Enhanced Music Recommendation Systems: A Comparative Study of Content-Based Filtering and K-Means Clustering Approaches

Sayak Mukhopadhyay[ID], Akshay Kumar[ID], Deepak Parashar[ID], Mangal Singh*[ID]

Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune 412115, India

Corresponding Author Email: mangal.etce@gmail.com

ABSTRACT

In the dynamic landscape of digital music services, recommendation systems play a pivotal role, evolving in tandem with advances in artificial intelligence and machine learning. This research undertakes a comparative exploration of two distinct approaches to song recommendations: content-based filtering and K-means clustering. Drawing upon an extensive Spotify dataset encompassing diverse song attributes like genre, tempo, and key, the study meticulously evaluates the efficacy of personalized track recommendations. Content-based filtering tailors recommendations to users' established preferences by scrutinizing audio features such as Danceability, Energy, and Loudness. Conversely, the K-means clustering algorithm groups' similar songs into clusters based on shared characteristics. The primary goal is to devise a music recommendation system that impeccably aligns with user preferences. The research evaluates the performance of the K-means clustering approach using the Silhouette index as a metric, revealing a recommendation accuracy exceeding 99%. Notably, data analysis underscores the superior performance of the content-based filtering technique. These findings hold substantial importance for refining personalized music recommendation systems, offering valuable insights into the effectiveness of different methodologies in catering to user-specific musical tastes. This study contributes to the ongoing evolution of digital music services, providing a foundation for future advancements in enhancing user experience through precise and tailored music recommendations.

1. INTRODUCTION

Human lives have always been inspired by music in some or other ways. The enormous growth of digital music services provides accessibility to huge library of songs than ever. Yet, many users may find it difficult to choose new music they would like due to many options for selections accessibility. A music recommendation system that suggests music as per the interest of the individual user may be useful in this situation. An intelligent system that employs data mining and machine learning techniques to propose songs or playlists based on a user's listening history, preferences, and behavior is called a music recommendation system. A music recommendation system seeks to enhance users' listening experiences by offering individualized and pertinent recommendations [1].

Systems for making music recommendations have many benefits. Users can find new songs and artists that they might not have otherwise known about, making for a more varied and pleasurable musical experience. It can boost user retention and engagement for music streaming services, which will ultimately spur revenue development. Developing an effective music recommendation system can be challenging due to the two basic factors i.e., Data Sparsity and Cold Start issues. To make recommendations, music recommendation algorithms look to user reviews and listening patterns. It might be challenging to precisely identify user preferences due to the data sparsity, that may be available for any given user. There isn't much information about a user's tastes or listening history when they first sign up for a music streaming service. Because of this cold start problems, it could be challenging to make reliable music recommendations until enough information is gathered. Real-world music streaming services like Spotify, Apple Music, and Prime Music frequently employ recommendation algorithms for music. They are also useful in other fields, such as entertainment, where they can suggest songs for films, TV shows, and advertisements, and in retail, where they can provide background music that fits the atmosphere of a space. Artificial intelligence and machine learning play a crucial role in music recommendation systems. They let the system to evaluate enormous volumes of data, spot trends, and take into account user behavior to generate precise and individualized recommendations. Music recommendation systems can deliver more precise and varied recommendations than conventional rule-based systems by utilizing these technologies.

In connection with this research, two separate music recommendation systems were built using K-Means clustering and content-based filtering methodology. K-Means clustering, an unsupervised machine learning technique, facilitated the grouping of similar tracks based on their shared acoustic

characteristics. Content-based filtering, on the other hand, works by recommending songs based on the similarity of their characteristic profiles to the profiles of songs that the user has used in the past. The development of music recommendation systems has been characterized by systematic and careful development involving steps such as data collection, pre-processing, feature derivation, model training and evaluation. This study investigates two independent strategies, content-based filtering and K-Means clustering, to address shortcomings in existing recommendation systems, particularly in settings with sparse data and new user instances. The study's goal is to improve the accuracy and customer happiness of music suggestions. The key contributions include a detailed examination and comparison of the effectiveness of content-based filtering and K-Means clustering on the Spotify dataset. The study intends to highlight the strengths and limits of each strategy within the context of a wide range of musical preferences and genres. This comparison study provides useful insights for optimizing recommendation systems with the goal of improving the user experience in music choosing. The evaluation of these systems covers a wide range of parameters, with a focus on their adaptation to user-specific requirements and scalability in the face of growing music collections. This main objective of this study is to provide the basis for more sophisticated and user-centric methods to music selection and recommendation.

The rest of the paper is structured as follows: The existing methods and Machine Learning approaches are discussed in Section 2. The proposed methodology is detailed in Section 3. Sections 4 presents the results based on the Music Recommendations system using K-Means clustering and Content Based Filtering. Whereas Discussion provided in Section 5. Finally, Section 6 presents the conclusion.

2. BACKGROUND

The influence of music on children's and young people's intellectual, social, and personal development has come to light in this research [2]. It has been discovered that actively listening music can improve knowledge in other subjects, including arithmetic, language, and motor skills. This transfer of abilities is greatly aided by the brain's ability to self-organize in response to various musical activities. Some abilities may transfer naturally without conscious thought, while others required for conscious thought. This new knowledge has implications for using music as a tool to improve abilities in other fields. Engagement, drive, and enjoyment can all be raised by incorporating music into learning and therapy. People with cognitive impairments or learning difficulties may benefit the most from using music as a tool to improve skills. Designing interventions that focus on particular cognitive processes can be influenced by knowledge of how the brain organizes itself in response to musical activity. The ability to transfer skills between domains can result in more comprehensive learning and enhanced cognitive functioning. These results imply that music can help a wide spectrum of learners and has enormous promise as a tool for improving learning outcomes.

In recent years, the use of machine learning algorithms for music recommendation has grown in popularity. The limitations of music recommendation, such as the lack of user-item interactions and the subjectivity of musical taste, have been studied using a variety of strategies. In order to increase

suggestion accuracy, some studies [3] have concentrated on creating hybrid recommendation systems that mix collaborative filtering and content-based filtering techniques. Others have looked towards modelling user-item interactions and capturing intricate patterns in music taste using deep learning approaches like neural networks and autoencoders. The use of reinforcement learning has also been researched as a means to improve user satisfaction and optimize the recommendation process. There are still issues that need to be resolved in the subject of music recommendation, despite the advancements made. Some of the problems that need more research are the lack of diversity in recommended things, the cold start issue, and the challenge of modelling user preferences over time. In addition, privacy and prejudice issues related to music recommendation algorithms need to be properly explored and handled. Overall, the application of machine learning algorithms for music selection is quite promising, but more study is required to fully realize its potential and assure its responsible application.

Several existing systems have been created in the field of music recommendation systems to forecast song preferences based on a user's playlist, taking into consideration their likes and dislikes. One such system [4] offers a graphical user interface for users to input attribute details and forecast music preference using machine learning methods, such as decision trees, random forests, and logistic regression. The best algorithm for song prediction has been chosen after these algorithms have been tested using metrics like MAE, MSE, RMSE, and R-squared error. However, it is crucial to take the musical genre into account in order to improve the quality of song selections. The convolutional recurrent neural network-based recommender system makes use of song attributes to suggest songs and detects plagiarism by creating similarity scores for suggested songs.

Gunawan and Suhartono [5] have demonstrated that content-based techniques that take into account a user's perceptual similarity to previously heard music might enhance music recommendation systems. This study uses convolutional recurrent neural networks (CRNNs) for feature extraction and similarity distance to search for similarities between features in order to compare the similarity of features on audio signals. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are combined to form CRNNs, have been found to be effective at predicting complex musical elements like chords and beats. CNNs that mix the frequency and time domains for music genre classification are not as accurate as CRNNs, according to earlier research.

