Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HYBRIDIZATION APPROACH TO ELIMINATE SPARSE DATA BASED ON NONNEGATIVE MATRIX FACTORIZATION & DEEP LEARNING

[1,2]HANAFI, [3]NANNA SURYANA, [4]ABD SAMAD BIN HASAN BASARI.

[1]Department of Computer Science, University Amikom of Yogyakarta, Sleman, Special Region of Yogyakarta, Indonesia
[2,3,4]Faculty of Information and Communication Technology, University Teknikal Malaysia Melaka, Melaka, Malaysia.
Emails: [1]hanafi@amikom.ac.id, hanafi@te.net.id, [3]nsuryana@utem.edu.my, [4]abdsamad@utem.edu.my

ABSTRACT

E-commerce company delivered product information to customer or customer candidate trough web portal. There is basic mechanism which a system has belong responsible to calculate and predict information that suitable to customers or customers candidate interested namely recommender system. Most successful approach to calculate customer/user interest are based on collaborative filtering. This approach relies on rating from customers to products or items as a basic approach aims calculate similarity of users responds about items to produce recommendation. In fact, just a little number of customers who giving the rating approximately less than 1 percent from all customer population in datasets. It's a reason of rising sparse data. In this research used 2 technical approach to deal with sparse data consist Non-Negative Matrix factorization to reduce dimensional reduction and involve deep learning to compute latent factor in a part of users, item and rating. This research consider dataset from MovieLens, many researchers believe to conduct experiment their approach algorithm to increase better performance. Final experiment we used RME (Root Mean Error) and RMSE (Root Mean Square Error) to measure accuracy of result experiment and according the result, our approach has obtained good result to reduce missing value.

Keywords: *Non-Negative, E-Commerce, Recommender System, Collaborative Filtering, Matrix Factorization, Deep Learning*

1. INTRODUCTION

Collaborative filtering is one of successful approach to produce product recommendation in case of online commerce (popular called e-commerce). Comparing to another approach, collaborative filtering having some major benefit as follow; Accuracy, relevant, serendipity and diversity to produce product recommendation [1]. This is one of reason why many e-commerce company adopted this approach. Useful implementation of recommender system has impact of increasing value of marketing target. The growth of e-commerce profit is influenced by the service quality of e-commerce company. The development of recommender system is aimed at improving the service satisfaction of e-commerce [2]. Began to develop in early 90's, Collaborative filtering bringing big problem issue, due collaborative filtering rely on explicit feedback from customer as response of level satisfied service, the problem often rises in this method caused the minimum explicit feedback from customer in the term of rating. According the evidence based on public dataset, only less than 1 percent who have giving the rating to e-commerce product [3]. This problem popular namely sparse data and in extreme sparse data condition called cold start problem. Sparse data and cold start problem that cannot be eliminated, causes inaccurate of product recommendation result. Even, there is no recommendation emerged from recommendation system. Figure 1 example of collaborative filtering table face sparse data.

Items

Users				
	5	?	?	3
	4	2	?	2
	?	?	5	1
	3	2	?	?

Figure 1. Table Matrix Collaborative Filtering

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Traditional collaborative filtering using statistical approach (popular called memory based) to develop recommender system such as cosine similarity, spearman rank, and etc. Thus, statistical approach has characteristic easy to implemented/simplicity, effective. Although, this approach having benefit in simplicity and effectivity, there are shortcoming heritage in scalability, sparse data, cold start, minimum accuracy. So, become another choice to dealing with this problem toward model based. To make a deal with several shortcomings in collaborative filtering, model based involve machine learning and data mining approach. Mathematical approach such as matrix factorization variant have been proposed by researcher for instance SVD (singular value decomposition)[1], Non Negative Matrix Factorization (NMF) [4], semi non negative matrix factorization [5]. Many researcher proposed for the same start of the art with different angle approach by using machine learning for example [6] enhance deep learning to extract feature content aims eliminate cold start items. [7] Involve hybrid approach between deep learning and matrix factorization to handle cold start problem. Following our best knowledge, the major problem of collaborative filtering is cold start in which consist 2 types as follow items cold start problem and user cold start problem. User cold start rise when new user that have no record activity have coming in the system, also item cold start rises when new item has coming in the system too [8].

This paper raised recommender system model which based on collaborative filtering combined with matrix factorization and deep learning approaches. Adopted [9] in collaborative filtering, the recommendation process can be inaccurate due to several problems [10], [11]. There are cold start, sparsity and scalability problems that are common measurable error in the system. The errors can be measured with RMSE. However, the RMSE is considered incapable of explaining the geographical structure of the neural network formed from the dataset of the recommender system [12]. Therefore, RMSE needs to be modified in order to overcome cold start, sparsity and scalability problems. To that end, we added a nonnegative matrix factorization (NMF) to reduce the error rate [4]. In addition, as the dataset gets larger, the error rate also increased due to scalability issues. In order the error can be reduced, it need to add another parameter. To achieve the goal, this paper will observe and examine the dataset formation to obtain the geometric structure through deep learning approach. We also use training repetition and measure the degree of stability of the structure by performing validation tests. There are other parameters which measured, e.g., target training, training outputs, validation targets, validation outputs, test targets, test outputs errors level, response time. We proposed a model that can reduce the error rating.

