**PENERAPAN ALGORITMA K-MEANS UNTUK SEGMENTASI PENGGUNA MEDIA SOSIAL BERDASARKAN AKTIVITAS**

**Jihad Akbar1**.

***ABSTRAK:*** *Penelitian ini bertujuan untuk mengelompokkan pengguna media sosial berdasarkan aktivitas mereka menggunakan algoritma K-Means Clustering. Aktivitas pengguna diukur melalui dua fitur utama, yaitu jumlah "likes" dan "shares". Proyek ini diimplementasikan dalam bentuk aplikasi web berbasis Flask yang memungkinkan pengguna mengunggah dataset dan melihat hasil clustering dalam bentuk tabel dan visualisasi. Hasil clustering dapat membantu memahami pola perilaku pengguna seperti pengguna aktif, moderat, dan pasif. Proses melibatkan preprocessing data, normalisasi, penentuan jumlah cluster optimal dengan metode Elbow, serta visualisasi hasil menggunakan scatter plot. Aplikasi ini menunjukkan bahwa algoritma K-Means efektif digunakan untuk segmentasi pengguna berdasarkan aktivitas digital.*

***Kata kunci :*** *clustering, K-Means, media sosial, Flask, segmentasi pengguna*

***ABSTRACT:*** *This research aims to cluster social media users based on their activity using the K-Means Clustering algorithm. User activity is measured using two main features: the number of likes and shares. The project is implemented as a Flask-based web application allowing users to upload datasets and view the clustering results in tables and visualizations. The clustering results help understand user behavior patterns such as active, moderate, and passive users. The process involves data preprocessing, normalization, determining the optimal number of clusters using the Elbow method, and visualizing results using scatter plots. This application shows that K-Means is effective for segmenting users based on digital activity.*

***Keywords:*** *clustering, K-Means, social media, Flask, user segmentation*

**PENDAHULUAN**

Media sosial telah menjadi bagian penting dalam kehidupan sehari-hari masyarakat modern. Setiap pengguna memiliki tingkat aktivitas yang berbeda-beda, mulai dari pengguna yang sangat aktif hingga pengguna yang jarang berinteraksi. Pemahaman terhadap pola aktivitas ini penting, khususnya bagi pihak yang ingin melakukan analisis perilaku pengguna untuk keperluan pemasaran, pengembangan produk, atau penelitian sosial.

Salah satu pendekatan yang digunakan untuk mengelompokkan pengguna adalah algoritma K-Means Clustering. K-Means adalah algoritma unsupervised learning yang efektif untuk menemukan pola dalam data tanpa label. Dalam penelitian ini, dilakukan penerapan K-Means untuk mengelompokkan pengguna media sosial berdasarkan dua indikator aktivitas utama, yaitu jumlah likes dan shares.

Studi ini juga bertujuan untuk mengembangkan aplikasi berbasis web menggunakan Flask yang dapat digunakan secara interaktif oleh pengguna untuk melakukan clustering terhadap data mereka sendiri. Dengan demikian, hasil penelitian ini tidak hanya bersifat analitis tetapi juga aplikatif.

**METODA**

Penelitian ini menggunakan pendekatan eksperimen dengan tahapan sebagai berikut:

1. Pengumpulan data: Dataset yang digunakan adalah "Social Media Behavior Dataset" dari Kaggle.
2. Preprocessing: Data dibersihkan dari nilai duplikat dan null, serta fitur yang tidak relevan disingkirkan.
3. Normalisasi: Fitur "Likes" dan "Shares" dinormalisasi menggunakan StandardScaler dari Scikit-Learn.
4. Penentuan jumlah cluster: Digunakan metode Elbow untuk menentukan jumlah cluster optimal.
5. Clustering: Dilakukan proses clustering menggunakan K-Means dengan jumlah cluster yang ditentukan.
6. Visualisasi: Hasil divisualisasikan dalam bentuk tabel dan scatter plot melalui aplikasi web berbasis Flask.
7. Implementasi Web: Aplikasi web mendukung upload dataset, input jumlah cluster, serta visualisasi hasil clustering.

