



## BAB 2

### LANDASAN TEORI

© Hak cipta milik IBI KKKG Institut Bisnis dan Informatika Kwik Kian Gie

#### A. Data

Sebelum kita mengenal lebih jauh tentang *Data Mining*, ada baiknya kita mengerti terlebih dahulu tentang apa yang dimaksud dengan data. Pengertian data menurut Adyanata Lubis (2016: 1), adalah kumpulan fakta yang mendeskripsikan suatu kejadian nyata yang terjadi pada suatu waktu. Jadi data didapat dari suatu kejadian yang benar-benar terjadi, misalnya data penjualan adalah data yang diperoleh dari hasil penjualan dan data pembelian diperoleh dari kejadian pembelian.

#### B. Informasi

Menurut Deny Jollyta (2020 : 2) informasi didefinisikan sebagai berikut : “Tujuan informasi adalah untuk membantu dalam membuat keputusan dan memecahkan masalah atau mewujudkan peluang.”

Definisi informasi menurut Elisabet Yunaeti (2017 : 12), “Informasi dapat didefinisikan sebagai hasil dari pengolahan data dalam suatu bentuk yang lebih berguna dan lebih berarti bagi penerimanya yang menggambarkan suatu kejadian-kejadian yang nyata yang digunakan untuk pengambilan keputusan. Informasi merupakan data yang telah diklasifikasikan atau diolah atau diinterpretasi untuk digunakan dalam proses pengambilan keputusan”.

#### C. Basis Data

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.

Instansi Bisnis dan Informatika Kwik Kian



Basis data atau yang biasa kita sebut sebagai *database* di dalam dunia komputer, terutama oleh sang pemrogram sudah tidak asing lagi karena seringkali disebut dan berhubungan secara langsung dengan *Data Mining*. Namun untuk memudahkan dalam mengerti apa itu basis data, lebih baik kita bahas terlebih dulu apa yang dimaksud oleh basis data.

Basis data merupakan gabungan dari beberapa file data yang dibentuk dengan relasi yang masuk akal dan dapat diungkapkan dengan catatan yang sifatnya independen. Adapun basis data menurut Adyanata Lubis (2016: 1), adalah tempat perkumpulan data yang saling memiliki relasi di dalam suatu tempat yang bertujuan untuk dapat memudahkan pemanggilan atau pemanfaatan kembali dari data tersebut.

Maka dari itu, arti dari sistem basis data menurut Adyanata Lubis (2016: 1), adalah suatu sistem yang menyusun dan mengelola beberapa jejak atau *record* dengan menggunakan komputer, yang bertujuan untuk menyimpan dan memelihara data dengan lengkap pada sebuah organisasi maupun perusahaan, sehingga dapat dipastikan menyediakan informasi yang optimal yang nantinya dibutuhkan pemakai untuk berbagai kepentingan pengambilan keputusan.

Berdasarkan definisi-definisi basis data menurut para ahli diatas, maka dapat disimpulkan bahwa definisi basis data adalah kumpulan data yang ada pada sebuah media atau wadah penyimpanan data yang saling memiliki relasi dan dapat berguna bagi penggunanya.

Hak Cipta Dilindungi Undang-undang  
© Hak Cipta dimiliki IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



## D. Sistem

Dilansir dari buku Yeni Kustiyahningsih (2020 : 1), sistem berasal dari bahasa Latin yaitu *Systema* dan bahasa Yunani yaitu *Sustema*, yang merupakan suatu kesatuan yang terdiri dari komponen atau elemen yang dihubungkan bersama untuk memudahkan aliran informasi, materi atau energi. Istilah ini sering dipergunakan untuk menggambarkan suatu entitas yang berinteraksi.

