

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Saat sekarang ini perkembangan teknologi informasi dan komputer semakin maraknya di seluruh dunia dan diberbagai bidang tentunya, yang dilakukan secara manual sekarang dilakukan oleh mesin atau sistem. Teknologi informasi adalah teknologi yang digunakan untuk mengolah data, termasuk mendapatkan, memproses, menyimpan data dalam berbagai cara untuk menghasilkan informasi yang berkualitas, yaitu informasi yang akurat dan tepat. Dalam toko Asrafi Jaya, penjualan Tupperware masih ada proses yang dilakukan secara manual sehingga sering terjadi kesalahan pada saat mencatat data-data yang ada dan kurangnya efisiensi waktu yang diperlukan. Sehingga pada toko terjadi penumpukan barang akan merugikan pemilik toko. Oleh karena itu untuk pengolahan data, penulis melakukan penelitian pada Toko Asrafi Jaya, Untuk penerapan Data Mining dalam mengolah data yang diambil dari sampel dari Toko Asrafi Jaya di Kabupaten Pasaman Barat.

Penemuan Pengetahuan dalam Basis Data (KDD) adalah proses non-sepele untuk mengidentifikasi pola baru, valid, berpotensi berguna, dan pada akhirnya dapat dipahami dalam data. Istilah "pola" mengacu pada subset data yang dinyatakan dalam beberapa bahasa atau model yang dieksploitasi untuk mewakili subset itu. KDD bertujuan untuk menemukan pola-pola yang tidak menghasilkan perhitungan langsung dari jumlah yang telah ditentukan (yaitu, non-trivial), dapat diterapkan pada data baru dengan tingkat kepastian tertentu (yaitu, apakah valid), belum diketahui sejauh ini (yaitu, novel), memberikan beberapa

manfaat bagi pengguna atau untuk tugas lebih lanjut (yaitu, berpotensi berguna), dan mengarah pada wawasan yang berguna, segera atau setelahnya beberapa pasca-pemrosesan (yaitu, dapat dimengerti) (Jaja dkk, 2021).

Kata Mining merupakan kiasan dari bahasa Inggris, mine. Jika *mine* berarti menambang sumber daya yang tersembunyi di dalam tanah, maka Data Mining merupakan penggalian makna yang tersembunyi dari kumpulan data yang sangat besar. Karena itu Data Mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan basis Data (Tjaya, P, P dkk, 2021).

Data Mining yaitu suatu proses penggalian data-data dari sebuah informasi yang sangat penting. Data Mining juga merupakan suatu proses untuk menggali pola-pola dari data. Pola-pola itu didapatkan dari berbagai jenis basis data seperti basis data relasional, data warehouse, data transaksi, dan data berorientasi objek. Penggunaan Data Mining dapat membantu para pebisnis dalam pengambilan keputusan secara cepat dan tepat (Normah, N, 2021).

Menentukan pola penjualan dengan melihat kecenderungan pembelian konsumen jika dianalisa dan diolah dengan baik maka dapat membantu mengetahui produk mana yang terlaris dan kurang laku terjual, sehingga kontrol stok persediaan dapat dilakukan, dan dapat dijadikan sebagai masukan bagi perusahaan dalam mengembangkan strategi pemasaran (Normah, N, 2021).

K-Means merupakan sebuah algoritma *Clustering* pada Data Mining untuk dapat menghasilkan kelompok dari data yang jumlahnya banyak dengan metode partisi yang berbasis titik dengan waktu komputasi yang cepat dan efisien. Selain itu definisi K-Means merupakan metode *Clustering* berbasis jarak yang membagi data ke dalam sejumlah *cluster* dan algoritma ini hanya bekerja pada atribut numerik. Dapat disimpulkan bahwa algoritma K-Means akan membuat data-data yang ingin dilakukan pengelompokan yang nantinya akan menuju beberapa pusat/titik dimana titik tersebut yang digunakan sebagai acuan dalam pengelompokan dan data-data yang dilakukan klusterisasi harus numerik. Pada dasarnya Algoritma K-Means hanya mengambil sebagian dari banyaknya komponen yang didapatkan untuk kemudian dijadikan pusat *cluster* awal, pada penentuan pusat *cluster* ini dipilih secara acak dari populasi data (Yuliani, 2021).

