JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

Effectiveness of Using Autoencoder Method on Recommender System in E-Commerce Domains

Jayana Citra Agung Pramu Putra, Z K A Baizal*

Faculty of Informatics, Informatics, Telkom University, Bandung, Indonesia
Email: [1]jayanacitra@student.telkomuniversity.ac.id, [2,*]baizal@telkomuniversity.ac.id
Correspondence Author Email: baizal@telkomuniversity.ac.id

Abstract− The growth of large data in the online market can cause problems for users, one of which is in finding products that are to their liking. Recommender systems can overcome this problem by providing specific product recommendations to be promoted and offered to buyers, for example with Collaborative Filtering. The Collaborative Filtering Paradigm consists of Memory-based and Model-based techniques. Model-based techniques are considered to be able to complement the shortcomings of memory-based because of their high scalability, accuracy, and reduced dimensions. The type of model-based that is best known for having good results is Singular Value Decomposition (SVD) and what has been frequently used recently is deep learning, especially Autoencoder. The both models are very popular for use in dimension reduction so they are suitable for making recommendations. The advantage of Deep Learning is this method can be done without preprocessing so it can minimize the process that must be done. The evaluation results show that the errors produced by SVD and Autoencoder are lower compared to other studies. RMSE is 0.7 and MAE is 0.5. Even though the RMSE and MAE on the Autoencoder are greater than the SVD, the results of the T-Test show that there is no significant difference in the two error results. Autoencoders have been shown to have good results without preprocessing and are more effective with shorter processes and there are no significant differences with SVDs. Thus, Autoencoder can be said to be worthy of use and more effective in giving recommendations.

Keywords: Recommender System; Singular Value Decomposition; Collaborative Filtering; Model-Based; Autoencoder

1. INTRODUCTION

In recent years, e-commerce systems have developed rapidly. Many people prefer to make buying and selling transactions online. This causes a huge flood of data in the online market system due to a significant increase in data starting from the number of customers, products, stores, to the features offered. The problem that often appear is that customers have difficulty finding products that suit their needs and sellers who try to promote their products that are frequently visited by buyers. Therefore, it takes the simplicity of the system offered by the marketplace to assist users in carrying out sales activities [1][2]. The recommender system has made an impact in business growth such as 80% of movies watched on Netflix, 60% of videos on Youtube, book recommendations, e-commerce and many more. The recommender system is proven to improve e-commerce platforms by increasing consumer engagement and purchases [3]. So, the recommender system can be a solution to overcome the problem of e-commerce data growth by providing recommendations to market users.

Recommender systems have the ability to analyze a user's history and come up with a model that better suits their needs, even if the user's interests are unclear. The recommendations provided by the system are based on an in-depth analysis of user preferences and behavior in the past, so that it can provide more relevant recommendations and according to the needs of the user [4]. The recommender system can automatically analyze buyers usage data to eliminate irrelevant webpage content, organize newsgroup messages into relevant categories and recommend information[5]. Recommender systems analyze data about products or user-product interactions to find relationships between products and users. The results obtained are then displayed as recommendations. One simple and frequently used recommender system method is Singular Value Decomposition (SVD).

SVD is part of the Matrix-Factorization technique and has been proven to have high quality in providing recommendation results [6]. In addition to using Matrix-Factorization techniques, recommender systems can also be done using Deep Learning. Deep Learning has been widely used in applications to help human activities. Deep Learning algorithms mostly mimic the human brain to make artificial intelligence then applied to computer systems. Deep learning can also be applied without preprocessing so that it can minimize the process in solving a problem [7].

Rouhia, et al, implemented a recommender system for the Large-scale Arabic Book Review (LABR) dataset using SVD and K-Nearest Neighbour (KNN) methods. Tests were conducted using RMSE and MAE with cross-validation techniques. KNN test results show RMSE of 1.1969 and MAE of 0.922, while SVD gets RMSE of 1.087 and MAE of 0.8077. This study shows that SVD is superior to KNN in overcoming scalability and sparsity in terms of RMSE, and minimal MAE [8].

Badr, et al, used the same method as Rouhia, et al, for the recommender system on the Movie dataset, the Movielens 100k dataset. In this study, SVD is better than when combined with KNN based on MAE, precision, and recall. SVD is also faster than KNN because the model is only created at the beginning and can be used for new data, which can potentially make the process faster and gain good scalability[9].

Wervyan, et al, conducted a recommender system research on sales data on Amazon's ModCloth using the SVD method. This study also used other matrix-factorization algorithms as a comparison including SVDpp,

Submitted: **13/06/2023**; Accepted: **31/07/2023**; Published: **31/07/2023**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

CoClustering, SlopeOne, and NMF. However, SVD proved to be the best algorithm compared to other algorithms because of its best accuracy with RMSE 1.05. The RMSE results were then improved by changing the parameters so that the RMSE obtained was better than the previous 1.04 [10].

