

Tugas 6

Nama : muhammad makhlufi makbullah

Kelas : TK 45 01

NIM : 1103210171

Berdasarkan hasil evaluasi dengan menggunakan berbagai metrik, berikut adalah kesimpulan untuk menentukan model clustering terbaik:

1. ****Silhouette Score****:

- Silhouette score digunakan untuk mengukur seberapa baik setiap data terkelompok dalam klusternya. Semakin tinggi nilai silhouette, semakin baik model dalam mengelompokkan data.
- Dari hasil di atas, model dengan nilai silhouette tertinggi dianggap sebagai yang terbaik dalam memisahkan kluster dengan jelas dan efektif.

2. ****Davies-Bouldin Index****:

- Davies-Bouldin Index mengukur jarak antara kluster yang berbeda dan variabilitas dalam kluster. Semakin rendah nilainya, semakin baik, karena ini menunjukkan bahwa kluster saling terpisah dengan baik dan tidak terlalu tersebar.
- Model dengan nilai Davies-Bouldin terendah menunjukkan bahwa kluster yang dihasilkan lebih terpisah dan lebih kompak, sehingga lebih berkualitas.

3. ****Calinski-Harabasz Index****:

- Calinski-Harabasz Index mengukur tingkat kepadatan dan pemisahan kluster, di mana semakin tinggi nilainya, semakin baik kualitas klustering yang dihasilkan.
- Model dengan nilai Calinski-Harabasz tertinggi dianggap memiliki pemisahan dan kepadatan kluster yang lebih baik, yang menunjukkan bahwa data dikelompokkan dengan baik.

Kesimpulan Akhir

Berdasarkan evaluasi ketiga metrik di atas (Silhouette Score, Davies-Bouldin Index, dan Calinski-Harabasz Index), kita dapat memilih model yang memberikan nilai terbaik secara keseluruhan. Berikut adalah kesimpulan untuk setiap dataset:

- ****Dataset Online Retail****:

- Model dengan nilai Silhouette Score tertinggi adalah [Nama Model Terbaik untuk Online Retail].

- Model dengan Davies-Bouldin Index terendah adalah [Nama Model Terbaik untuk Online Retail].

- Model dengan Calinski-Harabasz Index tertinggi adalah [Nama Model Terbaik untuk Online Retail].

- Dengan demikian, berdasarkan hasil evaluasi, model terbaik untuk dataset Online Retail adalah ****[Model Terbaik Berdasarkan Keseluruhan Metrik]****.

- ****Dataset Clustering****:

- Model dengan nilai Silhouette Score tertinggi adalah [Nama Model Terbaik untuk Clustering].

- Model dengan Davies-Bouldin Index terendah adalah [Nama Model Terbaik untuk Clustering].

- Model dengan Calinski-Harabasz Index tertinggi adalah [Nama Model Terbaik untuk Clustering].

- Berdasarkan evaluasi keseluruhan metrik, model terbaik untuk dataset Clustering adalah ****[Model Terbaik Berdasarkan Keseluruhan Metrik]****.

Dengan mempertimbangkan seluruh metrik evaluasi di atas, model yang dipilih untuk setiap dataset memberikan keseimbangan terbaik antara pemisahan kluster dan kepadatan internal kluster.