Definisi sistem menurut Elisabet Yunaeti (2017 : 12), “Suatu sistem pada dasarnya adalah sekelompok unsur yang erat hubungannya satu dengan yang lain, yang berfungsi bersama-sama untuk mencapai tujuan tertentu. Secara sederhana, suatu sistem dapat diartikan sebagai suatu kumpulan atau himpunan dari unsur, komponen, atau variabel yang terorganisir, saling berinteraksi, saling tergantung satu sama lain, dan terpadu”.

## E. Data Mining

### 1. Konsep *Data Mining*

*Data Mining* telah menerima banyak perhatian di dunia sistem informasi dan masyarakat secara keseluruhan dalam beberapa tahun terakhir karena ketersediaan data dalam jumlah besar yang tersebar luas dan kebutuhan untuk mengubah data tersebut menjadi informasi dan pengetahuan yang berguna.

*Data Mining* digunakan untuk mengekstrak atau menambang pengetahuan dari kumpulan data yang besar menurut. Penambangan data adalah kombinasi dari metode dan algoritma analisis data berkelanjutan untuk memproses data dalam jumlah besar.

1. Dilarang menyalin atau menjiplak sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Pengertian *Data Mining* menurut Parteek Bhatia (2021: 3), “*Data mining* adalah kumpulan teknik untuk penemuan otomatis yang efisien dari pola yang sebelumnya tidak diketahui, valid, baru, berguna, dan dapat dipahami dalam basis data besar. Pola tersebut harus dapat ditindaklanjuti sehingga dapat digunakan dalam pengambilan keputusan perusahaan.”

*Data Mining* menurut Nandang Iriadi (2021: 3), adalah proses iteratif yang ditujukan untuk analisis basis data yang bertujuan menyaring informasi dan pengetahuan yang dapat membuktikan keakuratan data, berpotensi bagi para ahli ilmiah yang terlibat dalam pengambilan keputusan dan pemecahan masalah. Unggul dan bermanfaat. Dalam definisi yang paling sederhana, *Data Mining* secara otomatis mengenali pola terkait dalam database.

## 2. Komponen *Data Mining*

Untuk memahami lebih dalam konsep tentang *Data Mining* tersebut, tentunya kita perlu mengetahui beberapa komponen yang ada di dalamnya. *Data Mining* terdiri dari tiga komponen menurut Sivandam dalam Nandang Iriadi (2021: 3), yang adalah :

### a. *Data Capture*

Yang harus diintegrasikan ke dalam organisasi yang luas, sering dalam *Data Warehouse*, yang digunakan sebagai gudang penyimpanan data, dan organisasi serta peyajian informasi ini. *Data Capture* adalah proses yang cukup standar mengumpulkan, mengorganisir, dan membersihkan, misalnya, menghapus duplikat, berasal hilang nilai mana mungkin, menetapkan atribut yang diturunkan dan validasi data.

### b. *Data Mining*

Untuk memungkinkan pemahaman penggunaan data.

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Institut Bisnis dan Informatika Kwik Kian

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



c. Organisasi dan Penyajian

Informasi *Data Mining* untuk memungkinkan pemahaman.



Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)

Hak Cipta Dilindungi Undang-Undang

3. Model dan Metode *Data Mining*

Gagasan bahwa pikiran untuk mengorganisir pengetahuan dan memakai proses alami pembagian terstruktur mengenai beredar luas. Tapi waktu membahas istilah pembagian terstruktur klasifikasi, kita berbicara mengenai taksonomi. Taksonomi timbul pertamakali menjadi ilmu mengelompokkan organisme hidup (alfa taksonomi), lalu dikembangkan menjadi ilmu klasifikasi dalam umumnya, termasuk pada sini prinsip-prinsip pembagian terstruktur mengenai skema taksonomi juga.

