ISSN 2959-6157

The Feature Importance Analysis of Music Recommendation System with K-means and K-Nearest Neighbors

Ruining Yang[1, *]

[1] NOIC Academy, ON, Canada
*Corresponding author: xi7919053@gmail.com

Abstract:
With the rapid development of science and technology, society has entered an era of high informatization. The recommendation system can alleviate the problem of information overload due to the vast and complex information on the Internet. Music recommendation is one of the main application fields of the recommendation system. This paper revolves around building a music recommendation system using Spotify's dataset. There are two main methods used in this paper to analyze the importance of features in music recommendation based on machine learning techniques. Specifically, this paper uses K-means clustering to identify similarities in the feature combination, bringing together songs with similar types. K-Nearest Neighbors (KNN) is used to find the nearest neighbours to a song by combining the selected features. In the evaluation part, to ensure the feature pairs are significant to KNN models, the accuracy is calculated and compared before and after one specific feature combination is removed. The results reveal that energy and valence are the most compelling feature combinations according to the cluster analysis. Besides, the accuracy after removing any feature is smaller than the accuracy using all features, which reflects that these features are essential for the KNN model. However, the feature combination (energy and valence) shows that the contradictory results indicate that any single analysis method for determining the importance of a feature is one-sided.

Keywords: Recommendation System; K-means; K-Nearest Neighbors; Machine Learning.

1. Introduction

In this era of big data, the redundant information on the Internet has significantly increased the cost for users to obtain valuable information. Currently, the recommendation system is essential in helping users get helpful information when faced with overloaded details in various fields, such as online shopping, book recommending, and music recommending. Specifically, a music recommendation system is essential to almost all music platforms. Users' experience and engagement can be enhanced. Personalized suggestions can better meet users' needs and improve customer satisfaction, creating a more enjoyable experience. For platforms, a recommendation system can drive business value. The recommended album will likely be bought because it fits the user's preferences and tastes.

There are three traditional types of recommendation systems: collaborative filtering, content-based filtering and hybrid approach. Collaborative filtering suggests items to users by analyzing the preferences and behaviours of other users. The fundamental concept is that individuals with similar tastes will likely have similar tastes based on previous information. There are two main types: user- and item-based filtering. Content-based filtering recommends items based on the characteristics of items that the user likes based on the user's past activities and interactions. Hybrid approaches combine two or more recommendation methods to take advantage of their advantages and compensate for their shortcomings [1-3].

Although these traditional methods have achieved success in building a music recommendation system, there exist some things that could be improved. For example, a significant issue with collaborative filtering is the cold-start problem. Besides, it is hard for collaborative filtering to work when there is a lack of user-item interaction data. While content-based filtering can avoid the cold-start problem, it may limit the novelty and diversity of user discovery because it tends to recommend songs that are too similar to songs that users already like. These limitations emphasize the need for more adaptable and complete strategies to build a music recommendation system, making more accurate and flexible recommendations. Previous works address the above challenge based on more complex models (e.g., neural networks) and external knowledge, ignoring the importance of existing features. Therefore, most of the existing work is computationally demanding and challenging to generalization [4-6].

This paper aims to understand the differences between different feature combinations for recommendation systems so researchers can build effective music recommendation systems based on limited features. Specifically, this paper leverages the unsupervised machine-learning algorithm, K-means, and groups data instances into clusters based on their features. Songs can be effectively clustered based on their specific feature combination, such as danceability and energy. When clustering with K-means, the Elbow Method is used to determine the optimal number of clusters. The scatter plot is visualized and compared to identify the most compelling feature pair for K-means clustering. In the evaluation part for K-means clustering, the clustering performance is assessed to ensure that the selected k value can achieve the best clustering effect. In this case, the recommender system can generate suggested songs with similar characteristics the user already likes. Meanwhile, the researchers can intuitively understand the effects of features. K-Nearest Neighbors (KNN) is a supervised learning algorithm used for classification and regression. In the music recommendation system, KNN can make customized suggestions by identifying the nearest neighbors to a specific user based on the distance between feature vectors. KNN is simple and does not need a training process; it just requires calculating the distance between two songs. In this paper, the importance of each feature combination is assessed by calculating and comparing the accuracy using all features and the accuracy after removing one feature.

The experimental results show that the feature combination (energy and valence) have the good performance in K-means, and the smallest performance effect in KNN classification. These contradictory results indicate that any single analysis method for determining the importance of a feature is one-sided.

