

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perpustakaan merupakan salah satu tempat paling penting dalam dunia Pendidikan dan pengembangan ilmu pengetahuan. Salah satu peran perpustakaan adalah menyediakan layanan peminjaman buku, layanan ini tidak hanya mencerminkan minat baca masyarakat tetapi juga sebagai tolak ukur dalam menilai efektivitas perpustakaan dapat memberikan manfaat kepada masyarakat. Untuk itu, petugas perpustakaan dapat menganalisis data peminjaman buku guna mengetahui tren, selera pemustaka, dan kebutuhan koleksi, yang nantinya dapat digunakan sebagai acuan dalam pengadaan koleksi. Sebagai pusat pengetahuan perpustakaan menyediakan berbagai sumber daya, termasuk buku, jurnal dan media digital [1].

Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur berperan sebagai *public administration* untuk melestarikan dan meningkatkan kegemaran membaca di Masyarakat. Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur setiap tahun akan melakukan evaluasi untuk mencari anggota yang aktif dalam peminjaman buku. Namun, proses pengevaluasian tersebut masih dilakukan secara manual tanpa dilakukan analisis yang mendalam, sehingga hasilnya belum mampu memberikan gambaran yang lengkap mengenai pola meminjam buku seluruh anggota. Selain itu, anggota yang dianggap aktif biasanya diberi hadiah berupa *souvenir*, tanpa adanya tindak lanjut yang strategis untuk meningkatkan keterlibatan anggota yang kurang aktif. Hal ini semakin diperparah oleh ketiadaan klasifikasi yang jelas mengenai tingkat keaktifan anggota dalam meminjam buku, sehingga perpustakaan cukup kesulitan dalam merancang strategi layanan yang sesuai dengan kebutuhan masing-masing kelompok anggota. Minat masyarakat dalam memanfaatkan koleksi buku di perpustakaan beragam, karena beberapa ada yang memiliki ketertarikan pada topik yang sama [2].

Berdasarkan permasalahan tersebut, penelitian ini difokuskan pada ketiadaan klasifikasi tingkat keaktifan anggota dalam peminjaman buku. Sampai saat ini, evaluasi yang dilakukan hanya sampai pada penentuan anggota paling aktif untuk diberi penghargaan berupa *souvenir*, belum menghasilkan informasi

terstruktur yang mampu mengelompokkan seluruh anggota secara sistematis. Akibatnya, pola perilaku anggota yang kurang aktif tidak terlihat dengan jelas, sehingga perpustakaan kesulitan dalam menyusun strategi layanan yang tepat sasaran. Kondisi ini menunjukkan perlunya metode analisis yang mampu mengelompokkan anggota berdasarkan tingkat keaktifan mereka, sehingga hasil evaluasi tidak hanya bersifat administratif, tetapi juga bisa menjamin kepuasan anggota perpustakaan. Permasalahan yang terjadi menunjukkan bahwa kebutuhan informasi pengguna sering kali belum terpenuhi secara optimal, sehingga informasi yang diperoleh belum sepenuhnya sesuai dengan harapan [3].

Beberapa penelitian terdahulu telah membahas cara mengelompokkan tingkat keaktifan anggota perpustakaan dengan metode *Clustering*. Penelitian yang berjudul “Pengelompokan Keaktifan Anggota Perpustakaan Menggunakan Algoritma *K-Means*” menganalisis data kunjungan dan peminjaman buku di Perpustakaan Rimba Baca Jakarta Selatan selama tahun 2023–2024, yang mencakup 81 data keanggotaan. Hasil penelitian menunjukkan adanya tiga kelompok anggota dengan karakteristik berbeda, yaitu anggota yang sangat aktif, aktif, dan kurang aktif. Temuan ini bisa membantu perpustakaan dalam merancang strategi pelayanan yang lebih baik. Namun, metode *K-Means* memiliki beberapa kekurangan, seperti harus menentukan jumlah kelompok sejak awal dan mudah terpengaruh oleh data *outlier* [4].

Penelitian lain yang dilakukan di Perpustakaan SMA Assisi Pematangsiantar menggunakan metode *K-Means* untuk mengelompokkan data peminjaman buku oleh siswa. Penelitian ini berhasil menemukan pola peminjaman yang berbeda di antara berbagai kelompok siswa, tetapi juga menghadapi masalah yang sama, yaitu sensitivitas terhadap data yang tidak normal dan keharusan menentukan jumlah kelompok sejak awal [5]. Di sisi lain, penelitian dengan judul “Pengelompokan Jumlah Kunjungan Mahasiswa ke Perpustakaan Kampus Menggunakan Algoritma *DBSCAN*” dalam jurnal G-Tech menggunakan metode *DBSCAN* untuk mengelompokkan mahasiswa berdasarkan jumlah kunjungan ke perpustakaan dan IPK. Hasil penelitian tersebut menunjukkan bahwa *DBSCAN* mampu membuat kelompok dengan pola yang lebih fleksibel dan dapat mengidentifikasi mahasiswa yang jarang berkunjung sebagai data *outlier*. Dari

berbagai penelitian tersebut dapat dilihat bahwa meskipun metode *K-Means* cukup banyak digunakan, namun metode *DBSCAN* memiliki kelebihan dalam menangani data yang tidak rapi dan adanya data yang tidak normal, sehingga dapat lebih cocok digunakan dalam menganalisis keaktifan anggota perpustakaan [6].

