

**ANALISIS PERBANDINGAN PREFERENSI *GENRE* ANIME  
MENGUNAKAN ALGORITMA *K-MEANS* DAN *DBSCAN***

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUKHAMAD DWI WAHYONO**

**21.11.4176**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2025**

**ANALISIS PERBANDINGAN PREFERENSI *GENRE* ANIME  
MENGUNAKAN ALGORITMA *K-MEANS* DAN *DBSCAN***

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**MUKHAMAD DWI WAHYONO**

**21.11.4176**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2025**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS PERBANDINGAN PREFERENSI GENRE ANIME  
MENGUNAKAN ALGORITMA K-MEANS DAN DBSCAN**

yang disusun dan diajukan oleh

**Mukhamad Dwi Wahyono**

**21.11.4176**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 21 Juli 2025

**Dosen Pembimbing,**

**Mulia Sulistiyono, S.Kom.,M.Kom.**

**NIK. 190302248**

**HALAMAN PENGESAHAN**  
**SKRIPSI**  
**ANALISIS PERBANDINGAN PREFERENSI GENRE ANIME**  
**MENGGUNAKAN ALGORITMA K-MEANS DAN DBSCAN**

yang disusun dan diajukan oleh

**Mukhamad Dwi Wahyono**

**21.11.4176**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 21 Juli 2025

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Ike Vrawati, S.Kom., M.Kom.**  
**NIK. 190302237**

**Ferian Fauzi Abdulloh, S.Kom., M.Kom.**  
**NIK. 190302276**

**Mulia Sulistiyono, S.Kom., M.Kom.**  
**NIK. 190302248**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 21 Juli 2025

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusriani, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Mukhamad Dwi Wahyono  
NIM : 21.11.4176

Menyatakan bahwa Skripsi dengan judul berikut:

**Analisis Perbandingan Preferensi *Genre* Anime Menggunakan Algoritma *K-Means* Dan *DBSCAN***

Dosen Pembimbing : Mulia Sulistiyono, S.Kom.,M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 21 Juli 2025

