

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Perkembangan teknologi digital di era modern telah membawa dampak signifikan terhadap berbagai aspek, salah satunya adalah di bidang hiburan. Platform untuk *streaming* musik seperti Spotify telah menjadi salah satu layanan yang paling populer di dunia. Dengan terdapat lebih dari 100 juta lagu yang tersedia, Spotify mencatat per September 2024 (Q2), Spotify memiliki 640 juta pengguna aktif bulanan. Jumlah konten yang sangat besar ini memberikan kemudahan akses bagi pengguna, namun di sisi lain juga menimbulkan masalah berupa informasi yang berlebihan (*information overload*), hal ini membuat banyak pengguna mengalami kesulitan dalam memilih atau menemukan musik yang sesuai dengan preferensi mereka. Pada penelitian (BAUMAN & RASOR, 2019) Penelitian ini mengkaji bagaimana pengguna mencari musik di platform *streaming*. Ditemukan bahwa Spotify memiliki katalog musik yang sangat besar, sehingga “menemukan musik yang tepat bisa sangat menantang”. Untuk mengatasi permasalahan ini, dibutuhkan sebuah sistem rekomendasi yang mampu mengolah dan menyaring informasi secara efektif.

Sistem rekomendasi adalah alat penyaring informasi yang menyajikan item kepada pengguna berdasarkan preferensi dan perilaku mereka, misalnya rekomendasi tentang makalah ilmiah atau musik yang mungkin disukai pengguna (Afoudi et al., 2021). Teknik *collaborative filtering* adalah teknik yang populer digunakan untuk memberikan rekomendasi kepada pengguna (Paleti et al., 2021). Sistem Rekomendasi Musik berarti sistem yang berfungsi sebagai mekanisme untuk menyaring dan menyajikan lagu-lagu yang relevan berdasarkan preferensi pengguna. Sistem Rekomendasi Musik juga meningkatkan keterlibatan dan retensi pengguna di platform mendengarkan musik.

Dalam penelitian yang dilakukan oleh (Kumar & Prabhu, 2020), Sebagian besar tantangan dalam sistem rekomendasi terletak pada masalah *cold start*, yang mencakup kebutuhan pengguna baru untuk memberikan *rating* pada sejumlah item yang cukup agar sistem dapat mengamati akurasi dengan tepat, dan kemudian sistem akan memberikan rekomendasi yang dipersonalisasi berdasarkan kepercayaan. Selain itu, tantangan lainnya adalah kebutuhan untuk mengelola volume data yang sangat besar dan memastikan bahwa proses rekomendasi berjalan dengan efisien tanpa mengorbankan akurasi.

Dalam penelitian ini *Matrix Factorization* berbasis algoritma *Alternating Least Squares (ALS)* dan *HDBSCAN Clustering* memiliki potensi yang besar. *Matrix Factorization* sendiri telah terbukti mampu membuat prediksi yang sangat baik bahkan pada matriks yang sangat *sparse* (Davagdorj et al., 2020). Menurut (Paleti et al., 2021), ALS merupakan salah satu teknik *matrix factorization* yang efisien dalam sistem rekomendasi. Algoritma ALS digunakan untuk memprediksi rekomendasi dengan cara menemukan entri yang hilang dalam matriks yang diberikan. Selain itu, ALS memiliki keunggulan dalam konvergensi cepat melalui pendekatan optimisasi iteratif yang dapat diterapkan secara efisien pada dataset berskala besar. Di sisi lain untuk mengatasi *Cold Start User*, *HDBSCAN Clustering* akan digunakan. Menurut (Yin et al., 2023), pendekatan model menggunakan clustering berkontribusi untuk mengatasi masalah umum seperti *cold start* dan data *sparsity* yang sering ditemui dalam sistem rekomendasi berbasis *machine learning*.

Penelitian tentang Sistem Rekomendasi yang dilakukan oleh (Rao et al., 2021), Menyebutkan bahwa *Matrix Factorization* yang dikembangkan untuk mengolah data yang besar seperti beberapa juta item dan pengguna, *Alternating Least Squares (ALS)* bisa digunakan dalam melatih model *Matrix Factorization* menggunakan perhitungan Algoritma *Alternating Least Squares* dengan memecah matriks pengguna dan item menjadi dua matriks laten. Proses dilakukan secara iteratif (berulang-ulang) dengan memperbarui satu matriks sementara matriks lainnya masih bernilai tetap, sampai model mencapai konvergen, teknik ini bertujuan untuk meminimalkan *Cost Function* atau nilai error dengan mengungkapkan faktor-faktor laten tersembunyi dari pengguna dan item.