Algoritma K-Means untuk mengetahui *cluster* dari data rekam medis pada penyakit DBD dalam Rumah Sakit Dharma Kerti dari tahun 2014 sampai 2018, hasil pengelompokan atau *cluster* data tersebut nantinya dapat dijadikan analisis untuk mengetahui pasien dengan variabel (usia, jenis kelamin, periode masuk RS, Kelas Kamar) seperti apa yang datang ke Rumah Sakit Dharma Kerti. Data hasil *cluster* dapat dimanfaatkan oleh Rumah Sakit dalam teknologi berbasis *website* sehingga dapat diakses secara fleksibel sebagai *decision support* atau membantu dalam membuat kebijakan terhadap fenomena penyakit DBD dalam Rumah Sakit Dharma Kerti (Adiputra ddk, 2022).

Peningkatan Algoritma K-Mean dalam mengelompokan data siswa kelas unggulan sehingga dalam proses pembelajaran di sekolah mudah untuk memfasilitasi pendidikan sesuai kemampuan siswa. Untuk itu diperlukan data mining agar dapat menyelesaikan masalah tersebut melalui pola-pola dari penumpukan data. Penerapan k-means dengan sampel 120 data dari nilai rapor siswa SMK Raksana 2 Medan, dapat mempercepat pengelompokan siswa kelas unggulan dengan menginputkan nilai centroid dan mencari nilai terdekat sehingga menghasilkan 3 *cluster* yaitu: *cluster* 1 penguasaan pemrograman android sebanyak 10 siswa, *cluster* 2 penguasaan pemrograman web sebanyak 62 siswa dan *cluster* 3 penguasaan pemrograman desktop sebanyak 48 siswa. Penerapan k-means berbasis web dilakukan untuk mengolah datanya agar lebih bermanfaat sebagai alat bantu mempermudah dalam mengambil keputusan bagi pendidik (Hutagalung, J, 2022).

Metode K-Means *clustering* digunakan agar mempermudah dalam menganalisis serta mengelompokan data untuk mengetahui beberapa daerah rawan bencana di wilayah Purbalingga. Data yang diperoleh sejumlah 5 *Cluster* pembagian daerah rawan bencana di Kabupaten Purbalingga dengan tingkat rawan sangat tinggi, tingkat rawan tinggi, tingkat rawan sedang, tingkat rawan rendah dan tingkat rawan sangat rendah. Terdapat 146 desa dengan potensi rawan sangat rendah, 7 desa dengan tingkat rawan rendah, 1 desa dengan tingkat rawan sedang, 3 desa dengan tingkat rawan tinggi dan 1 desa dengan tingkat rawan sangat tinggi. Hasil pengelompokan daerah tersebut diharapkan dapat menjadi antisipasi yang

dilakukan terhadap adanya bencana yang mungkin timbul dapat terus dilakukan secara tepat sehingga dapat meminimalisir efek bencana bagi masyarakat (Ramadhani dkk, 2022).

Data *Clustering* merupakan salah satu metode Data Mining yang bersifat tanpa arahan (*unsupervised*). K-Means merupakan salah satu metode data *clustering* nonhirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster/* kelompok. Metode ini mempartisi data ke dalam *cluster/*kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu Kelompok yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain (Yuliani, 2021).

Dengan menggunakan teknik *Clustering* maka data-datayang menyimpan informasi dapat dimanfaatkan. Informasi yang dihasilkan dapat dipergunakan untuk memprediksi informasi yang dibutuhkan oleh suatu toko sehingga kinerja jato komenja di efektif dan efisien dalam menjalankan usahanya.

