF with implicit rating can perform better than MF.

The rest of the paper is organized as follows. Some related work regarding the recommender system is discussed in section II. Section III described the detail of our proposed NCF with implicit rating methods. The experiment results are discussed in section IV. Lastly, we conclude the experiments in section V.

II. Related Research

A lot of research has been conducted on recommender systems, including that described in [5]–[7]. We discuss those research briefly in this section.

Erlangga and Sutrisno [5] conducted research on recommendation systems for beauty shops using collaborative filtering methods. They use explicit ratings by users to form the collaborative filtering model. The output of the recommender systems is the product with the highest predicted rating based on historical user data. The software feasibility test

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.

results show that the online beauty shop application with the recommendation system for the Collaborative Filtering method has met the appropriate criteria.

Safitri et al. [6] proposed a recommender system for skincare using content-based filtering and prior knowledge. The recommender system was created using an expert system based on several criteria and prior knowledge. The recommendation process is carried out by calculating the similarity value of the content of an item that produces the highest to lowest product rating and also by calculating the minimum support value and minimum confidence value to determine the association rules for an itemset combination. The results of applying the Content-Based Filtering method on a skincare recommendation system with 40 skincare product data obtained the highest rating of 0.447 and a minimum value of support=40%, a minimum value of confidence=40% forming an association rule with a confidence value of 88, 8%.

Adebo [7] conducted similar research to our research for skincare recommender systems using NCF with additional word embedding vector representation. The word embedding vector representation inputs were taken by feeding a word vector to the embedding layer and LSTM. The final NCF input is formed by concatenating the word embedding vector representation with other information embedding vectors. Adebo reports that NCF with additional word embedding vector representation achieved an MSE of 0.3 which results better compared with original NCF methods.

III. System Design

This section explains in detail the design of our proposed method, including how to obtain the primary data, data preprocessing, sentiment analysis, and NCF model.

A. Data Collecting and Preprocessing

Data Collection. Data collection is carried out using web scraping from the Female Daily website by utilizing Selenium and BeautifulSoup libraries. Selenium library was used to perform data scraping that may change dynamically using *JavaScript*. For static contents, BeatifulSoap was used to perform data scraping.

Data Preprocessing. After the data is obtained from the web scraping process, several preprocessing steps are required clean the raw data. Explanation regarding each preprocessing step is described as follows.

1) **Data Cleaning.** Data cleaning is done by removing unnecessary data, discarding empty data, duplicate data, and unnecessary columns.
2) **Data Extraction.** The data extraction process aims to extract the useful information used to perform NCF experiments. We use the tag pattern of the HTML source of the data to extract specific data. The tag pattern is very specific to the webpage and may change in the future. If we collected the data again on the same webpage in the near future, there might have different tag patterns to extract the same specific data.
3) **Data Encoding.** Data encoding aims to convert categorical data obtained into numeric types. We use ordinal encoding for data with few categories and label encoding for data with many categories. Ordinal encoding is a method for converting categorical data by mapping each category into a predetermined numerical value so that when performing Ordinal encoding, a map function will be created as a map reference. Label encoding is a method for converting categorical data into numeric by converting each category into unique numeric data.

NLP Preprocessing. The preprocessing steps were applied only for user review text data. The preprocessing steps are necessary before applying the sentiment analysis algorithm. Therefore, the review text data is preprocessed before sentiment analysis, including text cleaning, case-folding, tokenizing, and filtering.

Sentiment Analysis. The next stage is sentiment analysis. The purpose of sentiment analysis is to get the implicit rating obtained from the sentiment score of each review. Sentiment analysis is carried out using a lexicon-based approach. The lexicon used in this sentiment analysis is obtained from the InSet (Indonesia Sentiment Lexicon). The InSet contains a word column and a weighting column weight for each word. The way sentiment analysis works with a lexicon-based approach is that each review text in the data will be separated into words by tokenizing. Then the polarity of each word will be calculated based on the word score on the lexicon.

B. Modelling

In the experiments, we use NCF as our primary model with an additional MF model for comparison. Inputs of each model we used in the experiments can be described as follows.

Matrix Factorization. This model is used to see the performance of the Collaborative Filtering method without using a neural network approach. The data used as input for this model include user data and product data items.