Penelitian yang sama, (Kumar & Prabhu, 2020), yang membahas tentang penggunaan *Hybrid Clustering* untuk Sistem Rekomendasi menyebutkan bahwa, Algoritma *clustering* diterapkan pada dataset MovieLens untuk mengelompokkan pengguna ke dalam berbagai *cluster*. *Cluster* dipilih secara stokastik pada tahap awal, kemudian pengguna dianalisis satu per satu untuk memperkirakan perbedaan dalam penilaian dan pusat *cluster*. Hasil dari penelitian ini menunjukkan bahwa sistem rekomendasi menawarkan skalabilitas dan efisiensi dalam rekomendasi dengan mengurangi masalah *cold start*.

Penelitian sebelumnya dalam bidang segmentasi, seperti yang dilakukan oleh (Kristanti et al., 2024), telah berhasil mendemonstrasikan keefektifan algoritma K-Means clustering. Studi mereka, yang menganalisis segmentasi pelanggan retail berdasarkan usia, pendapatan, dan model RFM (Recency, Frequency, Monetary), menghasilkan

pengelompokan yang jelas dan dapat ditindaklanjuti. Kelebihan K-Means yang ditunjukkan dalam penelitian tersebut antara lain efisiensi komputasi yang tinggi pada dataset yang besar ($n=1135$) dan kemudahan interpretasi hasil cluster yang berbentuk spherical dan well-separated. Berangkat dari kelebihan dan keterbatasan K-Means inilah, penelitian ini mengusulkan penggunaan algoritma HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) untuk melakukan segmentasi pada data million song dataset.

Dengan mempertimbangkan keahlian kedua algoritma yang diperkenalkan, Peningkatan Personalisasi Sistem Rekomendasi Musik berbasis Ocassion menggunakan *Matrix Factorization* berbasis ALS dan *Clustering HDBSCAN* menawarkan perekomendasi Musik yang relevan meskipun data pengguna baru masih terbatas. Dengan pendekatan berbasis kepadatan, HDBSCAN mengelompokkan pengguna ke dalam *cluster* berdasarkan area dengan kepadatan tinggi, kemudian membangun hierarki cluster untuk mengevaluasi stabilitas masing-masing *cluster*. (Kumar & Prabhu, 2020), *Clustering* bisa menangani *Cold Start user*. Selain itu, pengoptimalan sistem rekomendasi yang datanya sering kali *Sparse* dapat ditutupi oleh ALS, dimana algoritma ALS mengisi nilai nilai kosong, dari pengguna dan item (Rao et al., 2021).

Dari Beberapa Penelitian terdahulu, dalam bidang Sistem Rekomendasi, berbagai pendekatan telah dilakukan untuk mengatasi data yang sering kali *sparse* dan tentunya *cold start*. Tetapi, dalam Penelitian ini memperkenalkan aspek keterbaruan, berupa penggabungan Algoritma Clustering dengan Algoritma Optimasi tanpa bantuan Platform kerja *big data*. Penggunaan data yang konstan dan tidak berubah sangat cocok untuk pengimplementasian kedua algoritma tersebut dalam pengaktifan pemberian rekomendasi dan penanganan *cold start* pada model sistem rekomendasi musik.

Di luar tantangan teknis *sparsity* dan *cold start*, relevansi rekomendasi musik sangat bergantung pada konteks kegiatan pengguna, di mana preferensi musikal dapat berubah drastis, seperti membutuhkan musik dengan tempo tinggi untuk berkendara mobil (Orsini et al., 2024), musik dengan tempo rendah untuk belajar (Mathiesen et al., 2022), atau lagu santai saat makan (Felszeghy et al., 2023). Sistem rekomendasi lama mengabaikan hal ini sering gagal menyajikan item yang sesuai dengan kebutuhan sesaat pengguna, meskipun secara historis profilnya telah disesuaikan. Oleh karena itu, penelitian ini mengintegrasikan analisis fitur audio ke dalam model untuk menyaring rekomendasi berdasarkan kategori kegiatan spesifik.