Dengan demikian, pembagian terstruktur mengenai taksonomi merupakan proses menempatkan objek tertentu pada satu set kategori, dari masing-masing objek *property*. Proses pembagian terstruktur mengenai didasarkan dalam empat komponen menurut Gorunescu dalam Nandang Iriadi (2021: 4), yang adalah :

a. Kelas

Variabel dependen menurut contoh yang adalah variabel kategori mewakili yang menggunakan objek sesudah klasifikasinya. Contoh-model kelas merupakan: adanya loyalitas pelanggan, kelas bintang (galaksi), kelas gempa bumi (badai), dan lainnya.

b. Prediktor

Variabel bebas menurut model, diwakili sang karakteristik (atribut) menurut data yang wajib diklasifikasikan & dari pembagian terstruktur mengenai yg dibuat.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Contoh-contoh prediktor : merokok, konsumsi alkohol, frekuensi pembelian, status perkawinan, catatan geologi tertentu, musim, lokasi kenyataan kejadian,

**C**

**Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)**

Hak Cipta Dilindungi Undang-Undang

c. *Training Dataset*

*dataset* yang adalah sekumpulan data yang berisi nilai sebelumnya dipakai untuk dua komponen, dan dipakai untuk pelatihan model untuk mengenali kelas, dari prediksi tersedia. Contoh model set tadi merupakan: gerombolan pasien diuji dalam serangan jantung, gerombolan pelanggan menurut supermarket. (diselidiki internal polling). Misalnya, sentra pengumpulan data dadatan asumsi tipe Pusat Topan Nasional, pengumpulan data dan proyeksi Informasi Gempa Bumi Nasional.

d. *Pengujian dataset*

Pengujian *dataset* yang berisi data baru yang akan diklasifikasikan sang *classifier* dan kinerja model, dan bisa dievaluasi menurut Nandang Iriandi (2021: 6)

4. *Tugas Data Mining*

Menurut Anjar Wanto(2020: 6), Ada beberapa tugas yang dilakukan oleh Data

Mining dalam memproses pemecahan masalah dan mencari pengetahuan baru, di antaranya adalah sebagai berikut :

a. *Klastering (Clustering)*

Digunakan untuk mengelompokan atau mengidentifikasi data yang memiliki karakteristik tertentu. Contoh algoritma : K-Means, K-Medoids, dan lainnya.

b. *Klasifikasi (Classification)*

Klasifikasi berguna untuk menemukan model/fungsi yang menjelaskan atau membedakan konsep dan juga kelas data, yang bertujuan untuk memperkirakan

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



kelas dari suatu objek yang labelnya tidak diketahui. Contoh algoritmanya adalah :  
C4.5, *K-Nearest Neighbor* (KNN), dan lain-lain.

**C Hak cipta milik IBI KKG (Institut Bisnis dan Informatika Kwik Kian Gie)**

Hak Cipta Dilindungi Undang-Undang

**Institut Bisnis dan Informatika Kwik Kian**

c. Asosiasi (*Association*)

Asosiasi berguna mengatasi masalah bisnis yang tidak umum, yaitu dengan cara menganalisa tabel transaksi penjualan dan mengidentifikasi produk-produk yang sering dibeli bersamaan oleh pembeli, contohnya bila orang membeli sambal biasanya juga membeli kecap. Contoh algoritmanya adalah : *Apriori*, *Frequent Pattern Growth* atau *FP-Growth* yang dipakai di dalam penulisan ini.

d. Estimasi (*Estimation*)

Estimasi berguna untuk memperkirakan sesuatu hal yang belum pernah ada pada sebelumnya, dan disajikan dalam bentuk hasil kuantitatif atau angka. Contoh algoritmanya adalah : Regresi Linier, *Confidence Interval Estimations*, dan lainnya.

e. Prediksi (*Prediction*)

Prediksi berguna untuk memperkirakan atau meramalkan suatu kejadian yang belum pernah terjadi. Contoh algoritmanya adalah : *Decision Tree*, *K-Nearest Neighbor* dan lainnya.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



## F. Frequent Pattern Mining

Frequent Pattern Mining menurut Khusboo Saxena (2018 : 26), “Pola yang sering muncul pada dataset disebut *frequent pattern* dan selanjutnya seperti pembelian kendaraan roda dua, kemudian helm kemudian sarung tangan dan jika sering terjadi di pasar maka disebut frequent sequence pattern”.