NCF with explicit rating. The model used is Collaborative Filtering with a neural network approach. The data that will be used as input for this model include user data and temporary item data that are treated as labels in training, namely the explicit rating obtained from the rating score data given by the user to the product.

NCF with implicit rating. This model will use the same method as the previous model, namely Neural Collaborative Filtering. The data used as input is the same, and the difference is the target label used in the training process. In this model, the target label is an implicit or indirect rating obtained from a sentiment analysis score. This model is used to compare the use of implicit rating and explicit rating as the target of model training.

NCF with additional input of skin type and age category. This model is a modification of the previous model. In this model, two inputs are added: age category data and skin type data. This input is added to compare the results when adding input data.

C. Training Processs

Before training, the dataset is divided into two: train data and test data. The division of train data and test data use the 80/20 principle, where 80% of the data is used for train data, and 20% of the data is used for test data. The training

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.

TABLE I
THE DESCRIPTION OF VARIABLE USED

Column	Data Type	Description
user	string	Username
age	string	User Age Category
skintype	string	User Skin Type
product	string	Product Purchased
category	string	Category Product Purchased
duration	string	Usage Duration
recommend	string	Recommend or no
rate	integer	Rating
date	integer	Date
text	string	User Text Review
date_scrape	integer	Date Scraped

TABLE II
SKINTYPE COLUMN SEPARATION RESULTS

Skintype	Skintype Separation
Normal, Medium Light, Neutral	Normal
Combination, Medium Light, Warm	Combination
Normal, Medium Light, Warm	Normal
Normal, Medium Light, Neutral	Normal
Oily, Medium Light, Neutral	Oily
...	...

process is done for ten epochs using a batchsize of 128. All hyperparameters are adjusted during the training process, and the best classifier will be used for testing.

D. Evaluation

The model will produce outputs of skincare product recommendations according to user preferences. Each model will be evaluated using the Root Mean Squared Error (RMSE) so that each model will have five RMSE values from the training results. The Root Mean Squared Error metric (RMSE), one of the most popular metrics, was adopted to assess the accuracy of the ranking prediction model.

IV. RESULTS AND DISCUSSION

A. Data Collection Results

TWe get 180,104 raw data with 36,022 distinct users and 1339 skincare products from the data collection process. We extract 11 data attributes consisting of 7 primary and 4 complementary attributes. Table I shows 11 column types that extracted from web scraping process.

B. Data Preprocessing

As explained before, the data preprocessing step aims to convert raw data into clean data that is ready to be used for training. The results of preprocessing step can be described as follows.

Skintype Column. The skintype column can consist of several combinations, and we pick the first mentioned skinntype because we think it is more important than others. Table II shows the skintype column separation process and the word chosen to represent the skintype.

Data Clearning. Data cleaning is done by removing duplicates and data with *NaN/Null* values. This step is very important to make data consistent and guarantee no data duplication.

TABLE III
SKIN TYPE ENCODING RESULTS

Skintype	Skintype Encoded
Normal	3
Combination	2
Normal	3
Normal	3
Normal	3
Oily	1
...	...

TABLE IV
AGE CATEGORY ENCODING RESULTS

AgeCat	AgeCat Encoded
30 - 34	3
19 - 24	1
19 - 24	1
19 - 24	1
35 - 39	4
...	...

Data Encoding. As mentioned in the previous section, two different encodings were used, ordinal and label encoding. We only encoded the seven primary attributes from eleven attributes, including user, age, skintype, product, category, rate, and text. Ordinal encoding is used to convert the skintype and age attributes due to the small number of unique values. Label encoding is used to change the user, and product attributes because these two fields have a lot of unique values. Table III, IV, V, and VI shows examples of encoding process for skitype, age, user, and product.

NLP Preprocessing. Standard NLP preprocessing is performed on user review data before the lexicon sentiment analysis method. The results of the NLP preprocessing steps can be viewed in Table VII, where the data turns into words and unnecessary words are removed from the data.

C. Sentiment Analysis

The next stage is lexicon-based sentiment analysis. Sentiment analysis is done by equating words in user reviews that have been tokenized and then calculating the number

TABLE V
USER ENCODING RESULTS

Username	User Encoded
v**n	33268
ca**ta	14514
Ik**92	4110
Wu**ov	9839
Ar**ha	1054
...	...