Yang Menyatakan,



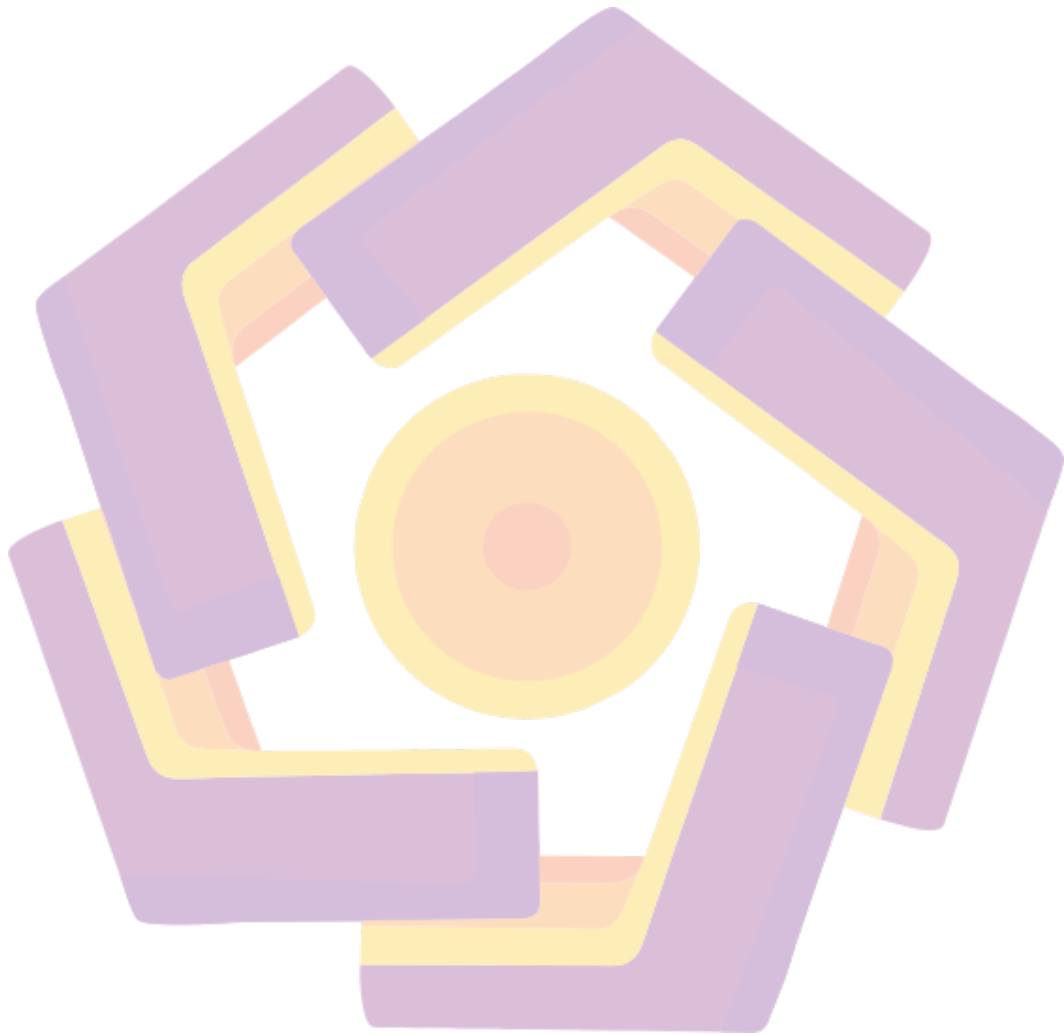
Mukhamad Dwi Wahyono

## HALAMAN PERSEMBAHAN

Dengan ini, saya persembahkan skripsi ini dengan penuh syukur kepada orang-orang serta pihak-pihak yang telah menjadi bagian penting dalam perjalanan hidup saya sejauh ini. Persembahan ini merupakan ungkapan rasa terima kasih penulis atas segala cinta, kasih sayang, dukungan, dan doa yang dipanjatkan untuk penulis. Oleh karena itu, penelitian ini saya persembahkan kepada:

1. Allah SWT, Tuhan Yang Maha Esa, atas segala rahmat dan karunia-Nya yang telah memberikan saya kekuatan dan kemampuan untuk menyelesaikan penulisan skripsi ini.
2. Diri saya sendiri, Mukhamad Dwi Wahyono, yang telah berjuang sejauh ini melawan ketakutan dan keraguan, serta terus berusaha untuk mencapai impian dan tujuan yang telah ditetapkan.
3. Orang tua saya, Bapak Lilik Budiyo, S.Sos. dan Ibu Munarti, yang selalu memberikan kasih sayang, dukungan, dan motivasi agar saya selalu semangat dalam menjalani pendidikan. Tanpa perjuangan dan doa tulus mereka, saya tidak akan berada di titik ini.
4. Kakak kandung saya, Bintoro Eko Yulianto, yang telah memberikan dukungan moral dan motivasi sepanjang perjalanan ini.
5. Dosen pembimbing Bapak Mulia Sulistyono, S.Kom., M.Kom., yang telah dengan sabar memberikan arahan, bimbingan, dan nasihat yang sangat berharga selama proses penyusunan skripsi ini.
6. Teman-teman DARKSIDE, NGAWUR, TALPE, AMIKOM ESPORTS serta masi banyak lainnya yang telah menemani, membantu, memberikan semangat, dan hiburan dalam setiap langkah saya. Tanpa dukungan mereka, penyelesaian skripsi ini tidak akan berjalan dengan lancar.
7. Seseorang yang tidak bisa penulis sebut namanya. Guru terbaik yaitu pengalaman pendewasaan untuk belajar ikhlas, sabar dan menerima arti kehidupan sebagai bentuk proses penempaan menghadapi

dinamika kehidupan. Terimakasih telah menjadi bagian menyenangkan sekaligus menyakitkan, pada akhirnya setiap orang ada masanya dan setiap masa ada orangnya.