Metode *Clustering* seperti *K-Means* dan *DBSCAN* (*Density-Based Spatial Clustering of Applications with Noise*) masing-masing memiliki kelebihan dan kekurangan. *K-Means Clustering* menjadi salah satu metode yang efektif untuk menganalisis data [7]. Metode *K-Means* memiliki beberapa kelebihan, seperti mudah diimplementasikan, proses perhitungannya cepat, dan hasil yang didapatkan bisa dipahami dengan mudah karena setiap kelompok data dinyatakan oleh titik pusatnya. Akan tetapi, *K-Means* juga memiliki beberapa kekurangan, yaitu harus menentukan jumlah kelompok ( $k$ ) sejak awal, rentan terhadap nilai yang tidak biasa atau gangguan, serta tidak mampu menangani kelompok dengan bentuk tidak teratur. Sementara itu, *DBSCAN* menawarkan fleksibilitas yang lebih besar dalam menangani data dengan bentuk *cluster* yang tidak beraturan dan dalam mengelola *outlier* [8]. *DBSCAN* lebih baik dalam membentuk kelompok dengan bentuk apa pun, tidak memerlukan jumlah kelompok yang sudah ditentukan sejak awal, serta mampu mengenali data yang tidak termasuk dalam kelompok secara otomatis. Namun, *DBSCAN* tergantung pada nilai parameter *epsilon* dan *MinPts* yang harus ditentukan dengan tepat agar hasilnya optimal [9].

Meskipun beberapa penelitian sebelumnya sudah membahas penggunaan metode *Clustering* seperti *K-Means* untuk mengelompokkan data perpustakaan, sebagian besar penelitian tersebut hanya fokus pada frekuensi peminjaman tanpa mempertimbangkan keterlambatan dan buku yang belum dikembalikan. Selain itu, metode yang digunakan biasanya memiliki kekurangan, seperti memerlukan penentuan jumlah kelompok sejak awal dan kurang efektif dalam menghadapi data yang memiliki *noise* atau variasi kepadatan. Dengan demikian, masih ada celah penelitian dalam pengelompokan Tingkat keaktifan anggota perpustakaan dengan metode yang mampu menangani *noise* seperti *DBSCAN* [10].

Dari literatur yang sudah penulis baca diatas, penulis mengambil topik mengenai pengelompokan Tingkat keaktifan anggota perpustakaan dengan memanfaatkan data peminjaman buku. Topik ini cukup penting karena peminjaman

buku merupakan salah satu indikator utama dalam melihat bagaimana anggota memanfaatkan layanan perpustakaan. Namun, masih ada *gap* dalam penelitian terkait pengelompokan tingkat keaktifan anggota perpustakaan. Sebagian besar penelitian sebelumnya hanya berfokus pada frekuensi pinjaman tanpa mempertimbangkan faktor lain seperti keterlambatan dan buku yang belum dikembalikan, padahal dua aspek tersebut juga mencerminkan tingkat kedisiplinan dan konsistensi anggota dalam memanfaatkan layanan perpustakaan. Selain itu, metode yang banyak digunakan seperti *K-Means* memiliki keterbatasan dalam menangani data dengan variasi kepadatan dan keberadaan *noise*. Hingga saat ini juga, belum banyak penelitian di bidang perpustakaan yang secara *eksplisit* menggabungkan *Principal Component Analysis (PCA)* dengan metode *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)*. Berdasarkan *gap* tersebut, penelitian ini mengusulkan penggunaan metode *Density-Based Spatial Clustering of Applications with Noise (DBSCAN)* yang mampu mengelompokkan data berdasarkan tingkat kepadatan dan mendeteksi data yang tidak termasuk kelompok (*outlier*). Dengan menambahkan atribut keterlambatan dan buku belum dikembalikan dan kombinasi *PCA-DBSCAN*, penelitian ini diharapkan bisa membuat klasifikasi anggota perpustakaan berdasarkan tingkat keaktifan dalam peminjaman buku secara objektif, terstruktur, dan menyeluruh. Sehingga hasil penelitian ini dapat dimanfaatkan oleh Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur dalam Menyusun strategi layanan dan meningkatkan partisipasi anggota perpustakaan.

