

## BAB 1

### PENDAHULUAN

#### 1.1 Latar Belakang

Dalam kehidupan sehari-hari kebutuhan akan informasi yang akurat sangat dibutuhkan dalam perkembangan masyarakat saat ini dan waktu mendatang. Namun kebutuhan informasi yang tinggi kadang kala tidak sesuai dengan informasi yang telah tersedia, sering kali informasi tersebut masih harus diolah kembali dari data yang jumlahnya sangat besar. Oleh karena itu sangat diperlukan sebuah metode yang dapat menghasilkan informasi yang masih tersimpan didalam sebuah data yang berkapasitas besar. Salah satu metode yang dapat mengatasi masalah penggalian informasi atau pola yang penting atau menarik dari data dalam jumlah besar disebut dengan data *mining*. Penggunaan teknik data *mining* diharapkan dapat memberikan pengetahuan-pengetahuan yang sebelumnya tersembunyi di dalam gudang data sehingga menjadi informasi yang berharga.

Data *mining* merupakan serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata *mining* sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu data *mining* sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan database.

Metode data mining sudah banyak digunakan oleh para peneliti. Nuqson (2010) membuat penelitian untuk menganalisa tingkat kelulusan mahasiswa dengan menggunakan algoritma apriori. Ruldeviyani dan Fahrian (2008), membuat penelitian dalam mengimplementasikan algoritma-algoritma *association rules* yaitu Apriori, FP-Growth, CT-Pro dan Apriori Cristian Borgelt. Demikian juga halnya dengan Dhivya

dan Kalpana (2010) yang dalam penelitiannya juga membandingkan algoritma apriori dan CT-Pro. Berdasarkan hal di atas, penulis mencoba membandingkan algoritma apriori dengan algoritma CT-Pro dengan menggunakan data operasional di bidang ekspor dan impor. Algoritma Apriori adalah algoritma paling terkenal untuk menemukan pola frekuensi tinggi. Pola frekuensi tinggi adalah pola-pola item di dalam suatu database yang memiliki frekuensi atau support di atas ambang batas tertentu yang disebut dengan istilah minimum support. Pola frekuensi tinggi ini digunakan untuk menyusun aturan asosiatif dan juga beberapa teknik data *mining* lainnya. Sedangkan Algoritma *CT-Pro* merupakan salah satu algoritma pengembangan dari *FP-Growth*. Perbedaannya terdapat pada langkah kedua dimana *FP-Growth* membuat *FP-Tree* sedangkan *CT-Pro* membuat *Compressed FP-Tree* (*CFP-Tree*). Pada tahap *mining* algoritma *CT-Pro* juga menggunakan pendekatan *bottom-up* dimana *item* pada *item tabel* dan *CFP-Tree* dilakukan *scan* dari jumlah terkecil hingga terbesar.

## 1.2 Rumusan Masalah

Setelah melihat latar belakang masalah yang telah diuraikan di atas maka dapat diperoleh rumusan masalah yaitu bagaimana menghasilkan sebuah pengetahuan ataupun informasi baru yang masih tersimpan dalam sebuah data yang besar.

## 1.3 Batasan Masalah

Dalam merancang dan mengimplementasikan sistem tersebut ada beberapa batasan masalah, yaitu :

1. Informasi yang ditampilkan adalah informasi tentang tingkat komoditas dengan negara tujuan terbanyak.
2. Data yang dipergunakan sebagai studi kasus yaitu data operasional di bidang tumbuhan, khususnya data Tindak Karantina Tumbuhan Ekspor dan Impor di Balai Besar Karantina Pertanian Belawan dan data yang digunakan yaitu data dalam waktu tiga bulan.
3. Implementasi program menggunakan PHP 5.3.8 dan Mysql 5.5.16
4. Parameter yang digunakan dalam penelitian ini adalah waktu dalam menghasilkan proses *mining* dari kedua algoritma.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah membandingkan kecepatan waktu proses pada Algoritma Apriori dan Algoritma *CT-Pro* untuk mengetahui frekuensi komoditas dengan negara tujuan terbanyak.

#### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini bagi penulis yaitu dapat lebih mengetahui cara menerapkan teknik data *mining*, dan manfaat lainnya yaitu dapat membantu instansi BBKP Belawan dalam menyajikan informasi tentang komoditas terbanyak dan negara tujuan terbanyak.

#### **1.6 Metodologi Penelitian**

Metodologi penelitian yang akan digunakan adalah:

##### **1. Studi Literatur**

Pada tahap ini dilakukan penelusuran terhadap berbagai macam literature seperti buku, referensi – referensi baik melalui perpustakaan maupun internet dan lain sebagainya yang terkait dengan judul penelitian ini yang berguna untuk pembelajaran bagi penulisan skripsi ini.

##### **2. Analisis dan Pengumpulan Data**

Mengumpulkan data dan informasi yang kemudian dijadikan bahan analisis untuk mengambil kesimpulan. Dalam tahap ini mengumpulkan semua kebutuhan yang diperlukan dalam membandingkan kedua algoritma, kebutuhan tersebut berupa data oprasional. Kemudian melakukan analisa pengolahan data untuk menghasilkan suatu informasi berupa pola (*pattern*) dengan teknik asosiasi (*association rule*)

##### **3. Perancangan dan Implementasi Sistem**

Pada tahap ini dilakukan analisis dan perancangan dalam membangun aplikasi *data mining*. Akan dilakukan analisis terhadap dua algoritma dengan menggunakan data yang diperoleh, sehingga aplikasi ini nantinya dapat diimplementasikan sesuai dengan yang diharapkan.

#### 4. Pengujian

Pada tahap ini dilakukan uji coba terhadap sistem yang telah dibangun, apakah sudah sesuai dengan yang diharapkan.

#### 5. Penyusunan Laporan

Tahap akhir dari proses penyusunan skripsi ini adalah pembuatan laporan atau dokumentasi secara lengkap dan menyeluruh dari semua kegiatan yang telah dilakukan.

### **1.7 Sistematika Penulisan**

Sistematika penulisan skripsi ini dibagi menjadi 5 (lima) bab yaitu sebagai berikut :

#### **Bab 1 : Pendahuluan**

Bab ini berisi tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian dan sistematika penulisan dari skripsi ini.

#### **Bab 2 : Tinjauan Pustaka**

Bab ini membahas tentang landasan teori yang mendukung pembahasan yang berhubungan dengan sistem yang akan dibuat.

#### **Bab 3 : Analisis dan Perancangan Sistem**

Bab ini membahas mengenai analisa data yang akan diproses berdasarkan analisis asosiasi dengan algoritma apriori dan algoritma CT-Pro kemudian membuat DFD, flowchart, struktur tabel dan merancang tampilan user interface dari aplikasi data mining.

#### **Bab 4 : Implementasi dan Pengujian Sistem**

Bab ini meliputi pembahasan tentang implementasi dari sistem yang dibuat berdasarkan hasil analisis dan perancangan sistem yang telah dibuat, dan pengujian sistem tersebut.

**Bab 5 : Kesimpulan dan Saran**

Bab ini meliputi kesimpulan dan saran yang diberikan untuk perbaikan sistem sehingga menjadi lebih baik dan bermanfaat bagi pengguna dan pembaca skripsi ini.