Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

Content-Based Music Recommender System Using Deep Neural Network

Rich Andiety, Z K A Baizal*

Informatika, Informatika, Telkom University, Bandung, Indonesia
Email: andietyrich@gmail.com, baizal@telkomuniversity.ac.id
Correspondence Author Email: baizal@telkomuniversity.ac.id
Submitted: **09/08/2024**; Accepted: **10/09/2024**; Published**: 12/09/2024**

Abstract−Music is one of the most popular forms of entertainment. Along with the development of information technology, music streaming platforms such as Spotify, Apple Music, and Deezer are increasingly popular among users. However, with thousands of songs available on these music streaming platforms, users often have difficulty finding songs that suit their tastes. Therefore, we design a music recommender system that can assist users in finding songs that are more in line with user preferences. In this research, we propose the development of a content-based music recommender system using a combination of Content-Based Filtering and Deep Neural Network (DNN) methods. The DNN used is Convolutional Neural Network (CNN) which serves to analyze the audio features of songs and learn user preferences to increase the percentage of accuracy in providing recommendations that match user needs. This recommender system works by extracting features from songs listened to by the user and then recommending other songs with similar features. We trained and evaluated our model on a dataset of 250 songs. This research aims to develop a music recommender system that can provide personalized recommendations to users according to the preferences of users. This research provides an accuracy result of 73.5%. From these results, it has been proven that the resulting music recommendations can be an alternative to the existing Collaborative Filtering-based recommender system.

Keywords: Recommender System; Music Recommender System; Content-based Filtering; Deep Neural Network; Convolutional Neural Network

1. INTRODUCTION

Music is one of the most popular forms of entertainment. Along with the development of information technology [1]. With the advancement of information technology, music streaming systems like Spotify, Apple Music, and Deezer are gaining popularity among users. Nevertheless, due to the vast number of songs accessible on these music streaming sites, customers frequently have challenges in discovering songs that align with their preferences. Thus, the implementation of a music recommender system is necessary as it can assist users in discovering songs that align with their interests [2]. A viable approach for a music recommender system is the utilization of the content-based filtering method [3]. This approach operates by examining the attributes of the music that users have expressed a preference for, and subsequently suggesting songs that possess comparable attributes. Content-Based Filtering approaches have been extensively utilized in numerous applications, such as music recommender systems.

Nevertheless, the Content-Based Filtering technique still possesses limitations in delivering precise recommendations. We address this issue by employing a Deep Neural Network (DNN) [4]. Deep Neural Networks (DNN) is a very efficient technique for handling extensive and intricate datasets. Therefore, it can be employed to analyze music data and offer more precise recommendations [5].

This project involves the development of a music recommender system by utilizing a blend of content-based filtering and Deep Neural Network methodologies. The objective of this project is to develop a customized music recommender system for users, taking into account their preferences. Additionally, the study aims to evaluate the precision of the developed music recommender system by employing a combination of content-based filtering and deep neural networks. We analyze the characteristics of the music that consumers enjoy and subsequently utilize Deep Neural Networks (DNN) to suggest tracks that align with their preferences. The method we plan to develop is expected to aid users in locating songs that align with their interests.

Multiple papers served as the primary sources of literature for this investigation. The writers utilized these references to get crucial information for the successful completion of their research. This research drew upon five primary studies, one of which was undertaken by Gunawan et al. [7]. Their study focused on the development of a genre-based music recommender system utilizing Convolutional Recurrent Neural Networks (CRNN). This study employs a combination of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) architectures within the CRNN network to tackle the issue of music recommendations. The Convolutional Neural Network (CNN) is employed to extract music features from audio recordings, whilst the Recurrent Neural Network (RNN) is utilized to simulate the intricate interactions among these elements. Niyazov et al. [8] conducted research on a content-based music recommender system. This system recommends songs to users based on the music content itself, including audio features and metadata. Content-based music recommender systems offer song suggestions that align with user tastes. Schedl [9] discusses the utilization of Deep Learning in music recommender systems. The study elucidates the utilization of advanced Deep Learning methodologies, namely Neural Networks and Convolutional Neural Networks (CNN), to enhance the caliber of music recommendations. Gharaei et al. [10] created a garment recommender system based on the content of the apparel using Deep Neural Network (DNN). The objective of this system is to provide

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

clothing recommendations to users by analyzing the characteristics of the apparel, such as product photos or descriptions. The research concludes that Deep Neural Network-based content-based clothing recommender systems have the capability to offer clothing recommendations that align with customer preferences. Martijn et al. [11] created a music recommender system that offers individualized explanations to users. The objective of this study is to augment users' comprehension of the suggestions offered by the system and to bolster users' confidence in the recommendations. The key references for constructing an optimal music recommender system involved five studies that utilized a combination of content-based filtering and deep neural network techniques.

