

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada masa teknologi komputasi dan media penyimpanan telah memungkinkan manusia untuk mengumpulkan dan menyimpan data dari berbagai sumber dengan jangkauan yang amat luas. Fenomena seperti ini terjadi dalam banyak bidang kehidupan, seperti bisnis, perbankan, pemasaran, produksi, sains, pendidikan, dan sebagainya. Meskipun teknologi basisdata modern telah menghasilkan media penyimpanan yang ekonomis bagi aliran data yang amat besar ini, teknologi untuk membantu user menganalisis, memahami, atau bahkan memvisualisasikannya belumlah tersedia. Hal inilah yang melatarbelakangi dikembangannya konsep Pengambilan Pengetahuan dari Basisdata (PPB) atau dikenal dengan data mining yang diterjemahkan secara sederhana sebagai 'penambangan data' (Firdaus, 2017).

Data mining atau disebut juga *Knowledge in Discovery* (KDD) adalah proses pengeksekusian pengetahuan dari data yang berukuran besar untuk menemukan pola-pola baru sehingga mendapatkan sebuah pengetahuan dan informasi baru. Teknologi Data Mining digunakan untuk menggali pengetahuan yang ada didalam database. Salah satu metode yang digunakan dalam Data Mining adalah metode clustering dengan algoritma *K-Means*. Algoritma *K-Means Clustering* adalah teknik dalam Data Mining yang mengelompokkan item-item data yang mirip kedalam satu kelompok atau cluster. Algoritma K-Means melakukan dua proses yaitu proses penentuan pusat cluster (*centroid*) dan proses pencarian anggota dari tiap-tiap cluster. Kerangka proses data mining tersusun atas tiga tahapan, yaitu pengumpulan data (*data collection*), transformasi data (*data transformation*), dan analisis data (*data analysis*). Proses tersebut diawali dengan *preprocessing* yang terdiri atas pengumpulan data untuk menghasilkan data mentah (*raw data*) yang dibutuhkan oleh data mining, yang kemudian dilanjutkan dengan transformasi data untuk mengubah

data mentah menjadi format yang dapat diproses oleh data mining, misalnya melalui filtrasi atau agregasi. Hasil transformasi data akan digunakan oleh analisis data untuk membangkitkan pengetahuan dengan menggunakan teknik seperti analisis statistik, machine learning, dan visualisasi informasi (Firdaus,2017).

Penelitian terdahulu mengenai Data Mining oleh Firdaus (2018) Penggunaan Data Mining dalam Kegiatan Sistem Pembelajaran Berbantuan Komputer dan dihasilkan pengetahuan yang dapat membantu guru dalam mengelola kelasnya dengan memahami cara belajar siswa, dan memberikan umpan balik proaktif kepada siswanya. Penelitian berikutnya oleh Eska (2016) untuk prediksi penjualan wallpaper dan dihasilkan pengetahuan berupa faktor-faktor yang mempengaruhi tingginya penjualan wallpaper. Penelitian berikutnya oleh Maharani et al (2017) penggunaan data mining untuk pengaturan layout minimarket dan menghasilkan Kombinasi barang untuk menyusun layout minimarket dapat ditemukan dengan mengolah data transaksi penjualan.

Metode yang digunakan dalam penelitian ini yaitu Algoritma *K-Means*. *K-means* merupakan algoritma clustering. *K-means Clustering* adalah salah satu “unsupervised machine learning algorithms” yang paling sederhana dan populer. *K-Means Clustering* adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. *K-means clustering* merupakan salah satu metode cluster analysis non hirarki yang berusaha untuk mempartisi objek yang ada kedalam satu atau lebih cluster atau kelompok objek berdasarkan karakteristiknya, sehingga objek yang mempunyai karakteristik yang sama dikelompokkan dalam satu cluster yang sama dan objek yang mempunyai karakteristik yang berbeda dikelompokkan kedalam cluster yang lain. Metode *K-Means Clustering* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Penelitian terdahulu mengenai Algoritma *K-Means* oleh Dawitri et al (2020) untuk menentukan tingkat penyebaran pandemi covid-19 di indonesia dan menghasilkan tiga cluster penyebaran utama covid-19. Penelitian selanjutnya oleh

Muliono (2019) untuk klasterisasi tingkat tridarma pengajaran dosen dengan hasil jumlah ketepatan prediksi yang dilakukan oleh algoritma K-means terhadap 15 data mengalami perbedaan sebanyak 53.33 % keakuratan prediksi.

Kehidupan yang didambakan oleh semua manusia di dunia ini adalah kesejahteraan. Baik yang tinggal di kota maupun yang di desa, semua mendambakan kehidupan yang sejahtera. Sejahtera lahir dan bathin. Secara umum, istilah kesejahteraan sosial sering diartikan sebagai kondisi sejahtera (konsepsi pertama), yaitu suatu keadaan terpenuhinya segala bentuk kebutuhan hidup, khususnya yang bersifat mendasar seperti makanan, pakaian, perumahan, pendidikan dan perawatan kesehatan. Pengertian kesejahteraan sosial juga menunjuk pada segenap aktifitas pengorganisasian dan pendistribusian pelayanan sosial bagi kelompok masyarakat, terutama kelompok yang kurang beruntung (*disadvantage groups*). Penyelenggaraan berbagai skema perlindungan sosial (*social protection*) baik yang bersifat formal maupun informal adalah contoh aktivitas kesejahteraan sosial (Suharto, 2009).