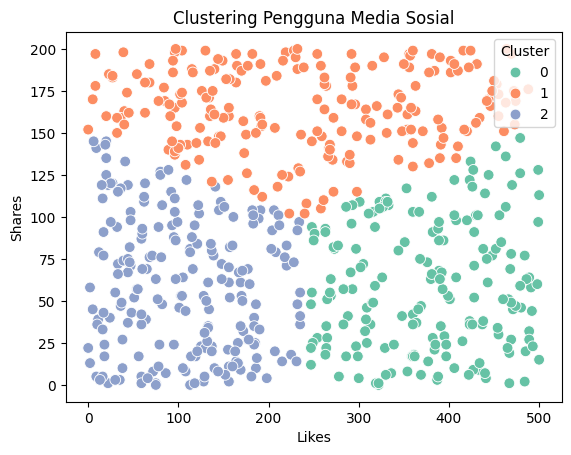
**HASIL DAN PEMBAHASAN**

1. *Preprocessing dan Normalisasi* Data yang digunakan berisi kolom Platform, Hashtag, Likes, dan Shares. Kolom yang digunakan dalam proses clustering adalah Likes dan Shares. Data dinormalisasi agar fitur memiliki skala yang sama.
2. *Penentuan Jumlah Cluster Optimal* Melalui metode Elbow, ditentukan bahwa jumlah cluster optimal berada pada K = 3. Grafik inertia menunjukkan titik siku pada angka tersebut.

*Hasil Clustering* Setelah proses clustering dilakukan, pengguna terbagi menjadi 3 segmen utama:

* Cluster 0: Pengguna dengan aktivitas rendah
* Cluster 1: Pengguna dengan aktivitas sedang
* Cluster 2: Pengguna dengan aktivitas tinggi

Visualisasi scatter plot menunjukkan pemisahan yang cukup jelas antara ketiga cluster tersebut berdasarkan nilai Likes dan Shares. Tabel hasil clustering juga ditampilkan dalam web.



1. *Implementasi Web* Aplikasi berbasis Flask memungkinkan pengguna untuk mengunggah file CSV, memilih jumlah cluster, dan melihat hasil clustering secara langsung. Tersedia juga grafik Elbow dan scatter plot.



**SIMPULAN DAN SARAN**

Penelitian ini berhasil mengimplementasikan algoritma K-Means Clustering untuk segmentasi pengguna media sosial berdasarkan aktivitas mereka. Hasil clustering menunjukkan adanya pola perilaku yang dapat dianalisis lebih lanjut untuk keperluan pemasaran atau penelitian sosial. Aplikasi web yang dibangun juga mempermudah pengguna dalam menjalankan proses clustering secara interaktif. Ke depan, sistem dapat dikembangkan dengan menambahkan fitur analisis teks atau integrasi dengan media sosial secara real-time.

**PUSTAKA ACUAN**

**Buku**

**Han, J., Kamber, M., & Pei, J. (2011).**  
*Data Mining: Concepts and Techniques.* Morgan Kaufmann.

**Manning, C. D., Raghavan, P., & Schutze, H. (2008).**  
*Introduction to Information Retrieval.* Cambridge University Press.

**Jurnal**

**Pedregosa, F., et al. (2011).**  
*Scikit-learn: Machine Learning in Python.* Journal of Machine Learning Research, 12, 2825–2830.

**Lain – lain**

Kaggle. (2023). Social Media Behavior Dataset. <https://www.kaggle.com/datasets/shibinshereef1/social-media-behavior-dataset>

**LINK YOUTUBE :** <https://youtu.be/yEwxpnvMNno>

**LINK GITHUB :**

<https://github.com/jihadakbar911/tugas6praktikumai>

**LINK WEB YANG SUDAH DI-DEPLOY :**

<https://clusteringmediasosial.up.railway.app/>