

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi saat ini telah memberikan dampak pada beberapa bidang intansi seperti di perguruan tinggi, dengan kemajuan yang ada setiap perguruan tinggi akan memberikan layanan fasilitas untuk mengakses informasi lebih cepat dan mendapatkan sumber yang jelas serta lengkap. Tempat atau wadah mencari sebuah informasi yang disediakan perguruan tinggi yaitu Perpustakaan dimana Perpustakaan masih banyak digemari oleh para pembaca buku dan menjadi pilihan alternatif bagi para mahasiswa untuk mencari referensi penelitian.

Perpustakaan merupakan salah satu sumber informasi yang dapat dimanfaatkan untuk memenuhi suatu kebutuhan informasi ilmiah oleh mahasiswa sehingga disetiap perguruan tinggi diwajibkan memiliki perpustakaan salah satunya Universitas Bina Darma Palembang. Perpustakaan Universitas Bina Darma Palembang memiliki tingkat pengunjung dan anggota yang cukup tinggi sehingga terdapat koleksi buku yang sudah lengkap dan telah terintegrasi dengan baik secara terkomputerisasi. Dari adanya kumpulan koleksi data buku yang sering di inputkan, Perpustakaan harus *intens update* data buku sehingga akan membutuhkan suatu pengelompokan buku untuk mengetahui informasi data setiap kelompok buku dan membantu mengelompokkan buku secara otomatis berdasarkan subjek yang di

butuhkan perpustakaan dan akan analisa menjadi suatu alat pengambilan sebuah keputusan.

Pengelompokan data buku akan di bagi menjadi beberapa klasterisasi sehingga menggunakan perbandingan dari dua metode yang berbeda untuk mendapatkan informasi yang lebih akurat mutlak dan sesuai kebutuhan. Pada penelitian ini akan menggunakan perbandingan metode K-Means dan K-Medoids karena kedua metode ini memiliki perbedaan pada proses implementasinya. K-means merupakan algoritma non hirarki yang berasal dari metode dan *clustering*, penerapan metode K-Means mempartisi data dalam suatu kumpulan kelompok data dimana data yang memiliki representatif persamaan dalam pengelompokan akan memiliki perbedaan dengan kelompok lain dan dimulai dengan pembentukan partisi klaster secara interaktif sehingga terjadi perubahan signifikan pada partisi cluster, prinsip utama dari proses metode ini adalah menyusun K buah partisi/rata-rata(mean) dari sekumpulan data.

Sedangkan metode K-Medoids merupakan salah satu algoritma *clustering* yang memiliki kelebihan yaitu dapat mengatasi sensitive terhadap outlier, K-Medoids menentukan *centroid* secara acak, menghitung jarak dari setiap data terhadap masing-masing *centroid* kemudian dalam perhitungan jarak menggunakan rumus *euclidean distance*, metode K-Medoids dikenal sebagai salah satu metode yang menghasilkan tingkat akurasi tinggi dan nilai lebih efisien untuk melakukan *clustering* data dataset yang lebih besar.

Dan dapat dilihat sudah ada dan cukup banyak para peneliti terdahulu menggunakan dua metode tersebut, berikut beberapa penelitian terdahulu yang mengimplementasikan metode K-Means dan K-Medoids sebagai klasterisasi data serta alat pengambil keputusan.

Penelitian (Sriliana and Sriningsih 2023) dari penelitian tersebut menggunakan dua metode, analisis cluster yang digunakan pada penelitian ini adalah K-Means dan K-Medoids dengan validitas Silhouette. Berdasarkan hasil dari penelitian analisis diketahui bahwa nilai Silhouette k-means (0.8018) lebih tinggi dibandingkan dengan nilai Silhouette k-medoids (0.7281). dalam hal ini dapat disimpulkan bahwa metode k-means lebih baik dibandingkan dengan k-medoids. Hasil analisis cluster menggunakan K-Means adalah 1) PTN dengan produktivitas publikasi ilmiah tinggi yaitu ITB, ITS, UGM, dan UI. Keempat PTN tersebut adalah PTN Badan Hukum (PTN-BH) yang berlokasi di Pulau Jawa, 2) PTN dengan produktivitas publikasi ilmiah sedang terdiri dari 16 PTN yang didominasi oleh PTN-BH dan PTN Badan Layanan Umum (PTN-BLU) dengan jumlah terbesar lokasi di Pulau Jawa, dan 3) PTN dengan produktivitas publikasi ilmiah rendah terdiri dari 102 PTN yang didominasi oleh PTN sebagai pengelola keuangan negara umum (PTN-Satker) dengan lokasi terbanyak di luar Pulau Jawa.

