

PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS UNTUK KLASTERISASI GAME PADA PLATFORM STEAM BERDASARKAN POPULARITAS, HARGA, DAN WAKTU BERMAIN

Fathur Rohman Mufid

Hindarto

Program Studi Informatika

Universitas Muhammadiyah Sidoarjo

2026



Pendahuluan

- Katalog game Steam sangat besar dan memiliki karakteristik yang heterogen.
- Genre saja belum cukup menjelaskan popularitas, harga, dan keterlibatan pemain.
- Klasterisasi digunakan untuk memetakan game berdasarkan kemiripan karakteristik numerik.
- K-Means dan K-Medoids dibandingkan karena keduanya memiliki pendekatan pusat klaster yang berbeda.
- Penelitian ini berfokus pada estimated owners, price, dan average playtime forever.

Pertanyaan Penelitian (Rumusan Masalah)

1. Bagaimana mengelompokkan game Steam berdasarkan estimasi jumlah pemilik, harga, dan rata-rata waktu bermain?
2. Bagaimana perbandingan performa K-Means dan K-Medoids berdasarkan Silhouette Score, Davies-Bouldin Index, dan distribusi anggota klaster?
3. Bagaimana profil klaster game Steam yang terbentuk berdasarkan popularitas, harga, dan keterlibatan pemain?

Metode

- Sumber data: Steam Games Dataset dari Kaggle.
- Data awal terdiri atas 136.080 entri dan 42 atribut.
- Penyaringan data dilakukan dalam empat tahap hingga menghasilkan 22.659 game.
- Fitur klasterisasi: estimated_owners, price, dan average_playtime_forever.
- Pra-pemrosesan: konversi midpoint, transformasi Yeo-Johnson, dan standardisasi.
- Jumlah klaster dievaluasi pada K=2 sampai K=10.
- Model yang dibandingkan: K-Means dan K-Medoids menggunakan FasterPAM.
- Evaluasi menggunakan Inertia, Silhouette Score, Davies-Bouldin Index, dan uji Kruskal-Wallis.

Hasil

- Populasi akhir penelitian terdiri atas 22.659 game setelah empat tahap penyaringan.
- Transformasi Yeo-Johnson menurunkan skewness fitur sehingga data lebih seimbang untuk perhitungan jarak.
- Nilai K optimal adalah $K=3$ karena menghasilkan Silhouette Score tertinggi sebesar 0,2913.
- K-Medoids memperoleh Silhouette Score 0,2922, sedikit lebih tinggi daripada K-Means sebesar 0,2913.
- K-Means memperoleh Davies-Bouldin Index 1,1234, sedangkan K-Medoids memperoleh 1,1493.
- K-Medoids menghasilkan distribusi anggota klaster yang lebih merata, yaitu 7.179, 7.672, dan 7.808 game.

Pembahasan

- Perbedaan nilai evaluasi antara K-Means dan K-Medoids relatif kecil sehingga tidak ada model yang unggul mutlak pada seluruh metrik.
- K-Medoids dipilih sebagai model utama karena memiliki Silhouette Score lebih tinggi dan distribusi anggota yang lebih seimbang.
- Medoid menggunakan observasi aktual sebagai pusat kluster sehingga lebih representatif pada data yang memiliki nilai ekstrem.
- Uji Kruskal-Wallis menunjukkan seluruh fitur berbeda signifikan antarkluster dengan $p\text{-value} < 0,001$.
- Hasil klusterisasi menunjukkan segmentasi game dapat dibaca melalui popularitas, harga, dan keterlibatan pemain.

Temuan Penting Penelitian

- Klaster 0: Populer-Premium-Engagement Tinggi dengan rata-rata 760.647 pemilik, harga US\$9,92, dan waktu bermain 2.238 menit.
- Klaster 1: Budget/F2P dengan rata-rata 288.355 pemilik, harga US\$1,17, dan waktu bermain 369 menit.
- Klaster 2: Niche-Engagement Rendah dengan rata-rata 13.734 pemilik, harga US\$6,15, dan waktu bermain 230 menit.
- Proporsi anggota klaster relatif seimbang, yaitu 31,68%, 33,86%, dan 34,46%.
- Segmentasi ini membantu menggambarkan posisi game berdasarkan jangkauan pasar, strategi harga, dan engagement pemain.

Manfaat Penelitian

- Memberikan perbandingan metodologis antara K-Means dan K-Medoids pada data Steam yang sama.
- Membantu pemetaan pasar game berdasarkan popularitas, harga, dan keterlibatan pemain.
- Menjadi dasar eksploratif untuk analisis pemosisian produk, strategi harga, dan portofolio game.
- Memberikan referensi bagi penelitian lanjutan dengan fitur, periode data, atau algoritma klasterisasi yang lebih beragam.

Referensi

- [1] Ikotun et al., 2023 - K-means clustering algorithms: A comprehensive review.
- [3] Sreevalsan-Nair, 2023 - K-Medoids.
- [4] Schubert & Rousseeuw, 2021 - FasterPAM for K-Medoids clustering.
- [8] Batool & Hennig, 2021 - Clustering with average silhouette width.
- [10] Ros et al., 2023 - Davies-Bouldin Index for clustering evaluation.
- [13] Riani et al., 2023 - Yeo-Johnson transformation.
- [15] Clark et al., 2023 - Kruskal-Wallis power studies.
- [20] Bartolomé et al., 2025 - Video games dataset research context.

Penutup

Demikian presentasi karya tulis ilmiah ini.

Terima kasih atas waktu dan perhatian Bapak/Ibu dosen pembimbing serta penguji.



Saya mohon arahan dan masukan untuk perbaikan penelitian ini.