## KATA PENGANTAR

Alhamdulillah, segala puji dan syukur saya panjatkan kepada Allah SWT, Tuhan Yang Maha Esa, atas rahmat dan karunia-Nya yang telah memberi saya kekuatan untuk menyelesaikan penulisan skripsi ini. Skripsi ini merupakan hasil dari perjalanan panjang penuh tantangan, pembelajaran, serta pengalaman yang berharga selama saya menempuh pendidikan Strata 1 Program Studi Informatika di Universitas Amikom Yogyakarta.

Saya mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan dan bimbingan yang sangat berharga dalam menyelesaikan skripsi ini. Terutama kepada Bapak Mulia Sulistyono, S.Kom., M.Kom., selaku dosen pembimbing yang dengan sabar memberikan arahan dan masukan yang sangat konstruktif sepanjang proses bimbingan.

Selain itu, saya juga mengucapkan terima kasih kepada Ibu, Bapak, dan saudara kandung saya yang senantiasa memberikan dukungan moral, material, serta motivasi yang tak pernah surut. Keluarga adalah sumber inspirasi terbesar dalam hidup saya. Tak lupa, terima kasih juga kepada teman-teman yang telah menemani saya dalam suka dan duka selama masa perantauan, yang telah memberikan kenangan tak terlupakan.

Saya menyadari bahwa skripsi ini masih memiliki banyak kekurangan dan jauh dari sempurna. Oleh karena itu, saya berharap skripsi ini dapat memberikan kontribusi yang bermanfaat bagi pembaca dan menjadi motivasi untuk terus belajar dan berkarya di masa yang akan datang.