TABLE VI
PRODUCT ENCODING RESULTS

Product	Product Encoded
Hydra Rose Petal Infused Toner	581
Serum Mask Hydra Bomb Lavender	1023
Light Complete Super UV Spot Proof Watery	655
Aqua Essence Sun Shield Serum SPF 50 PA+++	141
Complete No-Stress Physical Sunscreen	316
...	...

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.

TABLE VII
SAMPLE OF NLP PREPROCESSING RESULTS

Preprocess Text	*Text*
Befpre	biasa aja ah :(malah nyesel krna emg gada efek
After	[biasa, aja, ah, nyesel, krna, emg, gada, efek]

TABLE VIII
SAMPLE OF SENTIMENT ANALYSIS RESULTS

Text	**Sentiment Score**
biasa aja ah malah nyesel krna emg gada efek	-3
beneran menghilangkan blemishes mencerahkan	6

of weights per word according to the positive and negative words in the lexicon. The output of this sentiment analysis is the sum of the scores of the words. The results of these scores can determine the polarity of the neutral, positive, or negative reviews. If the score is equal to 0, then the review is considered neutral. If the score is above 0, then the polarity is positive, while if the score is below 0, then the polarity is negative. After the sentiment score is obtained, normalization of the data is carried out such that the value is between zero and one. Table VIII shows some examples of lexicon-based sentiment analysis results. The normalized sentiment score data are combined into data frames to form complete model inputs. The normalization process shows that the most positive sentiment score is 70 (normalize to 1), and the most negative sentiment score is -70 (normalize to -1). Examples of normalization results can be viewed in Table IX.

D. Model Architecture

We use two different NCF models, the NCF model with two inputs and NCF model with additional inputs. As described in [8], [9], the NCF model has proven superior to other collaborative filtering methods. Figure 1 and 2 show the architecture of the NCF model with two inputs and the NCF model with additional inputs.

E. Training Results

The training was carried out on four different models, with the dataset that had been divided in the previous stage using 80999 data trains. This training will be tested later to see the results of its performance. The training process for the Neural Collaborative Filtering model is carried out with *epoch* 10 times and *batch size* 128. Each model will be *training* 5 times. Meanwhile, for Matrix Factorization, a Singular Value Decomposition (SVD) model will be used, providing recommendations to users from the latent features of the user-item matrix.

TABLE IX
NORMALIZATION RESULTS OF SENTIMENT SCORE

Score	**Normalized Score**
-3	0.511450
6	0.580153
13	0.633588
-3	0.511450
-12	0.442748
...	...