Music is a genre of art that combines sound and rhythm to achieve artistic expression [6]. Music comprises various elements, such as melody, harmony, rhythm, lyrics, and instruments. Various civilizations and customs around the world have given rise to several styles of music. Common music genres encompass pop, rock, hip-hop, jazz, classical, country, and R&B, among other others. Every genre possesses distinct musical characteristics and techniques. Content-based recommender systems are a type of recommender system that utilizes data on products that consumers find enjoyable in order to suggest similar things [12]. Features or properties of things, such as genre, category, color, or material, play a crucial role in content-based recommender systems. The system acquires users' preferences by analyzing the items they express a liking for and then identifies similarities with the items stored in the database. The system thereafter suggests things that possess comparable characteristics to the items that customers prefer. A content-based recommender system is exemplified by an e-commerce website that suggests similar products to consumers based on their previous views or purchases [14]. Content-based recommender systems on music or video streaming sites might suggest songs or movies with similar genres or artists to the ones that users enjoy.

Deep Neural Network (DNN) is a machine learning model composed of multiple layers of neurons. Each neuron in the network receives input from the neurons in the previous layer and generates an output, which is then transmitted to the neurons in the next layer [15]. Deep Neural Networks (DNN) are widely adopted and very effective machine learning models utilized in several domains such as pattern recognition, facial detection, object identification, audio classification, and language translation. Deep neural networks (DNN) enable the gradual acquisition of more intricate feature representations at each layer of neurons, based on the input data[23]. The process ultimately leads to the creation of more abstract feature representations [16]. Subsequently, these feature representations are utilized to carry out classification or regression tasks on the provided data. The DNN learning process is conducted by the utilization of the backpropagation method, which computes the gradient of the error function and improves the model parameters to minimize the error.

Convolutional Neural Network (CNN) is a specialized Deep Learning model that is specifically built to handle grid-structured data, such as photographs or other spatial data [17]. CNN employs convolution operations to extract significant features from the input data and convolutional layers to progressively acquire intricate feature representations in a hierarchical manner [18]. CNN utilizes convolution techniques to extract fundamental attributes from input data, while convolutional layers acquire increasingly intricate feature representations in a hierarchical fashion. CNN possesses a superiority in the identification of intricate visual patterns and characteristics, as well as the ability to maintain consistency in the interpretation of incoming data[20]. CNN is highly valuable for computer vision tasks such as image classification, object identification, and picture segmentation[19].

2. RESEARCH METHODOLOGY

2.1 Research Stages

The music recommender system utilizes the Content-Based Filtering (CBF) approach, which is based on Deep Neural Network (DNN). This system offers song recommendations by analyzing the qualities and attributes of music that consumers have previously enjoyed [21]. The system primarily utilizes Deep Neural Network (DNN) technology for data processing and extraction of music features. A Deep Neural Network (DNN) is a type of machine learning algorithm that is capable of extracting intricate patterns from data, such as music features from individual songs. The CBF technique utilizes many music variables including as tempo, duration, genre, lyrics, and instruments to suggest new songs that share similar attributes and qualities. This system will assess each music and extract its characteristics using Deep Neural Network (DNN) technology. Subsequently, the computer will search for and suggest additional songs that possess music characteristics similar to those that the clients have found enjoyable.

The music recommender system utilizing the Content-Based Filtering (CBF) approach with Deep Neural Network (DNN) can offer highly individualized and precise recommendations by considering the user's own music likes and tastes. Moreover, this system has the capability to consistently enhance its recommendations by incorporating user comments and analyzing interactions with the suggested songs.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

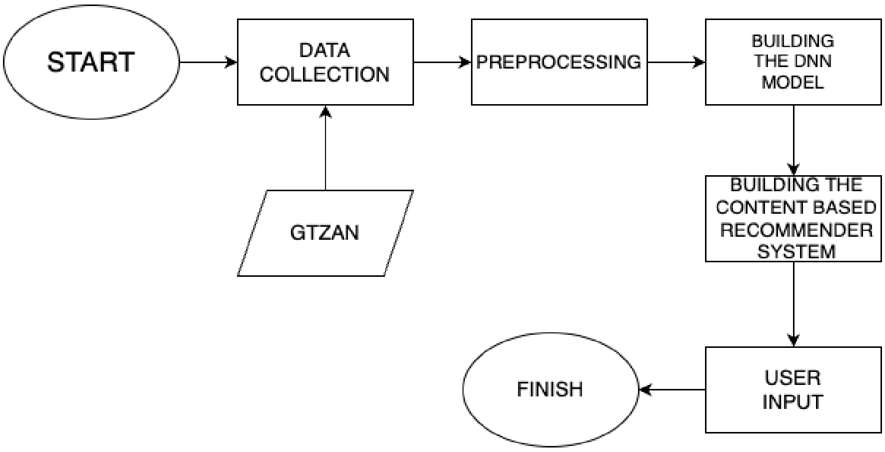


Figure 1. Research Stages

Figure 1 shows the flowchart of the entire research process and system design. The steps taken in developing this system are as follows:

a. Data Collection: This process involves gathering data to be used in the development of the music recommender system. The dataset comprises audio files and associated metadata, including genre, artist, and album details. The dataset utilized in this study is the GTZAN dataset acquired from a publicly available source (Kaggle).
b. Preprocessing: Following the acquisition of data, the subsequent phase involves the preparation of the data. This process entails cleansing the data, standardizing the audio files, and extracting characteristics from the audio data. Feature extraction involves utilizing techniques such as the Short-Time Fourier Transform (STFT) to transfer the audio signal from the time domain to the frequency domain. Audio data is analyzed to extract features such as Mel-frequency cepstral coefficients (MFCC), chroma feature, and spectral contrast.
c. Building the DNN Model: In this stage, a Deep Neural Network (DNN) model is created using the Convolutional Neural Network (CNN) approach. The CNN is employed to categorize music according to its genre by acquiring intricate feature representations from the audio input.
d. GTZAN: This step involves using the GTZAN dataset, which consists of 10 types of music genres with 25 audio files each. The dataset serves as the training and testing data for the DNN model.
e. Building the content-based recommender system: In this step, the framework for the content-based recommender system is created. The system analyzes the characteristics of the songs that users have liked and recommends songs with similar characteristics based on the extracted features.
f. User input: This step entails gathering input from users, such as their music tastes or songs they have expressed a liking for. The recommender system utilizes this input to deliver personalized music suggestions.

2.2 Dataset and Evaluation Method

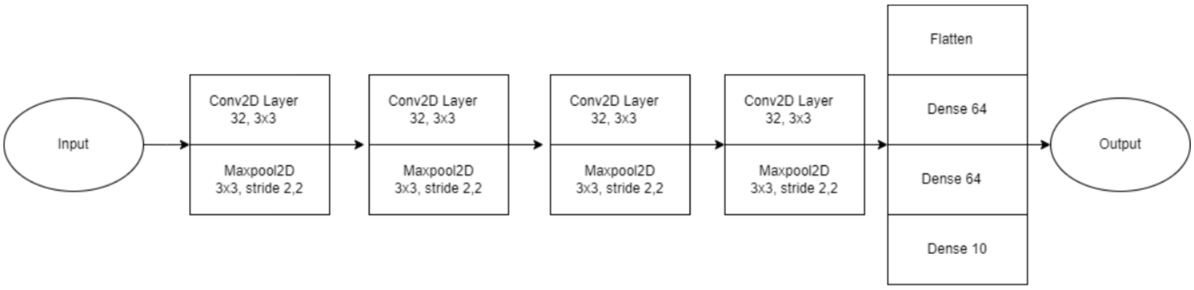
We used is the GTZAN dataset acquired from a publicly available dataset (Kaggle). The dataset comprises 10 distinct music genres, including classic, blues, alternative, hip-hop, disco, jazz, pop, metal, rock, and reggae. Each genre is represented by 25 audio recordings. Each audio file is a song that lasts for 30 seconds. It is recorded in mono channel, with a bit depth of 16 bits, and has a sample rate of 22,050 Hz. Table 1 explains the dataset.

Table 1. Dataset Explanation

No.	Genre	Number of Samples
1	Classic	25
2	Blues	25
3	Alternative	25
4	Hip-hop	25
5	Disco	25
6	Jazz	25
7	Pop	25
8	Metal	25
9	Rock	25
10	Reggae	25

Our research used a convolutional neural network (CNN) as the deep neural network method to assess and categorize music genres [22]. The CNN model begins by inputting the dataset into the classification process, which is composed of 4 convolutional layers and 4 pooling layers. The structure of the Convolutional Neural Network (CNN) method established in this study is illustrated in Figure 2.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762



Gambar 2. CNN Architecture [13]

The CNN architecture begins by processing the input dataset, which includes the outcomes of feature extraction from the audio data during the feature extraction stage. Afterwards, the procedure proceeds to the convolutional layer and applies the Rectified Linear Unit (ReLU) activation function. Next, the process proceeds to the max pooling layer and the dropout step. During the dropout process, the data will undergo dimensionality reduction as it is sent into the hidden layer, resulting in a conversion to a 1-dimensional representation. Subsequently, the audio data values will be fed into the softmax activation function. Following activation, they will proceed to the output layer, where data categorization is performed.

