

BAB I

PENDAHULUAN

1.1 Latar Belakang

Teknologi dan informasi yang semakin pesat, membuat tidak terhitungnya data informasi dalam kehidupan manusia, diamati secara jelas pada bidang pengolahan data yang berjumlah besar dalam penyimpanan datanya. Hal ini menjadi daya tarik besar pada perusahaan dan organisasi baik negeri ataupun swasta untuk memiliki penyimpanan data yang cukup besar kapasitasnya. Kemampuan teknologi dan informasi untuk mengumpulkan data menyimpan berbagai tipe data yang jumlahnya sangat besar, umumnya mendukung dalam segi pengolahan data internal maupun eksternal dalam transaksi perusahaan serta layanan yang didukung dan dikelola oleh teknologi informasi.

Pemanfaatan data dalam sebuah perusahaan untuk menunjang pengambilan keputusan tidak cukup dalam sistem operasi saja, diperlukan analisis data perusahaan untuk mendapatkan hasil kajian yang tepat dan akurat. Hal ini menjadi daya tarik dalam pemanfaatan ilmu yang dapat menyelesaikan masalah data dengan jumlah besar menjadi sebuah informasi. Data *mining* dapat dilakukan dengan sebuah aplikasi seperti *WEKA*, *SPSS Clementine*, *Mathlab*, *Dataminer*, *Orangecanvas*, *Dataengine*, *DBMiner*, *WebMiner* maupun *Rapidmine* yang dapat mempermudah dalam penggalian data. Menurut Witten, Frank & Hall, 2011. “*The Waikato Environment for Knowledge Analysis (WEKA)* adalah perkumpulan *ava* yang lengkap dan diimplementasikan *State-of-the-art* dalam pembelajaran dan algoritma di data *mining*”. *Statistical Product and Service Solution (SPSS)* dipilih sebagai aplikasi yang digunakan untuk pengolahan data dengan prosedur statistik yang digunakan dalam bidang bisnis mulai dari tingkat sederhana.

Analisis data *mining drug* merupakan sebuah pengamatan terhadap data *drug* yang didapat dari *website UCI Machine Learning Repository Education*

merupakan sebuah data yang berjumlah sangat besar dan sudah diolah dengan menggunakan teknik klasifikasi. Oleh karena itu, dengan pemanfaatan analisis data *mining* ini dimanfaatkan oleh penulis sebagai solusi penanganan informasi yang tersembunyi dalam penumpukan data tersebut. Penulis mengambil penyelesaian studi kasus analisis data *mining drug* dengan menggunakan aplikasi *SPSS Clementine* dengan metode Teknik *Clustering* menggunakan algoritma *K-Means*.

1.2 Rumusan masalah

Berdasarkan pada latarbelakang diatas, analisis akan diselesaikan adalah:

1. Bagaimana menerapkan data *mining* dengan menggunakan teknik *clustering* algoritma *K-Means* kedalam aplikasi *SPSS Clementine* pada sebuah *Drug Review Dataset*?
2. Bagaimana mengevaluasi hasil presentasi keakuratan prediksi masing-masing?
3. Bagaimana menginterpretasi presentasi prediksi yang dihasilkan menjadi sebuah informasi ?

1.3 Tujuan Penelitian

1. Mengetahui hasil dari penerapan *datamining* pada sebuah *Drug Review Dataset* yang menggunakan Teknik Clustering Algoritma K-Means.
2. Mengetahui hasil presentasi keakuratan.
3. Menginterpretasi presentasi hasil data menjadi informasi.

1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian, diharapkan dapat membantu peneliti dan peneliti selanjutnya dalam memudahkan mengelompokan *Drug Review Dataset*.

1.5 Batasan Masalah

Penulis membatasi permasalahan yang perlu dianalisis yaitu :

1. Data yang akan digunakan adalah *Drug Review Dataset* merupakan sebuah pengamatan terhadap data *drug review* diambil dari *website Universal Centre*

and Intelligent system (UCI) Machine Learning Repository Education 3 tahun sebelumnya.

2. Data yang digunakan dalam penelitian adalah data *drug*.

1.6 Metode Penelitian

Metode penelitian ini adalah :

1. Memahami teknik *K-Means clustering* dalam data mining keakuratan data pada penelitian sebelumnya.
2. Mengetahui data *Drug Review Dataset*.
3. Pengumpulan data melalui website *UCI Machine Learning Repository Education*
4. Implementasi algoritma ke dalam sistem.

1.7 Sistematika

BAB I PENDAHULUAN

Menjelaskan tentang latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, metode penelitian, dan sistematika penulisan

BAB II TINJAUAN PUSTAKA

Berisi tentang pengertian dari teori yang dapat menunjang dalam penelitian.

BAB III METODELOGI PENELITIAN

Bab ini menjelaskan tentang proses dalam pengelompokan pada *Drug* data yang akan dianalisis.

BAB IV HASIL DAN PEMBAHASAN

Deskripsi dari hasil penelitian, serta analisa pada data *drug* dengan menggunakan program *clementine* menggunakan metode *clustering* . Pembahasan dari hasil analisa akan memberikan penjelasan yang lebih detail.

BAB V KESIMPULAN DAN SARAN

Menjelaskan hasil akhir dari penelitian data *drug* dan memberikan saran untuk penelitian selanjutnya.