The rest of this paper is organized as follows: section 2 describes the methodology, including data collection and preparation, feature selection and scaling, clustering with K-means, recommendation with KNN and evaluation. Section 3 presents the experimental results and analysis. Section 4 discusses the limitations and the recommendations for further study. Section 5 concludes the paper with a summary of key findings and the significance of the research.

2. Method

2.1 The Overview of Dataset

This paper selects Spotify as experimental dataset [4]. The Spotify dataset consists of information about many features of songs (e.g., audio features, artists, ID.) The missing values are handled, and the duplicates are removed to clean the data. Then, the data are converted to numeric type. All numeric columns are standardized to adjust the data on a similar scale.

2.2 Clustering with K-Means

The Elbow Method is used to identify the optimal number of clusters (the value of k). This method is achieved by plotting the sum of the squared distance between the data point and its clustered centroid (SSE) when the value of k changes. The optimal number of clusters is the k-value corresponding to the point where the SSE curve starts to flatten [5]. After the optimal number of clusters is determined, K-means clustering is applied to group songs into different clusters. These four audio features: 'danceability', 'energy', 'acousticness' and 'valence' are selected because they can describe the musical qualities of songs excellently and greatly impact on user preferences. The scatter plots of six feature pairs (danceability-energy, danceability-acoustics, danceability-valence, energy-acoustics, energy-valence, acoustics-valence) are graphed. The clusters quality is assessed to ensure that the subjectivity caused by spotting the elbow point does not significantly influence clustering. The Silhouette Coefficient, Calinski-Harabasz Index and Davies-Bouldin Index are used to identify the best way to partition a set of objects and determine the optimal number of clusters.

2.3 Feature Importance Analysis with KNN

This paper analyses the importance of features based on KNN. Specifically, this paper removes the specified features and observes the performance of KNN, similar to the enumeration method. Due to the KNN being a kind of lazy learning method, the performance change of KNN will not be affected by model parameters like a neural network; it can intuitively demonstrate the feature importance. In experimental settings, this paper removes six kinds of feature combinations mentioned in section 2.2, including danceability-energy, danceability-acoustics, danceability-valence, energy-acoustics, energy-valence, and acoustics-valence. This is so that the importance of the feature can be understood better than a single view by combining the results of K-means and KNN.

3. Experiment Results and Analysis

3.1 The Results of Clustering

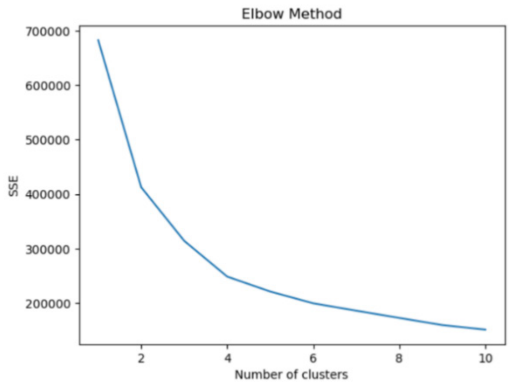

Fig. 1 The SSE results

This paper leverages the Elbow Method to identify the optimal number of clusters (the value of k). This method is achieved by plotting the sum of the squared distance between the data point and its clustered centroid (within-clusters sum of squared errors, SSE) when the value of k changes. The SSE can be computed by formulating (1):

$$SSE = \sum_{i=1}^{k} \sum_{p \in C} |p - m_i|^2 \tag{1}$$

where c_i is the i-th cluster, p is instance point in c_i, m_i is the average value of c_i and indicate the performance of cluster results. As shown in Figure 1, the value of k can be roughly determined as 5.

Table 1. Evaluation for Clustering

	Silhouette Coefficient	Calinski-Harabasz Index	Davies-Bouldin Index
k=5	0.2660	87467	1.2701
k=6	0.2611	82457	1.2275
k=7	0.2575	76132	1.2420

In addition, Table 1 compares the Silhouette Coefficient, the Calinski-Harabasz Index and the Davies-Bouldin Index when k = 5, 6 and 7, it is clear that the SC is closest to 1 and the CH Index is the biggest when k = 5. However, the DB Index is the smallest when k is 6. In comparison, the clustering performance is better when k = 5.

Figure 2 shows the visualization of different feature combinations (danceability-energy, danceability-acoustics, danceability-valence, energy-acoustics, energy-valence, and acoustics-valence) in feature space. Energy and valence are the most compelling features of K-means clustering. This paper further discusses the results of Figure 1 in section 3.2.