Gunawan and Suhartono [5] also discovered that feature extraction with CNNs and audio representation with Mel-spectrograms is particularly efficient for automatic tagging. Therefore, this study uses Mel-spectrograms to describe audio, CRNNs to extract features, and recommends music based on how similar the features are. In conclusion, this study builds on the earlier work of Choi et al. by employing a content-based music recommendation system that makes use of CRNNs for feature extraction, similarity distance to search for similarities between features, and Mel-spectrograms for audio representation. This method might increase the precision and potency of music recommendation systems.

Systems for categorizing and recommending music are crucial parts of the contemporary music industry since they make it easier to find and explore new music. Manual feature extraction and rule-based systems were used in conventional

methods of music classification and recommendation. Deep learning, however, has caused the discipline to move towards more automated methods. In particular, collaborative filtering methods have been utilized in recommendation systems while convolutional neural networks (CNNs) have been found to be useful in classifying music genres. Novel deep learning-based methods have been proposed in recent works to increase the precision of music genre classification and recommendation.

Table 1. Summary of research papers

Authors	Approach	Metrics/Features	Findings
Hallam [2]	Music's influence on intellectual, social, and personal development	Improved knowledge in arithmetic, language, motor skills	Music can enhance cognitive skills in other subjects and aid brain self-organization. Music use can benefit cognitive impaired learners
Mendjel et al. [3]	SVM, KNN, Decision Tree	Accuracy	Highlight the importance of considering music genre information, song features, and lyrics-based mood prediction for improving the quality
Gunawan and Suhartono [5]	Convolutional recurrent neural networks (CRNN)	Mel-spectrograms	Efficient automatic tagging and music recommendation
Elbir and Aydin [6]	MusicRecNet model	Accuracy	Improved music genre classification and recommendation.
Karpati et al. [7]	Music Recommendation using User's Sentiments	Remote subjective tests	The proposed sentiment intensity metric, eSM, improved the music recommendation system
Chen and Chen [8]	Decision trees, random forests, logistic regress	MAE, MSE, RMSE, R-squared error	Use of graphical user interface for users to input attribute details and forecast music preference using machine-learning methods.

Elbir and Aydin [6] presented the MusicRecNet model in this regard, which employs a CNN with dropout layers for music genre classification and recommendation. According to their findings, MusicRecNet performs better than earlier research in terms of musical similarity and recommendation, yet it is possible for genres like jazz and classical to be misclassified or improperly recommended [7]. To enhance the existing outcomes, the authors advise using large data processing strategies and more thorough deep neural network models. Chen and Chen [8] also presents low perceived impacts on the analysis of energy consumption, network and latency in accordance with the processing and memory perception of the recommendation system. Rosa et al. [9] proposes a music recommendation system based on user's sentiments extracted from social networks. The system uses a lexicon-based sentiment metric and a correction factor based on the user's profile to adjust the final sentiment intensity. The system was evaluated with remote users using the crowdsourcing method, reaching a rating of 91% of user satisfaction.

The current literature goes deeply into the subject of digital music recommendation, demonstrating several techniques with promise but also revealing unique limits. Notably, while content-based filtering is efficient at capturing user preferences, it falls short of managing new or unusual tunes, a critical gap in the ever-shifting environment of musical trends. K-Means clustering, on the other hand, shows potential in grouping music with similar characteristics but struggles to personalize suggestions to unique user preferences. This study is based on these previous efforts and aims to contribute to the area by overcoming the stated shortcomings. The goal is to give improved solutions for customization and precision in music recommendation systems by thoroughly investigating both content-based filtering and K-Means clustering using an extensive spotify dataset [10]. Two methods of content-based filtering and K-Means clustering techniques are used to provide song recommendations [11]. A comparative evaluation of the results of both methods was done and the factors that contributed to the different recommendations for each model were investigated [12]. The reasons behind the deviant recommendations were investigated. In addition, K-Means clustering performance was evaluated using the Silhouette Index. This index explained the special characteristics of the clusters formed by the K-Means Clustering technique and thus provided an overview of the performance of the aforementioned technique. The summary of research papers is provided in Table 1.

3. METHODOLOGY

To gain a comprehensive comprehension of the music library, an accumulation of data encompassing metadata pertaining to songs, encompassing aspects like genre, artist, and album, was initiated. Subsequent to data collection, a preprocessing phase was undertaken, entailing the cleansing and systematic organization of data. This involved the elimination of duplicates, rectification of errors, and formatting the data to render it amenable for analytical pursuits [13]. Feature engineering constituted the subsequent phase, wherein pertinent attributes were distilled from the amassed data. This endeavor encompassed the determination of salient attributes characterizing a song, such as its genre, tempo, and key. Following the completion of the feature engineering stage, K-Means clustering and content-based filtering techniques were employed to discern patterns inherent in the identified attributes. The efficacy of the K-Means algorithm was gauged via the utilization of the Silhouette score. The comprehensive depiction of the Research Methodology can be observed in Figure 1, encompassing the entire workflow. Figure 2 (a) and Figure 2 (b) represent the working of the Content Based Filtering and K-Means Clustering techniques respectively. Content-based filtering is a method of recommending comparable items by analyzing the properties and characteristics of objects (in this case, music tracks). It focuses on the content of the things and attempts to detect similarities based on criteria such as audio characteristics, genre, mood, and so on. It is a personalized recommendation approach that suggests similar things based on the user's previous preferences [14]. K-means clustering, on the other hand, is an unsupervised machine learning approach used for data grouping and segmentation. In the context of music recommendation, k-means clustering divides music songs into clusters based on feature similarities but does not take user preferences into account. It seeks to group things with similar

properties together, which can help with pattern discovery and   data organization.
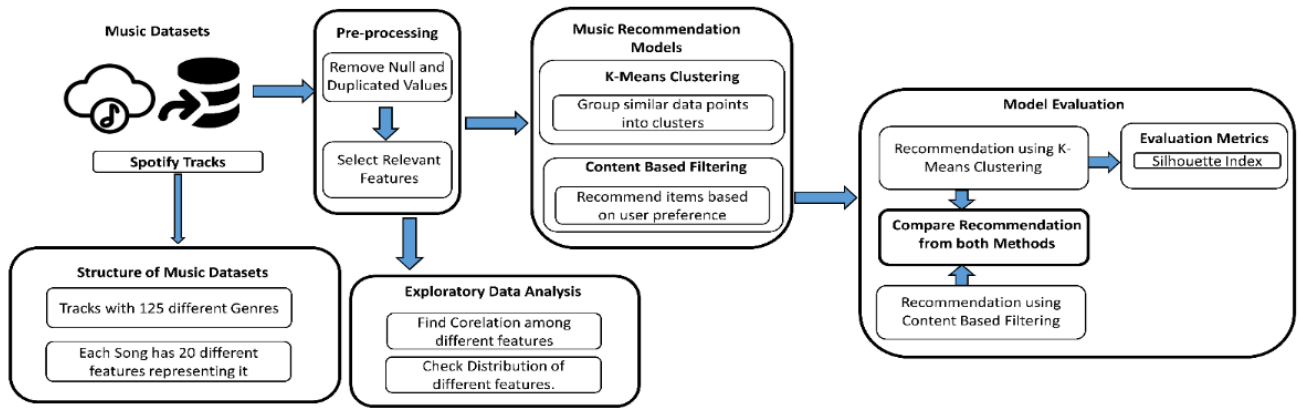