Penelitian berikutnya dari (Arinaldi, Soesanto, and Anggana 2023) pada penelitian menggunakan metode K-Means *Clustering* yang didukung dengan menggunakan visualisasi Interface R-Shiny di R-Studio dengan bahasa pemrograman R. Hasil dari penelitian ini adalah melakukan pencarian cluster optimal yang dilakukan sebanyak 8 kali, sehingga setelah dilakukan proses

perhitungan *silhouette* terhadap data, hasil *silhouette* didapatkan dengan menghasilkan 5 pilihan paket wisata alam di Kabupaten Rembang dengan karakteristik lokasi wisata yang berbeda- beda.

Penelitian selanjutnya (Mantik 2022) dari penelitian menggunakan metode K-Means dan Kmedoids hasil *clustering* menggunakan tools Rstudio. Hasil perbandingan tersebut diperoleh jumlah cluster yang optimal dan diperoleh hasil cluster tinggi, sedang, dan rendah. Hasil perhitungan K-Means berdasarkan rata-rata pada cluster 1 yang tertinggi adalah Tempo dengan nilai 118, pada cluster 2 tertinggi adalah Tempo dengan nilai 125, dan pada cluster 3 yang tertinggi adalah Tempo dengan nilai sebesar 123. Perhitungan K-Medoids berdasarkan rata-rata pada cluster 1 yang tertinggi adalah Tempo dengan nilai 129, pada cluster 2 tertinggi adalah Tempo dengan nilai 122, dan pada cluster 3 yang tertinggi adalah Tempo dengan nilai dari 110.

dari (Prasetyo et al. 2023) menggunakan metode K-Means, hasil dari penelitian ini yaitu Agar koleksi bahan pustaka di dalam Perpustakaan dapat dimanfaatkan secara maksimal, harus diterapkan suatu sistem untuk mengatur koleksi buku tersebut. Saat ini, sistem yang diterapkan di Perpustakaan berjalan kurang efektif, Pengelompokan yang baik adalah pengelompokan yang memiliki homogenitas yang tinggi. Model Perpustakaan yang diteliti, menggunakan metode K-Means *clustering*. Berdasarkan hasil uji sistem yang telah dirancang, sistem mampu memecahkan permasalahan di dalam penentuan minat baca pengunjung pada Perpustakaan STMIK Triguna Dharma Medan. Berdasarkan hasil *Clustering* k-Means permasalahan yang terjadi dengan menentukan minat baca pengunjung

pada Perpustakaan STMIK Triguna Dharma Medan dapat diselesaikan dengan menerapkan metode K-Means *Clustering*. Penerapan metode K-Means *clustering* dilakukan dengan cara menghitung nilai titik pusat centroid berdasarkan algoritma K-Means dan hasilnya diperoleh dengan nilai sangat minat, minat dan kurang minat.

Dari penelitian terdahulu di atas dapat diketahui bahwa perbandingan dua metode antara K-Means dan K-Medoids dapat diimplementasikan dengan baik menggunakan data klasterisasi yang ada secara terintegrasi sesuai kebutuhan setiap instansi, dan dapat diproses dari beberapa tahapan dari metode K-Means maupun K-Medoids, kedua metode ini juga memiliki beberapa kelebihan dan kekurangan yang bisa dipertimbangkan sebagai alat pengambil keputusan yang efektif sesuai fungsi dan kebutuhan objek/instansi.

Penerapan perbandingan dari dua metode tersebut akan memberikan dan menghasilkan metode mana yang cocok dan sesuai fungsi kebutuhan sehingga digunakan untuk pengelompokan klasterisasi buku. Oleh karena itu pada penelitian ini akan diimplementasikan metode tersebut dan dicari hasil akurasi dengan metode K-Means dan K-Medoids dan hasil pengelompokan dari tingkat akurasi klasterisasi tersebut, sehingga dapat menggali informasi mengenai hasil data mining pada pengelompokan buku Perpustakaan universitas Bina Darma Palembang. Maka dari permasalahan yang ada dan berdasarkan uraian dan fenomena diatas maka peneliti tertarik untuk menyusun penelitian dengan judul **“Analisa Dan Perbandingan Metode Klasterisasi Untuk Mengelompokkan Koleksi Buku Perpustakaan”**.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana perbandingan metode K-Means dan K-Medoids dalam pengelompokan klasterisasi buku dan pada Perpustakaan Universitas Bina Darma?
2. Bagaimana mengimplementasikan metode K-Means dan K-Medoids dalam pengelompokan klasterisasi buku pada Perpustakaan Universitas Bina Darma?
3. Bagaimana kinerja metode K-Means dan K-Medoids dalam pengelompokan klasterisasi buku pada Perpustakaan Universitas Bina Darma?