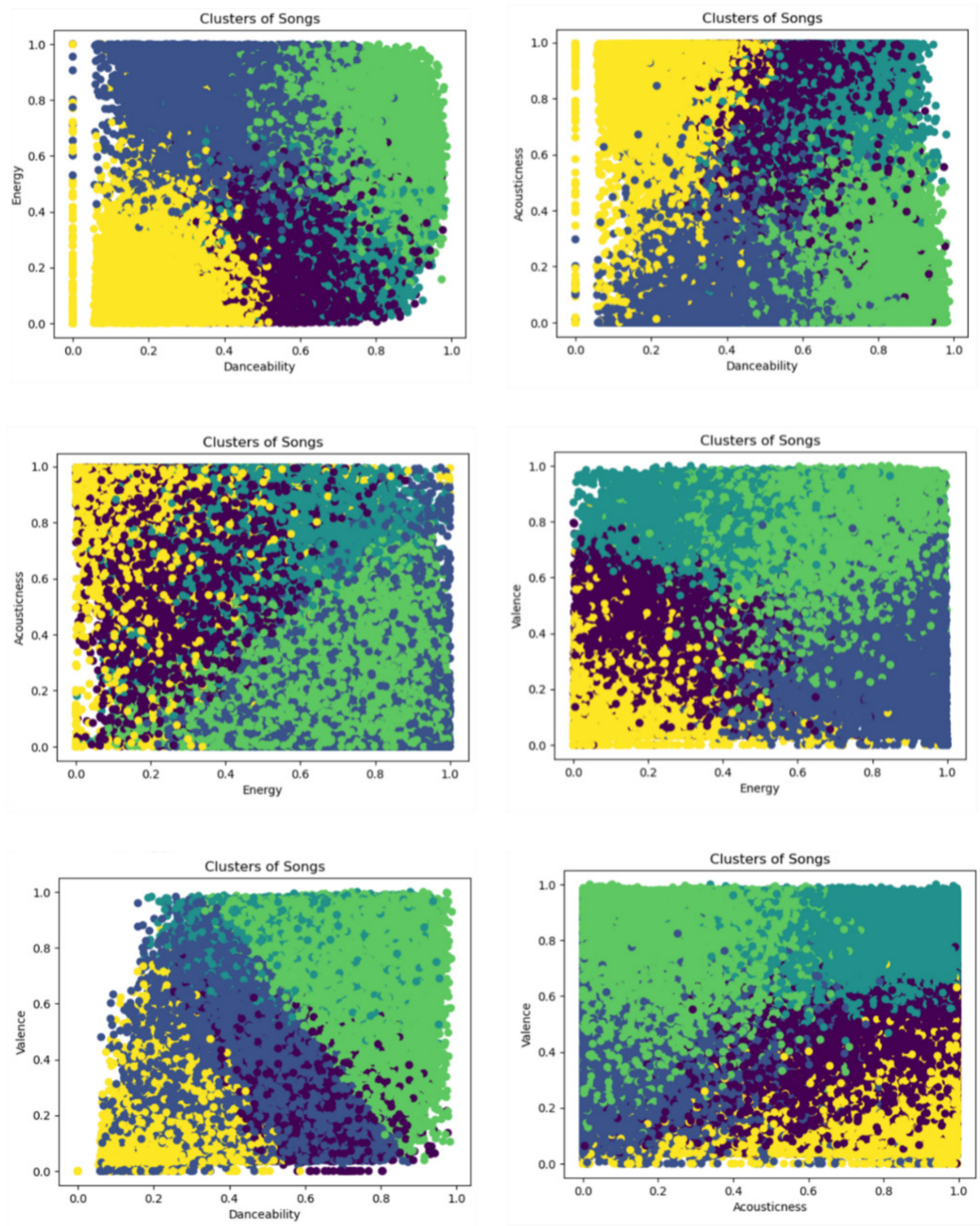

Fig. 2 Clusters of different feature combinations (danceability-energy, danceability-acoustics, danceability-valence, energy-acoustics, energy-valence, and acoustics-valence)

3.2 The Results of KNN

This section compares the accuracy using all features after removing the specified features combination mentioned in section 3.1 to check if the removed features are essential for the model. The feature genres is chosen as target variable (y) It was found that the accuracy after removing one feature pair was lower than the accuracy using all features, which means the selected features are essential for the model.