In this research, the system's performance is evaluated using the accuracy metric. The accuracy metric is employed to assess the precision of the recommendations generated by the system. During the review step, every user will be required to assign a rating to each recommendation using a scale ranging from 1 to 5. Recommendations that receive ratings from 1 to 3 will be categorized as unsuccessful, whilst recommendations that receive ratings of 4 or 5 will be labeled as successful. The classification results will be utilized to compute the accuracy value by applying equation

$$Accuracy = \frac{Successful\ Recomendations}{Sum\ of\ all\ recommendations} \quad (1)$$

3. RESULT AND DISCUSSION

Performance testing was conducted on the music recommender system that was developed in order to assess its accuracy, which serves as the principal metric for evaluating its efficacy. These testing procedures were conducted in accordance with the research phases and methodology that were designed. A comprehensive study was performed on each constructed and developed model to determine the most optimal model for our music recommender system.

3.1 Data Sample Visualization

The GTZAN dataset utilized in this work has a total of 250 audio samples, uniformly divided among 10 distinct music genres. Each genre is comprised of 25 typical samples. A visual representation of the genre distribution within the dataset may be shown in Table 1. Audio features, including Mel-frequency cepstral coefficients (MFCCs), chroma features, and spectral contrast, are derived from the audio samples through the utilization of the Librosa package. The dataset that has been extracted is presented in Table 2.

Table 2. Dataset Overview

filename	length	chroma_stft_mean	chroma_stft_var	rms_mean	…	label
blues.00000.wav	661794	0.3500881195	0.08875656873	0.1302279234	…	blues
blues.00001.wav	661794	0.3409135938	0.09498025477	0.09594780952	…	blues
blues.00002.wav	661794	0.3636371791	0.08527519554	0.1755704135	…	blues
blues.00003.wav	661794	0.4047847092	0.09399903566	0.1410930008	…	blues
blues.00004.wav	661794	0.3085260391	0.08784098178	0.0915287137	…	blues
blues.00005.wav	661794	0.3024562895	0.08753237873	0.1034936383	…	blues
blues.00006.wav	661794	0.2913279831	0.09398132563	0.1418741494	…	blues
blues.00007.wav	661794	0.3079547882	0.09290287644	0.1318221837	…	blues
blues.00008.wav	661794	0.4088792205	0.08651247621	0.1424164921	…	blues
…	…	…	…	…	…	…
rock.00098.wav	661794	0.3584013283	0.08588409424	0.05445402861	…	rock

The dataset is partitioned into training and test sets, wherein 80% of the data is allocated for training purposes and the remaining 20% is reserved for testing. This ensures that the performance of the model is assessed using data that has not been previously observed. Subsequently, four distinct models were constructed in order to evaluate our hypothesis.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

3.2 Performance Testing

The present analysis employs the GTZAN dataset obtained from Kaggle. The GTZAN dataset, sourced from a publicly accessible repository on Kaggle, comprises a heterogeneous assortment of 10 distinct music genres, namely classic, blues, alternative, hip-hop, disco, jazz, pop, metal, rock, and reggae. Each genre is represented by a total of 25 digital audio files. Each genre is comprised of 25 audio recordings. Every individual audio file consists of a 30-second song that is stored in a mono channel format, with a depth of 16 bits and a sampling rate of 22,050 Hz. This study aims to examine and evaluate four models in order to identify the most suitable model for the construction of a music recommender system. The specified conditions are as follows:

a. Model 1: Model 1 comprises four layers, each with a dense weight of (256, 128, 64, 10). The first to third layers employ the relu activator, while the fourth layer utilizes the softmax transformation. The model was trained for a total of 70 epochs.
b. Model 2: The Model 2 architecture comprises five layers, each with a dense block size of 512, 256, 128, 64, and 10. The first to fourth layers employ the relu activator, while the fourth layer utilizes the softmax activator. The dropout rate in the first to fourth layers is set at 0.2. Perform 100 epochs of model training using the Adam optimizer.
c. Model 3: The Model 3 architecture comprises five layers, each with a dense block size of 512, 256, 128, 64, and 10. The first to fourth layers employ the relu activator, while the fifth layer utilizes the softmax activator. The dropout rate in the first to fourth layers is set to 0.2. This model was trained for 700 epochs using the SGD optimizer.
d. Model 4: The Model 4 architecture comprises six layers, each with a dense layer size of 1024, 512, 256, 128, 64, and 10. The first to fifth levels employ the relu activator, while the sixth layer utilizes the softmax activator. The dropout rate in the first to fifth layers is set at 0.3. The model was trained iteratively for 500 epochs using the RMSProp optimizer.

Epochs denote the frequency at which the complete training dataset is iteratively processed by the model throughout the initial learning phase. During each epoch, the model is given the opportunity to adaptively modify its internal parameters with the aim of minimizing mistakes. Nevertheless, a suboptimal number of epochs might lead to underfitting, a phenomenon in which the model is unable to effectively capture the inherent patterns present in the data. In contrast, an excessive number of epochs may result in overfitting, a phenomenon in which the model becomes accustomed to the training data but encounters difficulties in generalizing to novel, unseen instances.