Gourav et al, in their paper implement the Autoencoder method for system recommendations on 3 different datasets, i.e. the Epinions, FilmTrust and Ciao datasets. The deep architecture of the Autoencoder helps learn about hidden users and items. Autoencoders provide significantly improved RMSE and MAE results compared to existing approaches. In this study, it was also proposed to conduct experiments on activation functions such as tanh and RelU [11].

Diana et al, used Autoencoder and SVD methods for 2 Movie datasets from Movielens with different amounts of data, i.e. 1M and 10M. The Autoencoder was improvised with some experiments that changed the layer, epoch, batch size, number of neurons and several other parameters. SVD is used as a comparison because it is considered a technique that is often used in recommendations. Although SVD has a faster execution time, the Autoencoder actually shows a smaller RMSE [12].

Miao et al, conducted system recommendation research with the Stacked Autoencoder (SAE) method which was compared with N-CF-U, PDM, and FunkSVD for 2 different movie datasets, i.e. EachMovie and MovieLens 1M. The SAE algorithm is superior to the other 3 because it can avoid overfitting problems and can learn relationships between features. This makes SAE usable for big data with its advantages to solve sparsity problems [13].

Some previous studies have shown that SVD and Autoencoder have good results in recommendation systems. The focus of the problem raised by previous study is to compare a method with other methods and then determine the best method based on the lowest error value. However, the focus of this study is to evaluate the method not only based on the lowest error value but by testing whether the difference in errors generated by SVD and Autoencoder is significant or not using the T-Test method. In other words, if Autoencoder has a higher error value than SVD and the T-Test results show that the difference is not significant, then Autoencoder can be considered equivalent to SVD whose quality has been proven to be good in many system recommendation problems.

In addition, in previous studies, the data used would be preprocessed even though Autoencoder can be used without preprocessing. In this study, the preprocessing stage will only be carried out on the SVD, while the autoencoder will not perform the preprocessing stage. The preprocessing stage will also be the main thing taken into account in this study because by skipping the stage, the process that must be done is less and the time used will be shorter. So, this study will discuss how the effectiveness of the Autoencoder method in the recommendation system in the e-commerce domain. Autoencoder will be applied without preprocessing and using several experiments with different parameters. Besides Autoencoder, another model-based method, SVD, will also be applied for comparison. SVD has been considered as a well-known method in recommendation for some time so it is expected that Autoencoder without preprocessing can be better or at least resemble SVD.

2. RESEARCH METHODOLOGY

This study will use two methods i.e. SVD and Autoencoder to create a recommender system in e-commerce. The SVD method will be used as a comparison to assess the effectiveness of the Autoencoder method. Before applying the SVD method, the data will go through a preprocessing stage. While the Autoencoder method will be applied without the preprocessing stage. The results of the application of the two methods were tested using RMSE and MAE. The test results were tested again using a statistical test i.e. the T-Test to assess the differences in the RMSE and MAE tests. An overview of this study process can be described according to

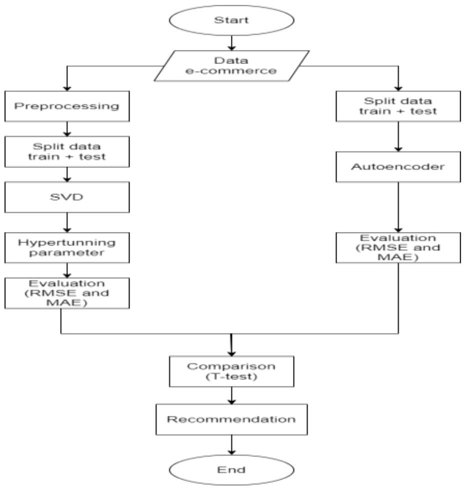

Figure 1

Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

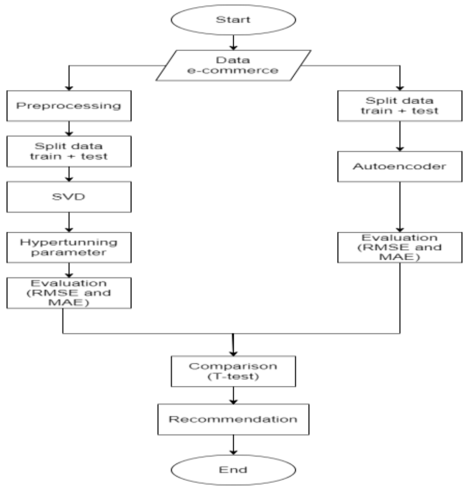

Figure 1. Research Overview

2.1. Recommender System

A recommender system is a type of information filtering system that considers purchase history or item ratings in predicting products that users like to provide product recommendations to users[2][8]. Recommender systems collect information that can be obtained explicitly (based on ratings) or implicitly (based on user behavior). The information is then studied and provides the most relevant recommendations. This can improve the adaptation of the application to users.