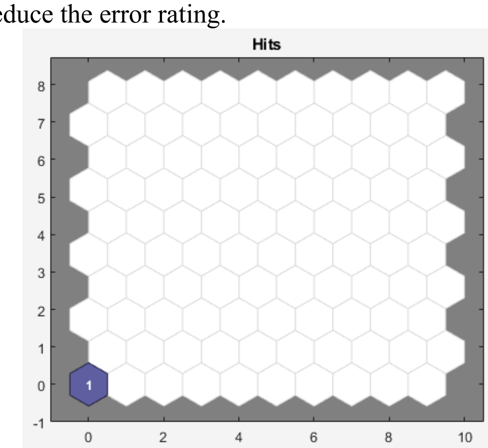

Figure 2. Neural Network With Sparsity Issue

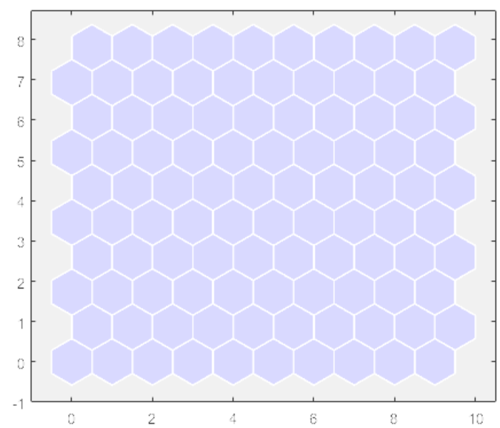
Figure 3. Neural Network With Full Value Rating

The scope of this study is to expand the collaborative filtering (CF) system by adding nonnegative matrix factorization (NMF) and deep learning (DL) approach [13]. We propose a novelty of combination of three approaches into an integration with advanced feed neural network using three steps of training, validation, testing, and all evaluation. It is done by activating cosine similarity between the vectors into the neural network algorithm. In the first phase, we perform collaborative filtering with nonnegative matrix factorization approach [14]. In the second stage we measure the impact of feature representation calculated by the quadratic polynomial regression formula to obtain more accurate latent feature by improving Item-average Clustering. Next, we classified a collaborative filtering algorithm based

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

on in-depth analysis which computed by the non-linearity formula of user items to produce a hybrid collaborative system. The use of neural networks to address implicit feedback issues is calculated by additional information attributes, such as item descriptions and in-depth content text to bridge the semantic gap in the Movielens' movie dataset. We also propose a new model by training deep neural networks to improve the prediction accuracy by using Mean Squared Error (MSE) estimation to represent the structure of neural network which tested in this study.

We propose the new model by developing techniques based on neural networks to overcome the problem of implicit feedback to obtain hybrid filter model with deep structure to solve cold start and sparsity problem in item-users rating matrix by modifying the rating matrix to predict user preferences to the item matrix.

2. LITERATURE REVIEW

Many researchers have been totally tried to address several problems regarding sparse data. Reference [15] consideration based on location use to eliminate sparse data by calculate similarity of location based. According study [16] strategy to handle sparse data consist 3 approach as follow; 1. enhance side information, 2. improve mathematical approach and 3. Conduct hybridization and enhance machine learning approach. On this table 1 in below shown the research result in equal research field.

Author [17] employed deep learning to handle items cold start problem, its deferent with our approach, the author involving side information in the term review feedback from customers. In this research, author classified the cold start problem into 3 kind class and finalized step by step from non-cold start to extreme cold start, then they calculate similarity between result of non-cold start and cold start item with nearest neighbor.

Another author exploits of deep learning [18] involving tags aware, in this research, author use stack denoising auto encoder (SDAE) to make a recommender system more robust over cold start by exploiting users tags. They addition tags layer in deep stack denoising.

Usage of matrix factorization on collaborative filtering cannot be effective face extreme cold start, so the result of recommendation is inaccurate. It is a background author [19] proposed a method by combining probabilistic matrix factorization with marginalized denoising stacked auto-encoders.

Author [20] propose in this paper a hierarchical Bayesian model called collaborative deep learning (CDL). which tightly couples a Bayesian formulation of the stacked denoising autoencoders and probabilistic matrix factorization.

3. OUR PROPOSED MODEL

3.1. Collaborative Filtering Approach

Generally collaborative filtering system uses input data in the form of a database that records information about the user's taste for predicting a topic or new products that may be favored by active users [21]. The input dataset for the collaborative filtering process is a 2-dimensional matrix with products as columns, user names as rows, and their intersection results as the rating ratings given by a given user against the designated product (user-item matrix). This form of dataset is better known as transactional matrix. Collaborative filtering is a system capable of rewarding active users about a particular item or product that might interest them [22].