Optimization techniques, such as Stochastic Gradient Descent (SGD) or Adam, dictate the manner in which the parameters of the model are modified throughout the training process. Various optimizers demonstrate distinct features with regards to their convergence speed, stability, and capacity to overcome local minima. The efficacy of a given optimizer may be impacted by the individual dataset and model architecture being evaluated.

The simultaneous variation of both the number of epochs and the optimization approach presents a challenge in disentangling their respective contributions to the performance of the model. The superior outcome observed could potentially be attributable to a lucky mix of epoch count and optimizer, rather than being solely determined by another ideal model configuration.

Through the implementation of an optimized optimization approach and a systematic increase in the number of epochs, it becomes possible to observe the evolution of the model's performance over time. This analysis enables the determination of the most suitable number of epochs, at which the model attains an optimal equilibrium between underfitting and overfitting..

In contrast, by maintaining a constant number of epochs and conducting experiments with various optimization methods, one can ascertain the comparative efficacy of each optimizer in relation to the specific task at hand. This process enables the identification of the most appropriate optimizer for the given dataset and model architecture.

The performance of each model will be assessed through testing on the dataset, with a focus on analyzing the metrics of Accuracy. An ideal model will be employed for the construction of the music recommender system.

3.3 Implementation and Testing

In this study, we successfully deployed the system and conducted performance testing on the developed music recommender system. The implementation results are detailed in the subsequent sub-sections.

3.3.1 Implementation

We utilized the findings from our research to create a Python program, which was executed using a Jupyter notebook. The purpose of this program was to optimize performance and accelerate the runtime of the Convolutional Neural Network (CNN) model. The model was responsible for processing dataset files that contained audio files from various music samples. The ultimate goal was to recommend music that aligns with the user's preferences. We utilize beats per minute (BPM) and audio wave analysis to identify commonalities in rhythm and accompaniment across different music samples, as each song possesses its own distinct qualities. The objective is to provide recommendation outcomes

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

that encompass all music genres present in the dataset, rather than being restricted to a single genre. Figure 3 illustrates the general sequence of steps involved in the program's implementation.

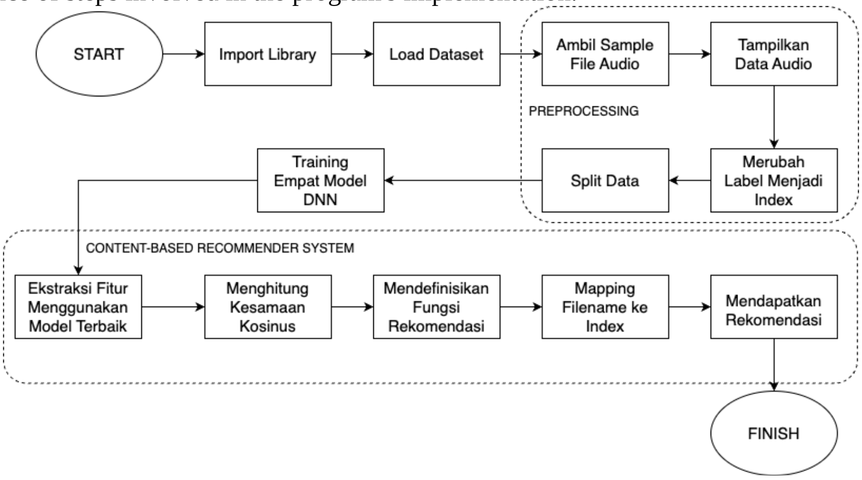


Figure 3. Implementation Flow

The method commences by importing diverse libraries necessary for data analysis and audio signal processing. The imported libraries comprise numpy for numerical calculation, pandas for data manipulation, matplotlib and seaborn for data visualization, librosa for audio signal processing, and sklearn for machine learning methods and evaluation measures. It is crucial to import these libraries to guarantee that all required functions and methods are accessible for use in the next operations. Now, we need to import the dataset that will be utilized for the analysis. The dataset typically consists of a CSV file that contains diverse attributes taken from the audio samples. The dataset is imported using the pandas library, which facilitates convenient manipulation and analysis of data presented in a tabular format. Verifying the accurate loading of the dataset is an essential preliminary measure before to conducting any further research. Once the dataset has been loaded, the subsequent phase involves selecting sample audio recordings for subsequent processing. The purpose of these audio files is to serve as illustrations for showcasing different signal processing techniques that will be implemented. The librosa library is utilized for the sampling of audio files, providing a convenient means of loading and manipulating audio data. Once the audio files have been sampled, the subsequent task is to provide the audio data. The audio data can be represented as waveforms or frequency spectra, which offer a comprehensive depiction of the audio signal's properties. The visualization is crucial for comprehending the patterns and characteristics inside the audio signal, which will subsequently be utilized in the process of extracting features and doing additional analysis.