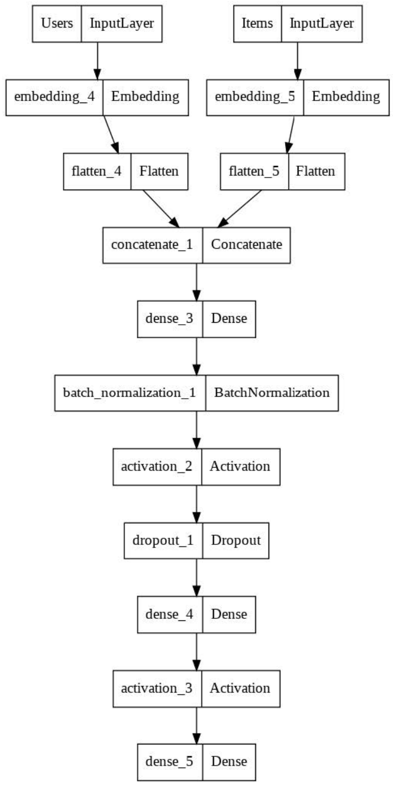

Fig. 1. Model Neural Collaborative Filtering 2 *Input*

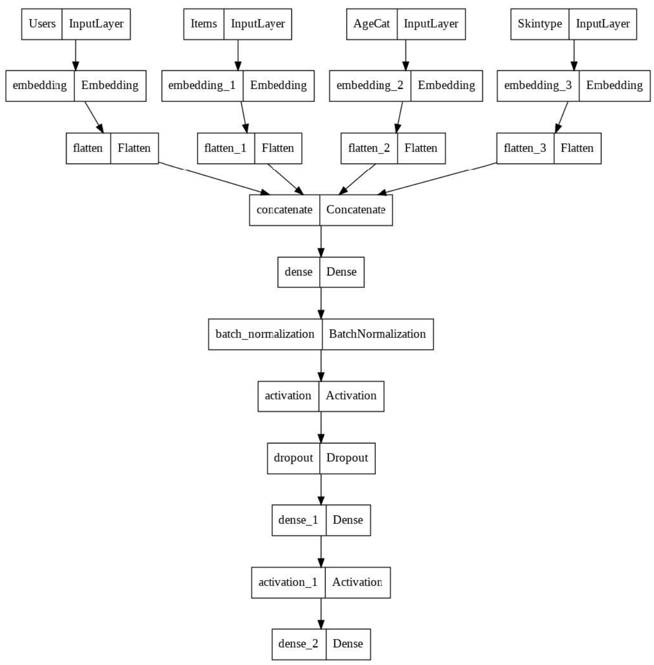

Fig. 2. Model Neural Collaborative Filtering 4 *Input*

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.

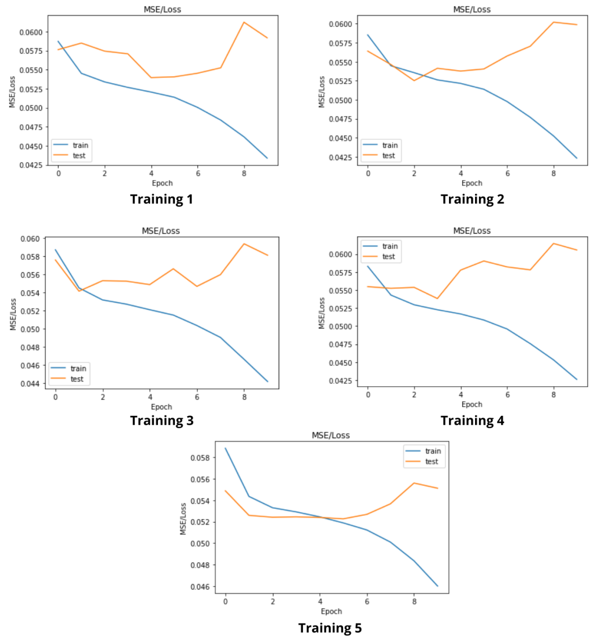

Fig. 3. Neural Collaborative Filtering With Explicit Rating Graph Training Results

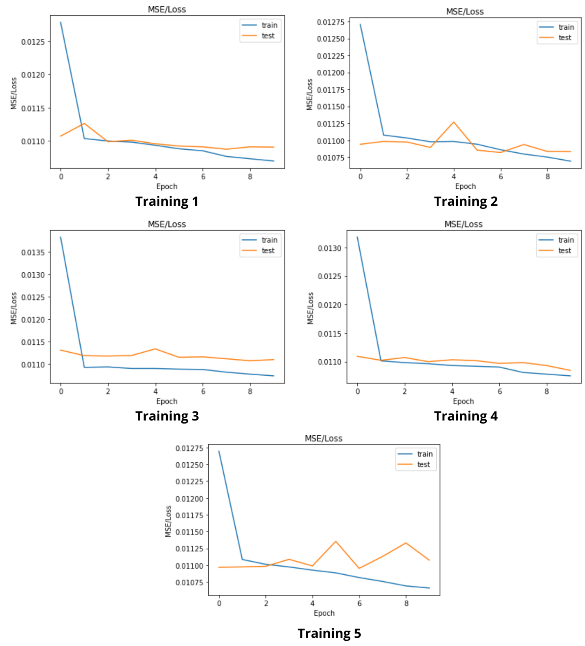

Fig. 4. Training Neural Collaborative Filtering With Implicit Rating Graph Training Results

TABLE X
NEURAL COLLABORATIVE FILTERING WITH EXPLICIT RATING TRAINING RESULTS

	Train Loss	*Val Loss*	RMSE
Training **1**	0.6310	0.6442	0.7944
Training **2**	0.6223	0.6428	0.7889
Training **3**	0.6367	0.6481	0.7979
Training **4**	0.6254	0.6347	0.7908
Training **5**	0.7110	0.7284	0.8432
Average	0.64528	0.65964	0.8033