In an effort to develop this product recommendations, Collaborative filtering uses collaboration of other user information that has similar tastes to the active user [15]. So, the main task of collaborative filtering is to look for a group of users with similar or similar tastes [16]. The usual CF formula uses Pearson correlation (equation 1).

$$r = \frac{n(\sum(xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2}}$$

(equation 1)

To compute r, it is assumed that there is an active user matrix as a reference for predicting an active user's appetite for a particular topic or product. Collaborative filtering as a recommendation system basically consisted of three important stages with the first step is to get profile of each user who have rated any item, that is by getting the rating value of the items that exist [23] [24]. The rating value of the user can be a number with certain conditions, for example item with rate number 5 indicates a very like and will continue to fall to the number 1 indicating a very dislike value, so among the numbers can be entered. For that use the formula $p_{u,i}$ (equation 2), and we applied the rule of relation from u to i.

$$p_{u,i} = \frac{n(\sum(u,i) - (\sum u)(\sum i))}{\sqrt{[n\sum u^2 - (\sum u)^2][n\sum i^2 - (\sum i)^2}}$$

(equation 2)

Assuming that the number n increases continuously in real time, it is assumed that the prediction rate also changes following rules of cosine similarities $S_{u,i}$ At this stage, once each user has obtained user profile that the system will compare the profile of active user with all the existing user profiles and make measurements of the similarity / correlation level. At this stage there will be an error difference due to computer limitations to measure the level of similarity between users so that the formula must be modified to be $S_{(u,i)}$ (equation 3).

$$S_{(u,i)} = \frac{\sum_{i \in I_u \bigcup I_v} (r_{u,v} - \hat{r}_v)(r_{u,v} - \hat{r}_v)}{\sqrt{\sum_{i \in I_u \bigcup I_v} (r_{u,v} - \hat{r}_v)(r_{u,v} - \hat{r}_v)^2} \sqrt{\sum_{i \in I_u \bigcup I_v} (r_{u,v} - \hat{r}_v)(r_{u,v} - \hat{r}_v)^2}}$$

(equation 3)

By modifying the formula into the decomposition of the matrix U, S, V a new formula is produced in the form of a singular value decomposition (equation 4).

$$P_{u,i} = r + u\sqrt{S_{u,i}^T(i)}.\sqrt{S_{u,i}}.V_{u,i}^T \text{....(equation 4)}$$

Singular value decomposition aims to connect the singular value of the rating matrix on the user rating matrix to the matrix in the descending order so that the smallest number will be above the entry into the priority list [25]. SVD helps the system to find a more accurate correlation value based on a group of users that will act as an advisor to active users [26]. The selection of the advisor is based on a high correlation value between active users and other users and uses the information they have to make recommendations for the active user. SVD works by running a singular value decomposition of matrix rating with the rule that $R = U * S * V'$ [27] Where U is the user matrix, S is the feature matrix U and V is the matrix of cosine similarities dot products between the cluster customer and the cluster item that has been rated by the users. Although this algorithm can work well in some cases, this algorithm still has its limitations. Generally, SVD is ineffective when there is overlap between user profiles [28]. In addition, the number of items that can be selected by the user is also limited which bring overlap issue between items to be rated is very little. In this case the calculation of SVD cannot be a guarantee in measuring the degree of similarity.

Another limitation of this algorithm is the level of efficiency in the preparation of recommendations for active users. As the database becomes larger, on-line calculations for a group of users become inefficient (always doing database readings). Customers are generally impatient to wait for results from recommendations. To overcome this problem, the SVD algorithm must be modified using nonnegative matrix factorization techniques although sometimes the method may compromise the accuracy factor of the recommendations [29]. However, this can be mend further by using a deep learning approach that will be explained below.

3.2. Nonnegative Matrix Factorization (NMF)

The new variable calculation step is done by multiplying the member of the matrix P by the matrix member Q into the matrix $P'Q'$. Thus, a new matrix $P'Q'$ (equation 6) is obtained [30].

$$(P'Q') = \arg\min_{(P^*,Q^*)} RMSE \text{ ....(6)}$$

Where P' and Q' are matrix P and matrix Q after containing matrix U, S, V. Since the matrix component $P'Q'$ still contains error e, For that we will integrate into the form $e'_{u,i}$ (equation 5).

$$e'_{u,i} = \frac{1}{2}(e'_{PQ}) \text{ so that } E_{error} = \sum e'_{u,i} = E'_{u,i}$$

(equation 5)

Non-negative matrix factorization works by searching for network weights θ and latent features (U,V) as the error function (rating loss) $F(U,V)$ following the general form of NMF function denoted by (equation 6) [31].

$$F(U,V) = \sum_{(n,m) \in j} (X_{(n,m)} - \hat{X}_{n,m})^2 -$$

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\sum_n |U_n|^2 + \sum_p |S_P|^2 + \sum_Q |V_Q|^2 + E'_{u,i}$$

(equation 6)

Where the E error matrix will be analyzed by deep learning of neural network framework with the number of training n in the next section.

3.3. Deep Learning Neural Network

Once the factorization matrix is obtained the function $F(U,V)$ as a function of loss rating based on the weight θ, latent features of matrix U and latent features of matrix V. While the matrix S turns into S' which is combination of all similarity cluster u to cluster i. The result of the factorization matrix produces a 10x10 layer as a SOM layer with the first stage output reaching 100 elements. This step continues to be repeated and continues to apply in subsequent processes. The next step is to obtain the weight values θ and the sum n of these two matrices. The last step is to get the SOM layer 10x10 into layer 100+w where obtained the output value 100 layers. If the subroutine layers are implemented entirely to each group, the results will be obtained as shown in the results figure (3, 4, 5)). To perform the process in the training phase as in figure 3, 4, 5, the time complexity is not predetermined. This is an advantage to affect the process of calculating the training between layers in the whole group in the dataset under test.