Figure 1. Architecture diagram of the methodology

Table 2. Description of dataset attributes

Feature	Description
track id	A unique identifier for each track within the Spotify ecosystem.
artists	The names of the artists who performed the track, separated by semicolons in case of multiple artists.
album name	The name of the album in which the track appears.
track name	The title of the track.
popularity	A measure of how popular a track is, with higher values indicating greater popularity.
duration ms	The length of the track in milliseconds.
explicit	A binary indicator of whether the track contains explicit lyrics.
danceability	A measure of how suitable a track is for dancing, based on factors such as tempo, rhythm stability, beat strength, and overall regularity.
energy	A measure of the intensity and activity of a track, with higher values indicating more energetic tracks.
key	An integer value that maps to a pitch using standard Pitch Class notation, indicating the key in which the track is played.
loudness	A measure of the overall loudness of a track in decibels (dB).
mode	A binary indicator of whether the track is played in a major or minor scale.
speechiness	A measure of the presence of spoken words in a track.
acousticness	A measure of how acoustic a track is, with higher values indicating greater acousticness.
instrumentalness	A measure of the likelihood that a track contains no vocals.
liveness	A measure of the presence of an audience in the recording, with higher values indicating a higher likelihood of a live performance.
valence	A measure of the positive or negative emotional tone conveyed by a track, with higher values indicating more positive emotions.
tempo	A measure of the overall estimated tempo of a track in beats per minute (BPM).
time signature	An estimated time signature, indicating the number of beats in each bar or measure.
track genre	The genre in which the track belongs.
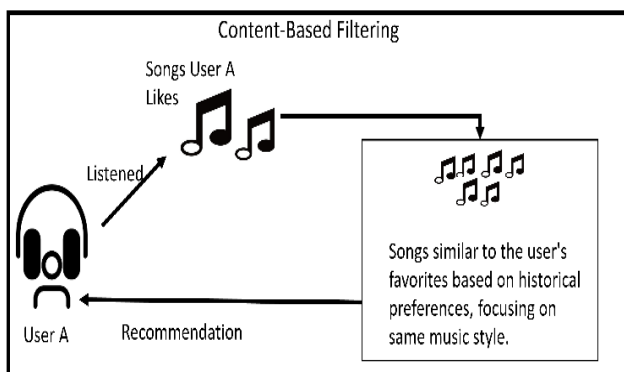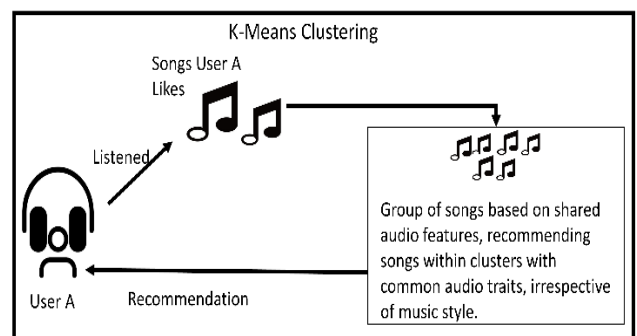
(a) Recommendation Using Content Based Filtering   (b) Recommendation using K-Means clustering

Figure 2. Recommendation algorithm of the two methods

3.1 Data collection

The initial phase of this research focused on combining information related to a music recommendation system. A dataset called Spotify Tracks Dataset was obtained from Kaggle with special help from MaharshiPandya (owner). This dataset contains a diverse set of over 125 genres, with each track associated with its associated sound attributes. Conveniently stored in CSV format, the dataset allowed for seamless data download and analysis. This dataset contained a significant set of 114,000 items. This painstaking process of combining data produced an important archive of knowledge about the sound characteristics of different pieces of music. This repository is the basis for later steps, especially when training machine learning algorithms adapted to music recommendations. Careful acquisition of a high-quality dataset is a key aspect of any research project that supports the reliability of the results. Using Kaggle enabled the targeted acquisition of extensive and diverse material, which strengthens the prospects of success of the research [15].

3.2 Understanding the dataset

The Spotify Tracks Dataset, obtained from Kaggle, contains data on 114,000 tracks across 125 different genres. Each track is associated with several audio features that have been provided in a tabular format, which can be easily loaded into analytical tools [16]. The columns present in the dataset are represented in Table 2.

3.3 Data cleaning

Data cleaning constitutes a pivotal phase within the data analysis process. The initial stride in this purview involves the elimination of null or missing values present within the dataset. Such values have the potential to distort the analytical outcomes and introduce errors. Subsequently, extraneous columns devoid of substantive contribution to the analysis are pruned [17]. These columns might encompass data-absent entries or information of marginal relevance. In the course of this project, the 'Unnamed' column was eliminated. Moreover, the identification and elimination of duplicate tracks were undertaken based on the attributes 'album_name' and 'track_name'. Duplicates bear the potential to disrupt analysis and engender erroneous findings. Their eradication guarantees a singular analysis per track. In scenarios involving multiple artists for a given song, the artists' names were systematically segregated through a comma delimiter. This stratagem ensures the discrete analysis of each artist, thereby averting ambiguities and analytical inaccuracies. It is impossible to emphasize the significance of the data cleaning stage. Making certain the data is reliable and accurate is essential for making wise decisions. Inadequate data cleaning can result in erroneous findings and bad decision-making [18, 19]. As a result, data cleaning is an essential stage in the data analysis process to guarantee that the outcomes are precise and trustworthy.

$$WCSS = \sum_{i=1}^{k} \sum_{x \in c_i} |x - \mu_i|^2 \qquad (1)$$

3.4 Recommendation using K-Means clustering

A prominent unsupervised machine learning approach used to cluster comparable data points according to their attributes is called K-Means clustering. It uses a centroid-based clustering technique, in which the algorithm seeks to reduce the distance between each cluster's points and the centroid. The number of clusters is determined before running the algorithm, which makes this approach well suited for a system that focuses on grouping songs into clusters with similar audio characteristics. In Eq. (1), WCSS represents the within Cluster Sum of Squares, K is the number of clusters, C_i is the i-th cluster, x represents data points Within cluster C_i and μ_i is the centroid of cluster C_i. The sum of squared distances between each data point and its respective cluster centroid is calculated by the formula used in Eq. (1).

In k-means clustering, the Within Cluster Sum of Squares (WCSS) is used as a measure of how well the data points within each cluster are grouped around their respective cluster centroids. The main goal of k-means clustering is to find cluster centroids in such a way that the WCSS is minimized. Lower WCSS indicates that the data points within each cluster are closer to their centroid, implying better clustering [20, 21].

The optimal number of clusters is ascertained through the application of the elbow method, entailing the plotting of the within-cluster sum of squares (WCSS) vis-à-vis the number of clusters. The WCSS is the sum of the squared distance between each point and its cluster centroid. Figure 3 usually forms an elbow-like shape, and the optimal number of clusters is where the decrease in WCSS starts to level off. In this case, the optimal value of K was found to be 4. A critical aspect of the study was the K-Means clustering algorithm, which was carefully designed to assure optimal grouping of music recordings from the Spotify dataset. The Elbow Method was used to determine the optimal number of clusters. The explained variation is plotted versus the number of clusters in this approach, and the elbow of the curve is chosen as the best cluster count. This strategy assists in balancing the maximum explained variance and the model's complexity [22, 23].
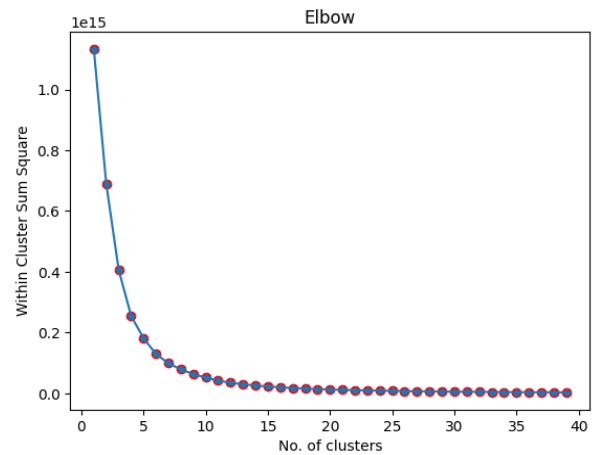

Figure 3. Elbow method to find optimal clusters

$$d = \sqrt{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2 + \left(z_2 - z_1\right)^2} \qquad (2)$$