Dengan mengintegrasikan kedua metode ini, diharapkan penelitian ini dapat memberikan kontribusi signifikan terhadap pengembangan sistem rekomendasi yang lebih efisien dan efektif. Selain itu, hasil penelitian ini diharapkan dapat menjadi dasar untuk pengembangan lebih lanjut dalam mengatasi berbagai tantangan dalam sistem rekomendasi modern, termasuk permasalahan *sparsity* dan *cold start*. Dengan demikian, penelitian ini memiliki potensi untuk memberikan dampak yang luas dalam meningkatkan kualitas layanan pada platform mendengarkan musik

1.2 RUMUSAN MASALAH

Berdasarkan penjabaran latar belakang di atas, pemberian rekomendasi musik sesuai dengan preferensi pengguna dan *ocassion* menjadi tantangan khusus, dengan terlalu banyaknya item yang dapat diakses dapat membuat pengguna kesulitan untuk memilih konten yang sesuai. Teknologi Rekomendasi Musik, khususnya yang berbasis *Matrix Factorization* berbasis *Alternating Least Squares*, telah menunjukkan potensi signifikan dalam memberikan rekomendasi dengan menyesuaikan kebiasaan pengguna maupun pengguna yang lain. Namun, teknologi ini masih rentan terhadap pengguna baru yang kurang atau bahkan belum melakukan interaksi terhadap aplikasi *streaming* musik, yang dapat membuat sistem sulit untuk melakukan rekomendasi. Dengan mengintegrasikan teknologi *High Density Based Clustering*, diharapkan dapat meningkatkan kemampuan sistem untuk merekomendasikan musik kepada pengguna sesuai dengan *ocassion* dan kebiasaan pengguna ataupun pengguna yang baru, maka dirumuskan masalah sebagai berikut :

1. Bagaimana algoritma *Alternating Least Squares* (ALS) dalam kerangka *Matrix Factorization* dan *High Density Based Clustering* dengan fitur audio dapat memberikan rekomendasi musik berdasarkan *ocassion* dan preferensi pengguna?
2. Bagaimana *Hierarchical Density-Based Clustering* dapat dimanfaatkan untuk mengelompokkan musik berdasarkan pola interaksi pengguna terhadap musik?

1.3 TUJUAN PENELITIAN

1. Mengembangkan model *Matrix Factorization* berbasis ALS dan *High Density Based Clustering* yang efektif untuk rekomendasi musik berbasis *ocassion*
2. Mengintegrasikan fitur audio (*tempo, energy, danceability, dan accousticness*) ke dalam model *Matrix Factorization* berbasis ALS untuk meningkatkan rekomendasi musik berbasis *ocassion*.

3. Membuat Sistem Personalisasi Rekomendasi Musik untuk kegiatan Berkendara, Belajar, dan Makan menggunakan Hybrid Matrix Factorization berbasis ALS dan HDBSCAN pada Million Song Dataset

1.4 MANFAAT PENELITIAN

1. Penelitian ini memberikan kontribusi teoretis dalam bidang sistem rekomendasi dengan mengintegrasikan *Matrix Factorization* berbasis ALS dan HDBSCAN Clustering sebagai solusi inovatif untuk mengatasi tiga tantangan utama: *data sparsity*, *cold start problem* dan Rekomendasi berbasis *ocassion* . Hasil penelitian diharapkan dapat menjadi referensi akademis untuk pengembangan sistem rekomendasi berbasis *density-based clustering* dan *latent factor models*.
2. Sistem rekomendasi yang dihasilkan dapat membantu pengguna menemukan lagu sesuai preferensi mereka secara lebih personal dan kebutuhan, Hal ini berdampak pada peningkatan *user engagement* dan kepuasan pengguna layanan *streaming* musik.

1.5 BATASAN MASALAH

Batasan-batasan dalam penelitian ini dirancang untuk memastikan fokus pada pengembangan dan evaluasi model *Matrix Factorization* berbasis ALS dan *High Density Based Clustering* sebagai solusi yang efisien untuk rekomendasi musik berbasis *ocassion* dengan bantuan filter fitur audio. Penelitian ini secara khusus tidak mencakup aspek-aspek di luar konteks pengenalan pola yang dapat memperluas ruang lingkup secara tidak terkendali. Adapun batasan-batasan penelitian ini sebagai berikut :

- a. Model merekomendasikan musik dalam bentuk daftar lagu (*Playlist*) berdasarkan konteks kegiatan yang diinginkan user (berkendara mobil, belajar, makan), tidak termasuk rekomendasi berdasarkan artis atau genre tertentu.
- b. Total lagu yang direkomendasikan kepada pengguna pada setiap playlist adalah sebanyak 50 lagu
- c. Pemberian Rekomendasi yang dilakukan melalui User Satisfaction dilakukan dengan similiaritas tertinggi antara preferensi pengisi kuisisioner dengan data user yang ada