The amount of information in the web system is one of the problems that makes it difficult for users to choose the information that best suits the needs and desires of users. This problem can be solved by the advantages of a recommender system that makes automation of recommendations personalized according to user tastes using machine learning techniques. The recommender system filters the information that directs the items to be used. Therefore, the level of accuracy in providing recommendations to users is the main problem of the recommender system. The recommender system has been implemented in several fields including health, industry, music, e-commerce, marketing and many other fields. In the field of e-commerce, recommender systems have increased turnover in providing relevant products to customers, increasing client loyalty, and making product suggestions to increase sales [1][6].

The most well-known and widely used recommender system method is Collaborative Filtering. This technique is based on observing past behavior where users who have similar behavior will have similar tastes or buying habits. The recommender system with collaborative filtering looks for rankings then the ratings between users are analyzed for similarity to get new recommendations from ranking comparisons between users. Collaborative Filtering methods consist of Memory-based and model-based.

Memory-based provides recommendations based on the size of similarity between users or items. However, memory-based is not suitable for data with high dimensions and sparse. Similarity values are based on common items so that data is sparse and items with general ratings are very small and makes the accuracy of recommendations low. While model-based provides recommendations by studying and adjusting the parameter model to the user-item rating matrix. This technique uses training data in finding rankings. Model-based is considered to overcome memory-based shortcomings due to its high scalability, accuracy, and dimensionality reduction [14][15].

2.2 Dataset

In the world of system recommender, it is common practice to use publicly available datasets from various application environments to evaluate and compare the performance of recommendation algorithms. In general, the performance evaluation of context-aware recommendation models is based on four popular contextualized real-world datasets from various domains: music, food, and movies. This variation allows us to assess the performance of the proposed models on various datasets that have different characteristics [16].The data used in this study came from one of the e-commerce in Indonesia. The data is taken from the Kaggle.com site. The data used amounted to 37301 e-commerce user reviews about products that have been purchased. Data size 10.46 MB.

2.3 Preprocessing

Before data splitting, SVD models require data preprocessing to clean the data to make it more ready for use. While the Autoencoder model Autoencoder can be directly split without preprocessing. Preprocessing that will be applied to datasets includes missing values, data scaling and feature selection. The missing value in the data will be filled with the number 0. Data scaling is done to make the number range not too far so that it is easier, especially in Autoencoder models. The data scaling technique used is Min-max using the sklearn library. The selection feature

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

functions to select the column to be used so that it only focuses on the data to be analyzed. After the data is ready to be applied to modeling techniques, the data will be divided into 2 types, i.e. data train and test data. The percentage of data used for this study is 80% data train and the remaining 20% as test data.

2.4 Modeling

There are 2 models used i.e. SVD and Autoencoder. The stages of development of the recommender system consist of several structured stages. The following is a sequence of stages that must be carried out in the development of the recommender system:

1) Model creation: This stage involves creating recommendation models with SVD and Autoencoder methods. This model becomes the basis for generating recommendations based on existing data.
2) Model training: The model created must be trained using training data. The training process aims to understand patterns and relationships in the data in order for the model to provide accurate recommendations.
3) Model testing: The final stage is to test the recommendation model by using data that is not used in training. This test is conducted to test the performance of the model in providing relevant and useful recommendations on new data. The SVD method will be tested with the K-Fold Cross Validation method with K = 5 where the data will be divided into 5 different parts. The first part will be used as a data train and the other data into test data. Next, the second part will be made for the data train and the other part will be used as the test data, and so on until the fifth part. In addition, SVD involves a process of parameter hypertuning aimed at finding the best parameters in the algorithm. The goal is that the results produced by SVD can be better in quality. In addition to SVD, Autoencoder will also involve a parameter experiment process so that there are several Autoencoder algorithms with different parameters to get the Autoencoder algorithm with the best parameters.

2.5 Singular Value Decomposition

The most popular model-based technique is Matrix-Factorizarion as evidenced by the results of the Netflix Prize Conteest [15]. One model-based method that uses dimension reduction techniques in recommender systems is SVD. By representing hidden features of users or items in a lower forecast matrix, SVD helps improve the efficiency and accuracy of recommendations. Thus, the space matrix N becomes the space K where $K < N$ [17]. The SVD formula is shown in the formula **Error! Reference source not found.**.

$$A = U \,.\, S \,.\, V^T \tag{1}$$

Where, A is the usability matrix, U and V represent orthogonal matrices that show the relationship between latent factors with users and items, and S is a singular or diagonal matrix that shows the strength of latent factors[10].