Next, we will proceed to train four Deep Neural Network (DNN) models utilizing the prepared data. This model training process entails utilizing training data to instruct the model in identifying patterns and characteristics within the data. Subsequently, the model that has undergone training will be utilized to provide predictions and offer recommendations by leveraging the available data. Training the deep neural network (DNN) model demands a substantial amount of time and computer resources, although it is essential for achieving a precise and dependable model. In order to ensure optimal performance of the trained model, the dataset is divided into two distinct portions: the training data and the test data. This partition is conducted to assess the efficacy of the model and guarantee that it does not excessively match the training data. The training data is utilized to train the model, whereas the test data is employed to assess the model's performance and evaluate its effectiveness. During this stage, the labels in the dataset are transformed into indexes to simplify subsequent processing. The indexes will serve to identify and map elements in the collection, specifically audio files and their respective attributes. Transforming labels into indexes also facilitates the computation of similarities and the generation of recommendations.

Once the model has been trained, the subsequent stage involves doing feature extraction using the model that has been chosen as the best. The collected features from the audio data are utilized to compute the similarity between items in the dataset. Precise feature extraction is essential for producing pertinent and top-notch recommendation outcomes. After obtaining the extracted features, the subsequent task involves computing the cosine similarity among the items present in the dataset. Cosine similarity is a metric utilized to quantify the degree of similarity between two objects by evaluating the derived features. The cosine similarity calculation findings will be utilized in the recommendation process. Once the cosine similarity has been computed, the subsequent task involves establishing the recommendation function. The function will accept input in the form of item indexes and provide recommendations for the most comparable items based on the computed similarity scores. The recommendation function serves as the central component of the recommender system, which will be utilized to offer ideas to consumers.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

In order to streamline the recommendation process, it is necessary to establish a correlation between filenames and indices inside the dataset. This mapping is achieved by constructing a dictionary that establishes a connection between filenames and their corresponding indices, as well as the reverse relationship. By utilizing this mapping, we may effortlessly acquire the index of a particular filename and employ it in the recommendation function. The last stage is acquiring suggestions for particular music or things based on their filenames. The filenames are transformed into indexes using the established mapping, which are subsequently utilized in the recommendation function to provide a list of suggested items. The recommendation results are subsequently displayed to the users, offering suggestions based on the computed feature similarities. We employ the music reggae.00019.wav as a reference to identify songs that have similarities with reggae.00019.wav, utilizing the most optimal model, model 3. Initially, we conduct feature extraction utilizing the keras library, followed by the computation of cosine similarity using the sklearn package. Subsequently, we input the song file reggae.00019.wav to identify five songs that exhibit the highest degree of similarity to the aforementioned song. Figure 4 displays the suggestion findings.

```
Recommended items indices: [208, 806, 987, 598, 197]
Recommended items filenames: ['country.00008.wav', 'reggae.00006.wav', 'rock.00087.wav', 'jazz.00098.wav', 'classical.00097.wav']
```

Figure 4. Recommender System

Figure 4 illustrates the songs that exhibit the most similarity to reggae.00019.wav, namely country.00008.wav, reggae.00006.wav, rock.00087.wav, jazz.00098.wav, and classical.00097.wav.

3.3.2 Performance Testing

During the music selection process, we evaluate performance by utilizing the Accuracy metric. This metric is selected to assess the precision of the recommendation outcomes generated by the utilized model. The accuracy results are displayed in Table 3.

Table 3. Performance Result

No	Model Name	Accuracy (%)
1	Model 1	70.588
2	Model 2	71.569
3	Model 3	73.529
4	Model 4	70.588

In table 3, there are two columns, namely Model and Accuracy. Model is the name of the test model. Accuracy is the percentage of correct predictions out of all predictions made by the model. The higher the accuracy value, the better the model performance. The table 3 explains that model 3 is the model with the lowest loss value and the highest accuracy. Therefore, model 3 is the best model in this test. In a recommender system, the more accurate and closer the recommendation value is to the desired actual value, the better the recommender system. With an accuracy value of 73.529, the model is suitable for providing music recommendations to users.

3.4 Discussion

The accuracy results demonstrate that Model 3, with an accuracy of 73.529%, outperforms the other models in providing accurate music recommendations. The effectiveness of a recommender system hinges on the closeness between recommended and actual user preferences. The high accuracy of Model 3 suggests its suitability for generating relevant music suggestions.This research can be compared to many prior studies to evaluate the benefits and contributions it has produced. For instance, the study conducted by N. Ula, C. Setianingsih, and R. A study conducted by A. Nugrahaeni in 2021, titled "Sistem Rekomendasi Lagu Dengan Metode Content-Based Filtering Berbasis Website," demonstrates the utilization of the content-based filtering technique for music suggestions. However, the previous research did not include deep learning models, unlike our research, which has demonstrated superior accuracy in providing suggestions.