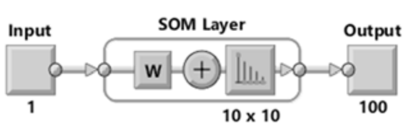

Figure 4. The First Step Of Deep Learning

Next, we will model $F(U,V)$ as a combination of all elemental functions of S,U,V with weights θ and number n. Thus, our model is a combination of hybrid function $F(U,V)_\theta$ and function $F(U,V)_n$. Where is the first function representing collaborative filtering model and the second function represents the deep learning approach with user to item matrix relationship [20].

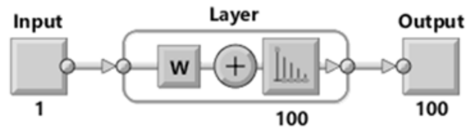

Figure 5. Step two deep learning

We can rewrite the relationship of the hybrid function as function $F(U,V)_\theta$ and function $F(U,V)_n$ the first and second functions are modified to form regularization factor λ_1 and λ_2 with notation *K* (equation 9).

$$K = \sum_{u,v} (p_{i,j} - \hat{p}_{i,j})^2 + \lambda_1 |\theta|^2 + \lambda_2 |n|^2$$

(equation 7)

The end result of our hybrid model becomes as shown below.

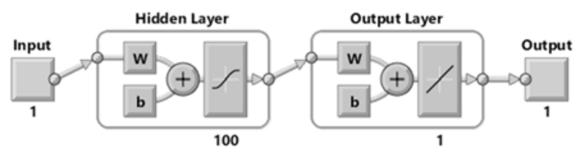

Figure 6. The Third Step Of Deep Learning

3.4. Use Data Sets

The dataset is taken from MovieLens (https://grouplens.org/datasts/movielens/). The dataset has been cleared and filtered to search for users who rated less than 10 with 100k counts with 786 users and 1429 movies. The dataset is then divided into ten clusters for training 10 times. The dataset is considered real, that is, the rating data actually provided by the user on a recommendation system [6]. In this system used MovieLens dataset. MovieLens is a recommendation system developed by GroupLens. MovieLens datasets have been widely used for research purposes related to the field of recommendation systems. This dataset contains the rating data provided by MovieLens users to various films genres where the ratings range from 1 to 5. This program does not use the whole of the MovieLens dataset, but only takes some of that dataset of 100,000 elements that includes 3952 movies and 2000 users. MovieLens provides some of its datasets to the public, for educational and research purposes.

3.5. Measure Metric Mean Squared Error (MSE)

The evaluation or quality measurement of our model is done using Mean Squared Error

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(MSE). MSE is a measure of recommendation deviation from true user-specified rating value. If $p_{u,i}$ the whole value of the rating prediction matrix given u user on item i, and $R_{u,i}$ the actual rating value, then, MSE explains how the structure of a neural network contains an average error rate, whereas RMSE aims to find the final value of an error from the system[9],[10]. To change the RMSE into MSE, it applied nonnegative matrix factorization (NMF) [32].

The squaring form aims to eliminate the negative sign. The squaring form also aims to increase the weight to the larger differences. MSE is also called the average of a set of errors. Mean Squared Error aims to calculate the degree of error mapping the matrix users with a certain value to the matrix items. MSE is a natural form of root mean squared error (equation 7) so that the MSE equation used in this study is not a RMSE form [33]. The general formula of MSE is given in equation 8.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(p_{u,i} - s_{u,i})^2}{n}} \quad \text{(equation 7)}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{N}(p_{u,i} - S_{u,i})^2 \quad \text{(equation 8)}$$

4. EXPERIMENTAL RESULTS

4.1. Neural Network Structure

At this stage, we test the dataset form into the neural network structure and obtain three clusters of red, black, and yellow clusters. The red-colored structure shows irreparable parts as the number of items that cannot be recommended and contains permanent errors. While the yellow and black structure is two clusters that have been rated by u user. The smaller the value of MSE, the more accurate the system in providing recommendations. We do an error measurement using MSE with the rating scale given at intervals [1,10] to determine the overall structure of the system.

