

## Analisis Perbandingan Algoritma Asosiasi Eclat dan FP-Growth

*Data mining* merupakan bidang ilmu yang digunakan untuk mendapatkan pola dan pengetahuan menarik, salah satunya association rule mining. Eclat dan FP-Growth merupakan algoritma yang populer yang memiliki keunikan tersendiri dalam melakukan pencarian hubungan asosiasi.

*Association rule mining* bertujuan untuk mencari *frequent itemset* dan menemukan hubungan di antara *item-item* dalam suatu data. Dalam pencariannya, tidak semua hubungan yang terbentuk merupakan *rules* yang valid. Sebuah *rule* dikatakan *rule* yang valid (*strong rule*) jika telah disaring menggunakan pengukuran *lift ratio*. Tetapi walaupun suatu *rules* sudah merupakan *strong rules*, tidak menjamin *rules* tersebut menarik. Oleh karena itu, paper ini meneliti mengenai *rule* yang ditemukan oleh kedua algoritma tersebut dengan pengukuran *interestingness measure*, yaitu *imbalance ratio*(IR) serta *Kulczynski* dengan nilai *support* dan *confidence* yang telah ditentukan. Penelitian ini juga melakukan uji coba kecepatan pembentukan *rules* pada algoritma Eclat dan FP-Growth dengan mengujikan sederet nilai *support* dan *confidence*.

Penelitian ini dilakukan dengan empat tahapan utama yaitu : 1) melakukan analisis nilai *threshold*(*support* dan *confidence*) yang dapat menghasilkan *interesting rules*; 2) pembentukan aturan asosiasi; 3) melakukan analisis terhadap *interesting rules* yang ditemukan; 4) melakukan analisis runtime dengan *threshold* yang telah ditentukan.

Penelitian pada makalah ini menemukan bahwa baik Eclat maupun FP-Growth, keduanya menemukan jumlah *rules* menarik yang sama pada nilai *support* dan *confidence* yang sama. Dalam pencarian *rules*-nya, algoritma FP-Growth selalu menghasilkan *rules* yang lebih sedikit, tetapi membutuhkan rata-rata waktu 4x lebih banyak dibandingkan algoritma Eclat.

Hasil penelitian makalah ini menunjukkan bahwa algoritma Eclat lebih baik dalam kecepatan pencarian hubungan asosiasi dibandingkan algoritma FP-Growth. Makalah penelitian ini juga menemukan bahwa *rule* yang menarik dibentuk dari *frequent itemset*, tetapi tidak semua *rules* yang dibentuk dari *frequent itemset* bisa menjadi hubungan asosiasi yang menarik.

## **Comparative Analysis Eclat and FP-Growth Algorithm for Association Rule Mining**

Data mining is a field of science that is used to get interesting patterns and knowledge, one of which is the science of finding association patterns. Eclat and FP-Growth are popular algorithms that have their own uniqueness in mining for frequent itemsets.

Association rule mining aims to find frequent itemsets and find relationships between items in a data. In the search, not all relationships formed are valid rules. A rule is said to be a valid rule (strong rule) if it has been filtered using lift ratio measurements. But even though a rule is already a strong rule, it does not guarantee that the rules are interesting. Therefore, this paper examines the rules found by the two algorithms by measuring the interestingness measure, namely the imbalance ratio (IR) and Kulczynski with predetermined support and confidence values. This study also tested the speed of forming rules on the Eclat and FP-Growth algorithms by testing a series of support and confidence values.

This research was conducted in four main stages, namely: 1) analyzing the threshold value (support and confidence) which can create interesting rules; 2) establishment of association rules; 3) conduct an analysis of the interesting rules found; 4) perform runtime analysis with a predetermined threshold.

The research in this paper found that both Eclat and FP-Growth both found the same number of interesting rules at the same support and confidence values. In the search for rules, the FP-Growth algorithm always produces fewer rules, but requires 4x more time on average than the Eclat algorithm.

The results of this paper show that the Eclat algorithm is better in finding association rules than the FP-Growth algorithm. This research paper also finds that interesting rules are formed from frequent itemset, but not all rules formed from frequent itemset can be interesting association relationships.