## **1.2 Rumusan Masalah**

### **1.2.1 Permasalahan**

Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur setiap tahun melakukan evaluasi untuk mengetahui seberapa aktif anggota perpustakaan dalam peminjaman buku. Namun, pada proses evaluasi tersebut dilakukan hanya sampai pada penentuan anggota paling aktif untuk diberi penghargaan berupa *souvenir*, tanpa ada analisis yang mampu mengelompokkan seluruh anggota secara sistematis. Hal ini membuat pola perilaku anggota tidak diketahui secara baik, sehingga sulit mengelompokkan anggota menjadi yang aktif, pasif, atau tidak aktif. Dari sisi informatika, masalah ini menunjukkan perlunya metode analisis data yang

lebih baik, sehingga bisa membagi anggota berdasarkan tingkat keseringan meminjam buku secara akurat, cepat, dan mampu menangani data yang tidak lengkap atau tidak normal.

### **1.2.2 Metode Usulan**

Untuk menyelesaikan masalah tersebut, penelitian ini menyarankan penggunaan metode *DBSCAN (Density-Based Spatial Clustering of Applications with Noise)*. Metode ini dipilih karena bisa mengelompokkan data tanpa perlu menentukan jumlah kelompok terlebih dahulu, mampu mendeteksi kelompok dengan bentuk yang tidak teratur, serta lebih kuat dalam menghadapi data yang tidak relevan atau gangguan dibandingkan metode *Clustering* lain seperti *K-Means*.

### **1.2.3 Pertanyaan Penelitian**

1. Bagaimana pengaruh penerapan metode *DBSCAN* dalam mengelompokkan data anggota perpustakaan berdasarkan tingkat kesamaan keaktifan?
2. Bagaimana kualitas hasil pengelompokan yang dihasilkan *DBSCAN* berdasarkan nilai *Silhouette Coefficient*?

## **1.3 Tujuan & Manfaat**

### **1.3.1 Tujuan Penelitian**

1. Mendapatkan hasil pengaruh penerapan metode *DBSCAN* dalam mengelompokkan anggota perpustakaan berdasarkan tingkat aktivitas mereka.
2. Mendapatkan kualitas hasil *Clustering* dengan menggunakan metrik *Silhouette Coefficient* untuk mengukur efektivitas metode *DBSCAN*.

### **1.3.2 Manfaat**

Dari penelitian ini, diperoleh model pengelompokan berbasis metode *DBSCAN* yang mampu mengelompokkan tingkat keaktifan anggota perpustakaan secara lebih tepat. Hasil pengelompokan tersebut memberikan manfaat berupa informasi analitis yang menggambarkan pola keaktifan anggota perpustakaan, sehingga evaluasi keanggotaan yang semula bersifat administratif dapat dilengkapi dengan hasil klasifikasi yang lebih terstruktur dan sistematis.

#### 1.4 Batasan-batasan

1. Data yang digunakan dalam penelitian ini dibatasi pada data peminjaman buku anggota yang tercatat di Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur.
2. Penelitian ini menggunakan data yang mencakup frekuensi peminjaman, rata-rata keterlambatan, serta jumlah buku yang belum dikembalikan selama periode tahun 2021 hingga 2025.
3. Analisis difokuskan pada keaktifan anggota tanpa mempertimbangkan jenis buku atau lama menjadi anggota perpustakaan.
4. Penelitian ini hanya berfokus sampai pada proses *clustering* saja tanpa adanya evaluasi lanjutan.
5. Metode *Clustering* yang digunakan dibatasi pada *DBSCAN* dengan bantuan *PCA* untuk reduksi dimensi.
6. Evaluasi hasil dilakukan menggunakan *Silhouette Coefficient*.

#### 1.5 Sistematika Penulisan

Penulisan skripsi ini disusun secara sistematis agar pembahasan penelitian dapat terstruktur dengan jelas dan terarah. Berikut sistematika penulisan dalam proposal ini:

##### BAB I PENDAHULUAN

Bab ini berisi uraian awal tentang latar belakang penelitian, rumusan masalah, batasan dalam penelitian, tujuan dari penelitian, serta manfaat dari penelitian ini.

##### BAB II KAJIAN PUSTAKA

Bab ini menjelaskan berbagai teori dan konsep dasar sebagai dasar dalam penelitian, termasuk Data Mining, *Clustering*, *PCA*, metode *DBSCAN*, *Silhouette Coefficient*, serta penelitian sebelumnya yang menjadi referensi dalam penelitian ini.

##### BAB III METODE USULAN

Bab ini membahas tentang rencana penelitian, tahapan penelitian, dataset, *preprocessing*, proses *Clustering DBSCAN*, validasi hasil, skema uji coba serta alat dan lingkungan yang digunakan dalam penelitian ini.

#### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi hasil penelitian yang telah dilakukan serta pembahasan dari hasil tersebut. Pada bab ini dijelaskan proses clustering menggunakan metode DBSCAN baik tanpa reduksi dimensi maupun dengan reduksi dimensi menggunakan PCA. Selain itu, pada bab ini juga disajikan perbandingan hasil clustering untuk mengetahui metode yang menghasilkan pengelompokan terbaik.

#### **BAB V PENUTUP**

Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan serta saran yang dapat diberikan untuk pengembangan penelitian selanjutnya.