2.6 Autoencoder

Autoencoder (AE) is a type of neural network and includes unsupervised learning that is very powerful in capturing key data features. AE is commonly used for generative modelling, efficient coding, dimension reduction [18].The Autoencoder has data training to project the input data into a latent or hidden space with low dimensions and then reconstructed into the output space so that the difference between input and output is expected to be as small as possible [19].

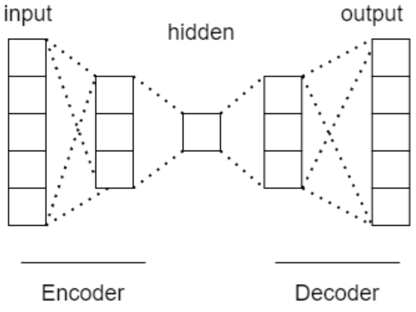

Figure 2 Architecture Autoencoder

Based on the Autoencoder architecture seen in Figure 2, the Autoencoder consists of a minimum of three layers, i.e. the input layer, the Hidden layer, and the output layer. The input layer to the hidden layer is called the encoder, which is responsible for compressing into the latent space. The hidden layer in the center represents the compression of the input and will be returned to the decoder. While the hidden layer to the output layer is called a decoder, which serves to reconstruct data from the hidden layer [12].

2.7 Evaluation

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

The test results of all models will be evaluated by looking at the error between the actual rating and the predicted rating. There are many ways to evaluate the output produced by including Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). RMSE is an evaluation technique where all differences between the initial rating and the predicted rating are squared and then the average value is calculated. After that, the average obtained is calculated square root. The RMSE formula is indicated by the formula **Error! Reference source not found.**.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} |p_i - q_i|^2} \tag{2}$$

While MAE is an evaluation technique in which the absolute value of all differences between the initial rating and the predicted rating is calculated on average. Generally, MAE ranges from 0 to infinity. The MAE formula is shown by formula **Error! Reference source not found.**.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |p_i - q_i| \tag{3}$$

Where
pi = initial/actual rating
qi = Prediction Rating
n = number of ratings[8]

In addition, a statistical test was also carried out, i.e. T-Test to see whether the performance of the Autoencoder can resemble SVD performance. T-Test is a test by looking at the average statistics of 2 different populations. This test looked at the differences between the 2 populations whether significant differences were indicated or not[20]. T-Test testing requires a T-Table, which is a table containing a list of critical T values. The table is used to compare the critical T value with the calculated T value. The important things that must be done for T-test testing are.

1) Determine the alpha level (acceptable error rate). In this study, the alpha level used is 0.05.
2) Calculate degrees of freedom for T.
3) Check the critical T value on the T-Table according to the calculated alpha level and degrees of freedom.

T-Tests can be classified into 2, i.e. one-tailed tests and two tailed tests. In one-tailed tests, the hypothesis is directed in one particular direction, while in two tailed tests, the hypothesis has no particular direction. Two tailed tests are used when we are not sure about the direction of the difference between two means, while one-tailed tests are used when we have a hypothesis about the direction of the difference [21]. So, in this study, the type of t-test that will be used is a one-tailed test and the hypotheses used in this test are as follows.

1) Null Hypothesis (H0): Average (RMSE and MAE) SVD algorithm and Average (RMSE and MAE) Autoencoder algorithm are almost the same or statistically, there is no significant difference between samples.
2) Hypothesis one (H1): Average (RMSE and MAE) SVD algorithm and Average (RMSE and MAE) Autoencoder algorithm are different or it can be said statistically, there is a significant difference between the samples[22].

The RMSE and MAE results will be retested using the T-Test where if the T-Count < T-Table, then the null hypothesis is accepted[20].

3. RESULT AND DISCUSSION

3.1 Data Preparation

The dataset used is data from one of the e-commerce in Indonesia which contains product review data in the e-commerce. The data consists of 40607 rows and 9 columns. There are 5 product categories in this data, i.e. electronics, fashion, sports, mobile phones, and carpentry. Available stores in the data are 158 and product types are 3664. The rating in this dataset ranges from 1-5 with the amount of data that has a dominant rating of 5 as shown in Figure 3.

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

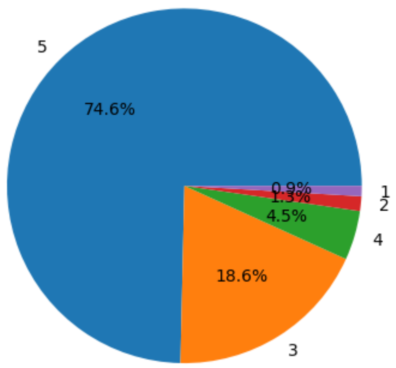

Figure 3. Rating Distribution

After studying the data in depth, it was found that there was 1 column of data that had a missing value, i.e. the sold column. This column has an integer type so that it can be filled using the value 0 without changing the relevance of the data, so products with empty sold columns are considered to have never been purchased. After that, 2 optional preprocessing were carried out with the aim of facilitating modeling work. Scaling data for user_id and product_id columns is done because the range of numbers is so high, especially columns product_id. This can reduce accuracy and can slow down the running process. Scaling is used with the Min-Max Scaling technique so that the data will be in the range of 0-1. Feature Selection aims to select columns to be used or columns that are considered to affect modeling so that the results are more accurate and the running process is faster.

