



IMPLEMENTASI DATA MINING MENGGUNAKAN K-MEANS CLUSTERING
UNTUK MENENTUKAN STRATEGI PROMOSI PROGRAM KB (STUDI
KASUS DI DINAS PEMBERDAYAAN PEREMPUAN, PERLINDUNGAN
ANAK, PENGENDALIAN PENDUDUK DAN KELUARGA BERENCANA
KABUPATEN PASAMAN)

TESIS

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Magister Komputer

JUPRIL DONI

212321015

PROGRAM MAGISTER (S2)
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PUTRA INDONESIA "YPTK" PADANG

2023

BAB I

PENDAHULUAN

1.1 Latar Belakang

Knowledge Discovery in Database (KDD) adalah proses terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah data set yang besar dan kompleks. KDD sering digunakan secara bergantian untuk menjelaskan proses penggalian informasi di dalam database (Siska dan Amiruddin, 2022). *Knowledge Discovery in Database* (KDD) didefinisikan sebagai ekstraksi informasi potensial, implisit dan tidak dikenal dari sekumpulan data. Proses KDD melibatkan hasil dari proses Data Mining (proses mengekstrak kecenderungan pola suatu data), kemudian mengubah hasilnya secara akurat menjadi informasi yang mudah dipahami. KDD sendiri diartikan sebagai keseluruhan proses non-trivial untuk mencari dan mengidentifikasi pola (*pattern*) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti (Zainudin, *et al.* 2022a).

Data Mining merupakan suatu proses penggalian data atau penyaringan data dengan memanfaatkan kumpulan data dengan ukuran yang cukup besar melalui serangkaian proses untuk mendapatkan informasi yang berharga dari data tersebut. Data Mining biasa juga dikenal dengan *Knowledge Discovery in Databases* (KDD), ekstraksi pengetahuan (*knowledge extraction*) analisa data atau pola, kecerdasan bisnis (*business intelligence*) dan merupakan alat yang penting untuk memanipulasi data untuk penyajian informasi sesuai kebutuhan user dengan tujuan untuk membantu dalam analisis koleksi pengamatan perilaku (Zainudin, *et al.* 2022b).

K-Means adalah metode yang ada pada data mining dengan proses awal mengambil sebagian komponen populasi untuk dijadikan titik pusat di cluster awal. Penetapan kelompok dalam satu cluster dapat dilakukan dengan menghitung jarak setiap objek dengan titik pusat. Algoritma *clustering K-Means* dapat membagi data berdasarkan jarak antar data pada kelompok yang telah ditetapkan. Algoritma ini bergantung pada fungsi untuk mengukur data yang mempunyai ciri khas sama. Jarak itu sendiri dihitung menggunakan fungsi *Euclidean*. Kemudian data dimasukkan dalam kelompok yang mempunyai jarak terdekat (Ayu dan Taufik, 2022).

Penelitian terdahulu *K-Means* digunakan sebagai metode untuk mengelompokkan data pada data obat. Hasil penelitian pertama dapat mengklasifikasikan tiga kategori yang terdiri dari penggunaan obat tertinggi, penggunaan obat sedang dan penggunaan obat rendah. Akurasi data diperoleh dengan syarat 99,45% catatan valid 53471, nama obat 100% dengan catatan valid 53766, hitung berguna 100% dengan catatan valid 53766 (April, *et al.* 2021). Penelitian berikutnya *K-means* digunakan untuk menentukan kelayakan penanaman jagung di Kabupaten Lampung Selatan. Data yang memiliki karakteristik yang sama dikelompokkan kedalam klaster lain sehingga data yang berada dalam satu klaster memiliki derajat variasi yang kecil. Dengan menerapkan Metode *K-Means Clustering*, tanaman jagung dua tahun terakhir dapat menghasilkan Informasi kelayakan dari masing-masing kecamatan (Aldino, *et al.* 2021).