The variation in epochs and optimization methods across the models was intentional, aiming to explore different training strategies and their impact on performance. However, a more controlled comparison could be achieved by keeping one of these factors constant while varying the other. For instance, training all models with the same number of epochs but different optimizers, or vice-versa, would provide a clearer understanding of the individual contributions of these factors.

This research contributes to the field of music recommender systems by demonstrating the effectiveness of deep learning models, particularly CNNs, in generating accurate and personalized recommendations. The integration of content-based filtering with DNNs allows for capturing complex relationships within music data, leading to improved recommendation quality.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

4. CONCLUSION

This study effectively created a music recommender system that utilizes deep learning methods based on content analysis. The system utilizes a Convolutional Neural Network (CNN) along with content-based filtering to analyze different music properties and provide individualized recommendations. The empirical evaluation findings clearly and indisputably illustrate the usefulness of the system, with Model 3 having an impressive accuracy rate of 73.529%. The remarkable performance of deep learning models highlights their potential to completely transform music recommender systems. Despite its success, the study acknowledges specific limits that offer many opportunities for future investigations. The GTZAN dataset has a limited amount of genres and samples, which could limit the model's capacity to properly generalize to a wider range of musical styles. Future study could involve expanding the dataset to include a wider range of genres and samples. This would enhance the training data of the algorithm and potentially improve the accuracy of recommendations. Another potential area to explore involves tackling the computational difficulties that arise when developing deep learning models. These models are well-known for their significant computational demands. Therefore, exploring optimization methods such as transfer learning or model compression could be crucial in reducing the computing load and speeding up the training process, thereby making the system more efficient and accessible. In addition, future study could enhance the evaluation of the model's performance by using supplementary assessment measures such as precision, recall, and F1-score, in order to provide a more thorough and detailed analysis. These measurements would provide a more detailed comprehension of the model's strengths and limitations, allowing for specific enhancements and adjustments. Assessing the user experience of the music recommender system is of great importance. Obtaining direct input from end-users regarding the quality and relevancy of recommendations would provide vital insights into the practical effectiveness of the system. Ensuring user satisfaction and engagement is of utmost importance, since this feedback can provide valuable insights for improving and refining the system in future iterations. Overall, this research provides a strong basis for future progress in the field of deep learning-based music recommender systems. The combination of content-based filtering with CNNs has shown significant potential in providing precise and tailored music suggestions. These developments can greatly enhance the user experience on music streaming platforms, allowing consumers to easily find music that aligns with their personal tastes and preferences. By acknowledging and resolving the constraints indicated in this study and persistently advancing in this captivating domain, we can expect the development of progressively refined and efficient music recommender systems that bring joy and inspiration to music enthusiasts throughout the globe.

REFERENCES

[1]. Z. R. Maulida and M. K. Gowi, "ANALISIS PENERAPAN MODEL BUSINESS APLIKASI STREAMING MUSIK DI INDONESIA," *Jurnal Ilmu Komunikasi*, vol. 15, no. 2, pp. 153-164, 2017.

[2]. R. Amanda, "Music Streaming Dalam Industri Musik Era Industri 4.0," *Jurnal Studi Komunikasi (Indonesian Journal Of Communications Studies)*, vol. 6, no. 1, pp. 358–382, Mar. 2022, doi: 10.25139/Jsk.V6i1.3772.

[3]. N. Ula Ula, Setia Ningsih, and R. A. Nugraha, "Sistem Rekomendasi Lagu Dengan Metode Content-Based Filtering Berbasis Website," *Jurnal Coding*, vol. 6, no. 1, pp. 358-362, Mar. 2022, doi: 10.25130/coding.v6i1.3772.

[4]. F. Aine Sanjaya, P. Pramunendar, A. Sujada, and A. Fergiani, "Prediksi penyakit jantung menggunakan metode deep neural network dengan algoritma backpropagation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 10, pp. 5574-5581, 2019.

[5]. K. Nugrohes, M. I. Fanany, and R. Rahim, "Sistem Rekomendasi Pemilihan Jurusan Menggunakan Metode Backpropagation Neural Network," in *Prosiding Seminar Nasional Teknologi Informasi dan Aplikasinya*, 2018.

[6]. K. Choi, A. Fazekas, and M. Sandler, "Automatic music tagging using deep convolutional neural networks," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, Suzhou, China, 2017, pp. 205-212.

[7]. Adiyansjah, A. A. S. Gunawan, and D. Suhartono, "Music Recommender System Based On Genre Using Convolutional Recurrent Neural Networks," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 99–109. doi: 10.1016/J.Procs.2019.08.146.

