

BAB I

PENDAHULUAN

1.1 Latar Belakang

Knowledge Discovery in Database (KDD) merupakan proses analisa yang terstruktur bertujuan mendapatkan informasi yang baru, benar, menemukan pola dari data yang kompleks, dan bermanfaat. *Data mining* merupakan inti dari proses KDD (Zanuardi & Suprayitno, 2018) dengan menggunakan algoritma-algoritma matematika yang bertujuan untuk mengekstrak informasi untuk menemukan pola, analisa, maupun prediksi kemungkinan tindakan di waktu mendatang (Muttaqin & Defriani, 2020).

Data mining merupakan sebuah proses penggalian pengetahuan yang baru, dengan menggali informasi data dalam jumlah yang banyak dengan teknik matematika dan statistik (Sugiono, et al. 2019). *Data mining* tidak hanya diimplementasikan untuk mendapatkan pengetahuan atau informasi baru, tetapi juga dapat meningkatkan pemahaman mengenai yang diketahui (Muttaqin & Defriani, 2020). Ataupun sebagai rangkaian proses untuk penentuan pola yang menarik dalam ruang lingkup yang besar (Sari, et al. 2018).

Clustering yaitu proses pengelompokan objek data yang memiliki kecenderungan satu sama lain kedalam *cluster* yang sama dan berbeda objek yang ada di kluster lain (Hidayad, et al. 2020). *Clustering* merupakan metode yang cocok untuk pengoptimalisasikan pelayanan perpustakaan dikarenakan dapat mengklusterisasikan buku dengan dengan efektif dan efisien, dengan algoritma *K-Means* data dapat di clustering dan informasi dari setiap nilai *centroid* dari setiap *cluster* (Sugiono, et al. 2019). *K-means* merupakan algoritma yang mengklusterisasikan berdasarkan partisi dan melakukan *clustering* melalui proses

iterasi berkelanjutan sampai dengan bertemu kondisi akhir, proses iterasi berhenti dan hasil *clustering* adalah *output* (Wang, et al. 2018)

K-Means clustering adalah pembelajaran mesin tanpa pengawasan yang mengelompokkan data menjadi k *cluster* eksklusif. *K-Means* beroperasi pada pengamatan dari dataset yang disediakan untuk membuat *cluster* (Wye, et al. 2019). *Algoritma K-Means* terdiri dari klasterisasi objek menjadi *cluster* yang saling eksklusif atau berbeda. Untuk menemukan clustering, algoritma ini menggunakan proses interaktif yang bertujuan meminimalkan jumlah jarak masing-masing tangki dalam kaitannya dengan *centroid* dari setiap kelompok yang akan menjadi tangki paling representatif dari setiap *cluster* (Oliveira, 2019).

K-Means tidak hanya di kembangkan pada bidang teknologi saja, akan tetapi pada bidang kedokteranpun bisa di implementasikan seperti pada topik *Mammography Images Segmentation* dan *Identification of spatial variations in COVID-19 epidemiological data*. Topik yang di angkat dari permasalahan yang marak saat ini dengan pemantaatan algoritma *K-Means Clustering* (Kamil & Salih, 2018) (Chandu, 2020).

Berdasarkan penelitian sebelumnya, penulis melakukan penelitian optimalisasi pelayanan perpustakaan terhadap minat baca dengan metode *K-Means clustering*, penitiliaan ini dilakukan di Pondok Pesantren Diniyah Limo Jurai sehingga perpustakaan dapat di *clustering* dengan optimal dan lebih terstruktur. Oleh karena itu di angkatlah judul “Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode *K-Means Clustering* (Studi Kasus di Pondok Pesantren Diniyah Limo Jurai Sungai Pua)”.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka dapat dirumuskan dua permasalahan sebagai berikut:

1. Bagaimana *K-Means Clustering* dapat membantu pihak perpustakaan dalam optimalisasi pelayanannya berdasarkan minat baca siswa?
2. Bagaimana metode *K-Means Clustering* dalam pengklasteran optimalisasi pelayanan perpustakaan?

1.3 Batasan Masalah

Agar penelitian ini lebih terarah maka penelitian diberi batasan-batasan sebagai berikut:

1. Data yang di olah hanya data perpustakaan pada Pondok Pesantren Diniyah Limo Jurai Sungai Pua dengan metode *K-means Clustering*.
2. Dalam menguji kebenaran hasil dari klusterisasi, penulis menggunakan *software KNIME Analytics Platform*.

1.4 Tujuan Penelitian

Tujuan yang ingin diperoleh dari penelitian ini agar lebih bermanfaat kedepannya adalah:

1. Menggali dan mendapatkan pengetahuan baru dari data perpustakaan pada Pondok Pesantren Diniyah Limo Jurai Sungai Pua
2. Optimalisasi pelayanan berdasarkan *cluster* yang telah di olah dengan metode *K-Means Clustering*.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberi manfaat kedepannya, yang beberapa diantaranya adalah:

1. Memberikan pengetahuan untuk pihak perpustakaan baik untuk optimalisasi pelayanan maupun pengadaan yang di prioritaskan.
2. Diharapkan dengan penerapan metode ini dapat optimalisasi pelayanan perpustakaan bisa lebih efektif dan efisien.

1.6 Sistematika Penulisan

Sistematika disesuaikan dengan template yang diatur dalam tata penulisan program studi masing-masing sebagai berikut:

Bab I PENDAHULUAN

Berisi Latar Belakang, Perumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penelitian.

Bab II LANDASAN TEORI

Pada bab ini membahas tentang landasan teori yang meliputi konsep dan penerapan *Knowledge Discovery in Databases (KDD)*, *Data Mining*, *Clustering*, *Algoritma K-means Clustering*.

Bab III METODE PENELITIAN

Bab ini membahas tentang analisa dan penggunaan secara matematis algoritma K-Means Clustering.

Bab IV ANALISA DAN PEMBAHASAN

Pada bagian ini akan dibahas tentang analisis sistem dan pembahasan sistem yang sedang berjalan.

Bab V IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas hasil implementasi algoritma K-Means Clustering pada Pondok Pesantren Diniyah Limo Jurai Sungai Pua.

Bab VI KESIMPULAN DAN SARAN

Bab ini membuat kesimpulan dan hasil penelitian Optimalisasi Pelayanan Perpustakaan terhadap Minat Baca Menggunakan Metode K-Means Clustering (Studi Kasus di Pondok Pesantren Diniyah Limo Jurai Sungai Pua).