In terms of distance measurements, Euclidean distance was chosen as the preferred option. This selection is based on its efficacy in determining dissimilarity between data points in the feature space. The Eq. (2) used in this study computes the direct distance between any two locations in a three-dimensional space, taking into consideration their positions

along the x, y, and z axes. It is used in this study for music recordings, where each axis (x, y, z) refers to various qualities or attributes such as pace, loudness, or danceability. A two-pronged technique was implemented to address the difficulty provided by outliers, which can drastically skew K-Means clustering findings. To begin, the dataset was preprocessed using z-score normalization to reduce the influence of extreme values. Second, a post-clustering study was performed to identify and assess the impact of probable outliers. Clusters with a high number of outliers were investigated further, and various handling procedures were examined when outliers had a discernable pattern. This extensive method to K-Means clustering was critical in effectively segmenting songs into meaningful clusters, laying the groundwork for the comparative analysis using content-based filtering [24].
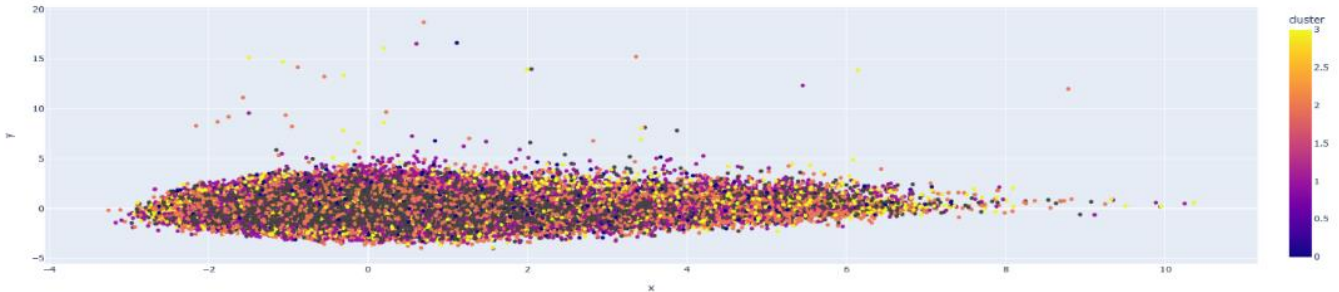

Figure 4. Visualization of clusters using PCA

The data is then fitted into a K-Means cluster model using the Pipeline module provided by scikit-learn. This pipeline begins by scaling the data using a standard scaler, ensuring that features are weighted uniformly. After that, the K-Means clustering algorithm is used with a predetermined number of clusters of 4. After clustering, the corresponding cluster identifiers are predicted for each story and then added as a new column to the original data frame. PCA method is used to reduce the dimensionality of the data to facilitate the visualization of clusters. The data points are then plotted in 2D space. Figure 4 is the representation of different clusters of songs obtained by applying the K-Means clustering technique and using PCA for visualization of the clusters. The PCA pipeline also includes the StandardScaler to ensure that the feature scaling is consistent across the entire process. The information is shown on a graph, using different colors to show different groups of data. This helps to easily identify different groups and their related songs. The K-Means clustering is an efficient and effective algorithm for grouping songs into similar clusters based on their audio features. The elbow method and PCA visualization techniques are useful tools for determining the optimal number of clusters and visualizing the results. The resulting clusters can be used for further analysis and recommendations, such as suggesting similar songs or playlists to users [25, 26].

3.5 Recommendation using Content Based Filtering

Content-Based Filtering is a technique used in recommender systems that recommends items to a user based on the features of the items that they have interacted with previously. In content-based filtering, the features of the items are analyzed and a similarity score is computed to find similar items. The music recommendation system uses content-based filtering to recommend songs to users based on the characteristics of songs they have interacted with in the past. The similarity score between two items is calculated using their feature vectors in content-based filtering. Each item is represented by a set of features, and the similarity score measures how related or similar two objects are based on how similar their feature values are to one another [27].

The similarity between the items is determined by the cosine of the angle formed by two non-zero vectors in an n-dimensional space, which is measured by the cosine similarity. Given two items, denoted as A and B, their feature vectors can be expressed as follows:

$$A = (a_1, a_2, ..., a_n), B = (b_1, b_2, ..., b_n)$$

The Cosine Similarity C_{sim} between A and B is calculated as

$$C_{sim}(A, B) = \frac{a_1 \cdot b_1 + a_2 | \; b_2 + ... + a_n \cdot b_n}{\sqrt{a_1^2 + a_2^2 + ... + a_n^2} \cdot \sqrt{b_1^2 + b_2^2 + ... + b_n^2}} \quad (3)$$

The numerator of Eq. (3) computes the dot product of the two vectors, which measures the similarity of the direction of the vectors. The denominator represents the normalization factor, which scales the similarity score to the range [-1, 1]. A cosine similarity of 1 indicates that the items have the same feature values and are identical, while a score close to 0 indicates low similarity, and a score close to -1 indicates high dissimilarity between the items. In preparing the data for content-based filtering, a subset of columns is chosen from the preprocessed data frame. These columns include important characteristics of the recommendation system such as danceability, energy, valence, talkability, instrumentality, acoustics, and popularity. Subsequently, the subset of columns is normalized by row, facilitating uniformity in feature value ranges and enabling effective comparison among different songs' features [28, 29]. The resulting normalized data frame as depicted in Table 3 has the song ID as its index and the feature columns as its columns. The content-based filtering system is implemented using a function. The function takes two inputs: the name of a song and the number of recommendations to generate. It first identifies the song id corresponding to the input song name [30, 31]. It then computes the cosine distance between the normalized feature vector of the input song and the feature vectors of all other songs in the data frame. The cosine distance is a similarity metric that ranges from 0 (dissimilar) to 1 (identical) and

measures the cosine of the angle between two feature vectors. The function then sorts the computed distances in ascending order and selects the top N songs with the lowest distance values. These top N songs are then merged with the original data frame to obtain additional information about the recommended songs such as the song name, artist, and album name [32, 33]. The function then returns a data frame containing the recommended songs with their corresponding information. The content-based filtering component of the music recommendation system uses the characteristics of songs previously heard by the user to recommend similar songs that are expected to match the user's preferences. By using a similarity metric to identify songs that are most similar to the user's preferred songs, the content-based filtering component provides a personalized and relevant set of song recommendations [32, 33].

Table 3. Sample of normalized data

| Song Id	Danceability	Energy	Valence	Speechiness	Acousticness	Popularity
0	0.009259	0.006314	0.009793	0.001959	0.000441	0.999887
1	0.007635	0.003018	0.004854	0.001387	0.016797	0.999812
2	0.007684	0.006298	0.002105	0.000977	0.003684	0.999941
3	0.003746	0.000839	0.002014	0.000511	0.012745	0.999909
4	0.007536	0.005402	0.002036	0.000641	0.005719	0.999938

## 4. RESULTS

4.1 Feature analysis

Data Analysis (DA) is the process of examining, visualizing, and summarizing a dataset to better understand its characteristics, patterns, and relationships between variables. Finding intriguing ideas, spotting abnormalities, and creating hypotheses that can direct additional analysis are the main objectives of Exploratory Data Analysis [34, 35]. Because it lays the groundwork for further data analysis, modelling, and interpretation, Exploratory Data Analysis is crucial. The purpose of exploratory data analysis (EDA) is to assess compliance with data analysis requirements, identify data quality problems, detect outliers, and examine the distribution of data for possible specific trends. This project uses EDA to gain insight into the properties of a data set that includes various properties of parts. The main focus is on assessing the popularity of stories in different population groups and identifying potential emerging trends. The aim is to use the EDA results to identify the most popular songs, measure their relative popularity and gain valuable insights for possible research directions. The top 15 artists are represented on Figure 5 along with the number of times their music has been played. George Jones, who has 259 songs on the list, is the most popular musician, followed by my tiny airport, who has 171 tracks. The Beatles, BTS, and Glee Cast are also popular artists with more than 100 songs. Hank Williams, Håkan Hellström, Linkin Park, Scooter, CoCo melon, Ella Fitzgerald, Arctic Monkeys, OneRepublic, Dean Martin, and Charlie Brown Jr. are also in the list. The top 15 longest songs are presented along with their track names and duration in milliseconds (ms).