Untuk menghadapi persaingan bisnis maka pihak toko dituntut agar dapat mengambil keputusan yang tepat dalam menentukan strategi pemasaran produk yang akan dijual. Tupperware adalah toko yang menjual suatu wadah plastik yang dipergunakan di rumah tangga untuk menyimpan makanan dan membuatnya kedapsuara.

Dari uraian di atas, maka penelitian ini akan menghasilkan informasi mengenai pengelompokan penjual tupperware dari yang “SANGAT LARIS”, “LARIS” dan “KURANG LARIS” yang selanjutnya akan diperhitungkan dari hasil penjualan yang ada dengan menggunakan metode K-Means *Clustering* dengan aplikasi Rapidminer. Hasilnya dapat dijadikan bahan atau dasar dalam penyediaan stok tupperware di Toko Asrafi Jaya. Berdasarkan latar belakang di atas, sehingga peneliti mengangkat judul: “Klasterisasi Menggunakan Metode Algoritma K-Means dalam Meningkatkan Penjualan Tupperware”.

## 1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan di atas, maka diperoleh rumusan seperti berikut :

1. Bagaimana mengklasterisasi penjualan produk Tupperware dengan menggunakan metode *K-Means Clustering*?
2. Bagaimana penerapan metode algoritma *K-Means Clustering*, untuk menentukan produk Sangat Laris, Laris dan Kurang Laris berdasarkan data penjualan Tupperware pada Toko Asrafi Raya dengan menggunakan aplikasi Rapid Miner?

## 1.3 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah :

1. Metode yang digunakan untuk melakukan *Clustering* data penjualan tupperware adalah algoritma K-Means.
2. Data yang diolah sebagai sampel adalah data penjualan produk Tupperware pada bulan Januari sampai Mei 2022.
3. Pengujian data untuk menentukan hasil *Clustering* menggunakan aplikasi Rapid Miner.

## 1.4 Tujuan Penelitian

Dalam melaksanakan penelitian ini terdapat beberapa tujuan yang ingin dicapai, adapun di antaranya :

1. Untuk melihat produk manayang paling diminatipada bulan Januari sampai Mei 2022 dengan menggunakan Algoritma k-means.
2. Mengimplementasikan penerapan metode K-Means untuk menentukan penjualan produk Tupperware.
3. Menguji hasil perhitungan manual dengan menggunakan aplikasi Rapid Miner untuk mengetahui minat pelanggan dari produk yang sudah dikelompokkan Sangat Laris, Laris dan Kurang Laris dengan metode K-Means.

## 1.5 Manfaat Penelitian

Manfaat dari penulisan tesis ini diuraikan sebagai berikut ini :

1. Dengan penerapan metode K-Means *clustering* dapat meningkatkan penjualan Tupperware pada toko Asrafi Raya.
2. Dengan Pengelompokan ini diharapkan dapat memudahkan penjual toko dalam menentukan stok barang yang harus dipenuhi.

## 1.6 Sistematika Penulisan

Sistematika penulisan yang akan ditulis dalam membuat laporan penelitian ini adalah sebagai berikut :

### BAB I PENDAHULUAN

Bab I berisi latar Belakang Masalah, Perumusan Masalah, Batasan Masalah, Tujuan Masalah, Manfaat Penelitian, Sistematika Penulisan.

### BAB II LANDASAN TEORI

Bab II berisi tentang teori dan penerapan Data Mining menggunakan metode K-Means.

### BAB III METODOLOGI PENELITIAN

Bab III akan menjelaskan tentang kerangka kerja penelitian dan langkah-langkah dalam penelitian.

### BAB IV ANALISA DAN PERANCANGAN

Bab IV membahas tentang prinsip kerja penerapan teknik Data Mining yaitu Algoritma K-Means dengan menggunakan aplikasi *RapidMiner*.

### BAB V IMPLEMENTASI DAN HASIL

Bab V berisi tentang hasil analisa dan pengujian dengan menggunakan aplikasi *RapidMiner*.

## BAB VI KESIMPULAN DAN SARAN

Padababiniakanmemberikanbeberapakesimpulandansaranyangdilakukan setelah dilaksanakannya penelitian.