1.3 Batasan Masalah

Berdasarkan uraian latar belakang diatas, maka batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Mengelompokan data buku pada Perpustakaan Universitas Bina Darma.
2. Menguji tingkat akurasi metode K-Means dan K-Medoids dalam klasterisasi pengelompokan buku dan Perpustakaan Universitas Bina Darma.
3. Data yang dipakai adalah data klasterisasi buku dan pada Perpustakaan Universitas Bina Darma periode tahun 2020-2021.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk menentukan perbandingan metode K-Means dengan K-Medoids pada pengelompokan klasterisasi data buku Perpustakaan Universitas Bina Darma.
2. Untuk mengetahui metode mana yang sesuai fungsi dan kebutuhan dari data buku Perpustakaan Universitas Bina Darma.
3. Untuk mengembangkan suatu metode pengambilan penunjang keputusan secara terintegrasi pada Perpustakaan Universitas Bina Darma.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian yang dilakukan oleh penulis adalah:

1. Dapat mengelompokkan klasterisasi buku dalam melengkapi dan sebagai informasi data guna untuk proses akreditasi pada Perpustakaan Universitas Bina Darma.
2. Dapat menyajikan secara multi dimensi dengan cepat dan tepat dari data buku sehingga mempermudah pelaporan dan analisa data pada Perpustakaan Universitas Bina Darma.
3. Dengan menerapkan metode K-Means dan K-Medoids sebagai perbandingan model baru digunakan untuk pengelompokkan klasterisasi data buku dan pada Perpustakaan Universitas Bina Darma.
4. Menghasilkan informasi data buku yang akan dijadikan sebagai analisa data mining pada Perpustakaan Universitas Bina Darma.

1.6 Ruang Lingkup

Untuk ruang lingkup pada penelitian ini akan berfokus pada beberapa hal yaitu pada pengelompokan klasterisasi menggunakan data buku Perpustakaan Universitas Bina Darma dan akan mengimplementasikan perbandingan antara dua metode yaitu K-Means dan K-Medoids.

1.7 Susunan Dan Struktur Tesis

Susunan dan struktur tesis dibuat dengan maksud dapat memberikan garis-garis besar dari penulisan sehingga hubungan antara bab satu dengan bab yang lain dapat terlihat dengan jelas. Adapun susunan dan struktur proposal tesis adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini penulis membahas mengenai latar belakang, identifikasi masalah, batasan masalah, perumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian, serta susunan dan struktur proposal tesis.

BAB II KAJIAN PUSTAKA

Pada bab ini membahas tentang kajian pustaka, penelitian terdahulu, kerangka berfikir, dan jadwal penelitian.

BAB III METODOLOGI PENELITIAN

Pada bab ini pembahasan terdiri dari desain dan jadwal penelitian, data penelitian meliputi jenis data, populasi dan sampel penelitian, kemudian konsep dan metode penelitian yang digunakan, metode pengumpulan data serta teknik pengelompokan data.

BAB IV ANALISIS DATA

Pada bab ini pembahasan terdiri dari pengolahan data terhadap dataset yang akan digunakan sebelum diolah dan di proses dengan bantuan *tools google colab*. Pada tahap ini dataset yang digunakan akan di proses mulai dari *selection, cleaning, transformation* sebelum akhirnya masuk pada proses data mining yang akan dibahas pada bab selanjutnya.

BAB V HASIL DAN PEMBAHASAN

Pada bab ini pembahasan merupakan hasil dan pembahasan terhadap dataset yang digunakan pada penelitian ini, dengan menggunakan perbandingan metode K-Means dan K-Medoids dan bantuan *tools google colab* akan menghasilkan akurasi dari penelitian ini.

BAB VI PENUTUP

Pada bab ini merupakan penutup dari rangkaian penelitian yang dilakukan terhadap penelitian yang berjudul analisa perbandingan metode klasterisasi untuk mengelompokkan koleksi buku Perpustakaan yang ada di Perpustakaan Universitas Bina Darma. Pada bab ini akan dibahas kesimpulan dan saran dari keseluruhan penelitian yang sudah dijalankan oleh peneliti.

LAMPIRAN

Berisi lampiran pendukung dari penelitian.