PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA *FP-GROWTH* UNTUK MENENTUKAN POLA PEMBELIAN PRODUK MAKANAN

SKRIPSI

ULI AUDINA NURSYANI

71153028

PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN
2020

PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA *FP-GROWTH* UNTUK MENENTUKAN POLA PEMBELIAN PRODUK MAKANAN

SKRIPSI

Diajukan untuk Memenuhi syarat Mencapai Gelar Sarjana Komputer

ULI AUDINA NURSYANI

71153028

PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN
2020

PERSETUJUAN SKRIPSI

Hal : Surat Persetujuan Skripsi

Lamp : -

Kepada Yth,

Dekan Fakultas Sains dan Teknologi

Universitas Islam Negeri Sumatera Utara Medan

Assalamu'alaikum Warahmatullahi Wabarakathu

Setelah membaca, meneliti, memberikan petunjuk, dan mengoreksi serta mengadakan perbaikan, kami selaku pembimbing berpendapat skripsi saudara,

Nama : Uli Audina Nursyani

NIM : 711530268

Program Studi : Ilmu Komputer

Judul : Penerapan Data Mining Menggunakan Algoritma *Fp – Growth* Untuk Menentukan Pola Pembelian Produk Makanan

Dapat disetujui untuk segera di*munaqasya*kan. Atas perhatiannya kami ucapkan terima kasih.

Medan, 26 Oktober 2020 M
09 Rabiul Awal 1442 H

Komisi Pembimbing,

Pembimbing I,	Pembimbing II,
Dr. Mhd. Furqan, S.Si., M.Comp.sc.	Sriani, S. Kom., M.Kom
NIP. 198008062006041003	NIB. 1100000108

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan dibawah ini:

Nama : Uli Audina Nursyani

Nim : 71153028

Program Studi : Ilmu Komputer

Judul : Penerapan Data Mining Menggunakan Algoritma *Fp-Growth* Untuk Menentukan Pola Pembelian Produk Makanan

Menyatakan dengan sebenarnya bahwa skripsi yang saya serahkan ini benar-benar merupakan hasil karya saya sendiri, kecuali kutipan-kutipan dari ringkasan- ringkasan yang semuanya saya jelaskan sumbernya.

Apabila kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil ciplakan, maka gelar dan ijazah yang diberikan oleh institut batal saya terima.

Medan, 26 Oktober 2020

Materai

6000

Uli Audina Nursyani

NIM. 71153028

KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA MEDAN
FAKULTAS SAINS DAN TEKNOLOGI
Jl. IAIN No. 1 Medan 20235
Telp. (061) 6615683-6622925, Fax. (061) 6615683
Url: https://saintek.unsu.ac.id, E-mail: saintek@uinsu.ac.id

PENGESAHAN SKRIPSI

Nomor: 191/ST/ST.V.2/PP.01.1/11/2021

Judul	: Penerapan Data Mining Menggunakan Algoritma *Fp-Growth* Untuk Menentukan Pola Pembelian Produk Makanan
Nama	: Uli Audina Nursyani
Nomor Induk Mahasiswa	: 71153028
Program Studi	: Ilmu Komputer
Fakultas	: Sains dan Teknologi

Telah dipertahankan di hadapan Dewan Penguji Skripsi Program Studi Ilmu Komputer Fakultas Sains dan Teknologi UIN Sumatera Utara Medan dan dinyatakan **LULUS.**

Pada hari/ Tanggal	: Senin, 26 Oktober 2020
Tempat	: Via Zoom Meeting

Tim Ujian Munaqasyah,
Ketua,

Dr. Mhd. Furqan, S.Si, M.Comp.Sc.
NIP. 198008062006041003

Dewan Penguji,

Penguji I,

Dr. Mhd. Furqan, S.Si, M.Comp.Sc.
NIP. 198008062006041003

Penguji II,

Sriani, S.Kom., M.Kom
NIB. 1100000108

Penguji III,

Abdul Halim Hasugian, M.Kom
NIB. 1100000113

Penguji IV,

Armansyah, M.Kom
NIB. 1100000074

Mengesahkan,
Dekan Fakultas Sains dan Teknologi
UIN Sumatera Utara Medan,

Dr. H.M Jamil, M.A.
NIP. 196609101999031002

ABSTRAK

Dalam membangun sebuah bisnis perusahaan distributor produk makanan sangat memerlukan penggunaan teknologi informasi, guna mendukung kelancaran penjualan produk-produk makanan yang disediakan. Masalah yang terjadi adalah bagaimana memeberikan pelayanan terbaik terhadap pasar atau pelanggan untuk menggunakan strategi bisnis yang tepat sasaran untuk menghindari kerugian pada perusahaan. Perusahaan dituntut untuk menemukan suatu strategi yang dapat meningkatkan penjualan produk yang dijual. Salah satu cara yang dapat dilakukan untuk penentuan strategi penjualan produk makanan ringan adalah dengan menggunakan teknik data mining. Adapun teknik yang digunakan dalam hal ini adalah algoritma *Fp-Growth*, yaitu sebuah algoritma yang menghasilkan *frequent itemset* yang nantinya akan digunakan dalam proses penentuan aturan yang dapat menghasilkan sebuah pilihan. Algoritma *Fp-Growth* merupakan pengembangan dari algoritma *Apriori*. Algoritma *Fp-Growth* menggunakan konsep *tree development* dalam pencarian jenis barang yang sering dibeli (*frequent itemsets*). Data yang digunakan yaitu 20 jenis produk makanan ringan dan 30 data transaksi. Pada penelitian ini ditentukan nilai *minimum support* sebesar 55% dan nilai *minimum confidence* sebesar 75%. Dari hasil pengujian yang dilakukan didapatkan sebuah aturan (*rule*) dengan nilai *confidence* 75% yaitu, jika konsumen membeli kopiko coffe shot cappuccino maka juga membeli beng-beng wafer cokelat.