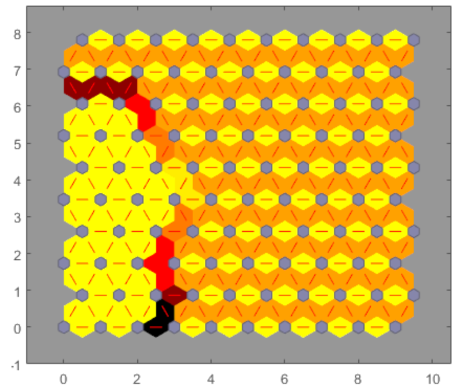

Figure 7. Before Implemented NMF and DL

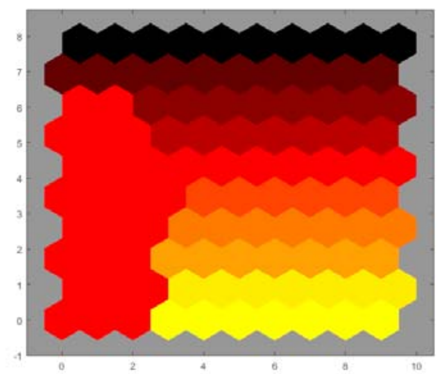

Figure 8. After Implemented NMF and DL

Analysis of test results on the user-item based collaborative filtering model with two different predictive formulas has obtained result that user-item based collaborative filtering method using prediction formula has a rough shape and has a predictive quality that is still rough [23]. This rough rating prediction results in inaccurate recommendation quality as well. The prediction formula $p_{u,i}$ decreases the average MSE value by 0.29% from the average MSE value generated by the RMSE prediction formula at 10% discharge rate. While at the discharge rate of 90%, the MSE prediction formula can only decrease the average MSE value by 3% of the average MSE value generated by the MSE prediction formula. The greater the rate of discharge rating, the greater the value of validation output. This suggests that the method of user-item based collaborative filtering has decreased the quality of prediction when handling data with large sparsity [34]. This can be seen from the average decrease in predictive error beyond the range generated by the two formulas. On the discharge rate of 50% to 80%, the MSE formula can reduce the average error number of predictions outside the range generated by the MSE prediction

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

formula. However, at the rate of discharge of 90%, the average number of predictions outside the range of numbers is the same for both of these predictive formulas. Based on the above analysis, the method of user-item based collaborative filtering with the prediction formula $p_{u,i}$ becomes better when using MSE than RMSE.

4.2. Reduce Errors Output NMF

Test results after Nonnegative Matrix Factorization (NMF) applied have shown a comparison values as an increase in accuracy marked by a decreased error. In this step, the parameters tested and changed are: target training, training outputs, validation targets, validation outputs, test targets, test outputs errors level, response time. It appears that the error rate will be minimum at the interval between 50-60 with the lowest close to 2. While the highest error value is achieved at positions 2.6, 3 and 2.5. using the test was performed on the matrix user-item rating data sample consisting of 2667 users and 1423 items with the number of rating cells filled with 82 ratings with an average sparsity level of 0.992. The data used for testing is only the data whose rated value is filled, or the value is not equal to zero.

The tests were conducted 30 times with randomly discharged ratings ranging from 10%, 20%, to 90%. After 30 tests, the collected data is calculated by the average MSE and the number of rating predictions whose value is outside the range for each level of rating discharge. The following chart shows the average comparison of MSE values before and after applied NMF and the average number of errors both inside and outside the range. The result test shown in figure 9 and figure 10 in below.

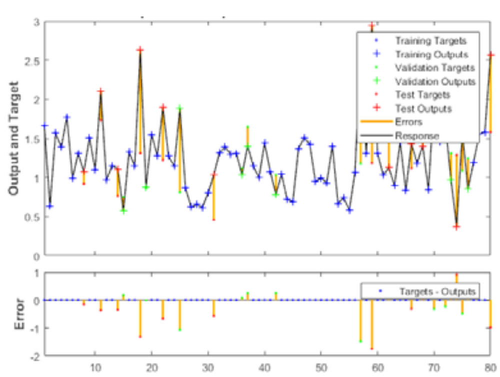

Figure 9. Before Deep Learning

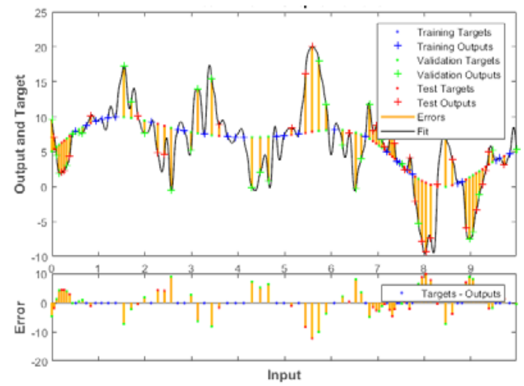

Figure 10. After Deep Learning

Note: Figure 9 The error rate before deep learning is applied, the modified parameters: target training, training outputs, validation targets, validation outputs, test targets, test outputs errors level, response time; 10 the error rate after deep learning is applied, the modified parameters: target training, training outputs, validation targets, validation outputs, test targets, test outputs errors level, response time.

After the collaborative filtering (CF) combined with the NMF is applied, the test results shown in figure 11 and figure 12. showed a reduced result of the error rate. We the tested and modified related parameters e.g., training target, training outputs, validation targets, validation outputs, test targets, test outputs errors level, response time. It appears that the error rate will be minimum at the interval between 50-60 with the lowest close to 2. While the highest error value is achieved at positions 2.6, 3 and 2.5.

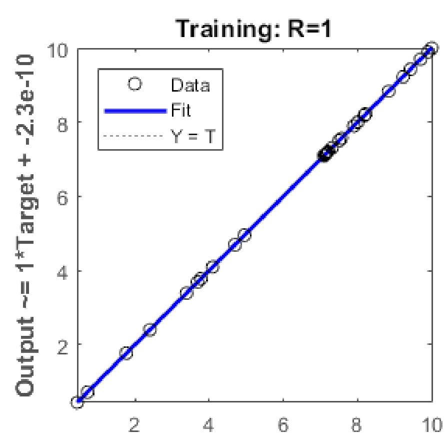

Figure 11. CF+NMF On Training Results

Journal of Theoretical and Applied Information Technology
31st July 2018. Vol.96. No 14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