3.2 Implementation Results

In this study, 2 methods will be carried out, i.e. SVD and Autoencoder. Data that has been preprocessed will be used for the SVD method, while the input data will be used for Autoencoder. Each data will be divided into training data and test data with a ratio of 80:20. Each method will be experimented then the error value will be found and the best experiment will be determined to be compared with other methods. The error value of the best results of the two methods will be tested using T-Test and determine the significance of the difference from the error results of the two methods according to the hypothesis.

3.2.1 SVD

The SVD method is carried out with 2 experiments, i.e. SVD has been tested using K-Fold and SVD with Hypertunning parameters. K used for K-Fold testing is 5. Hypertunning parameters aims to select parameters in the SVD model that are expected to improve SVD results. Hypertunning parameters is done with gridSearch method. The best parameter results are 'n_factors': 20, 'n_epochs': 10, 'lr_all': 0.002, 'reg_all': 0.4. Each experiment was calculated for its error (MAE and RMSE) and then the best experiment was found. The results of SVD modeling errors are shown in Figure 4.

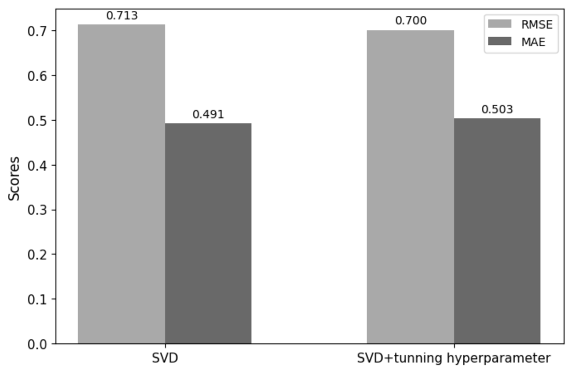

Figure 4. Testing results RMSE and MAE on SVD

Figure 4 shows a comparison of errors using SVD. The error results shown by the SVD model are not significant or can be said to be similar. Both models show RMSE error results of about 0.7 and MAE 0.5. However, the larger the RMSE value actually results in a smaller MAE value. SVD produces the lowest MAE value but the resulting RMSE value is higher than others. In contrast, SVD with Tunning Hyperparameters produces the lowest RMSE value and the highest MAE. In this case, RMSE and MAE differences between SVD and SVD with Tunning Hyperparameters will be found. If the RMSE difference is greater than the MAE difference, then the best model will follow the lowest RMSE, and vice versa. The results show that the RMSE difference is smaller, so that the

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

best SVD model will be SVD with Hyperparameter Tuning which has RMSE values, i.e. SVD with Hyperparameter Tuning with RMSE 0.7 and MAE 0.5.

3.2.2 Autoencoder

In addition to using the SVD model, an Autoencoder model was also carried out with 3 experiments with various different parameters.

Table 1. Autoencoder parameters used

Experiment to	Layer	Activation function	Learning rate	Batch size	Epochs
AE 1	3	Relu	0.01	75	25
AE 2	6	Linear	0.1	200	50
AE 3	3	Relu + Linear	0.001	16	5

Table 1 shows the parameters used in each Autoencoder experiment. The experiment was done by changing the number of layers and activation functions. The parameters that are always updated in the test are Learning rate, Batch_size, and epochs. Autoencoder modeling error results are shown in Figure 5.