Table 2. Accuracy comparison by removing the specified features combination

Removed Features	Accuracy
All features	0.2287
Danceability & Energy	0.2255 (-0.0032)
Danceability & Acousticness	0.2177 (-0.011)
Danceability & Valence	0.2184 (-0.0103)
Energy & Acousticness	0.2015 (-0.0272)
Energy & Valence	0.2207 (-0.008)
Acousticness & Valence	0.2173 (-0.0114)

As shown in Table 2, the accuracy using all features is higher than the accuracy after removing features, it presents that these features have a great influence on the accuracy of the model, which means they are important for the model. Besides, it's worth noting that Energy and Valence are the most unimportant feature combinations in KNN Classification, and Energy and Acousticness are the most crucial feature combinations. This means that any single analysis method for determining the importance of a feature is one-sided.

4. Discussion

K-means clustering is effectively leveraged to cluster songs based on their relevant characteristics, reducing the data's dimension and complexity. The recommender system can offer more personalized suggestions by focusing on a group with similar attributes based on clustered songs. Using K-means to cluster can simplify computation, enhance recommendation quality and discover the potential interests of new users, which assists in addressing the cold-start problem. Meanwhile, by calculating the distance between feature vectors to determine the closest neighbours, the KNN can generate highly customized recommendations. Incorporating K-means clustering and KNN can improve the quality and flexibility of the recommendation, reduce the computation and effectively enhance user experience. By initially clustering the data with K-means and then using KNN, some problems of traditional methods can be optimized, and the recommendation system can make more accurate suggestions.

This paper explores the essential factor of the music recommendation system from a feature view. This paper argues that building an effective recommendation system is more critical than a high-performance system depending on computation-demanding models such as deep learning. This paper employs K-means clustering to identify song similarities based on feature combinations. KNN algorithm leverages selected features to find songs' nearest neighbours. The evaluation assesses feature significance by comparing KNN accuracy before and after removing a feature combination. Cluster analysis highlights energy and valence as critical features. Removing any feature reduces accuracy, emphasizing their importance. However, contrasting results from the (energy & valence) combo suggest that the importance assessment of single-method features is limited. This paper will explore more comprehensive experiments (more datasets and analysis methods) to get robust results in future work.

5. Conclusion

To counteract the challenge of information overload stemming from the immensity and intricacy of data available on the internet, recommendation systems have emerged as vital tools. Music recommendation occupies a prominent position among these systems' application domains. This paper focuses on the feature importance analysis based on Spotify. The feature importance analysis within this framework utilizes two machine learning-based methodologies. One method involves K-means clustering, which serves to discern similarities in feature combinations, thereby grouping songs of kindred types. Besides, K-Nearest Neighbors (KNN) is employed to identify the songs closest to a given query by leveraging a combination of selected features. In the evaluation phase, a rigorous examination is conducted to ascertain the significance of feature pairs to the KNN models. This involves comparing the model's accuracy before and after removing a specific feature combination. The outcomes underscore the prominence of 'energy' and 'valence' as the most compelling feature pairing, as evident from the clusters analysis. Furthermore, the reduction in accuracy observed upon removing any feature underscores their indispensability to the KNN model's performance. However, the contrasting outcomes yielded by analyzing the 'energy' and 'valence' feature combination hint at the limitations of relying solely on a single analytical approach. It highlights the need for a comprehensive and multifaceted examination to accurately gauge the importance of individual features, as any isolated analysis risks being partial and incomplete.

References

[1] S. Bhaskaran, Raja Marappan. Enhanced personalized recommendation system for machine learning public datasets: generalized modeling, simulation, significant results and analysis. International Journal of Information Technology, 2023, 15(3): 1583-1595.
[2] S. Bhaskaran, Raja Marappan. Design and analysis of an efficient machine learning based hybrid recommendation system with enhanced density-based spatial clustering for digital e-learning applications. Complex & Intelligent Systems, 2023, 9(4): 3517-3533.
[3] Wangmei Chen, Zepeng Shen, Yiming Pan, et al. Applying Machine Learning Algorithm to Optimize Personalized Education Recommendation System. Journal of Theory and Practice of Engineering Science, 2024, 4(01): 101-108.
[4] Brijesh Bakariya, Arshdeep Singh, Harmanpreet Singh, et al. Facial emotion recognition and music recommendation system using CNN-based deep learning techniques. Evolving Systems, 2024, 15(2): 641-658.
[5] Yashar Deldjoo, Markus Schedl, Peter Knees. Content-driven music recommendation: Evolution, state of the art, and challenges. Computer Science Review, 2024, 51: 100618.
[6] Zhikang Dong, Xiulong Liu, Bin Chen, et al. Musechat: A conversational music recommendation system for videos//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 12775-12785.