Penelitian lainnya *K-Means* digunakan untuk pengelompokan kasus dan kematian Covid-19 di Asia Tenggara. Dengan menggunakan metode ini data yang telah didapatkan dapat dikelompokkan menjadi beberapa *cluster*, dimana *K-Means Clustering* diterapkan dengan menggunakan *tools Rapidminer*. Hasil yang diperoleh adalah terdapat empat negara dengan klaster tingkat tinggi (C1), satu negara dengan klaster tingkat sedang (C2), dan 6 negara dengan klaster tingkat rendah (C3) (Hutagalung, *et al.* 2021).

Penelitian berikutnya *K-Means* digunakan pada kependudukan sarana perdagangan dan akomodasi di Indonesia. Tiga klaster label pemetaan yang digunakan yaitu klaster tinggi (K1), klaster normal (K2) dan klaster rendah (K3). Hasil pengolahan *rapidminer* diperoleh dari data *centroid* untuk klaster tinggi yaitu

((1527), (810.4), (5865), (6655.3), (323), (315.1), kluster sedang yaitu ((286), (199.591), (1327), (2240.227), (93.227), (140.955)), dan kluster rendah yaitu (139.25), (122.5), (508.833), (919.222), (64.417), (94.444). Hasil kluster menunjukkan bahwa 5 provinsi tergolong kluster tinggi, 13 provinsi tergolong kluster sedang, dan 16 provinsi tergolong kluster rendah. Dari hasil kajian, sekitar 47% wilayah di Indonesia masih memiliki fasilitas perdagangan dan perumahan yang rendah. Dengan analisis tersebut diharapkan pemerintah dapat lebih memperhatikan daerah yang pendapatannya masih di bawah rata-rata (Munawar, *et al.* 2021).

Penelitian lainnya *K-Means* digunakan untuk meningkatkan siswa prestasi di kelas unggul. Aplikasi *K-Means* berbasis web dapat memberikan informasi dan solusi yang dibutuhkan sekolah untuk mengklasifikasikan dan menentukan kelas unggulan sehingga dapat meningkatkan prestasi siswa di sekolah. Hasil tersebut dapat digunakan pihak sekolah untuk menganalisis prestasi belajar siswa dan dapat membantu guru dalam membentuk kelas unggulan sehingga dapat memotivasi siswa untuk belajar lebih giat (Yopi dan Juniar, 2022).

Keluarga Berencana (KB) adalah program skala nasional untuk menekan angka kelahiran dan mengendalikan pertumbuhan penduduk di suatu negara. Program KB secara khusus dirancang agar menciptakan kemajuan, kestabilan, kesejahteraan ekonomi, sosial, serta spiritual setiap penduduknya. Keluarga Berencana (KB) adalah program yang juga diatur dalam UU No. 10 Tahun 1992 yang dijalankan dan diawasi oleh Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN). Wujud dari Keluarga Berencana (KB) adalah pemakaian alat kontrasepsi untuk menunda serta mencegah kehamilan.

Dalam menyebarkan dan mengajak masyarakat untuk mengikuti program KB, pihak Dinas Pemberdayaan Perempuan, Perlindungan Anak, Pengendalian Penduduk dan Keluarga Berencana (DP3AP2KB) melakukan promosi baik itu dengan poster KB yang di pasang di jalan-jalan, sosialisasi langsung ke masyarakat, dan lain-lain. Permasalahan yang sering timbul adalah tidak semua bentuk promosi itu tepat sasaran, pihak DP3AP2KB menyamaratakan promosi dilakukan, sehingga masyarakat yang mengikuti program KB ini tidak merata di seluruh daerah, ada yang satu desa sudah mengikuti KB tetapi ada desa lain masyarakatnya kurang mengikuti

KB. Dalam mengatasi masalah itu perlu adanya analisis data untuk mengidentifikasi program KB sudah berjalan dengan baik atau belum sehingga promosi yang dilakukan bisa tepat sasaran. Penelitian ini memiliki fokus masalah yang perlu diselesaikan berupa belum adanya strategi promosi yang tepat dalam melakukan promosi program KB.