Kata Kunci : *Penjualan, Data mining, Association rules, Frequent itemset, Fp-Growth*

ABSTRACT

In building a food product distributor company business, it is very necessary to use information technology, in order to support the smooth sales of the food products provided. The problem that occurs is how to provide the best service to the market or customers to use targeted business strategies to avoid losses to the company. Companies are required to find a strategy that can increase sales of products sold. One way that can be done to determine the sales strategy of snack products is to use data mining techniques. The technique used in this case is the Fp-Growth algorithm, which is an algorithm that produces a frequent itemset which will later be used in the process of determining the rules that can produce a choice. The Fp-Growth algorithm is a development of the Apriori algorithm. The Fp-Growth algorithm uses the concept of tree development in the search for frequently purchased items (frequent itemsets). The data used are 20 types of snack products and 30 transaction data. In this study, the minimum support value is 55% and the minimum confidence value is 75%. From the results of the tests carried out, it was found a rule with a confidence value of 75%, namely, if consumers buy kopiko coffee shot cappuccinos, they also buy chocolate wafers.

Keywords: Sales, Data mining, Association rules, Frequent itemset, Fp-Growth

KATA PENGANTAR

Puji dan syukur penulis ucapkan kepada ALLAH SWT yang telah memberikan nikmat berupa kesehatan, kesempatan kepada penulis sehingga mampu menyelesaikan Proposal skripsi ini.

Maka sehubungan dengan Proposal skripsi ini, segala bentuk hasil dituliskan kedalam bentuk Proposal skripsi ini sebagai persyaratan dalam menyelesaikan mata kuliah Proposal skripsi di Program Studi Ilmu Komputer. Tujuan utama dari penulisan Proposal skripsi ini adalah untuk memantapkan teori dan praktek yang telah dipelajari di kampus dan dapat diselesaikan dengan serta diaplikasikan.

1. Prof. Dr. Saidurrahman, M.Ag. Selaku Rektor Universitas Islam Negeri Sumatera Utara Medan yang telah memberikan kesempatan kepada penulis untuk pelaksanaan kegiatan penyusunan tugas akhir khusus tahun akademik 2019/2020.
2. Bapak Dr. H. M. jamil, M.A selaku Dekan Fakultas Sains dan Teknologi.
3. Bapak Dr. Mhd. Furqan, S.Si., M.Comp.Sc selaku Ketua Program Studi Ilmu Komputer sekaligus sebagai Dosen Pembimbing I penulis.
4. Ibu Sriani, S.Kom.,M.Kom selaku Dosen Pembimbing II penulis yang telah berkontribusi membantu dalam memberikan ide, saran, keritik, dan bimbingannya kepada penulis selama mengerjakan proposal skripsi ini.
5. Kedua orang tua yang sangat penulis sayangi dan cintai yang rela mengorbankan apapun demi kelancaran penulis dalam menyelesaikan pendidikan
6. Kepada saudara-saudari kandung yang penulis sayangi yang bayak memberikan semangat dan kepedulian yang luar biasa.
7. Kepada teman-teman sekelas Ilmu Komputer II yang telah memberikan semangat.
8. Dan semua pihak yang telah membantu penulis namun tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa Proposal skripsi ini masih jauh dari kata sempurna. Untuk itu penulis mengharapkan saran perbaikan dari para penguji, pembimbing dan pembaca demi kesempurnaan dari Proposal skripsi ini. Demikian penyusunan Proposal skripsi ini di tulis. Semoga Proposal skripsi ini dapat bermanfaat bagi pembaca dan dapat digunakan dengan semestinya.

Medan, 26 Oktober 2020
Penulis,

ULI AUDINA NURSYANI
NIM. 71153028

DAFTAR ISI

ABSTRAK ..... **i**
ABSTRACT ..... **ii**
KATA PENGANTAR ..... **iii**
DAFTAR ISI ..... **v**
DAFTAR TABEL ..... **vii**
DAFTAR GAMBAR ..... **viii**
DAFTAR LAMPIRAN ..... **x**
BAB I PENDAHULUAN **1**
1.1. Latar Belakang 1
1.2. Rumusan Masalah 3
1.3. Batasan Masalah 3
1.4. Tujuan Penelitian 3
1.5. Manfaat penelitian 3
BAB II TINJAUAN PUSTAKA **4**
2.1 Pengertian Data Mining 4
2.1.1. *Knowledge Discovery In Database* 7
2.1.2. Operasi Data Mining 8
2.1.3. Pengelompokan Data Mining 8
2.1.4. *Association Rules* 10
2.1.5. Langkah-Langkah Proses Aturan Asosiasi Rule 12
2.2. Algoritma *FP-Growth* 13
2.2.1. Pembangunan *FP-Tree* 13
2.3. *Flowchart* 19
2.4. *RapidMiner* 21
2.5. Kajian Terdahulu 22
BAB III METODOLOGI PENELITIAN **25**
3.1. Tempat dan Waktu 25
3.1.1. Tempat Penelitian 25
3.1.2. Waktu Dan Jadwal Pelaksanaan Penelitian 25

3.2. Bahan dan Alat Penelitian 25
3.2.1. *Hardware* (Perangkat Keras) 26
3.2.2. *Software* (Perangkat Lunak) 26
3.3. Teknik Pengumpulan Data 26
3.4. Penerapan Algoritma 27
3.5. Pemodelan 28
3.6. Pengujian 28
BAB IV HASIL DAN PEMBAHASAN 29
4.1. Pembahasan 29
4.1.1. Analisis Data 29
4.1.2. Representasi Data 30
4.2. Hasil 58
4.2.1. Implementasi Sistem 58
KESIMPULAN DAN SARAN 62
5.1. Kesimpulan 62
5.2. Saran 62
DAFTAR PUSTAKA 63

DAFTAR TABEL

Tabel 2.1 Tabel Data Transaksi Awal 14
Tabel 2.2 Frekuensi Kemunculan Tiap Karakter 14
Tabel 2.3 Tabel Data Transaksi 15
Tabel 2.4 Hasil *Frequent Itemset* 18
Tabel 2.5 Simbol *Flowchart* 19
Tabel 3.1 Waktu dan Jadwal Penelitian 25
Tabel 3.2 Perangkat Keras 26
Tabel 3.3 Perangkat Lunak 26
Tabel 4.1 Untuk Memberi Tanda Pada Setiap *Itemset* 32
Tabel 4.2 Data Transaksi 33
Tabel 4.3 Transaksi Penjualan Dalam Format Binominal 34
Tabel 4.4 Frekuensi Tiap Pembeli Melakukan Transaksi 35
Tabel 4.5 Frekuensi Kemunculan Tiap Item/Produk 36
Tabel 4.6 Tabel Frekuensi Dan Nilai *Support* 37
Tabel 4.7 Transaksi Yang Disesuaikan Dengan *Frequent List* 38
Tabel 4.8 Hasil *Frequent Itemset* 47
Tabel 4.9 *Conditional Pattern Base* 48
Tabel 4.10 *Conditional Fp-tree* 48
Tabel 4.11 Kombinasi 2 *Itemsets* 49
Tabel 4.12 *Rule* Yang Terbentuk 57