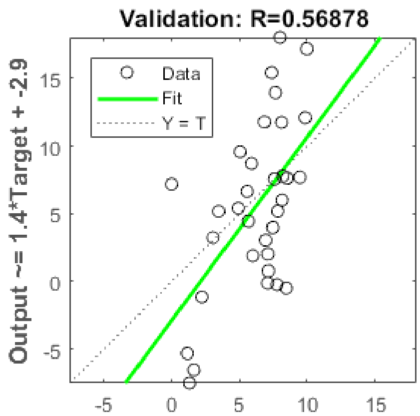

Figure 12. CF+NMF on validation results

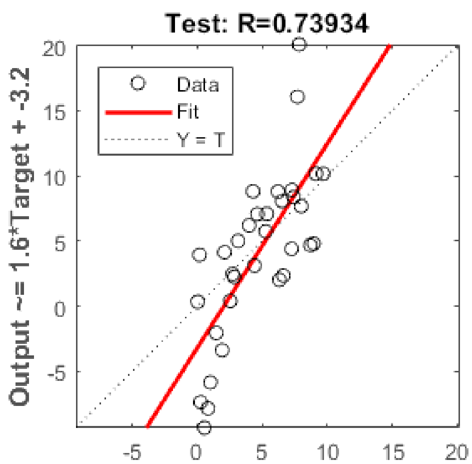

Figure 13. CF+NMF On Testing Results

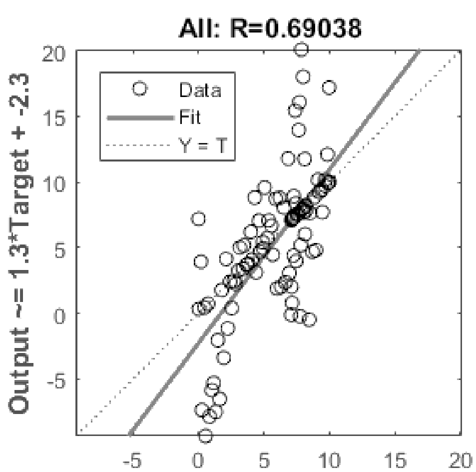

Figure 14. CF+NMF On All Results

4.3. Aggregation Function with Deep Learning (DL) Approach

The approach is used to estimate the overall system stability especially in providing similarity rating. Our model, it used aggregation function approach of Deep Learning following equation (9). This idea departed from the assumption that a multi-criteria rating system represents a user preference for several different components. Thus, the overall rating of an item is an aggregation function of the rating of all criteria, which can be written with the equation. In other words, the Deep Learning (DL) aggregation approach is based on the assumption of relationship between the overall rating and the multi criteria rating [35]. For example, in the recommender system for movie selection that places the story line criteria as a priority. So, a movie that has a high storyline rating will surely favor the user clusters and significantly affect the overall rating. The aggregation function approach is done by three stages, (a) predict the rating of each criterion, (b) estimate the overall rating relation with the multi criteria rating and (c) predict the overall rating.

In the first stage, it predicts the rating of each criterion. It begins by decomposing multi criteria into a single criterion. This means that multi-criteria problems have been transformed into classical collaborative filtering problems of k. Furthermore, a rating cluster must be predicted using similarity approaches such as equation formulas (7), (8) and (9). The second stage is to estimate the relation between the overall rating and the multi criteria rating into the neural network structure. The last step is to predict the overall rating value of $s_{u,i}$ and $p_{u,i}$ directly by using the multi-criteria rating value functions. It generated function $F(U,V)$ which estimated in $CF+NMF+DL$ equation. The experimental results are then performed on the training, validation, testing and all evaluation as in Figure on 15, 16, 17, 18