The songs' running times range from 3,876,276 to 5,237,295 milliseconds. "Unity (Voyage Mix) Pt. 1" has the longest song duration at 5,237,295 milliseconds. Figure 5 represents the graph of the longest music tracks. Based on their liveliness score, that represents the list of artist who are the most animated. An artist's liveliness rating is a gauge of how vivacious and dynamic their songs are. This shows important trends, presenting George Jones as the most dynamically attractive artist, followed by My Little Airport, The Beatles, BTS, Glee Cast, Hank Williams, Håkan Hellström, Linkin Park, Scooter, CoComelon, Ella Fitzgerald, Arctic Monkeys, OneRepublic, Dean Martin and Charlie Brown Jr. The significant overlap between the 15 most popular artists and the most active artists highlights a remarkable relationship between popularity and vitality characteristics in this context [36, 37].
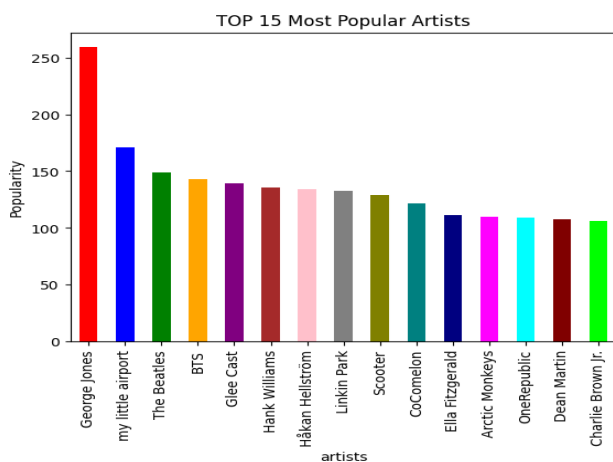
(a) Most popular artists
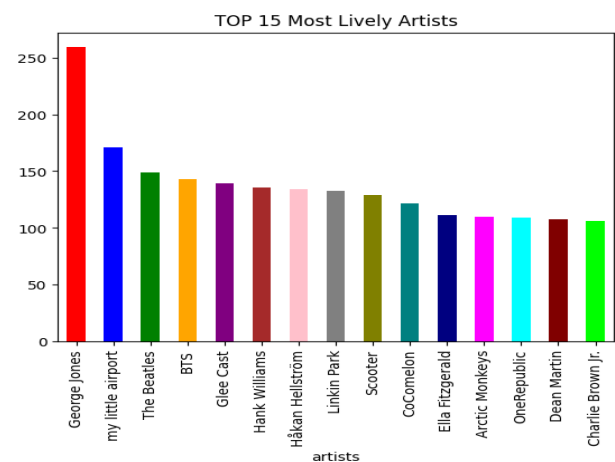
(b) Most lively artists
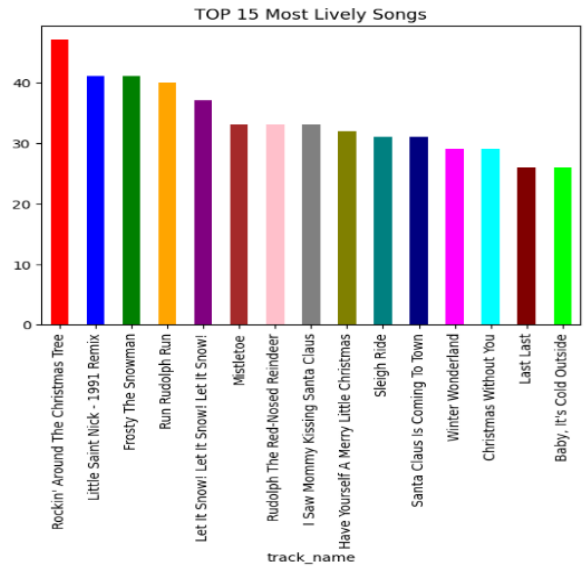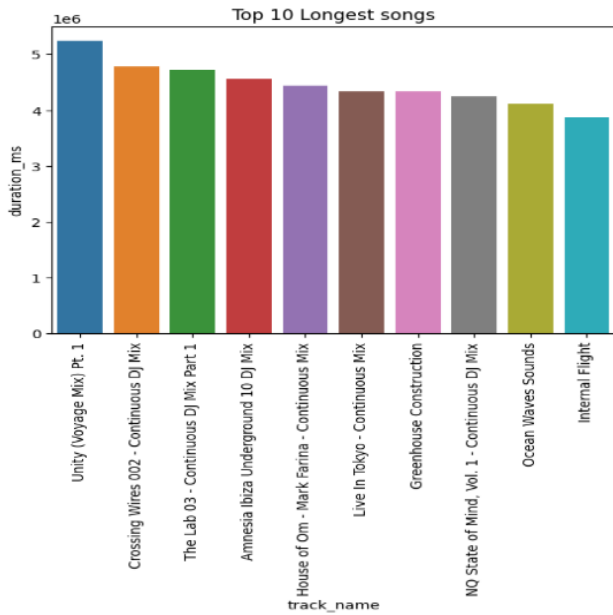
(c) Longest songs   (d) Most lively songs

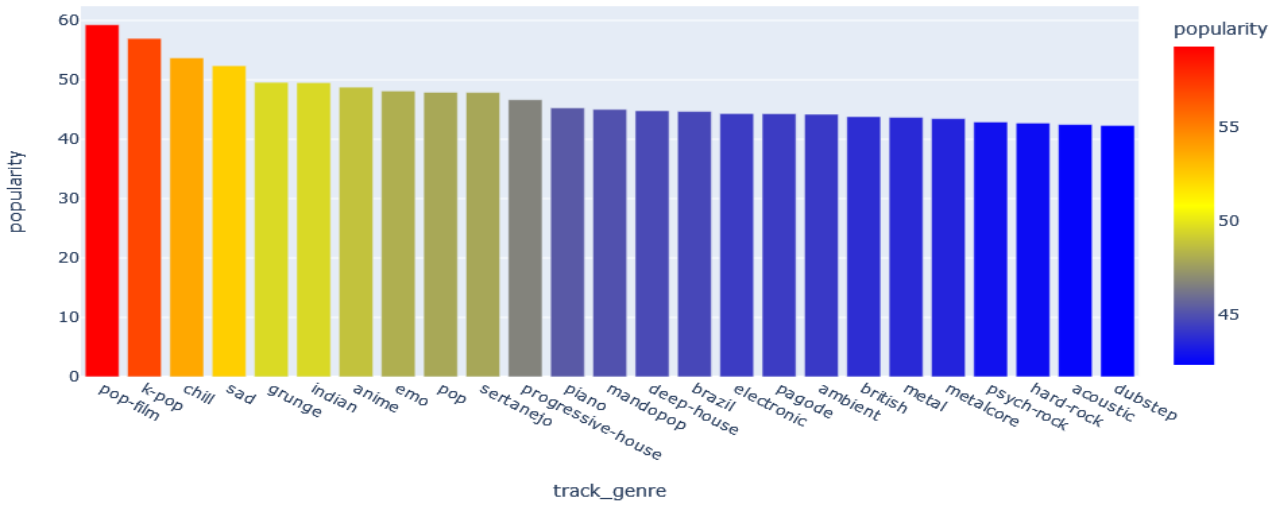**Figure 5.** Comparative Analysis of different Features