DAFTAR GAMBAR

Gambar 2.1 Tahapan Proses KDD 7
Gambar 2.2 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 1 15
Gambar 2.3 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 2 16
Gambar 2.4 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 3 16
Gambar 2.5 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 10 16
Gambar 3.1 Blok Diagram Algoritma *Fp-Growth* 27
Gambar 4.1 *Flowchart* Pengolahan Data Mining..... 31
Gambar 4.2 *Flowchart* Algoritma *Fp-Growth* 32
Gambar 4.3 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 1 40
Gambar 4.4 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 2 40
Gambar 4.5 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 3 41
Gambar 4.6 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 4 41
Gambar 4.7 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 5 42
Gambar 4.8 Hasil Pembentukan *Fp-Tree* Setelah Pembacaan TID 30 42
Gambar 4.9 Lintasan Yang Mengandung Simpul S 43
Gambar 4.10 Lintasan Yang Mengandung Simpul M..... 43
Gambar 4.11 Lintasan Yang Mengandung Simpul I..... 44
Gambar 4.12 Lintasan Yang Mengandung Simpul H 44
Gambar 4.13 Lintasan Yang Mengandung Simpul G 44
Gambar 4.14 Lintasan Yang Mengandung Simpul E..... 45
Gambar 4.15 Lintasan Yang Mengandung Simpul P 45
Gambar 4.16 Lintasan Yang Mengandung Simpul N 46
Gambar 4.17 Lintasan Yang Mengandung Simpul B..... 46
Gambar 4.18 Lintasan Yang Mengandung Simpul D 46
Gambar 4.19 Lintasan Yang Mengandung Simpul C..... 46
Gambar 4.20 Tampilan *Import Excel Sheet* 60
Gambar 4.21 *Fp-Growth* Pada *Rapidminer*..... 60
Gambar 4.22 *Support & Confidence Rapidminer* Dari Sample Data Penjualan . 61

Gambar 4.23 *Grafis Rule Sample* Data Penjualan Menggunakan *Rapidminer* ... 61
Gambar 4.24 *Association Rules* Dari *Rapidminer* 62

DAFTAR LAMPIRAN

Lampiran	Judul Lampiran
1.	Surat Keterangan Penelitian
2.	Daftar Riwayat Hidup
3.	Kartu Bimbingan Skripsi

BAB I
PENDAHULUAN

1.1. Latar Belakang

Persaingan dunia bisnis pada saat ini sangatlah ketat, terutama dibidang bisnis distributor produk makanan pokok dan kebutuhan sehari-sehari masyarakat. Salah satu distributor produk makanan adalah PT. Cipta Niaga Semesta (Mayora Group). Dalam membangun bisnis, perusahaan PT. Cipta Niaga Semesta (Mayora Group) yang menyediakan bahan pokok makanan tidak luput untuk menggunakan teknologi informasi guna mendukung kelancaran penjualan produk-produk makanan yang disediakan. Tidak dipungkiri bahwa penggunaan teknologi informasi menjadi sebuah keharusan pada saat ini, dengan melihat pesat dan ketatnya persaingan bisnis dibidang tersebut.

Masalah yang terjadi adalah dalam hal memberikan pelayanan terbaik terhadap pasar atau pelanggan perusahaan untuk menggunakan strategi bisnis yang tepat sasaran untuk menghindari kerugian perusahaan. Akan tetapi, terkadang pihak pengelola perusahaan terkendala dalam hal menentukan strategi. Banyak faktor yang menyebabkan hal tersebut. Salah satu faktor penyebab adalah sulitnya menghasilkan analisa terkait dengan pola data penjualan produk yang sudah ada, sehingga memberikan dampak negatif dari hasil penjualan produk yang naik dan turun. Berdasarkan dari hasil riset, data penjualan produk makanan pada perusahaan dari bulan Januari hingga Desember tahun 2019 mengalami siklus grafik naik turun, tentu hal ini tidak baik dalam strategi bisnis. Salah satu contohnya pada bulan Januari tahun 2019 tercatat produk yang keluar sebanyak 1580 dus, sedangkan pada 2 bulan berikutnya sebanyak 1384 dus dan 1491 dus.

Hal ini dikarenakan pada PT. Cipta Niaga Semesta (Mayora Group), produk-produk yang paling banyak terjual dan saling berhubungan masih belum terorganisir dengan baik, sehingga pola pembelian konsumen terhadap suatu produk yang dipilih tidak teranalisa dengan tepat untuk menentukan strategi penyetokan produk yang akan dipasarkan. Tentu hal ini menyebabkan strategi penyetokan barang tidak terencana dengan baik pada masa yang akan datang.

Dalam Al-Qur’an juga dijelaskan bahwa:

وَلَا تَأْكُلُوْٓا اَمْوَالَكُمْ بَيْنَكُمْ بِالْبَاطِلِ وَتُدْلُوْا بِهَآ اِلَى الْحُكَّامِ لِتَأْكُلُوْا فَرِيْقًا مِّنْ اَمْوَالِ النَّاسِ بِالْاِثْمِ وَاَنْتُمْ تَعْلَمُوْنَ

Artinya: “*Dan janganlah kamu makan harta di antara kamu dengan jalan yang batil, dan (janganlah) kamu menyuap dengan harta itu kepada para hakim, dengan maksud agar kamu dapat memakan sebagian harta orang lain itu dengan jalan dosa, padahal kamu mengetahui*”. (QS. Al-Baqarah Ayat 188)