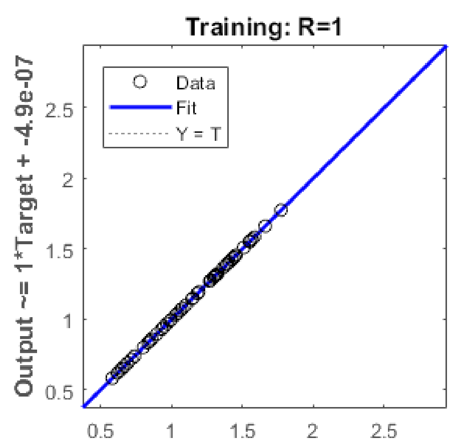

Figure 15. CF+NMF+DL On Training Results

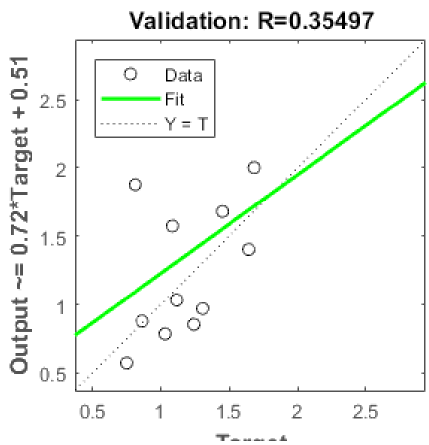

Figure 16. CF+NMF+DL On Validation Results

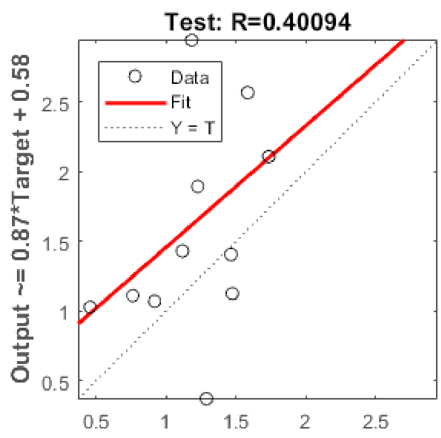

Figure 17. CF+NMF+DL On Testing Results

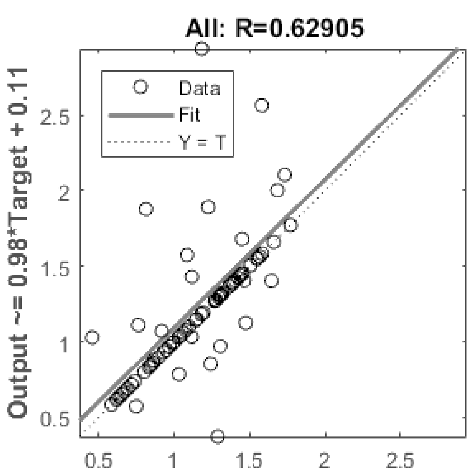

Figure 18. CF+NMF+DL On All Results

Finally, our Approach have benefit when compared with other approach, they have proven to eliminate missing value due cold start problem, but also it has shortcoming with another too. According our best knowledge, involving just mathematical approach deal to completion rating matrix, however they often losing in contextual aware. Then influence the recommendation result unsuitable.

Collaborative filtering is the most successful recommendation approach to predict the user want. Indeed, sparse data and cold start become major challenge in this approach that should be clear to addressed. Traditional collaborative filtering (memory based) rely on statistical approach to calculate similarity between users interesting about items. Once again, this approach is not robust in sparse data and cold start. Our proposed in this study to improve cold start and sparse data involving deep learning approach and NMF to handle dimensional reduction.

Deep learning involves in first test on figure 9 and 10. According the result figure 10, involving deep learning could reduce degree of error due sparse rating.

5. CONCLUSION AND FUTURE WORK

Based on the discussion in the previous section, it can be drawn some conclusions that similarity collaborative filtering algorithm can be extended to the form matrix factorization is actually an effort to increase the similarity of matrix users to matrix items. In this study both matrix was analyzed to obtain non-negative form with similarity of user to user matrix. By using the similarity matrix, it obtained a quite well result.

Deep learning and NMF play important role to reduce sparse data (figure 9 and 10) in contrast just involving collaborative filtering without deep learning still dominant in missing value due rely on similarity use Cosine, it has been better result when included deep learning and NMF more 50 percent missing value can be reduced.

For future research, it will be better final recommendation result if involving auxiliary information to predict rating value for example in the term of user feedback, product description, comment, review. It is a make sense reason because review product can be collect easily. Another reason why embedding review to predict rating is a good decision? In our best knowledge, mathematical approach success to reduce missing value but fail to detect contextual aware to improve meaningful of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

review product due the final recommendation result will be inaccurate.

REFERENCE:

[1] Y. Koren, R. Bell, And C. Volinsky, "C Ov Er F E At U Re Matrix Techniques For," Pp. 42–49, 2009.

[2] Hanafi, N. Suryana, and A. Sammad, "Evaluation of e-Service Quality, Perceived Value on Customer Satisfaction and Customer Loyalty: A Study in Indonesia." Medwell Journal, pp. 1892–1900, 2017.

[3] F. M. Harper and J. A. Konstan, "The MovieLens Datasets: History and Context," *ACM Trans. Interact. Intell. Syst.*, vol. 5, no. 4, p. 19:1--19:19, 2015.

[4] "Combining Non-Negative Matrix Factorization And Deep Neural Networks For Speech Enhancement And Automatic Speech Recognition School of Computer Engineering , Nanyang Technological University , Singapore Rolls-Royce @ NTU Corporate Lab , Nanyang Technological University , Singapore," pp. 499–503, 2016.

[5] Q. Mo and B. A. Draper, "Semi-nonnegative matrix factorization for motion segmentation with missing data," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7578 LNCS, no. PART 7, pp. 402–415, 2012.

[6] A. van den Oord, S. Dieleman, and B. Schrauwen, "Deep Content-Based Music Recommendation," pp. 2643–2651, 2013.

[7] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative Filtering and Deep Learning Based Hybrid Recommendation for Cold Start Problem," *2016 IEEE 14th Intl Conf Dependable, Auton. Secur. Comput. 14th Intl Conf Pervasive Intell. Comput. 2nd Intl Conf Big Data Intell. Comput. Cyber Sci. Technol. Congr.*, pp. 874–877, 2016.

[8] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems Handbook*, vol. 54. 2015.

[9] "Department of Electrical and Computer Engineering , University of Illinois at Urbana-Champaign , USA Department of Computer Science , University of Illinois at Urbana-Champaign , USA Adobe Research , USA."

[10] Y. Ar and E. Bostanci, "PT US CR," *Expert Syst. Appl.*, 2016.