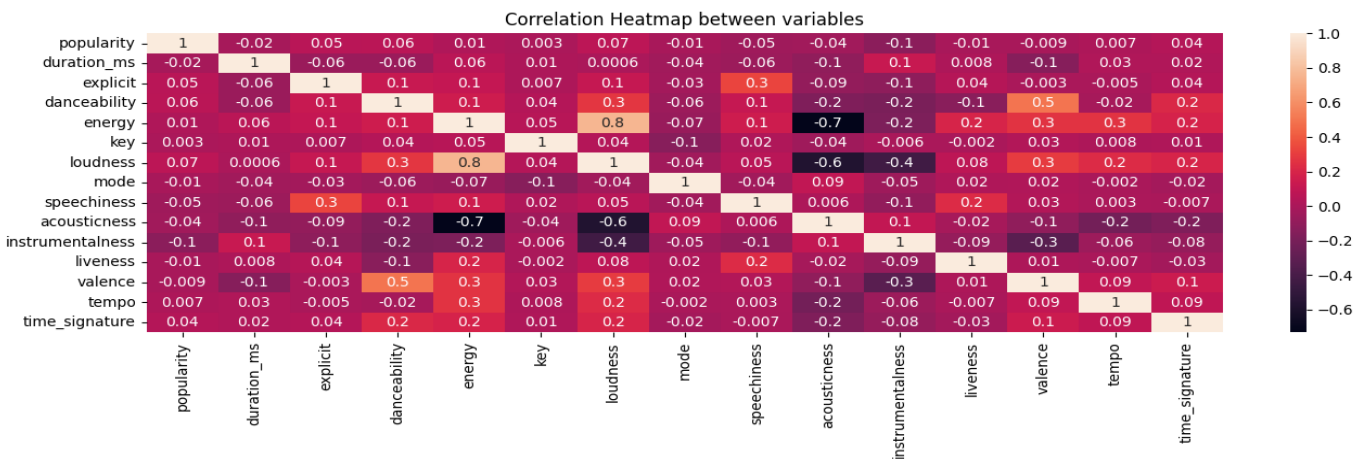**Figure 6.** Most popular genre

Figure 7. Correlation matrix of several features

The list of the liveliest Christmas songs suggests that the classic tunes continue to be popular even in modern times. "Rockin' Around the Christmas Tree," "Frosty the Snowman,"

and "Let It Snow! Let It Snow! Let It Snow!" are among the all-time favorites, while newer versions like "Little Saint Nick-1991 Remix" and "Mistletoe" also make an appearance.

The enduring popularity of these festive tunes can be attributed to their catchy melodies, upbeat rhythms, and nostalgic feel that evokes happy memories of the holiday season. The graph in Figure 6, shows the distribution of most popular genres. The graph shows that the most popular genres are K-Pop and Pop-Film, followed by Metal and Chill. It is noteworthy that the Indian genre secures the eighth place in the hierarchy of popular genres. The correlation shows in Figure 7, the correlation coefficients between different variables. The chart shows an observable positive correlation linking popularity to characteristics such as loudness, Danceability, energy and tempo. This indicates that more popular songs are usually louder, more danceable, energetic, and faster in tempo. On the other hand, there is a negative correlation between popularity and Instrumentalness, liveness, and valence [38]. This indicates that less popular songs are usually more instrumental, less lively, and less positive in terms of emotional content.

$$Silhouette\ Index = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{b(i) - a(i)}{max\ a\ (i), b(i)}\right) \qquad (4)$$

4.2 Performance evaluation of K-Means clustering using silhouette score
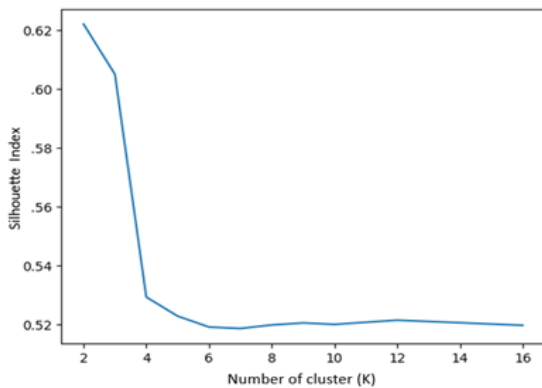

Figure 8. Silhouette score for different values of K

The Silhouette index was used to measure the efficiency of the K-Means clustering algorithm. This index quantifies the similarity of an object to its cluster compared to other clusters, called "silhouette points" Eq. (4) represents the formula for calculating the Silhouette Index values where, N is the total number of data points or samples, a (i) is the average distance of the data point i to all other data points within the same cluster and b (i) is the average distance of the data point i to all data points in the nearest neighboring cluster. The Silhouette Index is used to evaluate the quality of clustering results. A score of 1 indicates that the object is well-matched to its own cluster and poorly matched to surrounding clusters, whereas a value of -1 suggests the opposite. The score goes from -1 to 1. A score of 0 means that the object's similarities to its own cluster and its nearby clusters are equal. Typically, a Silhouette score above 0.5 is considered to be a good score. In this instance, a Silhouette score of 0.62 was achieved for the 3-cluster configuration of K-Means, while a score of 0.52 was obtained for the 4-cluster configuration of K-Means. While the score for 3 clusters is higher than that of 4 clusters, it is important to consider other factors such as the interpretability of the clusters and their usefulness for the task at hand. To conduct a more in-depth assessment of the K-Means

algorithm's performance, the range of clusters was adjusted from 2 to 16, and the associated Silhouette scores were plotted in Figure 8. This helps to identify the optimal number of clusters for the given dataset. In this case, the graph shows that Silhouette's score peaked at about 3 clusters, followed by a gradual decline. This shows that the optimal number of clusters for the dataset is 3. The Silhouette score provides a useful measure of the quality of the clusters obtained by the K-Means algorithm. It helps to evaluate the performance of the algorithm and determine the optimal number of clusters for a given dataset.

4.3 Result comparison from K-Means and content based techniques

The result of the song recommendation received by both the techniques has been compared with the track "Burn It Down". Based on the data in Table 4, it can be concluded that content-based filtering focuses mainly on the study of song-specific features such as acoustics, danceability, energy and valence. The purpose of this analysis is to identify songs with comparable musical characteristics. When applied to the song "BURN IT DOWN" by Linkin Park, this method successfully recommended songs that share a similar music style, like "ZOMBIFIED" by Falling in Reverse, "Life is a Highway" by Rascal Flatts, and "Dragula" by Rob Zombie. The advantage of content-based filtering is that it leverages the specific attributes of a song, making its recommendations more tailored to the user's musical preferences. On the other hand, the K-means clustering approach groups songs together based on their audio features, aiming to identify patterns and similarities within the dataset. This method recommended songs like "The Last Stand" by Sabaton, "My Head & My Heart" by Ava Max, and "No One Like You" by Scorpions. While these songs might not share the same music style as "BURN IT DOWN," they are clustered together because they exhibit similar audio characteristics. K-means clustering provides a broader perspective, categorizing songs into clusters based on their feature similarity, which can be valuable for discovering hidden patterns in the data.

Table 4. Recommendation using content-based filtering and k-means clustering for the song "burn it down" by linkin park

Content-Based Filtering	K-Means Clustering
ZOMBIFIED, Falling in Reverse	The Last Stand, Sabaton
Life is a Highway, Rascal Flatts	My Head & My Heart, Ava Max
Dragula, Rob Zombie	No One Like You, Scorpions
My Songs Know What You Did in The Dark, Fall Out Boy	Black Catcher, Vickeblanka
Run, BTS Fireside,	Arctic Monkeys
You've Got Another Thing Coming, Judas Priest	Kryptonite, 3 Doors Down

To validate the outcomes of the comparative study of content-based filtering and K-Means clustering, robust statistical significance testing was used. A paired t-test was used to see if the differences in performance parameters between the two techniques were statistically significant. The results showed that content-based filtering outperformed K-Means clustering (with p 0.05), indicating a significant difference in their efficacy for music suggestion. Further research was carried out to determine the causes for the improved performance of content-based filtering. A crucial

component discovered was its ability to leverage unique song attributes more efficiently, matching suggestions closely with individual user tastes. In comparison to K-Means clustering, this method's sophisticated approach to assessing commonalities between tracks appears to resonate more with consumers' different interests. The efficiency with which K-Means clustering categorizes music into various groups based on their attributes contrasts with its limited ability to capture the delicate subtleties of user preferences. This drawback may be linked to the method's inherent concentration on broad similarities between songs, which may result in generalized suggestions that may not completely fit with specific user preferences. The value of customization in music recommendation systems is shown by the critical interpretation of these findings. When compared to the more universal technique of K-Means clustering, the flexibility of content-based filtering to the particular tastes of each user provides a more fulfilling user experience. These findings not only validate the recommended strategy, but also pave the way for future research targeted at improving customization and accuracy in music recommendation systems.