Berdasarkan penjelasan ayat Al-Qur’an di atas, maka dibutuhkanya teknik strategi bisnis yang baik untuk mendapatkan keuntungan yang maksimal dengan cara yang benar. Salah satu teknik yang dapat dilakukan dalam menentukan strategi pola pembelian konsumen adalah dengan teknik data mining. Data mining merupakan suatu proses otomatis untuk menemukan informasi (*knowledge*) baru yang memiliki potensi dari sekumpulan data (Kurniasih et al., 2012). Teknik data mining memerlukan algoritma agar perhitungan semakin kompleks, salah satunya adalah algoritma *FP-Growth*. Algoritma *FP-Growth* merupakan pengembangan dari algoritma *Apriori*. Sehingga kekurangan dari algoritma *Apriori* diperbaiki oleh algoritma *FP-Growth.*

Frequent Pattern Growth (FP-Growth) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data (Fahrin & Maulana, 2018). Penelitian yang dilakukan oleh Herasmus dengan menggunakan algoritma *FP-Growth* dalam menganalisa *Costemer Service System*, memberikan kesimpulan bahwa nilai *support* yang diberikan rata-rata 80 persen dan nilai *confidence* diatas 80 persen (Herasmus, 2017). Sehingga berdasarkan uraian masalah dan penelitian terdahulu, maka penelitian ini akan menerapkan algoritma *FP-Growth* dalam melakukan proses data mining untuk mengetahui pola pembelian konsumen PT. Cipta Niaga Semesta (Mayora Group). Dari pola yang dihasilkan tersebut akan menghasilkan sebuah informasi yang dapat dimanfaatkan pihak PT. Cipta Niaga Semesta (Mayora Group) untuk meningkat mutu penjualan, mutu pelayanan serta keuntungan dengan meminalisir kerugian.

1.2. Rumusan Masalah

Berdasarkan uraian dari latar belakang diatas, maka dapat diambil suatu rumusan masalah yaitu bagaimana menerapkan algoritma *FP-Growth* ini untuk menentukan pola pembelian produk makanan.

1.3. Batasan Masalah

Dalam penulisan skripsi ini diperlukan batasan-batasan, agar apa yang akan diteliti sesuai dengan tujuan penelitian. Adapun batasan masalahnya adalah sebagai berikut :

1. Informasi yang diperoleh berupa produk yang sering di order oleh konsumen secara bersamaan berdasarkan persentase nilai minimum *support* dan *confidence*.
2. Data yang diambil merupakan data transaksi penjualan dari bulan januari sampai dengan bulan desember 2019 di PT. Cipta Niaga Semesta (Mayora Group).
3. Penelitian ini khusus membahas tentang pola pembelian konsumen pada produk makanan di PT. Cipta Niaga Semesta (Mayora Group).
4. *Tools* yang digunakan dalam penerapan data mining ini adalah *Rapidminer*.

1.4. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk menerapkan algoritma *Fp-Growth* untuk menentukan pola pembelian produk makanan.

1.5. Manfaat penelitian

Adapun manfaat penelitian yang dilakukan adalah :

1. Mengetahui produk makanan apa saja yang sering di beli oleh konsumen sehingga dapat diperbanyak pengadaannya di gudang.
2. Dapat meningkatkan penjualan produk makanan karena produk makanan yang disediakan adalah produk yang laris dijual.
3. Sebagai bahan informasi bagi peneliti untuk menambah keilmuan tentang algoritma *fp-growth*.

BAB II
TINJAUAN PUSTAKA

2.1 Pengertian Data Mining

Data mining adalah proses untuk mendapatkan sebuah informasi dengan melakukan pencarian pola dan relasi-relasi yang tersembunyi di dalam timbunan data yang banyak. Data mining atau sering disebut sebagai *knowledge discovery in database* (KDD) adalah kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola atau hubungan dalam data yang berukuran besar. Keluaran Data mining ini bisa dipakai untuk membantu pengambilan keputusan di masa mendatang. (Sriani et al., 2015)

Pengembangan KDD ini menyebabkan penggunaan *pattern recognition* semakin berkurang karena telah menjadi bagian data mining. Data mining berisi pencarian trend atau pola yang diinginkan dalam database untuk membantu pengambilan keputusan yang akan datang. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunkan perangkat pendukung keputusan lainnya. Pemanfaatan data mining dapat dilihat dari dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, pemanfaatan data mining dapat digunakan dalam menangani meledaknya volume data, bagaimana menyimpannya, mengekstraknya serta memanfaatkannya. Data mining tidak hanya digunakan untuk menangani persoalan menumpuknya data atau informasi dan bagaimana menggunakannya atau menyimpan tanpa kehilangan informasi yang penting (*warehousing*). (Amalia, 2018)

Ada beberapa definisi data mining menurut beberapa ahli yaitu sebagai berikut :

1. Menurut Davies, Data mining adalah penambangan atau penemuan informasi baru dengan mecari pola atau aturan tertentu dari sejumlah data yang sangat besar.

2. Menurt Pramudiono, Data mining merupakan serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data.
3. Menurut Han, Data mining adalah egiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya.
4. Menurut Witten, Data mining adalah proses menemuan pola-pola dalam data, pola yang ditemukan harus penuh arti dan pola tersebut memberikan keuntungan, biasanya keuntungan secara ekonomi.
5. Menurut Turban, Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar.
6. Menurut Gartner Group, Data mining adalah proses menemukan korelasi baru yang bermakna, pola dan Tren oleh memilah-milah sejumlah besar data yang disimpan dalam repositori, menggunakan pola teknologi pengenalan serta teknik matematika dan statistik.
7. Menurut Abdul Kadir, Data Mining adalah kombinasi secara logis antara pengetahuan data dan analisa statistik yang dikembangkan dalam pengetahuan bisnisatau suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, tiruan dan *machine learning* untuk mengekstaksi dan mengidentifikasi informasi yang bermanfaat bagi pengetahuan yang terkait dari berbagai database besar.

Berdasarkan beberapa defenisi tersebut dapat ditarik kesimpulan bahwa data mining adalah suatu teknik menggali informasi berharga yang terpendam atau tersembuyi pada suatu koleksi data (*database*) yang sangat besar, sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. (Pramudiono, 2007).