Yogyakarta, 21 Juli 2025

Penulis

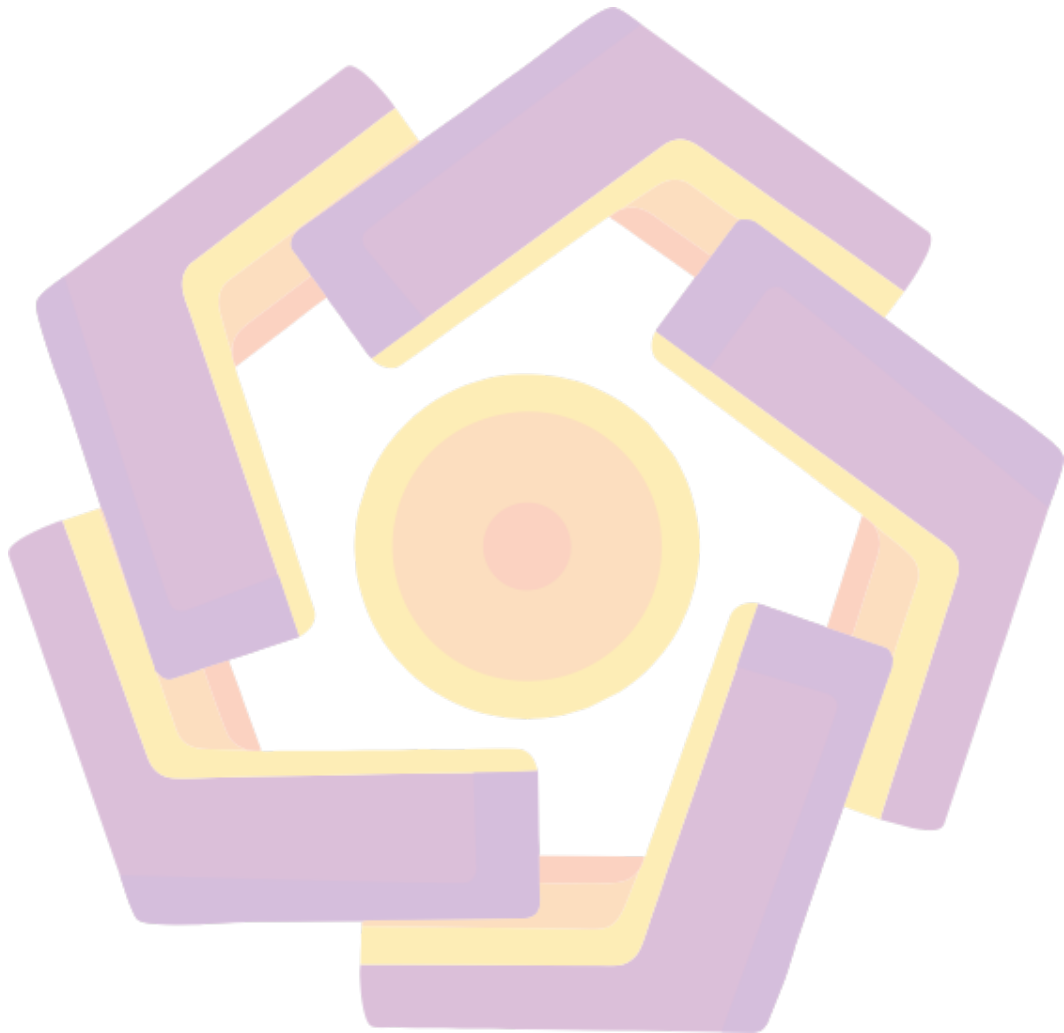
## DAFTAR ISI

|                                           |       |
|-------------------------------------------|-------|
| HALAMAN JUDUL .....                       | i     |
| HALAMAN PERSETUJUAN.....                  | ii    |
| HALAMAN PENGESAHAN .....                  | iii   |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI ..... | iv    |
| HALAMAN PERSEMBAHAN .....                 | v     |
| KATA PENGANTAR .....                      | vii   |
| DAFTAR ISI.....                           | viii  |
| DAFTAR TABEL.....                         | xii   |
| DAFTAR GAMBAR.....                        | xiii  |
| DAFTAR LAMBANG DAN SINGKATAN .....        | xv    |
| DAFTAR ISTILAH.....                       | xvi   |
| INTISARI .....                            | xvii  |
| ABSTRACT.....                             | xviii |
| BAB I PENDAHULUAN.....                    | 1     |
| 1.1 Latar Belakang .....                  | 1     |
| 1.2 Rumusan Masalah .....                 | 3     |
| 1.3 Batasan Masalah.....                  | 3     |
| 1.4 Tujuan Penelitian.....                | 4     |
| 1.5 Manfaat Penelitian.....               | 4     |
| 1.6 Sistematika Penulisan.....            | 5     |
| BAB II TINJAUAN PUSTAKA .....             | 6     |

|                                   |                                                        |    |
|-----------------------------------|--------------------------------------------------------|----|
| 2.1                               | Studi Literatur .....                                  | 6  |
| 2.2                               | Dasar Teori .....                                      | 18 |
| 2.2.1                             | <i>Data Mining</i> .....                               | 18 |