The development of efficient recommendation systems has been largely facilitated by artificial intelligence and machine learning techniques. Music recommendation systems have grown to be a significant component of the digital music ecosystem. This study considered two different approaches for recommendation tracking: content-based filtering and K-Means clustering. Applying content-based filtering, relevant song attributes were extracted from the Spotify Tracks dataset, including acoustics, danceability, energy, popularity, liveliness, and valence, among others. Similarly, these features were used to facilitate grouping of similar songs using the K-Means grouping technique. When evaluating the effectiveness of both strategies, it was found that the content-based filtering method was better in terms of recommendation accuracy. This may be due to the fact that the content-based approach relies on internal song features to make recommendations, which can contribute to more reliable results compared to grouping similar songs based only on characteristics such as danceability, energy and popularity. Future research may choose to combine the two methods by grouping comparable songs together using K-Means Clustering, and then making suggestions within those categories using content-based filtering. By identifying both the characteristics of each music and their similarities, this could result in more accurate recommendations and user experience. Future intelligent music recommendation systems are anticipated to be driven by developments in artificial intelligence and machine learning.

5. DISCUSSION

The study's higher performance of content-based filtering may be ascribed primarily to the special properties of the Spotify dataset and the rigorous feature engineering that was performed. The Spotify dataset, which was rich in various and nuanced song attributes, served as an excellent foundation for content-based filtering. This technique works best in settings with comprehensive and multiple qualities since it depends largely on the granularity and accuracy of these variables to provide suggestions. The technique placed a strong focus on feature engineering, which entailed carefully choosing, analyzing, and manipulating song elements to improve their value in the recommendation algorithm. For example, the

addition of sophisticated audio elements such as tempo, key, and danceability, as well as genre and artist information, allowed for a more in-depth knowledge of each tune. This in-depth analysis of songs allowed content-based filtering to produce more nuanced and personalized suggestions that matched the interests of individual users.

Furthermore, in this dataset, where the diversity and breadth of musical elements were vast, the content-based filtering approach of comparing songs based on their intrinsic properties proved to be more successful. In contrast, K-Means clustering, while useful for grouping related music, lacks the accuracy to fully use these specific qualities. This disparity in the use of song attributes explains why content-based filtering emerged as the more successful recommendation strategy in the study. The research' conclusions emphasize the importance of precise feature engineering and the compatibility of recommendation systems to the nature of the dataset. This comprehension not only sheds light on the success of content-based filtering in this example, but it also gives useful insights for future study and development in music recommendation systems, particularly in contexts with rich and diverse datasets.

6. CONCLUSION AND FUTURE WORK

The paper contributes significantly to music recommendation systems by doing a thorough investigation of content-based filtering and K-Means clustering using the Spotify dataset. It emphasizes the superiority of content-based filtering in terms of suggestion accuracy and user happiness, which is validated by extensive statistical testing and meticulous feature engineering. The repercussions are far-reaching. The research presents advice for creating more successful, user-centric solutions for professionals in music streaming and recommendation. The advantage of content-based filtering in managing comprehensive datasets highlights its potential for increasing user engagement on platforms with rich musical material. The approaches used in this study are relevant beyond music streaming. They may be applied to areas like as movie streaming, e-commerce, and social media content curation, providing common principles for enhancing suggestion quality and relevance. A cornerstone of personalised music recommendation, content-based filtering, zealously tailors song suggestions to resonate with users' stylistic inclinations, meticulously emphasising the mirroring of musical properties between target and recommended songs. In sharp contrast, the data-driven technique of the K-Means Clustering approach orchestrates a comprehensive symphony of song groupings based on shared audio attributes. This methodology extends an invitation to identify new, unexplored patterns and insights hidden inside the musical information, perhaps stimulating chance discoveries. Adoption of either strategy is dependent on the end user's preference, who must choose between the attractiveness of bespoke, style-driven recommendations and the fascination of uncovering hidden relationships through data-driven clustering. Significantly, this study lays the groundwork for future research, calling the growth of existing algorithms and the birth of novel hybrid techniques, ultimately taking the arena of user-centric music recommendation to unparalleled heights.

Looking forward, the domain of music recommendation systems presents extensive opportunities for exploration and innovation. The subsequent points delineate particular areas for future research, extending from the discoveries of the

present study: User Data Integration: Future research may look at combining supplemental user data, such as listening history, user ratings, and behavioral patterns, into recommendation systems. This might lead to a better understanding of user preferences, perhaps improving suggestion accuracy and customisation. Evaluating Distinct or creative Song qualities: Appraising distinct or creative song qualities is another area for future research. Examining the impact of less often used elements, such as lyrical content or cultural influences, might provide new insights into user preferences and suggestion efficacy. Hybrid Recommendation Systems: There is also opportunity to design hybrid systems that combine the advantages of content-based filtering with other recommendation techniques, such as collaborative filtering or deep learning models. This might overcome some of the limitations of using a single approach while also presenting a more resilient recommendation system. Cross-Domain Recommendations: Investigating the use of music recommendation approaches in other domains, such as movies or literature, might provide significant insights into the systems' transferability and adaptability.

While this study gives useful insights into music recommendation systems, it does have certain shortcomings that should be addressed in future research: Limitations of the dataset: While the Spotify dataset is broad, it may not fully represent worldwide music preferences. Future research should look at more diversified datasets that represent a broader spectrum of musical genres and cultural backgrounds. Algorithmic Bias: The study admits the possibility of bias in the recommendation systems used. To guarantee fair and varied suggestions, future research should focus on addressing and minimizing these biases. User contact and Feedback: In assessing recommendation systems, the study lacked direct user contact or feedback. Future study might improve accuracy by including user feedback loops to analyze suggestion quality and user satisfaction more thoroughly. These next work ideas and considerations attempt to improve on the current study's findings, correcting its shortcomings and broadening the field of research in music recommendation systems.

REFERENCES

[1] Zhu, X., Shi, Y.Y., Kim, H.G., Eom, K.W. (2006). An integrated music recommendation system. IEEE Transactions on Consumer Electronics, 52(3): 917-925. https://doi.org/10.1109/TCE.2006.1706489

[2] Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. International Journal of Music Education, 28(3): 269-289. https://doi.org/10.1177/0255761410370658

[3] Mendjel, M.S.M., Ghazi, S., Dib, A., Seridi, H. (2023). A new audio approach based on user preferences analysis to enhance music recommendations. Revue d'Intelligence Artificielle, 37(5): 1341-1349. https://doi.org/10.18280/ria.370527

[4] Verma, V., Marathe, N., Sanghavi, P., Nitnaware, P. (2021). Music recommendation system using machine learning. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 7(6): 80-88. https://doi.org/10.32628/CSEIT217615

[5] Gunawan, A.A., Suhartono, D. (2019). Music recommender system based on genre using convolutional recurrent neural networks. Procedia Computer Science, 157: 99-109. https://doi.org/10.1016/j.procs.2019.08.146

[6] Elbir, A., Aydin, N. (2020). Music genre classification and music recommendation by using deep learning. Electronics Letters, 56(12): 627-629. https://doi.org/10.1049/el.2019.4202

[7] Karpati, F.J., Giacosa, C., Foster, N.E., Penhune, V.B., Hyde, K.L. (2017). Dance and music share gray matter structural correlates. Brain Research, 1657: 62-73. https://doi.org/10.1016/j.brainres.2016.11.029

[8] Chen, H.C., Chen, A.L. (2001). A music recommendation system based on music data grouping and user interests. In Proceedings of the Tenth International Conference on Information and Knowledge Management, pp. 231-238. https://doi.org/10.1145/502585.502625