Hal penting yang terkait dengan data mining adalah sebagai berikut :

1. Data mining merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.

3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

Data mining bukanlah suatu bidang yang sama sekali baru. Salah satu kesulitan untuk mendefenisikan data mining adalah kenyataan bahwa data mining mewarisi banyak aspek dan teknik dari bidang-bidang ilmu yang sudah mapan terlebih dahulu. Teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang sudah ada untuk membangun sebuah model, kemudian dengan menggunakan model tersebut dapat dikenali pola data lain yang tidak berada dalam basis data yang tersimpan. Dalam data mining pengelompokan data juga bias dilakukan. Tujuannya untuk dapat mengetahui pola universal data-data yang ada. Anomali data transaksi juga perlu di deteksi untuk dapat mengetahui tindak lanjut berikutnya yang dapat diambil. Hal ini bertujuan untuk mendukung kegiatan operasional perusahaan, sehingga tujuan akhir diharapkan dapat tercapai. (Prasetyo, 2012)

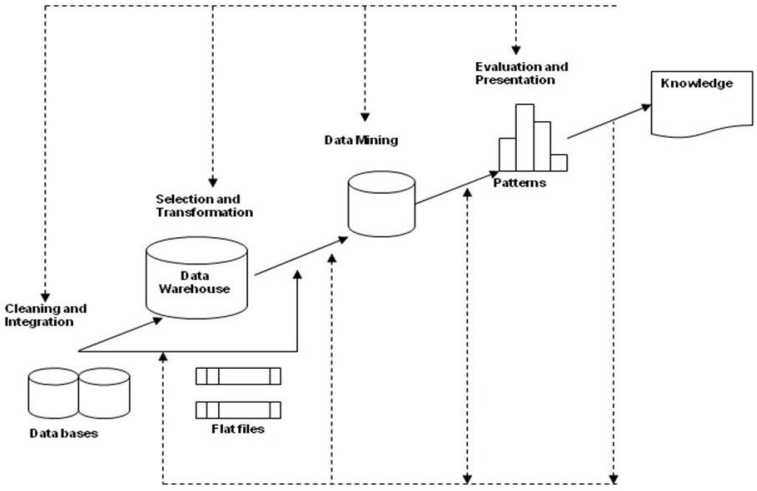
Data mining dilakukan dengan *tool* khusus, yang mengeksekusi operasi data mining yang telah didefinisikan berdasarkan model analisis. Data mining merupakan proses analisis terhadap data dengan penekanan menemukan informasi yang tersembunyi pada sejumlah data besar yang disimpan ketika menjalankan bisnis perusahaan. Kemajuan luar biasa yang terus berlanjut dalam bidang data mining didorong oleh beberapa faktor antara lain: (Larose, 2013)

1. Pertumbuhan yang cepat dalam kumpulan data.
2. Penyimpanan data dalam data *warehouse*, sehingga seluruh perusahaan memiliki akses ke dalam *database* yang andal.
3. Adanya peningkataan akses data melalui navigasi *web* dan internet.
4. Tekanan kompetisi bisnis untuk meningkatkan penguasaan pasar dalam globalisasi ekonomi.
5. Perkembangan teknologi perangkat lunak untuk data mining (ketersediaan teknologi).
6. Perkembangan yang hebat dalam kemampuan komputasi dan pengembangan kapasitas media penyimpanan.

2.1.1. *Knowledge Discovery In Database*

Istilah data mining dan *knowledge discovery in database* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Salah satu tahapan dalam keseluruhan proses KDD adalah data mining.

Knowledge Discovery in Database (KDD) adalah proses untuk menggali dan menganalisis sejumlah besar himpunan data dan mengekstrak informasi dan pengetahuan yang berguna. Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan pontensial bermanfaat. Data mining merupakan salah satu langkah dari serangkaian proses iterativ KDD. Tahapan proses KDD dapat dilihat pada gambar 2.1. (Haikal, 2017).



Gambar 2.1 Tahapan Proses KDD

(Sumber: Aditya, 2015)

Adapun komponen-komponen yang terdapat dalam KDD adalah sebagai berikut :

1. *Data Cleaning*, untuk menghilangkan noise data yang tidak konsisten. Dan Data Integration dimana sumber data yang terpecah dapat disatukan.
2. *Data selection*, yaitu dimana data yang relevan dengan tugas analisis dikembalikan ke dalam database.

3. *Data Transformation*, dimana data berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresi.
4. *Data Mining*, proses esensial dimana metode yang intelijen digunakan untuk mengekstrak pola data.
5. *Pattern Evolution*, untuk mengidentifikasi pola yang benar-benar menarik yang mewakili pengetahuan berdasarkan atas beberapa tindakan yang menarik.
6. *Knowledge Presentation*, dimana gambaran teknik visualisasi dan pengetahuan digunakan untuk memberikan pengetahuan yang telah ditambang kepada user.

2.1.2. Operasi Data Mining

Operasi data mining menurut sifatnya terbagi menjadi 2, yaitu : (Hermawati, 2013).

1. Bersifat prediksi (*prediction driven*) untuk menjawab pertanyaan apa dan sesuatu yang bersifat abstrak atau transparan. Operasi prediksi digunakan untuk validasi *hipotesis*, *querying*, dan pelaporan.
2. Penemuan (*discovery driven*) bersifat transparan, operasi penemuan digunakan untuk analisis data eksplorasi, pemodelan prediktif, segmentasi *database*, analisis keterkaitan, dan deteksi defiasi.

2.1.3. Pengelompokan Data Mining

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu sebagai berikut. (Ikhwan et al., 2015)

1. Deskripsi

   Terkadang peneliti dan analis secara sederhana ingin mencoba mencari cara untuk

   menggambarkan pola dan kecenderungan yang terdapat dalam data. Sebagai contoh, petugas pengumpulan suara mungkin tidak dapat menentukan keterangan atau fakta bahwa siapa yang tidak cukup profesional akan sedikit didukung dalam pemilihan presiden.

2. Estimasi

   Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih kearah numerik daripada kearah kategori. Sebagai contoh akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, indeks berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi. Model estimasi yang dihasilkan dapat digunakan untuk kasus baru lainnya.

3. Prediksi

   Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Contoh prediksi bisnis dan penelitian adalah:

   a. Prediksi harga beras dalam tiga bulan yang akan datang.
   b. Prediksi persentasi kenaikan kecelakaan lalu lintas tahun depan jika batas

4. Klasifikasi

   Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang dan pendapatan rendah.