TABLE XI
NEURAL COLLABORATIVE FILTERING WITH IMPLICIT RATING TRAINING RESULTS

	Train Loss	*Val Loss*	RMSE
Training **1**	0.2378	0.239	0.4876
Training **2**	0.2418	0.2415	0.4917
Training **3**	0.2543	0.2544	0.5043
Training **4**	0.2507	0.2501	0.5007
Training **5**	0.2309	0.2314	0.4805
Average	0.24310	0.24328	0.4931

1) Neural Collaborative Filtering With Explicit Rating: In this Neural Collaborative Filtering model, 2 inputs (x) are given, namely user and item with the target (y) being explicit rating. The training was carried out 5 times. The results of Neural Collaborative Filtering training model with explicit rating from 5 times of training, the average level of loss for training is 0.64528, the average level of loss for validation is 0.65964, and the average RMSE is 0.8033. loss overall training can be seen in Table X. and the graph of 5 times training can be seen in Figure 3.

2) Neural Collaborative Filtering dengan Implicit Rating: In this Neural Collaborative Filtering model, 2 inputs (x) will be given, namely *user* and *item* with the target (y) being *implicit rating*. The training was carried out 5 times. The results of Neural Collaborative Filtering training model with *implicit rating* from 5 times of training obtained an average level of *loss* for *training* of 0.24310 and an average level of *loss* for validation of 0.24328. The overall *loss* level of training can be seen in Table XI, and the graph of the 5 times training can be seen in Figure 4.

3) Neural Collaborative Filtering With Additional Input: Unlike the two previous models, this Neural Collaborative Filtering model will be given an additional two inputs (x) besides user and item, namely skin type and age category. With the same target (y) as the previous model, namely implicit rating.

The training was carried out 5 times. The results of Neural Collaborative Filtering training model with additional inpput from 5 times of training obtained an average level of loss for training of 0.24310 and an average level of loss for validation of 0.24328. The overall loss level of training can be seen in Table XII and the graph of 5 times training can be seen in Figure 5.

4) Matrix Factorization: The next model being trained is the Collaborative Filtering model, namely Matrix Factorization using Singular Value Decomposition. *Rating* used is *explicit rating*. *Training* was performed 5 times, the accuracy of the model was evaluated using *Mean Squared Error* (MSE) and *Root Mean Squared Error* (RMSE). The results

TABLE XII
NEURAL COLLABORATIVE FILTERING WITH ADDITIONAL INPUT TRAINING RESULTS

	Train Loss	*Val Loss*	RMSE
Training **1**	0.2378	0.239	0.4876
Training **2**	0.2418	0.2415	0.4917
Training **3**	0.2543	0.2544	0.5043
Training **4**	0.2507	0.2501	0.5007
Training **5**	0.2309	0.2314	0.4805
Average	0.24310	0.24328	0.4931

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.

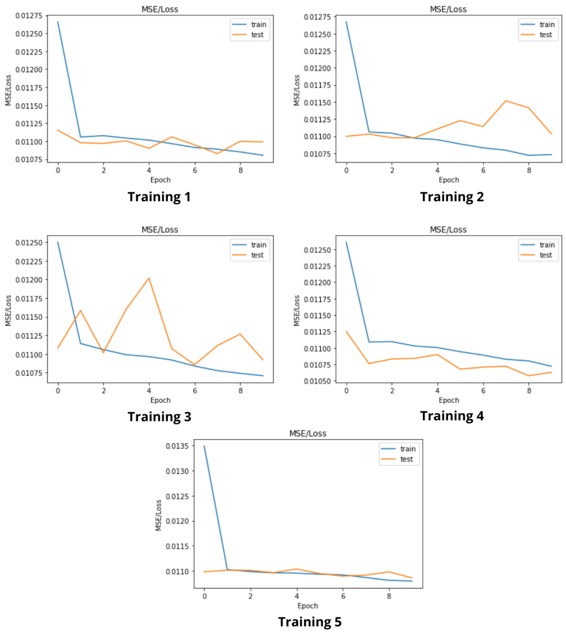

Fig. 5. Neural Collaborative Filtering With Additional Input Graph Training Results

TABLE XIII
MATRIX FACTORIZATION TRAINING RESULTS

	MSE	*RMSE*
***Training* 1**	0.8664	0.9308
***Training* 2**	0.8686	0.9320
***Training* 3**	0.8687	0.9321
***Training* 4**	0.8643	0.9297
***Training* 5**	0.8647	0.9299
Average	0.8665	0.93090

of *training* can be seen in the table XIII.

