**JURNAL MEDIA INFORMATIKA BUDIDARMA**
**Volume 7, Nomor X, Bulan 2023, Page 999-999**
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.v5i1.2293

# Music Recommender System using Autorec Method for Implicit Feedback

**Muhamad Faishal Irawan, Z.K.A. Baizal**
School of Computing, Informatics, Telkom University, Bandung, Indonesia
Email: [1]faishalirawan@student.telkomuniversity.ac.id, [2]baizal@telkomuniversity.ac.id
Correspondence Author Email: baizal@telkomuniversity.ac.id

**Abstract**− As the number of music and users in music streaming services increases, the role of music recommender systems is getting important to make it easier for users to find music that matches their tastes. The collaborative filtering paradigm is the most commonly used paradigm in developing recommender systems. Many studies have proven that deep learning is able to improve the performance of matrix factorization. One such method in deep learning that has been adapted for use in Recommender Systems is Autorec, which is a variation of the Autoencoder technique. Autorec shows that it performs better than the baseline matrix factorization using Movielens and Netflix datasets. Therefore, in this study we propose the use of Autorec to develop a recommender system for music. The experimental results show that Autorec performs better than Singular Value Decomposition (SVD), with an RMSE difference of 0.7.

**Keywords**: Recommender System; Autoencoder; Deep learning; Music recommender system; Autorec

## 1. INTRODUCTION

Music streaming services such as Spotify, Apple Music, and YouTube Music are increasingly used by many users around the world, along with the amount of music that continues to grow. The increase in consumable content will make it difficult for users to find suitable music to listen to [1]. This makes the role of recommender systems increasingly important in music search [2]. With the existence of a music recommender system, the problem of finding products that suit user tastes and information overload can be overcome. Paradigms for building recommender systems are collaborative filtering, content-based filtering and knowledge-based filtering [3].

Collaborative filtering (CF) is the most commonly used paradigm in recommender systems [4]. The idea of this paradigm is to predict a user's rating of an item based on other users who have similar tastes. One of the methods used in CF is matrix factorization. Matrix factorization is a classic method that has been used and won the Netflix Prize competition, so it has been proven to give good results [5]. This method is a latent factors method or a method that seeks to find hidden factors and predict ratings by characterizing items and users on many factors inferred from rating patterns. Previous studies have proven that the recommendation results of matrix factorization approaches can be improved with the use of deep learning [6].

In recent years, deep learning has become the dominant approach for many ongoing tasks in the field of machine learning [7]. Deep learning has also started to enter the realm of recommender systems. Where previously, deep learning has given excellent results on other machine learning problems such as computer vision [6] and speech recognition [8]. This research seeks to use a deep learning approach with an autoencoder architecture that has been customized for a recommender system called Autorec.

Compared to traditional models such as matrix factorization that can only use one data source such as ratings, the models using autoencoder can utilize multiple data sources such as ratings, audio, visual, and video. Additionally, autoencoders may offer better recommendation outcomes than traditional models due to their improved grasp of user preferences and item characteristics [9].

Sedhain et al. [10] proposed the use of autoencoder to perform collaborative filtering named Autorec. The model got better results than the state-of-the-art CF technique in that year. We examined the model's application in the field of music by constructing a music recommendation system and comparing the outcome to that of a conventional matrix factorization method.

The music domain itself is interesting to use in this research because the dataset contains implicit feedback. Implicit feedback is a type of feedback that does not directly indicate the user's preference for an item, as opposed to explicit feedback. For example, in explicit feedback users can rate an item between 1 and 5, 1 means the user dislikes the item and 5 means the user really likes an item. Meanwhile, implicit feedback has various forms, such as how long the user views an item, how many times the user clicks on an item, and how many times the user listens to a piece of music. In this research, we use the frequency information of each user in listening to each music in making recommendations.

The primary sources for this research include one of the initial studies that applied autoencoders in the area of recommendation systems. This method is called AutoRec [10], this method was proposed by Sedhain et al. in 2015. Using the Movielens and Netflix datasets, the study showed that AutoRec managed to provide better results than other matrix factorization methods and Restricted Boltzmann machine for collaborative filtering methods using the root-mean-square error (RMSE) evaluation metric. Movielens and Netflix datasets are among the datasets that use explicit feedback.