5. Pengklusteran

   Pengklusteran merupakan pengelom-pokan *record*, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

   Contoh pengklusteran dalam bisnis dan penelitian adalah:

   a. Mendapatkan kelompok-kelompok konsumen untuk target pemasaran dari satu suatu produk bagi perusahaan yang tidak memiliki dana pemasaran yang besar.
   b. Untuk tujuan audit akuntansi, yaitu melakukan pemisahan terhadap perilaku *financial* dalam baik dan mencurigakan.
   c. Melakukan pengklusteran terhadap ekspresi dari *gen*, untuk mendapatkan kemiripan perilaku dari *gen* dalam jumlah besar.

6. Asosiasi

   Tugas asosiasi dalam *Data Mining* adalah menemukan *attribut* yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja. Contoh asosiasi dalam bisnis dan penelitian adalah:

   a. Meneliti jumlah pelanggan dari perusahaan telekomunikasi seluler yang diharapkan untuk memberikan respon positif terhadap penawaran *upgrade* layanan yang diberikan.
   b. Menentukan barang dalam supermarket yang dibeli secara bersamaan dan yang tidak pernah dibeli secara bersamaan.

2.1.4. *Association Rules*

Association rules merupakan suatu proses pada data mining untuk menentukan semua aturan asosiatif yang memenuhi syarat minimum untuk *support* (*minsup*) dan *confidance* (*minconf*) pada sebuah *database*. Kedua syarat tersebut akan digunakan untuk *interesting association rules* yang dibandingkaan dengan batasan yang telah ditentukan, yaitu *minsup* dan *minconf*. (Eska, 2016)

Association rule memiliki dua tahap pengerjaan, yaitu sebagai berikut. (Wijaya, 2018)

1. Mencari *frequent itemset*
2. Mendefenisikan *condition* dan *result* (untuk *conditional association rule*)

Dalam menentukan suatu *association rule* terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapat dari hasil pengolahan data dengan perhitungan tertentu. Umumnya ada dua ukuran, yaitu :

1. *Support*, merupakan suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu itemset layak untuk dicari *confidence*-nya (misal, dari keseluruhan transaksi yang ada, seberapa besar tingkat dominasi suatu item yang menunjukkan bahwa item A dan item B dibeli bersamaan).
2. *Confidence*, merupakan suatu hubungan yang menunjukkan hubungan Antara dua item secara *conditional* (misal, menghitung kemungkinan seberapa sering item B dibeli oleh pelanggan jika pelanggan tersebut membeli sebuah item A).

Kedua ukuran ini nantinya bergunan dalam menentukan kekuatan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua paremeter tersebut yang ditentukan oleh pengguna. Bila suatu pola memenuhi kedua nilai minimum parameter yang sudah ditentukam sebelumnya, maka pola tersebut dapat disebut sebagai *interesting rule* atau *strong rule*.

Dalam menentukan nilai minimum support sebuah item dapat diperoleh dengan menggunakan rumus persamaan (1) seperti di bawah ini :

$$Support\ (A) = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total Transaksi}} \qquad (1)$$

$$Support\ A \cap B = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}} \qquad (1)$$

Sedangkan dalam menentukan *minimum confidence* dapat ditentukan dengan rumus persamaan (2) seperti dibawah ini :

$$Confidence\ (\text{A} \rightarrow \text{B}) = P\ (A|B)$$
$$= \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}} \qquad (2)$$

Penggalian *association rules* adalah suatu prosedur untuk mencari hubungan antar *item* dalam suatu *dataset*. Dimulai dengan mencari *frequent itemset*, yaitu kombinasi yang paling sering terjadi dalam suatu *itemset* dan harus memenuhi minimal support (*minsup*). Penggalian *itemset* yang sering muncul dari database transaksional yang besar adalah salah satu masalah yang paling menantang dalam *data mining*. Dalam banyak skenario dunia nyata, data tidak diekstraksi darisumber data tunggal tetapi dari data yang terdistribusi dan heterogen. Pengetahuan yang ditemukan diharapkan bisa membantu operasional bisnis yang lebih baik. Dalam metode *data mining*, *association rules* adalah salah satu yang paling populer. Namun, penggalian pola informasi yang menggunakan *association rules* seringkali menghasilkan pola individu yang sangat besar, sehingga

membiarkan analis menyelesaikan tugas dengan semua peraturan dan menemukan salah satu yang menarik. (Kurniawan, et al, 2018)

2.1.5. Langkah-Langkah Proses Aturan Asosiasi Rule

Proses Aturan Asosiasi terdiri dari beberapa tahap sebagai berikut. (Tyas.2018)

1. Sistem men-*scan* database untuk mendapat kandidat 1-*itemset* (himpunan item yang terdiri dari 1 item) dan menghitung nilai supportnya. Kemudian nilai supportnya tersebut dibandingkan dengan minimum support yang telah ditentukan, jika nilainya lebih besar atau sama dengan minimum support maka *itemset* tersebut termasuk dalam *large itemset*.
2. Itemset yang tidak termasuk dalam *large itemset* tidak diikutkan dalam *iterasi* selanjutnya (di prune).
3. Pada *iterasi* kedua sistem akan menggunakan hasil *largeitemset* pada *iterasi* pertama (L1) untuk membentuk kandidat *itemset* kedua (L2). Pada *iterasi* selanjutnya sistem akan menggunakan hasil *large itemset* pada iterasi selanjutnya akan menggunakan hasil large itemset pada *iterasi* sebelumnya (Lk-1) untuk membentuk kandidat *itemset* berikut (Lk). Sistem akan menggabungkan (join) Lk-1 dengan Lk-1 untuk mendapatkan Lk, seperti pada *iterasi* sebelumnya sistem akan menghapus (prune) kombinasi *itemset* yang tidak termasuk dalam *large itemset*.
4. Setelah dilakukan operasi *join*, maka pasangan *itemset* baru hasil proses *join* tersebut dihitung supportnya.
5. Proses pembentuk kandidat yang terdiri dari proses *join* dan *prune* akan terus dilakukan hingga himpunan kandidat *itemsetnya* null, atau sudah tidak ada lagi kandidat yang akan dibentuk.
6. Setelah itu, dari hasil *frequent itemset* tersebut dibentuk *association rule* yang memenuhi nilai support dan *confidence* yang telah ditentukan.
7. Pada pembentukan *association rule*, nilai yang sama dianggap sebagai satu nilai.
8. *Assosiotion rule* yang terbentuk harus memenuhi nilai minimum yang telah ditentukan.