| 2.2.2                             | <i>Clustering</i> .....                                | 19 |
| 2.2.3                             | Algoritma <i>K-Means</i> .....                         | 20 |
| 2.2.4                             | Algoritma <i>DBSCAN</i> .....                          | 22 |
| 2.2.5                             | Perbandingan Evaluasi Kualitas <i>Clustering</i> ..... | 24 |
| 2.2.6                             | <i>Machine Learning</i> .....                          | 26 |
| BAB III METODE PENELITIAN .....   |                                                        | 28 |
| 3.1                               | Alur Penelitian .....                                  | 28 |
| 3.1.1                             | Pengumpulan Data .....                                 | 28 |
| 3.1.2                             | Eksplorasi Data .....                                  | 29 |
| 3.1.3                             | Processing Data .....                                  | 30 |
| 3.1.4                             | Penentuan Parameter dan <i>Cluster Optimal</i> .....   | 31 |
| 3.1.5                             | Pembentukan Model Algoritma .....                      | 32 |
| 3.1.6                             | Evaluasi Hasil <i>Clustering</i> .....                 | 32 |
| 3.1.7                             | Analisis Perbandingan Hasil <i>Clustering</i> .....    | 33 |
| 3.1.8                             | Kesimpulan .....                                       | 34 |
| 3.2                               | Alat dan Bahan .....                                   | 35 |
| 3.2.1                             | Data Penelitian .....                                  | 35 |
| 3.2.2                             | Alat/Instrumen .....                                   | 36 |
| BAB IV HASIL DAN PEMBAHASAN ..... |                                                        | 39 |