[11] I. FernÃ¡ndez-TobÃ-as, M. Braunhofer, M. Elahi, F. Ricci, and I. Cantador, "Alleviating the new user problem in collaborative filtering by exploiting personality information," *User Model. User-Adapted Interact.*, vol. 26, no. 2–3, pp. 221–255, 2016.

[12] X. Feng, Q. Li, Y. Zhu, J. Hou, L. Jin, and J. Wang, "Artificial neural network forecasting of PM2 . 5 pollution using air mass trajectory based geographic model and wavelet transformation Arti fi cial neural networks forecasting of PM 2 . 5 pollution using air mass trajectory based geographic model and wavelet transformation," no. April, 2015.

[13] Z. Wang and L. Bing, "Neural Rating Regression with Abstractive Tips Generation for," pp. 345–354, 2017.

[14] X. Luo, M. Zhou, Y. Xia, and Q. Zhu, "An Ef fi cient Non-Negative Matrix-Factorization-Based Approach to Collaborative Filtering for Recommender Systems," vol. 10, no. 2, pp. 1273–1284, 2014.

[15] Hanafi, N. Suryana, A. Samad, and B. I. N. Hasan, "Paper Survey And Example Of Collaborative Filtering Implementation In Recommender," vol. 95, no. 16, 2017.

[16] Hanafi, N. Suryana, and A. Sammad, "An Understanding And Approach Solution For Cold Start Problem Associated With Recommender System : A Literature Review," vol. 96, no. 9, 2018.

[17] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative Filtering and Deep Learning Based Hybrid Recommendation for Cold Start Problem," *2016 IEEE 14th Intl Conf Dependable, Auton. Secur. Comput. 14th Intl Conf Pervasive Intell. Comput. 2nd Intl Conf Big Data Intell. Comput. Cyber Sci. Technol. Congr.*, pp. 874–877, 2016.

[18] Y. Zuo, J. Zeng, M. Gong, and L. Jiao, "Tag-aware recommender systems based on deep neural networks," *Neurocomputing*, vol. 204, pp. 51–60, 2016.

[19] Y. Fu, "Deep Collaborative Filtering via Marginalized Denoising," *CIKM 2015 Proc. 24th ACM Int. Conf. Inf. Knowl. Manag.*, pp. 811–820, 2015.

[20] H. Wang, N. Wang, and D.-Y. Yeung, "Collaborative Deep Learning for Recommender Systems," 2014.

[21] M. K. Najafabadi, M. Naz, S. Chuprat, and H. Sarkan, "Computers in Human Behavior Improving the accuracy of collaborative fi ltering recommendations using clustering and association rules mining on implicit data," *Comput. Human Behav.*, vol. 67, pp. 113–128, 2017.

[22] U. States, "(12) Patent Application Publication (10) Pub . No .: US 2014 / 0280251 A1," vol. 1, no. 19, 2014.

[23] M. Nilashi, N. Ithnin, and N. Haniza, "Electronic Commerce Research and Applications A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA – ANFIS," *Electron. Commer. Res. Appl.*, 2015.

[24] M. Zhang, J. Tang, X. Zhang, and X. Xue, "Addressing cold start in recommender systems," *Proc. 37th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. - SIGIR '14*, pp. 73–82, 2014.

[25] M. Nilashi, O. Ibrahim, and N. Ithnin, "Multi-criteria collaborative filtering with high accuracy using higher order singular value decomposition and Neuro-Fuzzy system," *KNOWLEDGE-BASED Syst.*, 2014.

[26] L. Candillier, K. Jack, F. Fessant, and F. Meyer, *Recommender systems*, vol. 40, no. 3. 1997.

[27] W. Pan, S. Xia, Z. Liu, X. Peng, and Z. Ming, "Mixed factorization for collaborative recommendation with heterogeneous explicit feedbacks," *Inf. Sci. (Ny).*, vol. 332, pp. 84–93, 2016.

[28] J. N. Cappella, S. Yang, and S. Lee, "Constructing dation Systems for Effective," no. May, pp. 290–306, 2015.

[29] J. Clark and F. Provost, "Matrix-Factorization-Based Dimensionality Reduction in the Predictive Modeling Process : A Design Science Perspective," 2016.

[30] N. Lopes and B. Ribeiro, *Machine Learning for Adaptive Many-Core Machines-A Practical Approach*. Springer Berlin Heidelberg, 2015.

[31] W. Chin, Y. Zhuang, Y. Juan, and C. Lin, "A Fast Parallel Stochastic Gradient Method for Matrix Factorization in Shared Memory Systems," vol. 6, no. 1, pp. 1–24, 2015.

[32] I. Tc, I. Conference, and D. Hutchison, *Information Systems and*. 2017.

[33] A. N. Sarve, M. N. Varma, and S. S. Sonawane, "RSC Advances network modeling of biodiesel production from," *RSC Adv.*, vol. 5, no. August, pp. 69702–69713, 2015.

[34] X. Amatriain, A. Jaimes, N. Oliver, and J. M. Pujol, "Data Mining Methods for Recommender Systems," pp. 1–44.

[35] A. Khoshkbarchi and H. R. Shahriari, "Coping with unfair ratings in reputation systems based on learning Coping with unfair ratings in reputation systems based on learning approach," *Enterp. Inf. Syst.*, vol. 11, no. 10, pp. 1481–1499, 2016.