9. Untuk setiap large *itemset* L, kita cari himpunan bagian L yang tidak kosong. Untuk setiap himpunan bagian tersebut, dihasilkan rule dengan bentuk aB (L-a) jika supportnya (L) dan supportnya (a) lebih besar dari minimum support. itu yang termaksud kedalam langkah-langkahnya.

2.2. Algoritma *FP-Growth*

Algoritma *FP-Growth* merupakan pengembangan dari algoritma *Apriori*. Algoritma *Frequent Pattern Growth* (*FP-Growth*) adalah salah satu alternatif algoritma yang dapat digunakan untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data. *FP-Growth* dapat menemukan frekuensi itemset dengan hanya sedikit mengakses pada database aslinya, dan pendekatannya adalah yang paling efisien. (Wiyana,2018)

Algoritma *FP-Growth* juga dapat menghindari permasalahan jika jumlah calon itemset nya terlalu besar. *FP-Growth* menggunakan konsep pembangunan tree dalam pencarian frequent itemset. Hal tersebut yang menyebabkan algoritma *FP-Growth* lebih cepat dari algoritma Apriori. Dengan menggunakan *FP-Tree*, algoritma *FP-Growth* dapat langsung mengekstrak *frequent itemset* dari *FP-Tree*. Penggalian *itemset* yang *frequent* dengan menggunakan algoritma *FP-Growth* akan dilakukan dengan cara membangkitkan struktur data tree. (Erwin, 2009)

2.2.1. Pembangunan *FP-Tree*

FP-Tree merupakan struktur penyimpanan data yang dimampatkan. *FP-tree* dibangun dengan memetakan setiap data transaksi ke dalam setiap lintasan tertentu dalam *FP-tree*. Karena dalam setiap transaksi yang dipetakan mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data *FP-tree* semakin efektif. Kelebihan dari *FP-tree* adalah hanya memerlukan dua kali pemindaian data transaksi yang terbukti sangat efisien. *FP-tree* adalah sebuah pohon dengan defenisi sebagai berikut.

1. *FP-tree* dibentuk oleh sebuah akar yang diberi label *null*, sekumpulan upapohon yang beranggotakan *item-item* tertentu, dan sebuah table *frequent header*.
2. Setiap simpul dalam *FP-tree* mengandung tiga informasi penting yaitu label item menginformasikan jenis item yang dipresentasikan simpul tersebut, dan pointer penghubung yang menghubungkan simpul-simpul dengan label item sama antar lintasan, ditandai dengan garis panah putus-putus. (Wijaya, 2018)

Tahapan dalam pembentukan FP-tree adalah sebagai berikut : (Samuel, 2008)

1. Tentukan data transaksi dengan minimum *support count*ç=2

Tabel 2.1 Table Data Transaksi Awal

TID	Transaksi
1	a,b
2	b,c,d,g,h
3	a,c,d,e,f
4	a,d,e
5	a,b,z,c
6	a,b,c,d
7	a,r
8	a,b,c
9	a,b,d
10	b,c,e

Frekuensi kemunculan tiap item dapat dilihat pada table berikut :

Tabel 2.2 Frekuensi Kemunculaan Tiap Karakter

Item	Frekuensi
a	8
b	7
c	6
d	5
e	3

f	1
r	1
z	1
g	1
h	1

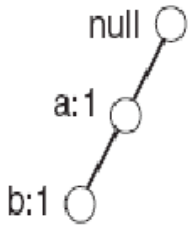
Setelah dilakukan pemindaian pertama didapat *item* yang memiliki frekuensi di atas *support count*ç=2 adalah a,b,c,d dan e. kelima *item* inilah yang akan berpengaruh dan akan dimasukkan kedalan *FP-tree*, selebihnya r,z,g,h dapat dibuang karena tidak berpengaruh signifikan.

Table berikut mendata kemunculan *item* yang *frequent* dalam setiap transaksi, diurut berdasarkan yang frekuensinya paling tinggi.

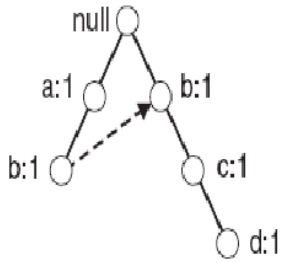
Tabel 2.3 Table Data Transaksi

TID	Transaksi
1	{a,b}
2	{b,c,d}
3	{a,c,d,e}
4	{a,d,e}
5	{a,b,c}
6	{a,b,c,d}
7	{a}
8	{a,b,c}
9	{a,b,d}
10	{b,c,e}

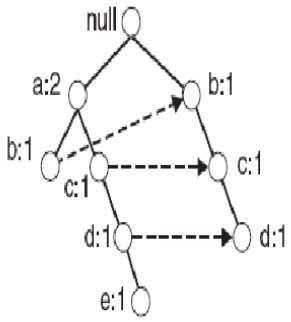
Gambar dibawah ini adalah ilustrasi pembentukan *FP-tree* setelah pembacaan TID 1.



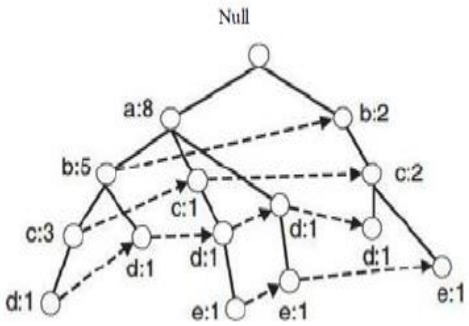
Gambar 2.2 Hasil pembentukan *FP-tree* setelah pembacaan TID 1



Gambar 2.3 Hasil pembentukan *FP-tree* setelah pembacaan TID 2



Gambar 2.4 Hasil pembentukan *FP-tree* setelah pembacaan TID 3



Gambar 2.5 Hasil pembentukan *FP-tree* setelah pembacaan TID 10

Diberikan 10 data transaksi dengan 5 jenis *item* seperti pada tabel di atas. Gambar 1 – 4 menunjukkan proses terbentuknya *FP-tree* setiap TID dibaca. Setiap simpul pada *FP-tree* mengandung nama sebuah item dan *counter support* yang berfungsi untuk menghitung frekuensi kemunculan item tersebut dalam tiap lintasan transaksi. (Arsyad, 2013)