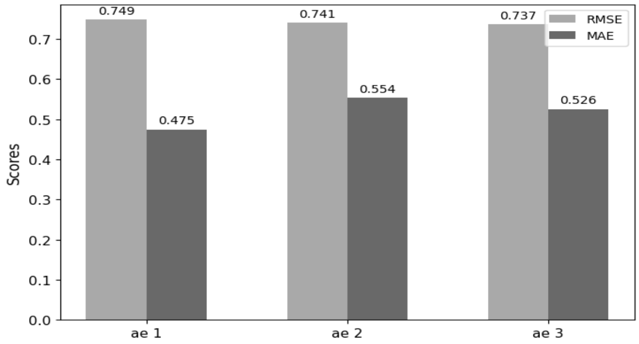

Figure 5. Testing result RMSE and MAE on Autoencoder

Figure 5 shows an error comparison using an Autoencoder model with 3 different parameter experiments as shown in the table. The error results shown by the Autoencoder model are the same as SVD whose differences are not significant or can be said to be similar. All models show RMSE error results of around 0.74 and MAE 0.4-0.5. However, the larger the RMSE value actually results in a smaller MAE value. The first AE experiment produced the lowest MAE value but the RMSE value produced was the highest compared to the others. In contrast, the second AE experiment yielded the lowest RMSE value and the highest MAE. Just like SVD, RMSE and MAE difference between the 2 experiments will be found. If the RMSE difference is smaller than the MAE difference, then the best model will follow the lowest RMSE, and vice versa. The results show that the difference in MAE is smaller, so that the best Autoencoder model will be used as an Autoencoder experiment that has an MAE value, i.e. the first experiment with RMSE 0.75 and MAE 0.47 with 3 layers, activation function = RelU, learning rate = 0.01, batch size = 75 and epochs = 25.

3.2.3 Analysis and Evaluation

After that, the two best models will be compared to find the best model between SVD and Autoencoder. The best SVD and Autoencoder modeling error results are shown in Figure 6.

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

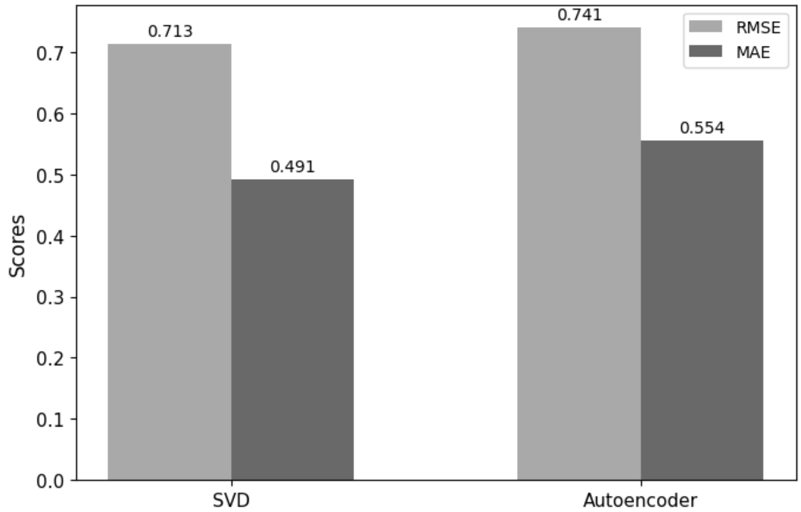

Figure 6. Testing result RMSE and MAE on SVD and Autoencoder.

Figure 6 shows a comparison of errors using SVD and Autoencoder models after filtering from various experiments. SVD has a smaller RMSE result than Autoencoder, but the resulting MAE is actually larger than Autoencoder. This is different from the study conducted by Diana et al, in their study, the error calculation used is only RMSE and the RMSE results of Autoencoder are smaller than SVD for 2 different data. So, Diana et al. can directly conclude that Autoencoder has better results than SVD [12]. However, in this study, there are differences in the results of RMSE and MAE which results in conclusions that cannot be drawn based on the error value alone but requires another test, which is the T-Test. The RMSE and MAE obtained will be assessed for differences and conclusions will be drawn based on the hypotheses that have been made previously. The RMSE and MAE T-Test results are shown as follows.

```
RMSE t-test result
========================================
T-statistic value:  -117.21342186645083
P-Value:  3.0319301187356225e-50
```

Figure 7. RMSE T-Test result

```
MAE t-test result
========================================
T-statistic value:  -291.70599602253117
P-Value:  2.8416080588887542e-65
```

Figure 8. MAE T-Test result

Figure 7 and Figure 8 show the results of the T-Test on SVD and Autoencoder, where Figure 7 is the RMSE T-Test and Figure 8 is the MAE **Error! Reference source not found.**. The t-statistic results above will b e compared with the t-values in the T-Table to obtain hypothetical results. N samples used as many as 20, so df $= 2(n - 1) = n1 + n2 - 2 = 20 + 20 - 2 = 38$. The alpha used is 0.05, so the t-value obtained from the T-Table is 2.02439. The value of t in the T-Test is smaller than t in the T-Table, so the null hypothesis is accepted. Therefore, it can be concluded, there is no significant difference between the average (both RMSE and MAE) Autoencoder and the average (both RMSE and MAE). The results of SVD and Autoencoder recommendations are shown in Figure 9 and Figure 10

```
Top 5 products recommendations for user 172:

USB HUB 4 port USB 3.0 / USB HUB 3.0 &#40;4port&#41; DIGIGEAR HIGH QUALITY
Casio Jam Tangan Unisex &#40;MQ-71-4BDF&#41;
TOSHIBA FLASH AIR WIRELESS SD CARD 32GB ORIGINAL 100%
TINTA / CATRIDGE HP 703 BLACK / COLOR ORIGINAL 100%
Adaptor Charger Laptop Asus N43SL A42J A43S A43SJ N46VJ 19V 4.74A ORI
```

Figure 9. SVD Recommendation Results for user 172.

Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

```
Top 5 products recommendations for user 172:

Isi Staples Tembak 8 mm best guard
Baterai Original Lenovo Thinkpad7 T420 L410 L412 L510 SL410 T410 E420
Adaptor Charger Original Asus A456 A456U~19V 3.42A &#40;4.0mm*1.35mm&#41;
Adaptor Charger Original Asus X441S ~19V 1.75A
Baterai Original Apple Macbook pro A1322 A1278 -13 inch
```

Figure 10. Autoencoder Recommendation Results for user 172.

Based on the results of this study and Diana et al, it can be seen that the error results are often used as a reference to determine the best method, but it turns out that this study found that there is an insignificant error difference between these methods, so the method with a lower error value is not always considered better [12]. This shows that in choosing an appropriate recommendation method, it is important to consider factors other than error rate. One example is the advantage of Autoencoder, where it can be used without the need for complex preprocessing steps. This advantage can result in a more efficient time in the recommendation process, as the data inputted into Autoencoder can be directly used without the need for complex preprocessing steps. If preprocessing such as SVD is necessary, then an in-depth analysis of the data is required to identify the issues that need to be addressed and determine the appropriate preprocessing technique. However, Autoencoder provides the advantage of reducing the complexity of preprocessing steps that are often time- and resource-consuming.

The result of the T-Test which shows no significant difference between the methods in the performance of the recommendation system, provides strong evidence that Autoencoder is a successful method and produces excellent results in providing recommendations. Overall, this study reveals that Autoencoder is a very viable option in addressing the problems in recommendation systems. By reducing the number of processes and time required, Autoencoder is able to provide excellent results in providing recommendations to users. The findings provide a deeper understanding of the importance of considering factors other than error rate in selecting an appropriate recommendation method, and provide direction in optimizing the efficiency and effectiveness of recommendation systems by leveraging the advantages of Autoencoder.

4. CONCLUSION

Based on the results of the experiments that have been done above, RMSE and MAE generated by SVD and Autoencoder models are not significantly different values based on the T-Test. Overall, the model showed smaller error results than previous studies, i.e. RMSE ranging from 0.7 and MAE at 0.4-0.5. The best SVD model is SVD by tuning hyperparameters, i.e. RMSE 0.7 and MAE 0.5. Meanwhile, the best Autoencoder Model is Autoencoder by tuning Hyperparameter, i.e. RMSE 0.75 and MAE 0.47. However, after statistical testing of the T-Test, the results showed no significant difference between the average (both RMSE and MAE) Autoencoder and the average (both RMSE and MAE) SVD. These results mean that the Autoencoder which provides recommendations without preprocessing is capable of resembling SVD. If you use an Autoencoder, the entered data can be processed directly without having to do preprocessing. In contrast to SVD which has to go through various stages of preprocessing such as handling missing values, feature selection, etc. This is very beneficial in making recommendations because it can minimize the process carried out and have good results. Thus, the Autoencoder is proven to be effective and has good results even without preprocessing so it is feasible to use in making recommendations.

REFERENCES

[1] P. M. Alamdari, N. J. Navimipour, M. Hosseinzadeh, A. A. Safaei, and A. Darwesh, “A Systematic Study on the Recommender Systems in the E-Commerce,” IEEE Access, vol. 8, pp. 115694–115716, 2020, doi: 10.1109/ACCESS.2020.3002803.

[2] F. T. A. Hussien, A. M. S. Rahma, and H. B. Abdulwahab, “An e-commerce recommendation system based on dynamic analysis of customer behavior,” Sustain., vol. 13, no. 19, Oct. 2021, doi: 10.3390/su131910786.

[3] M. Khoali, A. Tali, and Y. Laaziz, “Advanced Recommendation Systems Through Deep Learning,” ACM Int. Conf. Proceeding Ser., 2020, doi: 10.1145/3386723.3387870.

[4] G. A. Fadhlullah, Z. K. A. Baizal, and N. Ikhsan, “Conversational Recommender Systems Based on Mobile Chatbot for Culinary,” J. MEDIA Inform. BUDIDARMA, vol. 5, no. 4, p. 1233, Oct. 2021, doi: 10.30865/mib.v5i4.3242.

[5] K. K. Singh, M. Aggarwal, M. Pathak, and P. Dubey, “Recommender System for an E-commerce Web application,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 5, pp. 4747–4750, 2022, doi: 10.22214/ijraset.2022.43494.

[6] T. A. W. Tyas, Z. K. A. Baizal, and R. Dharayani, “Tourist Places Recommender System Using Cosine Similarity and Singular Value Decomposition Methods,” J. MEDIA Inform. BUDIDARMA, vol. 5, no. 4, p. 1201, Oct. 2021, doi: 10.30865/mib.v5i4.3151.