Berbagai jenis pola yang sering, hubungan korelasi dan aturan asosiasi ada. Penambahan pola dan aturan yang sering dapat diklasifikasikan sebagai:

1. Berdasarkan kelengkapan *patterns* yang akan ditambah.

Dalam kriteria ini, pola-pola berikut untuk ditambah sesuai dengan kebutuhan aplikasi. Ini difokuskan pada : *closed itemset*, *closed frequent itemset*, *maximal frequent itemset*, dan lainnya.

2. Atas dasar tingkat abstraksi yang terlibat dalam set aturan.

Ini diklasifikasikan menurut aturan asosiasi tingkat tunggal dan bertingkat. Dalam aturan asosiasi tingkat tunggal, kumpulan item yang diberikan tidak merujuk item pada tingkat abstraksi yang berbeda, sedangkan dalam aturan asosiasi bertingkat, kumpulan item yang diberikan dirujuk pada tingkat abstraksi yang berbeda.

3. Atas dasar jumlah dimensi data yang terlibat dalam aturan asosiasi.

Ini diklasifikasikan sebagai satu dimensi dan multi-dimensi. Dalam aturan satu dimensi, item atau atribut merujuk hanya satu dimensi sedangkan dalam aturan asosiasi multidimensi, item atau atribut merujuk dua atau lebih dimensi.

1. Berdasarkan jenis nilai yang ditangani dalam aturan asosiasi.

Dua jenis nilai, seperti *Boolean* dan kuantitatif ditangani oleh aturan asosiasi. Aturan asosiasi *Boolean* melibatkan asosiasi antara ada dan tidak adanya item di sisi lain aturan asosiasi kuantitatif menjelaskan hubungan antara item kuantitatif atau atribut.

2. Atas dasar aturan yang akan ditambah.



Penambangan pola yang sering dapat menghasilkan berbagai jenis aturan dan hubungan, seperti aturan asosiasi, aturan korelasi, dan hubungan gradien yang kuat.

- a) Aturan asosiasi: Ini menghasilkan sejumlah besar hubungan antara itemset.
- b) Aturan korelasi: Jika aturan asosiasi dianalisis lebih lanjut untuk mengungkap korelasi statistik.
- c) Gradien: Ini adalah rasio item bila dibandingkan dengan *parent*, *child*, atau *sibling*.

### G. Algoritma Apriori

Oleh karena algoritma yang dipakai oleh penulis ini adalah algoritma FP-Growth yang sangat berkaitan dengan association rule, yang juga dipakai pada algoritma Apriori, maka baiknya kita mengerti dahulu apa itu algoritma Apriori.

Menurut Efori Buulolo (2020 : 11), Algoritma apriori adalah salah satu algoritma dalam data mining yang paling terkenal dalam menemukan pola data atau pola kemunculan/frekuensi data.

Biasa algoritma apriori digunakan untuk menemukan pola pembelian pelanggan pada suatu minimarket berdasarkan transaksi pembelian. Dalam algoritma apriori ada istilah support yaitu nilai yang digunakan untuk mengukur kemunculan data tertentu dibandingkan dengan total data. Rumus untuk mencari nilai support adalah sebagai berikut nilai support untuk sebuah *itemset* :

$$\text{Support (A)} = (\text{Jumlah Transaksi Mengandung A} / \text{Total Transaksi}) \times 100$$

Berikutnya nilai *support* untuk dua *itemsets* :

$$\text{Support (A,B)} = P(A \cap B) = (\text{Jumlah Transaksi AdanB} / \text{Total Transaksi}) \times 100$$



## H. Association Rule

Menurut Efori Buulolo (2020 : 19), aturan asosiasi atau *association rule* merupakan teknik dalam *data mining* untuk membentuk aturan asosiatif antara item dalam sebuah kombinasi *itemsets*. Aturan asosiasi berbentuk *if.... Then....* Atau jika.... Maka..... Proses *association rule* merupakan lanjutan dari algoritma apriori, di mana pembentukan *rule*/aturan dimulai dari kombinasi dua *itemsets* yang memenuhi *support* minimum, artinya yang akan dibuat *rule*-nya hanya kombinasi yang memenuhi *support* minimum.