F. Analysis Results

Based on the training results of the three Neural Collaborative Filtering models and one Collaborative Filtering model, the performance of the models can be compared by comparing the RMSE obtained. Below, the average RMSE for each model can be seen in the XIV table.

The first model, namely NCF with input (x) index user and product index using explicit rate as a target (y) gets poor training results, besides that the graph also shows that the model has overfitting after epoch is above 2, so it gets poor accuracy.

TABLE XIV
RMSE RESULTS

	NCF+*Explicit Rate*	NCF+*Implicit Rate*
MSE	0.49944	**0.2431**
RMSE	0.8033	**0.4931**

TABLE XV
RMSE RESULTS

	NCF+*more input*	MF
MSE	0.24944	0.8665
RMSE	0.4994	0.9309

Then the second is NCF with input (x) index user, and product index that uses implicit rate as a target (y) gets pretty good results and is the best among the four models with an RMSE of 0.4931 where it can be said that the RMSE value between 0.2 and 0.5 indicates that the relative model can predict the data accurately.

The third model is NCF with the addition of input, this model has input (x) index user, product index, skin type, and age category, where *implicit rate* is used as targets (y). This third model gets results similar to the previous model, so it can be said that it is good too but relatively smaller.

Then the last model is a Collaborative Filtering model, namely Matrix Factorization. The results of this last model training are not good. This is in accordance with the [8] paper which concludes that the NCF model has better performance than the CF model.

Based on the overall results, the NCF model with *implicit rating* has the best performance among all the *training* models. This validates that the *deep learning* model performs better in developing the *skicnare* product recommendation system.

V. CONCLUSION

We have presented a skincare recommender system using Neural Collaborative Filtering (NCF) with implicit rating. The implicit rating was formed using lexicon-based sentiment analysis method on user review data. Our experimets show that NCF with implicit rating can achieved an RMSE of 0.4931 which superior to NCF with explicit rating and Matrix Factorization method. Further investigation on other attributes may increase the classifier performance.

REFERENCES

[1] Z. Mo, Y.-F. Li, P. Fan *et al.*, "Effect of online reviews on consumer purchase behavior," *Journal of Service Science and Management*, vol. 8, no. 03, p. 419, 2015.
[2] F. Latief and N. Ayustira, "Pengaruh online costumer review dan customer rating terhadap keputusan pembelian produk kosmetik di sociolla," *Jurnal Mirai Management*, vol. 5, no. 3, pp. 139–154, 2020.
[3] G. Lee, "A content-based skincare product recommendation system," 2020. [Online]. Available: https://portfolios.cs.earlham.edu/wp-content/uploads/2020/05/Gyeongeun_Lee_Paper.pdf
[4] D. C. G. Putri, J.-S. Leu, and P. Seda, "Design of an unsupervised machine learning-based movie recommender system," *Symmetry*, vol. 12, no. 2, p. 185, 2020.
[5] E. Erlangga and H. Sutrisno, "Sistem rekomendasi beauty shop berbasis collaborative filtering," *EXPERT: Jurnal Manajemen Sistem Informasi Dan Teknologi*, vol. 10, no. 2, pp. 47–52, 2020.
[6] D. A. N. Safitri, R. Halilintar, and L. S. Wahyuniar, "Sistem rekomendasi skincare menggunakan metode content-based filtering dan algoritma apriori," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, vol. 5, no. 2, 2021, pp. 242–248.
[7] A. Adebo, "A natural language processing approach to a skincare recommendation engine," Ph.D. dissertation, MSc Thesis, School of Computing National College of Ireland, 2020.
[8] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proceedings of the 26th international conference on world wide web*, 2017, pp. 173–182.
[9] S. Rendle, W. Krichene, L. Zhang, and J. Anderson, "Neural collaborative filtering vs. matrix factorization revisited," in *Fourteenth ACM conference on recommender systems*, 2020, pp. 240–248.

Authorized licensed use limited to: Zhejiang University. Downloaded on November 11,2024 at 11:33:47 UTC from IEEE Xplore. Restrictions apply.