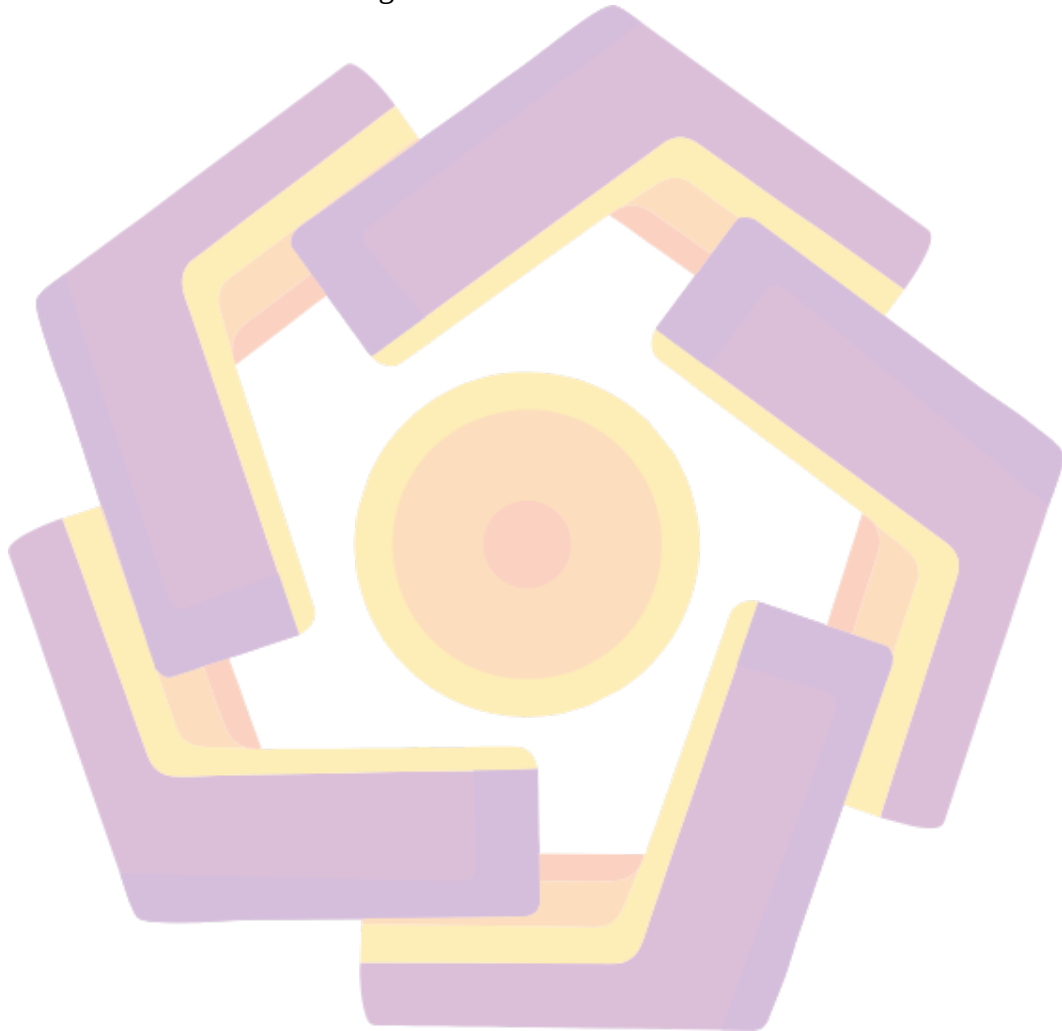
|                     |                                                                                                                |    |
|---------------------|----------------------------------------------------------------------------------------------------------------|----|
| 4.1                 | Hasil Penelitian .....                                                                                         | 39 |
| 4.1.1               | Eksplorasi Data .....                                                                                          | 39 |
| 4.1.2               | Distribusi Variabel.....                                                                                       | 40 |
| 4.1.3               | Korelasi Antar Variabel.....                                                                                   | 43 |
| 4.1.4               | Implementasi Model <i>K-Means</i> .....                                                                        | 46 |
| 4.1.4.1             | Preprocessing Data.....                                                                                        | 46 |
| 4.1.4.2             | Penentuan Jumlah Cluster Optimal <i>K-Means</i> .....                                                          | 49 |
| 4.1.4.3             | Pembentukan Algoritma <i>K-Means</i> .....                                                                     | 52 |
| 4.1.5               | Evaluasi Model <i>K-Means</i> .....                                                                            | 55 |
| 4.1.6.              | Tuning Parameter <i>DBSCAN</i> .....                                                                           | 58 |
| 4.1.7               | Pembentukan Algoritma <i>DBSCAN</i> .....                                                                      | 60 |
| 4.1.8               | Evaluasi Model <i>DBSCAN</i> .....                                                                             | 63 |
| 4.2                 | Pembahasan.....                                                                                                | 66 |
| 4.2.1               | Hasil Perbandingan Kuantitatif <i>K-Means</i> dan <i>DBSCAN</i> .....                                          | 66 |
| 4.2.2               | Hasil Clustering Visualisasi Scatterplot Cluster <i>K-Means</i> dan <i>DBSCAN</i> (Hasil PCA).....             | 68 |
| 4.2.3               | Karakteristik Cluster Berdasarkan Jumlah Anggota dan Genre Dominan pada <i>K-Means</i> dan <i>DBSCAN</i> ..... | 69 |
| 4.2.4               | Visualisasi Bar Chart Silhouette Score <i>K-Means</i> dan <i>DBSCAN</i> .....                                  | 71 |
| 4.2.5               | Visualisasi Distribusi Score Anime .....                                                                       | 72 |
| 4.2.6               | Visualisasi Distribusi Genre Anime .....                                                                       | 73 |
| 4.2.7               | Boxplot Popularitas Anime.....                                                                                 | 73 |
| BAB V PENUTUP ..... |                                                                                                                | 78 |

