

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

*Data mining* bertujuan untuk menemukan pola-pola yang valid, baru, mempunyai nilai guna, dan mudah dipahami dari data yang ada. Jenis pola yang dihasilkan ditentukan oleh fungsionalitas *data mining* yang digunakan.

Penambangan *frequent pattern* merupakan salah satu bagian yang sangat penting dalam *data mining*. *Frequent pattern* pertama kali diperkenalkan oleh R.Agrawal [2] sebagai bagian dari fungsionalitas *association*. Hampir keseluruhan proses penambangan *association rule* didominasi oleh proses penambangan *frequent pattern*. Sehingga dibutuhkan suatu algoritma *mining frequent pattern* yang efektif dan efisien.

Salah satu algoritma *mining frequent pattern* yang paling banyak dikembangkan adalah Apriori. Apriori melakukan penambangan *frequent pattern* dengan cara membangkitkan sejumlah kandidat *frequent pattern* kemudian menguji setiap kandidat tersebut dengan cara mencocokkannya dengan data yang ada. Dengan cara kerja tersebut Apriori mampu menemukan semua *frequent pattern*. Akan tetapi cara kerja seperti ini sangat mahal, baik dari segi ruang *memory* maupun waktu.

Beberapa algoritma telah dikembangkan untuk mengatasi kekurangan yang ada pada Apriori, diantaranya adalah FP-growth [8] dan Tree Projection [1]. Hasil penelitian dalam jurnal [1,8] menunjukkan bahwa kedua algoritma ini memiliki kemampuan lebih baik dari Apriori. Untuk mengetahui performansi serta kelebihan dan kekurangan kedua algoritma tersebut dibutuhkan suatu studi analisis perbandingan performansi.

## 1.2 Perumusan Masalah

Dengan mengacu pada latar belakang masalah di atas, maka permasalahan yang akan dibahas dan diteliti adalah:

- 1) Cara kerja algoritma FP-growth dan Tree Projection
- 2) Penerapan algoritma FP-growth dan Tree Projection untuk menambang *frequent pattern*.
- 3) Performansi algoritma FP-growth dan Tree Projection dalam menghasilkan *frequent pattern*.

## 1.3 Maksud Tugas Akhir

Berdasarkan rumusan masalah di atas, penyusunan Tugas Akhir ini dimaksudkan untuk :

- 1) Melakukan studi tentang dua algoritma *mining frequent pattern* yaitu FP-growth dan Tree Projection.
- 2) Menganalisa kompleksitas algoritma FP-growth dan Tree Projection
- 3) Membangun perangkat lunak untuk mengimplementasikan algoritma FP-growth dan Tree Projection pada suatu perangkat lunak untuk kebutuhan analisa.
- 4) Melakukan studi perbandingan performansi terhadap kedua algoritma di atas berdasarkan parameter analisa yang ada pada batasan masalah dengan menggunakan perangkat lunak pada poin 3.

## 1.4 Tujuan Tugas akhir

Dari Tugas Akhir ini diharapkan hal-hal sebagai berikut :

- 1) Mengetahui kelebihan dan kekurangan algoritma FP-growth dan Tree Projection.
- 2) Mengetahui kompleksitas algoritma FP-growth dan Tree Projection.
- 3) Mengetahui performansi kedua algoritma berdasarkan parameter-parameter yang ada pada batasan masalah poin 1.

## 1.5 Batasan Masalah

Batasan masalah untuk Tugas Akhir ini adalah sebagai berikut :

- 1) Parameter-parameter yang digunakan untuk mengetahui kinerja kedua algoritma tersebut yaitu :
  - a. *Minimal support* dan waktu untuk menghasilkan *frequent pattern*.
  - b. Jumlah transaksi dan waktu untuk menghasilkan *frequent pattern*.
- 2) Analisis kompleksitas algoritma hanya dilakukan untuk kompleksitas waktu.
- 3) Pembahasan Tugas Akhir ini tidak sampai menghasilkan *association rules*.
- 4) Jenis *dataset* yang digunakan adalah *dataset* transaksi, dan nilai setiap atribut telah disesuaikan dengan kebutuhan aplikasi.
- 5) Lexicographic tree pada Tree Projection dibangun menggunakan metode *breadth first* dan pemroyeksian transaksi menggunakan metode *depth first*.
- 6) Analisis performansi FP-growth dan Tree Projection hanya meliputi performansi FP-growth dan Tree Projection pada struktur data yang bisa dimuat di *main memory* sehingga perangkat lunak yang dibangun tidak menangani struktur data FP-tree dan Lexicographic tree yang melebihi kapasitas *main memory*.

## 1.6 Metodologi Penyelesaian Masalah

Metode - metode yang digunakan dalam pelaksanaan tugas akhir ini adalah:

- 1) Studi literatur :
  - a. Pencarian referensi, mencari referensi yang berhubungan dengan *data mining*, *frequent pattern*, algoritma FP-growth dan Tree Projection, dan kompleksitas algoritma.
  - b. Pendalaman materi, mempelajari dan memahami materi yang berhubungan dengan Tugas Akhir ini.
- 2) Analisis kompleksitas algoritma FP-growth dan algoritma Tree Projection.

- 3) Perancangan dan pembuatan perangkat lunak *mining frequent pattern* menggunakan algoritma FP-growth dan Tree Projection sebagai media pengumpul data untuk analisis.
- 4) Melakukan berbagai percobaan dengan nilai *minimal support* dan ukuran *dataset* yang bervariasi pada kedua algoritma.
- 5) Melakukan studi perbandingan performansi kedua algoritma.
- 6) Penyusunan laporan dalam bentuk tertulis sebagai laporan Tugas Akhir.

### 1.7 Sistematika Penulisan

Tugas Akhir ini disusun berdasarkan sistematika sebagai berikut :

#### **BAB I            PENDAHULUAN**

Bab ini meliputi latar belakang masalah, perumusan masalah, maksud dan tujuan Tugas Akhir, batasan masalah, metodologi penyelesaian masalah, dan sistematika penulisan.

#### **BAB II            DASAR TEORI**

Bab ini memuat berbagai dasar teori yang mendukung dan mendasari penulisan Tugas Akhir ini, yaitu mengenai konsep dari *data mining*, *frequent pattern*, FP-growth, dan Tree Projection.

#### **BAB III            ANALISIS DAN PERANCANGAN SISTEM**

Dalam bab ini diuraikan tentang analisis algoritma FP-growth dan Tree Projection dari segi teori dan kompleksitas, analisis dan perancangan perangkat lunak sebagai alat bantu dalam proses analisa.

#### **BAB IV            IMPLEMENTASI DAN HASIL PENGUJIAN**

Bab ini memuat tentang analisis terhadap performansi algoritma FP-growth dan algoritma Tree Projection, membandingkan kinerja kedua algoritma berdasarkan

parameter pembandingan yang telah disebutkan pada batasan masalah.

## **BAB V**

### **KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan yang diambil dari pembahasan bab-bab sebelumnya serta saran-saran untuk pengembangan selanjutnya.