After the autoencoder method for recommender systems was proposed, many researchers developed this method to improve its performance. The DeepRec method [11] was one of them, it was developed by Kuchaiev et


Submitted: **dd/mm/yyyy**; Accepted: **dd/mm/yyyy**; Published: **dd/mm/yyyy**

JURNAL MEDIA INFORMATIKA BUDIDARMA
Volume 7, Nomor X, Bulan 2023, Page 999-999
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.v5i1.2293

al. of NVIDIA team in 2017. The method extends AutoRec by using a deeper network, Scaled Exponential Linear Unit (SELU) activation function, high dropout rate, and iterative output re-feeding. Meanwhile, another type of autoencoder called variational autoencoders emerged that are developed specifically in the CF area, such as Mult-VAE [12], this method was created by Liang et al. in 2018. It introduces a Bayesian inference approach in recommending items and was developed specifically for implicit feedback data. Other commonly used autoencoder types include the denoising autoencoder [13], this method was built by Wu et al. in 2016. It introduces noise in the input rating to improve the recommendation of the system.

Mult-VAE is an autoencoder that has a problem formulation close to this research. The research specializes in creating a variational autoencoder to solve collaborative filtering problems using implicit feedback. While Autorec's research only uses explicit feedback datasets. Therefore, in this research we tried to explore the use of Autorec, specifically U-Autorec, in different domain from the original research and demonstrate how it performs in that particular domain. We chose music domain because the nature of most music dataset in recommender system is in the form implicit feedback. This poses an intriguing problem as the Autorec was only tested in the movie domain that have an explicit feedback type. The baseline that we are going to use, in accordance to the previous research, is matrix factorization type, in this case Singular Value Decomposition (SVD).

## 2. RESEARCH METHODOLOGY

### 2.1 Research Stages

The flow of music recommender system development can be seen in figure 1. First, we acquire the publicly available Million Song Dataset (MSD) The Echo Nest Taste Profile Subset and preprocess it using the method that will be explained in the subsequent section. Following that, we are going to build and train the Autorec model for recommending music. Finally, the model will be evaluated using root-mean-square error and compared with baseline method.

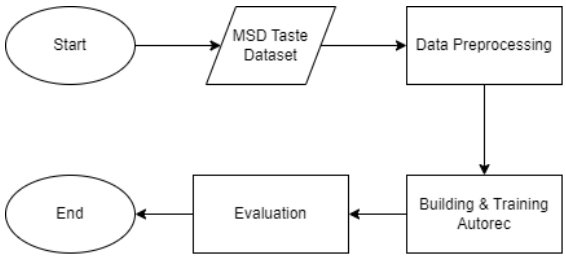


**Figure 1.** Music recommender system development flow

### 2.2 Autoencoder

An autoencoder (AE) is a type of neural network architecture that is typically used for unsupervised learning tasks, such as generating new data, reducing the number of features, and encoding data efficiently [14]. This method is better in learning the representation of latent features in recognizing images [15], speech [16], and in computer vision [17], among others. AE is considered an effective method to obtain nonlinear features [18]. Autoencoders are used to transform input into a more compact and informative representation, and then decode it back, so that the reconstructed input is as similar as possible to the original one [19].

AE is divided into an encoder and a decoder, the encoder consists of the input and hidden layer. Meanwhile, the decoder consists of that same hidden layer and an output layer. The input goes into the encoder where it's feature will be reduced, then the decoder tries to reconstruct the original input using that reduced form.

The encoder uses the function $f$ to convert the input data $x$ that's high-dimensional into a hidden representation $h$ that has a lower dimension. The formula for this process can be found in equation 1. Where $s_f$ is an activation function, meanwhile $W$ representing a matrix of weight, and $b$ representing the vector of bias.

$$h = f(x) = s_f(Wx + b) \quad (1)$$

The decoder utilizes another function $g$ to convert the hidden representation $h$ back to a reconstruction of $x'$. The formula for this process can be observed in equation 2. With $s_g$ is the activation function, $W'$, $b'$ representing the weight matrix and bias vector respectively.

$$x' = g(h) = s_g(Wh + b') \quad (2)$$


Submitted: **dd/mm/yyyy**; Accepted: **dd/mm/yyyy**; Published: **dd/mm/yyyy**