Implementasi Data Mining dengan metode Algoritma K-Means Clustering diatas memiliki kesamaan dengan masalah yang akan penulis bahas. Klasterisasi penduduk berdasarkan tingkat kesejahteraan ekonomi di kelurahan Ganting dilakukan secara manual oleh Pekerja Sosial Masyarakat (PSM). Algoritma K-Means dapat melakukan klasterisasi secara sistematis. Penulis mengangkat judul tentang “Optimalisasi Pengelompokkan Tingkat Kesejahteraan Ekonomi Penduduk Dalam Program Pembinaan Masyarakat (Studi Kasus di Kelurahan Ganting Kota Padang Panjang)”.

1.2 Perumusan Masalah

Berdasarkan permasalahan yang ada, supaya tesis ini sesuai dengan tujuan yang ingin dicapai. Maka penulis merumuskan beberapa permasalahan sebagai berikut :

1. Bagaimana cara melakukan klasterisasi pengelompokkan tingkat kesejahteraan ekonomi penduduk dengan metode K-Means?
2. Bagaimana menerapkan Algoritma K-Means untuk klasterisasi pengelompokkan penduduk yang optimal?

3. Bagaimana implementasi Algoritma K-Means dan software Rapid Minner Studio Versi 9.2 untuk mengelompokkan penduduk secara optimal guna mendapatkan komposisi kelompok penduduk berdasarkan tingkat kesejahteraan ekonomi penduduk.

1.3 Batasan Masalah

Agar pembahasan pada penelitian ini tidak menyimpang maka penulis membatasi ruang lingkup objek penelitian. Adapun ruang lingkup penelitian antara lain :

1. Penelitian dilakukan di Kelurahan Ganting Kota Padang Panjang dimana yang akan diteliti adalah cara klasterisasi penduduk berdasarkan tingkat kesejahteraan ekonomi dalam program pembinaan masyarakat yang optimal. Data yang di olah adalah data seluruh penduduk di Kelurahan Ganting.
2. Sistem yang digunakan adalah Data Mining dengan menggunakan metode klasterisasi K-Means.
3. Implementasi dari algoritma K-Means ini akan dapat dilihat dengan Software Rapid Minner Studio Versi 9.2

1.4 Tujuan Penelitian

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

1. Pekerja Sosial Masyarakat (PSM) dapat melakukan penempatan penduduk secara optimal dengan metode K-Means untuk mendapatkan komposisi penduduk sesuai dengan tingkat kesejahteraan Ekonomi melalui program pembinaan masyarakat.
2. Menerapkan cara kerja Algoritma K-Means untuk mengambil keputusan yang tepat dalam klasterisasi penempatan penduduk secara optimal.

3. Menguji metoda Algoritma K-Means untuk penempatan penduduk secara optimal guna mendapatkan komposisi kelas penduduk yang seimbang dengan menggunakan Rapid Minner Studio Versi 9.2
4. Untuk pembuktian metoda Algoritma K-Means untuk mengambil keputusan yang tepat dalam klasterisasi, penempatan penduduk secara optimal dilakukan perancangan program dengan bahasa pemrograman PHP.

1.5 Manfaat Penelitian

Penelitian ini diharapkan memberikan manfaat kedepannya yang beberapa diantaranya adalah :

1. Sebagai acuan dan pedoman bagi kelurahan dan Pekerja Sosial Masyarakat (PSM) dalam melakukan klasterisasi kelompok penduduk.
2. Mempermudah Pekerja Sosial Masyarakat (PSM) dan pihak kelurahan Ganting dalam menentukan kelompok penduduk berdasarkan tingkat kesejahteraan ekonomi.
3. Bagi peneliti sendiri, bisa menambah pengetahuan yang bermanfaat untuk dapat diaplikasikan dalam proses penjangingan masyarakat sesuai kategorinya.

1.6 Sistematika Penulisan

Sistematika yang digunakan dalam penyusunan tesis ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi kerangka kerja, mendefinisikan ruang lingkup masalah, analisa masalah, menentukan tujuan, mempelajari literature, mengumpulkan data dan informasi, menganalisa dan menerapkan Data Mining.

BAB II LANDASAN TEORI

Bab ini berisikan tentang penerapan Knowledge Discovery in Database (KDD), Data Mining, Algoritma K-Means Clustering dan sekilas tentang cluster masyarakat sesuai indikator pengelompokannya.

BAB III METODE PENELITIAN

Bab ini berisi kerangka kerja, mendefinisikan ruang lingkup masalah analisa masalah, menentukan tujuan, mempelajari literature, 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 system Algoritma K-Means Clustering, menentukan variabel dalam perhitungan hasil untuk penentuan klasterisasi.

BAB V IMPLEMENTASI SISTEM

Bab ini menjelaskan tentang implementasi Algoritma K-Means Clustering, kebutuhan perangkat keras dan lunak, pengujian dengan Data Testing.

BAB VI KESIMPULAN DAN SARAN

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