

ABSTRAK

Jutaan lagu yang tersedia pada platform musik digital menyulitkan pengguna dalam menemukan lagu yang sesuai dengan preferensi mereka. Pendekatan rekomendasi yang hanya mengandalkan genre, bahasa, atau tahun rilis belum mampu menggambarkan kemiripan karakteristik musikal secara objektif. Penelitian ini bertujuan mengembangkan sistem rekomendasi lagu berbasis fitur audio menggunakan pendekatan *Content-Based Filtering*, *K-Means Clustering*, dan *Cosine Similarity*. Dataset yang digunakan merupakan data lagu Spotify dari Kaggle yang terdiri atas 232.725 lagu dan menghasilkan 176.774 data bersih setelah proses pembersihan. Setiap lagu direpresentasikan menggunakan sembilan fitur audio dan dinormalisasi dengan metode *Min-Max Scaling*. Selanjutnya, lagu dikelompokkan menggunakan *K-Means Clustering* dan kemiripan antar lagu dihitung menggunakan *Cosine Similarity*. Evaluasi jumlah cluster dilakukan menggunakan *Elbow Method* dan *Silhouette Score*, yang menghasilkan jumlah *cluster* optimal sebanyak empat *cluster* dengan nilai *Silhouette Score* sebesar 0,3506. Keempat cluster tersebut merepresentasikan karakteristik audio yang berbeda, yaitu *Energetic/Upbeat*, *Spoken Word/Live*, *Instrumental/Ambient*, dan *Acoustic/Melancholic*. Sistem kemudian diimplementasikan dalam bentuk aplikasi web yang menyediakan fitur rekomendasi lagu, visualisasi cluster, dan analisis statistik fitur audio. Hasil *Black Box Testing* menunjukkan seluruh fungsi sistem berjalan sesuai harapan, sedangkan hasil *beta testing* terhadap 20 responden memperoleh indeks usability sebesar 83,33% dengan kategori sangat baik. Hasil penelitian menunjukkan bahwa sistem mampu memberikan rekomendasi lagu berdasarkan kemiripan karakteristik audio secara objektif tanpa bergantung pada genre, bahasa, atau tahun rilis.

Kata kunci: Sistem Rekomendasi Lagu, *Content-Based Filtering*, *K-Means Clustering*, *Cosine Similarity*, Fitur Audio.

ABSTRACT

Millions of songs available on digital music platforms make it difficult for users to find music that matches their preferences. Recommendation approaches that rely solely on genre, language, or release year are often unable to represent musical similarity objectively. This study aims to develop a song recommendation system based on audio features using the *Content-Based Filtering* approach combined with *K-Means Clustering* and *Cosine Similarity*. The dataset used consists of Spotify songs obtained from Kaggle, containing 232,725 records and resulting in 176,774 clean records after the data cleaning process. Each song is represented using nine audio features and normalized using the *Min-Max Scaling* method. Subsequently, songs are grouped using *K-Means Clustering*, while song *similarity* is measured using *Cosine Similarity*. The optimal number of clusters was determined through the *Elbow Method* and *Silhouette Score* evaluation, resulting in four optimal clusters with a *Silhouette Score* of 0.3506. The four clusters represent distinct audio characteristics, namely *Energetic/Upbeat*, *Spoken Word/Live*, *Instrumental/Ambient*, and *Acoustic/Melancholic*. The proposed system was implemented as a web - based application providing song recommendations, cluster visualization, and audio feature statistical analysis. *Black Box Testing* results showed that all system functions operated as expected, while *beta testing* involving 20 respondents produced an average usability index of 83.33%, categorized as very good. The results indicate that the developed system is capable of providing song recommendations based on objective audio feature *similarity* without depending on genre, language, or release year.

Keywords: Song Recommendation System, Content-Based Filtering, K-Means Clustering, Cosine Similarity, Audio Features.