



MODUL PERTEMUAN 10

ALGORITMA FP-GROWTH

A. TUJUAN

1. Praktikan mampu memahami yang dimaksud dengan Algoritma FP-Growth
2. Praktikan mampu mengimplementasikan Algoritma FP-Growth dalam bahasa pemrograman python
3. Praktikan mampu memahami cara kerja Algoritma FP-Growth

B. Landasan Teori

ALGORITMA FP-GROWTH

Algoritma FP-Growth merupakan pengembangan dari algoritma Apriori. Algoritma Frequent Pattern Growth (FPGrowth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data. Pada algoritma FP-Growth menggunakan konsep pembangunan tree, yang biasa disebut FP-Tree, dalam pencarian frequent itemset bukan menggunakan generate candidate seperti yang dilakukan pada algoritma Apriori. Dengan menggunakan konsep tersebut, algoritma FP-Growth menjadi lebih cepat daripada algoritma Apriori. FP-growth adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (frequent itemset) dalam sebuah kumpulan data. FPGrowth menggunakan pendekatan yang berbeda dari paradigma yang digunakan pada algoritma Apriori.

Metode FP-Growth dibagi menjadi tiga tahapan utama, yaitu:

1. Tahap pembangkitan conditional pattern base

Conditional Pattern Base merupakan subdatabase yang berisi prefix path (lintasan prefix) dan suffix pattern (pola akhiran). Pembangkitan conditional pattern base didapatkan melalui FP-tree yang telah dibangun sebelumnya.



2. Tahap pembangkitan conditional FP-Tree

Pada tahap ini, support count dari setiap item pada setiap conditional pattern base dijumlahkan, lalu setiap item yang memiliki jumlah support count lebih besar sama dengan minimum support count akan dibangkitkan dengan conditional FPtree.

3. Tahap pencarian frequent itemset

Apabila Conditional FP-tree merupakan lintasan tunggal (single path), maka didapatkan frequent itemset dengan melakukan kombinasi item untuk setiap conditional FP-tree. Jika bukan lintasan tunggal, maka dilakukan pembangkitan FP-Growth secara rekursif.

• Association Rule

Association rule merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk support (minsup) dan confidence (minconf) pada sebuah database. Kedua syarat tersebut akan digunakan untuk interesting association rules dengan dibandingkan dengan batasan yang telah ditentukan, yaitu minsup dan minconf.

Association Rule Mining adalah suatu prosedur untuk mencari hubungan antar item dalam suatu dataset. Dimulai dengan mencari frequen itemset, yaitu kombinasi yang paling sering terjadi dalam suatu itemset dan harus memenuhi minsup.

• Support

Support dari suatu association rule adalah presentasi kombinasi item tersebut, dimana jika mempunyai item A dan item B maka support adalah proporsi dari transaksi yang mengandung A dan B. Rumus untuk menghitung nilai support dari dua item tersebut adalah sebagai berikut:

$$\text{Support (A,B)} = P(A \cap B) = \frac{\text{Jumlah transaksi yang mengandung A dan B}}{\text{Total transaksi}}$$



- **Confidence**

Confidence dari association rule adalah ukuran ketepatan suatu rule, yaitu presentasi transaksi yang mengandung A dan mengandung B. Dengan adanya confidence dapat diukur kuatnya hubungan antar-item dalam association rule. Rumus untuk menghitung nilai confidence dari dua item tersebut adalah sebagai berikut:

$$\text{Confidence}(A \rightarrow B) = P(A|B) = \frac{\text{Jumlah transaksi yang mengandung A dan B}}{\text{Jumlah transaksi yang mengandung A}}$$