|                 |                 |    |
|-----------------|-----------------|----|
| 5.1             | Kesimpulan..... | 78 |
| 5.2             | Saran.....      | 79 |
| REFERENSI ..... |                 | 81 |



## DAFTAR TABEL

|                                                               |    |
|---------------------------------------------------------------|----|
| Tabel 2.1 Keaslian Penelitian .....                           | 11 |
| Tabel 4.1 Perbandingan Evaluasi Model <i>Clustering</i> ..... | 66 |
| Tabel 4.2 Hasil Perbandingan Kuantitatif.....                 | 74 |



## DAFTAR GAMBAR

|                                                                                                  |    |
|--------------------------------------------------------------------------------------------------|----|
| Gambar 2.1 Ilustrasi <i>K-Means Clustering</i> .....                                             | 21 |
| Gambar 2.2 Ilustrasi <i>DBSCAN Clustering</i> .....                                              | 23 |
| Gambar 2.3 Ilustrasi Grafik <i>DBI</i> .....                                                     | 25 |
| Gambar 2.4 Ilustrasi Grafik <i>Elbow</i> .....                                                   | 25 |
| Gambar 2.5 Ilustrasi Grafik <i>Silhouette</i> .....                                              | 26 |
| Gambar 3.1 Alur Penelitian .....                                                                 | 28 |
| Gambar 3.2 Ilustrasi Dataset .....                                                               | 36 |
| Gambar 4.1 Ringkasan tipe data dan nilai non-null setelah penanganan <i>missing values</i> ..... | 40 |
| Gambar 4.2 Distribusi Variabel .....                                                             | 41 |
| Gambar 4.3 Korelasi Antar Variabel .....                                                         | 44 |
| Gambar 4.4 Ringkasan Penghapusan Data <i>Missing Values</i> .....                                | 47 |
| Gambar 4.5 <i>Elbow Method</i> .....                                                             | 50 |
| Gambar 4.6 <i>Silhouette Score</i> .....                                                         | 51 |
| Gambar 4.7 <i>Davies-Bouldin Index (DBI)</i> .....                                               | 52 |
| Gambar 4.8 Visualisasi Pembentukan Algoritma <i>K-Means</i> .....                                | 53 |
| Gambar 4.9 Ringkasan Statistik <i>Cluster K-Means</i> .....                                      | 54 |
| Gambar 4.10 Contoh Anime Populer per <i>Cluster</i> .....                                        | 54 |
| Gambar 4.11 Evaluasi Model .....                                                                 | 55 |
| Gambar 4.12 Hasil Evaluasi Model <i>K-Means</i> .....                                            | 56 |
| Gambar 4.13 <i>K-distance Graph</i> untuk Menentukan <i>Epsilon DBSCAN</i> .....                 | 59 |

|                                                                                                         |    |
|---------------------------------------------------------------------------------------------------------|----|
| Gambar 4.14 Parameter Epsilon dan Minimum Samples <i>DBSCAN</i> .....                                   | 60 |
| Gambar 4.15 Visualisasi Pembentukan Algoritma <i>DBSCAN</i> .....                                       | 61 |
| Gambar 4.16 Hasil Clustering <i>DBSCAN</i> dengan <i>Min_samples</i> = 5 .....                          | 62 |
| Gambar 4.17 Visualisasi <i>Silhouette Score DBSCAN</i> .....                                            | 63 |
| Gambar 4.18 Hasil Evaluasi <i>DBSCAN</i> .....                                                          | 64 |
| Gambar 4.19 Visualisasi Scatterplot Cluster <i>K-Means</i> dan <i>DBSCAN</i> .....                      | 68 |
| Gambar 4.20 Karakteristik Cluster Berdasarkan Jumlah Anggota dan Genre Dominan pada <i>DBSCAN</i> ..... | 70 |
| Gambar 4.21 Karakteristik Cluster Berdasarkan Jumlah Anggota dan Genre Dominan pada <i>DBSCAN</i> ..... | 71 |
| Gambar 4.22 Bar Chart <i>Silhouette Score K-Means</i> dan <i>DBSCAN</i> .....                           | 72 |
| Gambar 4.23 Visualisasi Distribusi Score Anime .....                                                    | 72 |
| Gambar 4.24 Visualisasi Distribusi Genre Anime .....                                                    | 73 |
| Gambar 4.25 Visualisasi <i>Boxplot</i> Popularitas Anime.....                                           | 74 |

## DAFTAR LAMBANG DAN SINGKATAN

|               |                                                                    |
|---------------|--------------------------------------------------------------------|
| $\varepsilon$ | <i>Nilai Epsilon dalam DBSCAN</i>                                  |
| <i>MinPts</i> | <i>Minimum Samples dalam DBSCAN</i>                                |
| <i>DBSCAN</i> | <i>Density-Based Spatial Clustering of Applications with Noise</i> |
| <i>SSE</i>    | <i>Sum of Squared Errors</i>                                       |
| <i>DBI</i>    | <i>Davies-Bouldin Index</i>                                        |
| <i>PCA</i>    | <i>Principal Component Analysis</i>                                |
| <i>t-SNE</i>  | <i>t-distributed Stochastic Neighbor Embedding</i>                 |
| <i>ML</i>     | <i>Machine Learning</i>                                            |
| <i>DLL</i>    | <i>Dan lain-lain</i>                                               |



