LAPORAN TUGAS INDIVIDU DATA MINING

Zaka Zaidan Azminur

**Latar Belakang**

Data Mining atau dalam bahasa Indonesia diartikan sebagai memanen data adalah sebuah aktivitas yang bertujuan untuk mendapatkan sebuah informasi dari sekumpulan data yang jumlahnya sangat banyak. Aktivitas Data Mining saat ini menjadi sangat dibutuhkan, mengingat tingkat kompetisi antar perusahaan ataupun organisasi semakin ketat karena produk atau jasa yang ditawarkan hampir benar – benar sama persis antar satu sama lain. Untuk mengatasi hal ini, perusahaan atau organisasi perlu melakukan perencanaan strategi dengan baik dan terstruktur. Oleh karena itu, perusahaan perlu mengetahui selera konsumen di pasaran, salah satunya dengan menggunakan kumpulan data transaksi yang dimiliki oleh perusahaan atau organisasi.

Dari banyaknya jumlah data transaksi yang dimiliki oleh perusahaan, mereka bisa melakukan analysis terkait selera konsumen di pasaran dengan acuan data transaksi tersebut. Lalu, tidak hanya itu perusahaan juga bisa melakukan prediksi terhadap tingkah laku konsumen selanjutnya dengan cara kembali melihat pola atau tren yang terjadi di data transaksi mereka. Sehingga dari data transaksi yang sebelumnya hanya digunakan sebagai pencatatan saja, sekarang menjadi sebuah asset yang sangat berharga, karena dari data tersebut bisa dihasilkan sebuah informasi yang akan sangat berpengaruh terhadap keputusan suatu perusahaan.

**Masalah**

Diketahui sebuah perusahaan toko *online* memiliki riwayat data transaksi dari setiap pelanggannya yang disimpan dalam sebuah file bernama ***data\_train.txt*.** Lalu, dari data riwayat tersebut dilakukanlah sebuah analsisis terhadap data transaksi baru yang belum terprediksi apakah akan berakhir dengan sebuah pembelian atau pembatalan.

**Analisis Masalah**

Dapat dilihat bahwa dalam melakukan sebuah prediksi terhadap sebuah data dapat dilakukan dengan metode *unsupervised learning* ataupun *supervised learning*. Namun pada soal diminta bahwa metode yang harus dilakukan adalah *classification* atau dengan kata lain metode ini adalah metode yang  *superviced learning*. Lalu pada soal diketahui bahwa dalam melakukan prediksi, diminta menggunakan metode *Decision Tree* dimana *criterion* atau kriteria pada *Decision Tree* tidak ditentukan.

**Eksperimen**

Alat – alat yang digunakan :

1. RapidMiner Studio, untuk melakukan aktivitas Data Mining
2. Microsoft Excel, untuk melihat hasil prediksi dari Data Mining

Langkah Eksperimen :

1. Pertama