Dari uraian latar belakang yang telah dijelaskan, maka peneliti melakukan penelitian untuk mengkaji dan menyusun penelitian dalam bentuk tesis dengan judul penelitian: **“Implementasi Data Mining Menggunakan K-Means Clustering Untuk Menentukan Strategi Promosi Program KB (Studi Kasus di Dinas Pemberdayaan Perempuan, Perlindungan Anak, Pengendalian Penduduk dan Keluarga Berencana Kabupaten Pasaman)”**.

1.2 Perumusan Masalah

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

1. Bagaimana Data Mining dapat membantu pihak DP3AP2KB Kabupaten Pasaman?
2. Bagaimana menguji metode *K-Means* untuk menentukan strategi promosi program KB pada DP3AP2KB Kabupaten Pasaman?
3. Bagaimana Metode *K-Means Clustering* dalam pengklasteran optimalisasi strategi promosi program KB?

1.3 Batasan Masalah

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

1. Penelitian ini hanya terfokus pada strategi promosi Program KB.
2. Algoritma yang digunakan *K-Means Clustering*.

3. Penelitian ini difokuskan pada Dinas Pemberdayaan Perempuan, Perlindungan Anak, Pengendalian Penduduk dan Keluarga Berencana Kabupaten Pasaman.
4. *Tools* yang digunakan dalam penelitian ini adalah aplikasi Rapidminer.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut:

1. Untuk membantu pihak DP3AP2KB Kabupaten Pasaman dalam promosi program KB.
2. Untuk menguji metode *K-Means* dalam menentukan strategi promosi program KB pada DP3AP2KB Kabupaten Pasaman.
3. Untuk mengimplementasikan Metode *K-Means Clustering* dalam pengklasteran optimalisasi strategi promosi program KB.

1.5 Manfaat Penelitian

Manfaat yang bisa diperoleh dari penelitian ini adalah sebagai berikut:

1. Memberikan beberapa pengetahuan untuk menjadi pedoman kepada pihak DP3AP2KB Kabupaten Pasaman dalam melakukan promosi program KB.
2. Diharapkan dengan penerapan metode ini pihak DP3AP2KB Kabupaten Pasaman terbantu dalam memberikan keputusan untuk melakukan promosi program KB.

1.6 Sistematika Penulisan

Agar penelitian ini terarah dan sistematis, penulis membuat kerangka tulisan yang dituangkan dalam sistematika penulisan. Sistematika yang digunakan dalam penulisan ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang masalah, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penelitian.

BAB II LANDASAN TEORI

Bab ini berisikan tentang penerapan *Knowledge Discovery in Database (KDD)*, Data Mining, Algoritma *K-Means Clustering*.

BAB III METODOLOGI PENELITIAN

Bab ini berisi kerangka kerja, mendefinisikan ruang lingkup masalah, analisa masalah, menentukan tujuan, mempelajari literatur, mengumpulkan data dan informasi, menganalisa dan menerapkan *Data Mining*, mengimplementasikan Algoritma *K-Means Clustering*, menguji data dan menarik kesimpulan.

BAB IV ANALISA DAN PERANCANGAN

Bab ini berisi tentang analisa permasalahan, analisa kebutuhan sistem, menerapkan sistem Algoritma *K-Means Clustering*, menentukan variabel dalam perhitungan hasil untuk penentuan klasterisasi strategi promosi program KB.

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan tentang implementasi Algoritma *K-Means Clustering*, kebutuhan perangkat keras dan lunak, pengujian dengan *Data Testing*, untuk klasterisasi strategi promosi program KB.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan tentang penerapan algoritma *K-Means Clustering*, menggambarkan kekurangan dan kelemahan metode yang digunakan.