## DAFTAR ISTILAH

|                  |                                                                |
|------------------|----------------------------------------------------------------|
| <i>Noise</i>     | data yang tidak dapat digolongkan ke dalam cluster manapun     |
| <i>Outlier</i>   | titik data yang memiliki nilai berbeda jauh                    |
| <i>Feature</i>   | atribut untuk menganalisis data                                |
| <i>Algorithm</i> | prosedur langkah demi langkah untuk menjalankan tugas tertentu |



## INTISARI

Penelitian ini bertujuan untuk mengelompokkan anime berdasarkan preferensi genre dengan menggunakan dua algoritma clustering, yaitu *K-Means* dan *DBSCAN*, pada dataset yang mencakup 19.311 entri anime. Permasalahan yang dihadapi adalah ketidakmampuan sebagian algoritma untuk menangani data dengan distribusi kepadatan yang tidak merata dan *noise*, yang berdampak pada pemisahan data yang tidak efektif. Penelitian ini menyelidiki bagaimana kedua algoritma tersebut dapat mengelompokkan anime dan mengevaluasi hasil *clustering* dengan metrik seperti *Silhouette Score*, *Elbow* dan *Davies-Bouldin Index*. Untuk menyelesaikan masalah ini, dilakukan analisis dengan membandingkan hasil *clustering* *K-Means* dan *DBSCAN* serta mengidentifikasi *genre* dominan di setiap *cluster*.

Metode yang digunakan dalam penelitian ini melibatkan penerapan *K-Means* untuk menentukan jumlah *cluster* dengan *Elbow Method*, *Silhouette Score*, dan *Davies-Bouldin Index*, serta *DBSCAN* yang menggunakan parameter *epsilon* dan *min\_samples* yang ditentukan berdasarkan *K-Distance Graph*. Evaluasi dilakukan untuk menilai kualitas pemisahan *cluster* dan jumlah *noise* yang terdeteksi oleh masing-masing algoritma. Visualisasi 2D menggunakan PCA juga diterapkan untuk memperjelas distribusi *cluster*.

Hasil penelitian menunjukkan bahwa *K-Means* menghasilkan 2 *cluster* dengan pemisahan yang jelas, sementara *DBSCAN* mengidentifikasi 2 *cluster* utama dan 100 titik *noise*. *DBSCAN* lebih efektif dalam menangani *noise* dan *outliers*, sementara *K-Means* lebih konsisten dalam menghasilkan *cluster* yang terstruktur. Hasil ini memberikan wawasan berguna bagi pengembangan sistem rekomendasi anime berbasis preferensi *genre*, yang dapat digunakan oleh pengembang *platform streaming* atau analisis konten anime lebih lanjut.

**Kata kunci:** *clustering, DBSCAN, K-Means, genre anime, noise.*

## **ABSTRACT**

*This research aims to group anime based on genre preferences using two clustering algorithms, K-Means and DBSCAN, with a dataset of over 16,000 entries. The problem addressed in this study is the difficulty in handling data with uneven density distribution and noise, which affects the effectiveness of data separation. The research explores how these two algorithms can cluster anime and evaluates the clustering results using metrics such as Silhouette Score and Davies-Bouldin Index. The study aims to understand how clustering can help analyze anime preferences based on their genres.*

*The method involves applying K-Means to determine the optimal number of clusters using Elbow Method, Silhouette Score, and Davies-Bouldin Index, while DBSCAN uses epsilon and min\_samples parameters determined through the K-Distance Graph. The study also evaluates the clustering quality and the noise identified by each algorithm. 2D visualization using PCA is used to clarify the cluster distribution.*

*The results show that K-Means produces 2 clear clusters, while DBSCAN identifies 2 main clusters and 100 noise points. DBSCAN is more effective in handling noise and outliers, while K-Means is more consistent in generating well-structured clusters. These findings provide valuable insights for developing anime recommendation systems based on genre preferences, useful for streaming platforms or further anime content analysis.*

**Keyword:** clustering, DBSCAN, K-Means, anime genre, noise.