

ABSTRAK

Dalam era digital yang terus berkembang, platform berbagi video seperti YouTube telah menjadi sumber utama informasi dan hiburan global. Peningkatan signifikan dalam jumlah unggahan dan interaksi pengguna menuntut metode analisis data yang efisien untuk memahami preferensi audiens. Penelitian ini bertujuan mengelompokkan video trending di YouTube berdasarkan metrik interaksi (*views*, *likes*, *comments*), tahun, dan kata kunci menggunakan algoritma K-Means Clustering. Dengan teknik ini, video dikelompokkan ke dalam beberapa cluster berdasarkan kemiripan karakteristiknya.

Hasil analisis menunjukkan bahwa pengelompokan paling optimal terjadi pada $k = 2$, yang divalidasi oleh nilai Silhouette Score tertinggi sebesar 0,76. Model ini secara efektif memisahkan video menjadi dua kelompok utama: Cluster 0 yang mewakili video dengan interaksi rendah, dan Cluster 1 yang terdiri dari video berinteraksi tinggi. Analisis lebih lanjut pada kata kunci mengungkapkan bahwa video dalam Cluster 1 cenderung memiliki judul yang mengandung kata-kata populer atau sensasional, seperti “viral”, “official”, atau “trending”.

Temuan ini membuktikan bahwa K-Means Clustering dapat secara efektif mengidentifikasi pola dan karakteristik video populer, serta memberikan wawasan strategis bagi kreator konten, pengiklan, dan peneliti untuk mengoptimalkan strategi mereka di platform media sosial.

ABSTRACT

In the continuously evolving digital era, video-sharing platforms like YouTube have become a primary source of global information and entertainment. The significant increase in uploaded content and user interactions demands efficient data analysis methods to understand audience preferences. This research aims to cluster trending YouTube videos based on interaction metrics (views, likes, comments), year, and keywords using the K-Means Clustering algorithm. Using this technique, videos are grouped into clusters based on their characteristic similarities.

The analysis shows that the optimal clustering occurs at $k = 2$, validated by the highest Silhouette Score of 0.76. The model effectively separates videos into two main groups: Cluster 0, representing videos with low interaction, and Cluster 1, comprising high-interaction videos. Furthermore, keyword analysis reveals that videos in Cluster 1 tend to feature titles containing popular or sensational keywords, such as “viral,” “official,” or “trending.”

These findings demonstrate that K-Means Clustering can be effectively used to identify patterns and characteristics of popular videos, providing strategic insights for content creators, advertisers, and researchers to optimize their strategies on the social media platform.