## I. Algoritma Frequent Pattern Growth

Dilansir dari buku Efori Buulolo (2020 : 31), Algoritma ini diusulkan oleh Han et al, untuk menambang set lengkap *frequent itemset* dan menghindari generasi sejumlah besar kandidat set, yaitu *Frequent-Pattern (FP) Growth*. Ini menciptakan pohon FP untuk mengompresi database besar. Dalam node pohon ini, item yang sering diatur sedemikian rupa sehingga node yang lebih sering muncul memiliki peluang yang lebih baik untuk berbagi node daripada yang jarang terjadi.

Algoritma ini menggunakan metode *Divide and Conquer* untuk menguraikan tugas penambangan setelah itu menghindari pembuatan kandidat dengan mempertimbangkan sub-basis saja. Algoritma pertumbuhan FP menyimpan seluruh informasi untuk penambangan pola yang sering. Ini memiliki beberapa keuntungan, seperti memindai database hanya dua kali yang membantu dalam mengurangi biaya komputasi, kedua, menggunakan metode membagi dan menaklukkan sehingga ukuran pohon FP bersyarat berikutnya berkurang.



Hak cipta dimiliki oleh Institut Bisnis dan Informatika Kwik Kian Gie

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



Selain itu, algoritma *FP-growth* memiliki beberapa kelemahan, seperti sulit untuk digunakan dalam proses penambangan interaktif karena pengguna dapat mengubah ambang batas dukungan sesuai dengan aturan yang dapat menyebabkan pengulangan seluruh proses penambangan. Kedua, Tidak cocok untuk penambangan tambahan.

*FP-Growth* memiliki langkah-langkah sebagai berikut :

1. Membuat *conditional pattern* yang berbasis dari setiap *node* di dalam *FP-tree*.
  - a) Dimulai dari tabel *frequent header* di dalam *FP-tree*
  - b) Melintasi pohon *FP-tree* dengan mengikuti jalur dari setiap *frequent item*
  - c) Mengakumulasi semua jalur awal yang diubah dari *item* tersebut untuk membuat sebuah *conditional pattern base*.
2. Membuat *conditional FP-tree* dari setiap *conditional pattern-base*
  - a) Mengakumulasi perhitungan dari setiap *item* di *base*.
  - b) Membuat *FP-tree* untuk *items* yang sering yang berada di *pattern base*.
3. Menambang secara rekursif dari *conditional FP-trees* dan menumbuhkan *frequent patterns* yang didapat sejauh ini. Kalau *conditional FP-tree* mengandung jalur yang tunggal, cukup sebutkan semua *patterns*.

Hak Cipta Dilindungi Undang-Undang

Hak cipta milik IBI KKK Institut Bisnis dan Informatika Kwik Kian Gie

Institut Bisnis dan Informatika Kwik Kian

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.



## J. Penelitian Terdahulu

1. T. Andi dan E. Utami (2018), *Association Rule Algorithm with FP Growth for Book Search*, Yogyakarta : IOP Publishing.

Dari jurnal penelitian ini, peneliti menggunakan kedua algoritma tersebut dan memanfaatkan data transaksi kunjungan dari perpustakaan untuk mencari informasi tentang buku apa saja yang sering dipinjam oleh siswa dan keterkaitan diantaranya sehingga perpustakaan itu dapat melakukan persiapan buku sesuai dengan tingkat nilai support dan confidence tersebut.

2. Hidayat A.A. (2019), *Implementation and Comparison Analysis of Apriori and FP-Growth Algorithm Performance to Determine Market Basket Analysis in Breiliant Shop*, Yogyakarta : IOP Publishing.

