

# BAB I PENDAHULUAN

## 1.1 Latar Belakang Masalah

Pemanfaatan teknologi informasi telah merambah di berbagai bidang. Hal tersebut disebabkan begitu banyaknya tawaran kemudahan dalam aspek kehidupan manusia, seperti kemudahan mendapatkan dan menyebarkan informasi serta peningkatan layanan personal. Peningkatan layanan dilakukan dengan memanfaatkan berbagai teknik dan media agar perusahaan mampu bersaing. Salah satu media tersebut adalah sistem on-line, seperti *E-Learning*, *E-Government*, *E-Business*, *E-Library*, *E-Tourism* dan *E-Commerce* (B. Patel et al, 2018). Kelebihan dari sistem *on-line* yakni mampu memasarkan produk dengan cepat tanpa terkendala tempat dan waktu. Namun, seiring berjalannya waktu dan berkembangnya perusahaan maka semakin banyak dan beragam jumlah produk yang dipasarkan. Hal tersebut berdampak kepada pengguna yang mengalami kebingungan dalam memilih produk yang diinginkan sehingga menyita waktu (C. Liao & S. Lee, 2016). Untuk mengatasi hal tersebut, salah satu solusi yang digunakan adalah sistem rekomendasi (Y. Hu, 2014).

Sistem rekomendasi bekerja dengan cara mencari informasi paling relevan yang diambil dari sejumlah besar data untuk menghasilkan rekomendasi yang sesuai dengan minat *user*. Rekomendasi paling relevan terus digali dengan berbagai pendekatan, seperti *content-based filtering* (L. You et al, 2015) (Y. Xu & J. Yin, 2015), *demographic filtering*, *collaborative filtering* dan *hybrid filtering*.

*Collaborative filtering* (CF) bekerja dengan cara menganalisis pola data rating untuk membuat prediksi (A. Approach, G et al, 2015). Prediksi diperoleh setelah melakukan beberapa tahapan yakni mengumpulkan dan menganalisis informasi dalam jumlah besar mengenai perilaku, aktivitas, dan kecenderungan pengguna. Hasil analisis selanjutnya digunakan untuk memprediksi apa yang disukai oleh pengguna berdasarkan kemiripan dengan pengguna yang lainnya. CF memiliki beberapa kelebihan diantaranya mudah diimplementasikan dan dapat menyaring segala jenis informasi atau barang tanpa harus menganalisis komentar-komentar dari pengguna. Selain itu juga menghasilkan rekomendasi kualitas tinggi daripada sistem rekomendasi berdasarkan konten dan demografis (D. Lien & N. Phoung, 2014). Namun, CF menghadapi masalah krusial yaitu *cold start*, *sparsity* dan *scalability* (S. Sharma & A. Mahajan, 2017). *Cold start* adalah suatu kondisi di mana pengguna baru yang belum pernah memberikan rating terhadap suatu produk, sehingga informasi yang didapatkan untuk arah peminatan pengguna sulit diketahui. Jika arah peminatan pengguna tidak diketahui maka tidak bisa merekomendasikan produk. Oleh karena itu maka penelitian ini mengusulkan pendekatan *clustering* untuk meringankan permasalahan *cold start*. Adapun algoritma yang akan digunakan untuk clustering adalah DBSCAN dan K-Means. Dilanjutkan dengan perhitungan *similarity* sehingga diperoleh rekomendasi yang sesuai dengan peminatan *user*. Berdasarkan latar belakang yang telah diuraikan diatas dapat disimpulkan bahwa dibutuhkan sebuah penelitian yang berjudul ”Komparasi DBSCAN dan K-Means Dalam Mengatasi Masalah *Cold Star* Pada *Collaborative*”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka ditemukan rumusan masalah yaitu bagaimana menyelesaikan masalah *cold start* yang diderita oleh *collaborative filtering* dengan menggunakan algoritma *clustering* DBSCAN dan juga K-Means. Permasalahan krusial pada *Collaborative Filtering* adalah *cold start* dimana *user* baru belum pernah memberikan rating terhadap produk yang ada, sehingga belum bisa diketahui arah peminatan dari *user* tersebut.

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Algoritma yang digunakan adalah algoritma clustering K-Means dan DBSCAN
2. Dataset didapatkan dari Grouplens.com

## 1.4 Tujuan Penelitian

1. Menyelesaikan masalah *cold start* yang diderita oleh *Collaborative Filtering* dengan pendekatan *clustering*, sehingga rekomendasi yang dihasilkan lebih berkualitas.
2. Melakukan perbandingan kinerja dari pendekatan *clustering* yang digunakan yaitu DBSCAN dan K-Means.

### 1.5 Manfaat Penelitian

Permasalahan krusial pada penelitian ini adalah *cold start* dimana *user* baru belum pernah memberikan rating terhadap produk yang ada, sehingga belum bisa diketahui arah peminatan dari *user* tersebut. Kondisi ini akan mempengaruhi terhadap kualitas layanan karena tidak bisa memberikan rekomendasi produk dengan tepat. Oleh karena itu penelitian ini mengusulkan pendekatan *clustering* untuk meringankan permasalahan *cold start* yang terjadi. *Clustering* akan dilakukan terhadap data demografis yaitu umur dengan algoritma DBSCAN dan K-Means yang selanjutnya akan dilakukan perhitungan *similarity* sehingga menghasilkan rekomendasi yang berkualitas sesuai dengan preferensi *user*.

### 1.6 Sistematika Penulisan

Uraian singkat mengenai sistematika penulisan tugas akhir ini adalah.

1. Bab 1 Pendahuluan

Pada bab ini tercantum latar belakang, perumusan masalah, ruang lingkup penelitian, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

2. Bab 2 Tinjauan Pustaka

Pada bab ini memuat penelitian terdahulu dan teori-teori yang mendukung penelitian yang akan dilakukan oleh penulis mengenai DBSCAN, K-Means, Data mining dll.

3. Bab 3 Metodologi Penelitian

Pada bab ini memuat metodologi penelitian yang digunakan oleh penulis pada penelitian ini

#### 4. Bab 4 Hasil dan Pembahasan

Pada bab ini berisi tentang pemaparan hasil analisa persoalan mengenai rekomendasi film berdasarkan umur user

#### 5. Bab 5 Simpulan dan Saran

Pada bab ini berisi tentang rangkuman dari penelitian yang telah dilakukan oleh penulis yang berisikan apakah penelitian ini memecahkan permasalahan yang diangkat atau tidak