[8]. A. Niyazov, E. Mikhailova, and O. Egorova, "Content-based music recommendation system," in *2021 29th Conference of Open Innovations Association (FRUCT)*, May 2021, pp. 274-279. doi: 10.23919/FRUCT52343.2021.9487881.

[9]. M. Schedl, "Deep Learning In Music Recommendation Systems," *Frontiers In Applied Mathematics And Statistics*, vol. 5. Frontiers Media S.A., Aug. 29, 2019. doi: 10.3389/Fams.2019.00044.

[10]. N. Yarahmadi Gharaei, C. Dadkhah, and L. Daryoush, "Content-Based Clothing Recommender System Using Deep Neural Network," in *26th International Computer Conference, Computer Society Of Iran, Csicc 2021*, Institute Of Electrical And Electronics Engineers Inc., Mar. 2021. doi: 10.1109/Csicc52343.2021.9420544.

[11]. M. Martijn, C. Conati, and K. Verbert, "'Knowing me, knowing you': personalized explanations for a music recommender system," User Modeling and User-Adapted Interaction, vol. 32, no. 1–2, pp. 215–252, Jan. 2022, doi: 10.1007/s11257-021-09304-9.

[12]. R. Restu Hadi And R. Ferdian, "Analisis Komponen Melodi, Ritme, Dan Harmoni Dalam Lagu 'Tan Malaka' Karya Geliga: Tinjauan Musik Konvensional," *Jurnal Sendratasik*, Vol. 12, No. 4, P. 497, Dec. 2023, Doi: 10.24036/Js.V12i4.125764.

[13]. H. H. Arfisko and A. T. Wibowo, "Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan Content-based Filtering," *eProceedings of Engineering*, vol. 9, no. 3, 2022.

Building of Informatics, Technology and Science (BITS)
Volume 6, No 2, September 2024, Page: 1111−1119
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v6i2.5762

[14]. Y. Mu and Y. Wu, "Multimodal movie recommendation system using deep learning," Mathematics, vol. 11, no. 4, p. 895, 2023.

[15]. A. Tjahyanto And F. J. Atletiko, "Peningkatan Kinerja Pengklasifikasi Objek Bawah Laut Dengan Deep Learning," Matrik : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer, Vol. 21, No. 3, Pp. 753–760, Jul. 2022, Doi: 10.30812/Matrik.V21i3.1466.

[16]. S. S. Hosseini, M. Yamaghani, and S. P. Arabani, "A Review of Methods for Detecting Multimodal Emotions in Sound, Image and Text," *Journal of Applied Research on Industrial Engineering*, 2024.

[17]. R. B. Tosi, H. D. Mbura, and Y. R. Kaesmetan, "Implementasi CNN Dalam Mengidentifikasi Kematangan Cabai Berdasarkan Warna," *Indonesian Journal of Education And Computer Science*, vol. 2, no. 1, pp. 34-42, 2024.

[18]. A. S. Riyadi, I. P. Wardhani, and S. Widayati, "Klasifikasi citra anjing dan kucing menggunakan metode convolutional neural network (CNN)," in *Prosiding Seminar SeNTIK*, vol. 5, no. 1, pp. 307-311, Sep. 2021.

[19]. A. A. Santosa, R. Y. N. Fu'adah, And S. Rizal, "Deteksi Penyakit Pada Tanaman Padi Menggunakan Pengolahan Citra Digital Dengan Metode Convolutional Neural Network," *Journal Of Electrical And System Control Engineering*, Vol. 6, No. 2, Pp. 98–108, Feb. 2023, Doi: 10.31289/Jesce.V6i2.7930.

[20]. K. Rega, I. Christianto O, And H. Setiawan, "Implementasi Convolutional Neural Network Untuk Sistem Prediksi Pigmen Fotosintesis Pada Tanaman Secara Real Time," Pp. 2443–2229, 2018, Doi: 10.28932/Jutisi.V4i2.812.

[21]. F. Sofiyana, R. Andrian, And R. Safe'i, "Klik: Kajian Ilmiah Informatika Dan Komputer Mobilenet Untuk Identifikasi Skala Kerapatan Dan Transparansi Tajuk Pohon Daun Lebar," *Media Online*, Vol. 4, No. 3, Pp. 1850–1859, 2023, Doi: 10.30865/Klik.V4i3.1476.

[22]. A. Tiarasari And E. Haryatmi, "Penerapan Convolutional Neural Network Deep Learning Dalam Pendeteksian Citra Biji Jagung Kering," *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, Vol. 5, No. 2, Pp. 265–271, Apr. 2021, Doi: 10.29207/Resti.V5i2.3040.

[23]. J. Pons and X. Serra, "musicnn: Pre-trained convolutional neural networks for music audio tagging," arXiv preprint arXiv:1909.06654, 2019.