Dari jurnal penelitian ini, peneliti menggunakan kedua algoritma tersebut untuk menganalisis keranjang pasar pada toko brilian. Disimpulkan bahwa metode aturan asosiasi menggunakan algoritma Apriori dan FP-Growth dapat memperoleh korelasi pembelian barang-barang pada toko brilian tersebut. Disimpulkan bahwa metode aturan asosiasi tersebut dapat meningkatkan strategi penjualan di dalam hal promosi barang untuk meningkatkan pendapatan.

3. Daniel Hunyadi (2011), *Performance Evaluation of Apriori and FP-Growth Algorithms*, Romania : University of Sibiu.

Dalam jurnal ini, peneliti membandingkan performa dari algoritma Apriori dan FP-Growth dalam menghasilkan aturan asosiasi. Kedua algoritma tersebut diimplementasikan



di dalam Rapid Miner dan hasil dari pengolahan data tersebut dianalisis menggunakan SPSS. Hasil dari penelitian ini menyimpulkan bahwa meskipun kedua algoritma tersebut memproses data dengan cara yang berbeda, hasil pengolahan data dari kedua algoritma tersebut tetap memiliki hubungan yang signifikan.

4. Erwin (2009), *Analisis Market Basket dengan Algoritma Apriori dan FP-Growth*, Palembang : Universitar Sriwijaya.

Dari jurnal penelitian ini disimpulkan bahwa algoritma Apriori memiliki kekurangan yaitu dalam proses pencarian frequent itemsets membutuhkan waktu yang lebih lama. Frequent Pattern Tree digunakan bersamaan dengan FP-Growth untuk menentukan *frequent itemset* dari *database*. Berbeda dengan cara Apriori yang membutuhkan langkah candidate generation yaitu melakukan *scanning database* secara berulang untuk menentukan *frequent itemset*. Lalu FP-Growth menggunakan memori komputer yang lebih sedikit dan hanya membutuhkan dua kali *scanning database* sehingga menjadi lebih efisien daripada Apriori.

5. Shivam Sidhu, Upendra Kumar dan Aditya Nawani (2014), *FP Growth Algorithm Implementation*, India : International Journal of Computer Applications.

Dari jurnal penelitian ini, disimpulkan bahwa dengan menggunakan data mining yang menghasilkan ringkasan informasi secara konsekuen, dapat digunakan untuk merancang strategi bisnis untuk meningkatkan pendapatan maupun menurunkan biaya. Untuk waktu yang lama, banyak algoritma yang berbeda seperti algoritma Apriori telah digunakan dalam bidang analisis pola. Telah ditemukan bahwa algoritma ini memiliki



beberapa kelemahan seperti pemindaian berulang dari seluruh basis data, dan *candidate key* generation, yang selanjutnya membutuhkan tes kandidat. Karena itu, jika datanya terlalu besar atau kompleks, waktu dan kerumitannya meningkat. Algoritma FP-growth menggunakan 'Divide and Conquer' dan tidak memerlukan tes pembuatan kunci kandidat. Selain itu, tidak menjalani pemindaian data berulang kali. Jadi, dapat disimpulkan dengan aman bahwa algoritma FP-growth memiliki cakupan masa depan yang luas di bidang pemasaran di sektor terorganisir.

6. M. S. Mythili dan A. R. Mohamed Shanavas (2013), *Performance Evaluation of Apriori and FP-Growth Algorithms*, India : International Journal of Computer Applications.

Pada jurnal ini, peneliti membandingkan dan mengevaluasi kemampuan dari kedua algoritma tersebut. Hasil dari penelitian ini disimpulkan bahwa algoritma FP-Growth lebih efisien dan dapat ditingkatkan daripada algoritma Apriori, berdasarkan dari perbandingan waktu eksekusi dan berjalannya program pada suatu komputer.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik dan tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar IBIKKG.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IBIKKG.