[7] R. G, “A Study to Find Facts Behind Preprocessing on Deep Learning Algorithms,” J. Innov. Image Process., vol. 3, no. 1, pp. 66–74, 2021, doi: 10.36548/jiip.2021.1.006.

[8] R. M. Sallam, M. Hussein, and H. M. Mousa, “An Enhanced Collaborative Filtering-based Approach for Recommender Systems,” Int. J. Comput. Appl., vol. 176, no. 41, pp. 9–15, Jul. 2020, doi: 10.5120/ijca2020920531.

[9] B. Hssina, A. Grota, and M. Erritali, “Recommendation system using the k-nearest neighbors and singular value decomposition algorithms,” Int. J. Electr. Comput. Eng., vol. 11, no. 6, pp. 5541–5548, 2021, doi:

Submitted: **13/06/2023**; Accepted: **31/07/2023**; Published: **31/07/2023**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor 3, Juli 2023, Page 1494-1502
ISSN 2614-5278 (media cetak), ISSN 2548-8368 (media online)
Available Online at https://ejurnal.stmik-budidarma.ac.id/index.php/mib
DOI: 10.30865/mib.v7i3.6346

10.11591/ijece.v11i6.pp5541-5548.

[10] W. Shalannanda, R. F. Mulia, A. I. Muttaqien, N. R. Hibatullah, and A. Firdaus, "Singular value decomposition model application for e-commerce recommendation system," JITEL (Jurnal Ilm. Telekomun. Elektron. dan List. Tenaga), vol. 2, no. 2, pp. 103–110, Sep. 2022, doi: 10.35313/jitel.v2.i2.2022.103-110.

[11] G. Bathla, H. Aggarwal, and R. Rani, "AutoTrustRec: Recommender System with Social Trust and Deep Learning using AutoEncoder," Multimed. Tools Appl., vol. 79, no. 29–30, pp. 20845–20860, 2020, doi: 10.1007/s11042-020-08932-4.

[12] D. Ferreira, S. Silva, A. Abelha, and J. Machado, "Recommendation system using autoencoders," Appl. Sci., vol. 10, no. 16, pp. 1–17, 2020, doi: 10.3390/app10165510.

[13] M. Yu, T. Quan, Q. Peng, X. Yu, and L. Liu, "A model-based collaborate filtering algorithm based on stacked AutoEncoder," Neural Comput. Appl., vol. 34, no. 4, pp. 2503–2511, 2022, doi: 10.1007/s00521-021-05933-8.

[14] Y. Afoudi, M. Lazaar, and M. Al Achhab, "Collaborative filtering recommender system," in Advances in Intelligent Systems and Computing, 2019, vol. 915, pp. 332–345. doi: 10.1007/978-3-030-11928-7_30.

[15] R. Mehta and K. Rana, "Evolution of singular value decomposition in recommendation systems: a review," Int. J. Bus. Intell. Data Min., vol. 14, no. 4, p. 528, 2019, doi: 10.1504/IJBIDM.2019.099962.

[16] A. J. Chiheb-Eddine Ben Ncir, Ibraheem Alharbi, "Machine Learning and Data Analytics for Solving Business Problems," Springer International Publishing, 2022, p. 126.

[17] B. Srikanth and V. Nagalakshmi, "Songs Recommender System using Machine Learning Algorithm: SVD Algorithm," 2020. [Online]. Available: https://labrosa.ee.columbia.edu/millionson/

[18] G. Zhang, Y. Liu, and X. Jin, "A survey of autoencoder-based recommender systems," Front. Comput. Sci., vol. 14, no. 2, pp. 430–450, 2020, doi: 10.1007/s11704-018-8052-6.

[19] Z. Noshad, A. Bouyer, and M. Noshad, "Mutual information-based recommender system using autoencoder," Appl. Soft Comput., vol. 109, p. 107547, 2021, doi: 10.1016/j.asoc.2021.107547.

[20] S. K. Panda, S. K. Bhoi, and M. Singh, "A collaborative filtering recommendation algorithm based on normalization approach," J. Ambient Intell. Humaniz. Comput., vol. 11, no. 11, pp. 4643–4665, 2020, doi: 10.1007/s12652-020-01711-x.

[21] S. J. S. Jerry R. Thomas, Philip Martin, Jennifer L. Etnier, "Research Methods in Physical Activity," Britania Raya: Human Kinetics, Incorporated, 2022, p. 165.

[22] H. Singh and S. Bawa, "Predicting COVID-19 statistics using machine learning regression model: Li-MuLi-Poly," Multimedia Systems, vol. 28, no. 1. pp. 113–120, 2022. doi: 10.1007/s00530-021-00798-2.

Submitted: **13/06/2023**; Accepted: **31/07/2023**; Published: **31/07/2023**