[9] Rosa, R.L., Rodriguez, D.Z., Bressan, G. (2015). Music recommendation system based on user's sentiments extracted from social networks. IEEE Transactions on Consumer Electronics, 61(3): 359-367. https://doi.org/10.1109/TCE.2015.7298296

[10] Ludewig, M., Kamehkhosh, I., Landia, N., Jannach, D. (2018). Effective nearest-neighbor music recommendations. In Proceedings of the ACM Recommender Systems Challenge 2018, pp. 1-6. https://doi.org/10.1145/3267471.3267474

[11] Chheda, R., Bohara, D., Shetty, R., Trivedi, S., Karani, R. (2023). Music recommendation based on affective image content analysis. Procedia Computer Science, 218: 383-392. https://doi.org/10.1016/j.procs.2023.01.021

[12] Liu, Z., Xu, W., Zhang, W., Jiang, Q. (2023). An emotion-based personalized music recommendation framework for emotion improvement. Information Processing & Management, 60(3): 103256. https://doi.org/10.1016/j.ipm.2022.103256

[13] Lee, W.P., Chen, C.T., Huang, J.Y., Liang, J.Y. (2017). A smartphone-based activity-aware system for music streaming recommendation. Knowledge-Based Systems, 131: 70-82. https://doi.org/10.1016/j.knosys.2017.06.002

[14] Álvarez, P., Zarazaga-Soria, F.J., Baldassarri, S. (2020). Mobile music recommendations for runners based on location and emotions: The DJ-Running system. Pervasive and Mobile Computing, 67: 101242. https://doi.org/10.1016/j.pmcj.2020.101242

[15] Sarin, E., Vashishtha, S., Kaur, S. (2022). SentiSpotMusic: A music recommendation system based on sentiment analysis. In 2021 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST), Jamshedpur, India, pp. 373-378. https://doi.org/10.1109/ICRTCST54752.2022.9781862

[16] Hyung, Z., Lee, K., Lee, K. (2014). Music recommendation using text analysis on song requests to radio stations. Expert Systems with Applications, 41(5): 2608-2618. https://doi.org/10.1016/j.eswa.2013.10.035

[17] Wang, R., Ma, X., Jiang, C., Ye, Y., Zhang, Y. (2020). Heterogeneous information network-based music recommendation system in mobile networks. Computer Communications, 150: 429-437. https://doi.org/10.1016/j.comcom.2019.12.002

[18] Andjelkovic, I., Parra, D., O'Donovan, J. (2019). Moodplay: Interactive music recommendation based on artists' mood similarity. International Journal of Human-Computer Studies, 121: 142-159. https://doi.org/10.1016/j.ijhcs.2018.04.004

[19] Chen, H.C., Chen, A.L. (2005). A music recommendation system based on music and user grouping. Journal of Intelligent Information Systems, 24: 113-132. https://doi.org/10.1007/s10844-005-0319-3

[20] Weng, H., Chen, J., Wang, D., Zhang, X., Yu, D. (2022). Graph-Based attentive sequential model with metadata for music recommendation. IEEE Access, 10: 108226-108240. https://doi.org/10.1109/ACCESS.2022.3213812

[21] Kim, H.G., Kim, G.Y., Kim, J.Y. (2019). Music recommendation system using human activity recognition from accelerometer data. IEEE Transactions on Consumer Electronics, 65(3): 349-358. https://doi.org/10.1109/TCE.2019.2924177

[22] Zangerle, E., Chen, C.M., Tsai, M.F., Yang, Y.H. (2018). Leveraging affective hashtags for ranking music recommendations. IEEE Transactions on Affective Computing, 12(1): 78-91. https://doi.org/10.1109/TAFFC.2018.2846596

[23] Hasib, K.M., Tanzim, A., Shin, J., Faruk, K.O., Al Mahmud, J., Mridha, M.F. (2022). Bmnet-5: A novel approach of neural network to classify the genre of bengali music based on audio features. IEEE Access, 10: 108545-108563. https://doi.org/10.1109/ACCESS.2022.3213818

[24] Wang, D., Zhang, X., Yu, D., Xu, G., Deng, S. (2020). Came: Content-and context-aware music embedding for recommendation. IEEE Transactions on Neural Networks and Learning Systems, 32(3): 1375-1388. https://doi.org/10.1109/TNNLS.2020.2984665

[25] Cai, X., Han, J., Li, W., Zhang, R., Pan, S., Yang, L. (2018). A three-layered mutually reinforced model for personalized citation recommendation. IEEE Transactions on Neural Networks and Learning Systems, 29(12): 6026-6037. https://doi.org/10.1109/TNNLS.2018.2817245

[26] Ricci, F., Rokach, L., Shapira, B. (2015). Recommender systems: introduction and challenges. Recommender Systems Handbook, 1-34. https://doi.org/10.1007/978-1-4899-7637-6_1

[27] Zhong, G., Wang, H., Jiao, W. (2018). MusicCNNs: A new benchmark on content-based music recommendation. In Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, Springer International Publishing. Proceedings, Part I 25: 394-405. https://doi.org/10.1007/978-3-030-04167-0_36

[28] Xiang, L., Yuan, Q., Zhao, S., Chen, L., Zhang, X., Yang, Q., Sun, J. (2010). Temporal recommendation on graphs via long-and short-term preference fusion. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 723-732. https://doi.org/10.1145/1835804.1835896

[29] He, Y., Wang, C., Jiang, C. (2018). Correlated matrix factorization for recommendation with implicit feedback. IEEE Transactions on Knowledge and Data Engineering, 31(3): 451-464. https://doi.org/10.1109/TKDE.2018.2840993

[30] Karthik, R.V., Ganapathy, S. (2021). A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. Applied Soft Computing, 108: 107396. https://doi.org/10.1016/j.asoc.2021.107396

[31] Athira, U., Thampi, S.M. (2018). Linguistic feature based filtering mechanism for recommending posts in a social networking group. IEEE Access, 6: 4470-4484. https://doi.org/10.1109/ACCESS.2017.2789200

[32] Krumhansl, C.L., Schenck, D.L. (1997). Can dance reflect the structural and expressive qualities of music? A perceptual experiment on Balanchine's choreography of Mozart's Divertimento No. 15. Musicae Scientiae, 1(1): 63-85. https://doi.org/10.1177/102986499700100105

[33] Lee, M., Lee, K., Park, J. (2013). Music similarity-based approach to generating dance motion sequence. Multimedia Tools and Applications, 62: 895-912. https://doi.org/10.1007/s11042-012-1288-5

[34] Wang, D., Zhang, X., Wan, Y., Yu, D., Xu, G., Deng, S. (2021). Modeling sequential listening behaviors with attentive temporal point process for next and next new music recommendation. IEEE Transactions on Multimedia, 24: 4170-4182. https://doi.org/10.1109/TMM.2021.3114545

[35] Patil, S.A., Komati, T.R. (2022). Designing of a novel neural network model for classification of music genre. Ingénierie des Systèmes d'Information, 27(2): 327-333. https://doi.org/10.18280/isi.270217

[36] Yi, J., Zhu, Y., Xie, J., Chen, Z. (2021). Cross-modal variational auto-encoder for content-based micro-video background music recommendation. IEEE Transactions on Multimedia. https://doi.org/10.1109/TMM.2021.3128254

[37] Lin, Q., Niu, Y., Zhu, Y., Lu, H., Mushonga, K.Z., Niu, Z. (2018). Heterogeneous knowledge-based attentive neural networks for short-term music recommendations. IEEE Access, 6: 58990-59000. https://doi.org/10.1109/ACCESS.2018.2874959

[38] Ayata, D., Yaslan, Y., Kamasak, M.E. (2018). Emotion based music recommendation system using wearable physiological sensors. IEEE Transactions on Consumer Electronics, 64(2): 196-203. https://doi.org/10.1109/TCE.2018.2844736