FP-tree yang merepresentasikan data transaksi pada tabel 2.1 dibentuk dengan cara sebagai berikut: (Samuel, 2008)

1. Kumpulan data dipindai pertama kali untuk menentukan *support count* dari setiap item. *Item* yang tidak *frequent* dibuang, sedangkan *frequent item* dimasukkan dan disusun dengan urutan menurun, seperti yang terlihat pada tabel 2.1.
2. Pemindaian kedua, yaitu pembacaan TID pertama {a,b} akan membuat simpul a dan b, sehingga terbentuk lintasan transaksi Null→a→b. *Support count* dari setiap simpul bernilai awal 1
3. Setelah pembacaan transaksi kedua {b,c,d}, terbentuk lintasan kedua yaitu Null→b→c→d. *Support count* masing-masing *count* juga bernilai awal 1. Walaupun b ada pada transaksi pertama, namun karena prefix transaksinya tidak sama, maka transaksi kedua ini tidak bisa dimampatkan dalam satu lintasan.
4. Transaksi keempat memiliki *prefix* transaksi yang sama dengan transaksi pertama, yaitu a, maka lintasan transaksi ketiga dapat ditimpakan di a, sambil menambah *support count* dari a, dan selanjutnya membuat lintasan baru sesuai dengan transaksi ketiga.
5. Proses ini dilanjutkan sampai *FP-tree* berhasil dibangun berdasarkan tabel data transaksi yang diberikan.

Algoritma *FP-Growth* memiliki tahapan-tahapan yang harus dilewati agar dapat memberikan hasil yang maksimal, tahapan-tahapan tersebut, yaitu: (Anggraeni, 2014)

1. Tahap pembangkitan *conditional pattern base*

Pembangkitan *conditional pattern base* merupakan sub database yang berisi *prefix path* dan *suffix pattern*. Pembangkitan *conditinal 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 atau sama dengan minimum *support count* yang akan dibangkitkan dengan *conditional FP-Tree*.

3. Tahap pencarian *frequent itemset*.

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

Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapatkan *frequent itemset*. Setelah memeriksa *frequent itemset* untuk beberapa akhiran (*suffix*), maka didapat hasil yang dirangkum dalam tabel berikut:

Tabel 2.4 Hasil *Frequent Itemset*

Suffix	***Frequent Itemset***
e	{e},{d,e},{a,d,e},{c,e},{a,e}
d	{d},{c,d},{b,c,d},{a,c,d},{b,d},{a,b,d},{a,d}
c	{c},{b,c},{a,b,c},{a,c}
b	{b},{a,b}
a	{a}

Dengan metode *divide and conquer* ini, maka pada setiap langkah rekursif, algoritma *FP-growth* akan membangun sebuah *conditional FP-tree* baru yang telah diperbaharui nilai *support count*, dan membuang lintasan yang mengandung item-item yang tidak *frequent* lagi. (Astrina, 2019)

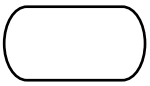
2.3. *Flowchart*

Flowchart menggambarkan urutan logika dari suatu procedure pemecahan masalah, sehingga flowchart merupakan langkah-langkah penyelesaian masalah yang dituliskan dalam simbol-simbol tertentu. (Sitorus, 2015). *Flowchart* biasanya mempermudah penyelesaian suatu masalah khususnya masalah yang perlu dipelajari dan dievaluasi lebih lanjut. *Flowchart* adalah bentuk gambar/diagram yang mempunyai aliran satu atau dua arah secara sekuensial. *Flowchart* digunakan untuk merepresentasikan maupun mendesain program. Oleh karena itu *flowchart* harus bisa merepresentasikan komponen-komponen dalam bahasa pemrograman.

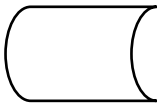
Pedoman dalam menggambar suatu Flowchart atau bagan alir, analissistem atau pemrograman sebagai berikut. (Lestari, 2015)

1. Bagan alir sebaiknya digambar dari atas ke bawah dan mulai dari bagian kiri dari suatu halaman.
2. Kegiatan didalam bagan alir harus ditunjukan dengan jelas.
3. Harus ditunjukan darimana kegiatan akan dimulai dan dimana akan berakhirnya.
4. Masing-masing kegiatan didalam bagan alir sebaiknya digunakan suatu kata yang mewakili suatu pekerjaan, misalnya; “persiapkan” dokumen “hitung” gaji.
5. Masing-masing kegiatan didalam bagan alir harus didalm urutan yang semestinya.
6. Kegiatan yang terpotong dan akan disambung ketempat lain harus ditunjukan dengan jelas menggunakan symbol penghubung.
7. Gunakanlah simbol-simbol bagan alir yang standar.

Tabel 2.5 Simbol Flowchart

No	Simbol	Keterangan
1	 Terminator (Simbol Terminal)	Permulaan atau akhir program.

2	Garis Alur (*Line Follower*)	Arah Aliran Program.
3	Simbol Persiapan	Proses instalasi atau pemberian nilai awal.
4	Simbol Proses	Proses Penghitungan atau proses pengolahan data.
5	Input/Output	Proses Input/Output data, parameter, informasi.
6	Simbol Sub Proses	Permulaan sub program/ proses menjalankan sub program.
7	Simbol Keputusan	Perbandingan pernyataan, penyeleksian data yang memberikan pilihan untuk langkah sebelumnya.
8	*Connector*	Penghubung bagain flowchart yang berada pada satu halaman.
9	Simbol Penghubung	Penghubung bagian flowchart yang berada pada halaman yang berbeda.

10	 *Document*	Input/Output dalam format yang dicetak.
11	 *Disk Storage*	Input/Outputyang menggunakan penyimpanan akses langsung.
12	 *Punched card*	Menyatakan input berasal dari kartu atau output ditulis ke kartu.

2.4. *RapidMiner*

RapidMiner merupakan perangakat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *RapidMiner* memiliki kurang lebih 500 operator *data mining*, termasuk operator untuk *input*, *output*, *data preprocessing* dan visualisasi. *RapidMiner* merupakan *software* yang berdiri sendiri untuk analisis data dan sebagai mesin *data mining* yang dapat diintegrasikan pada produknya sendiri. (Bruno, 2019)

RapidMiner merupakan perangkat lunak yang dibuat oleh Dr. Markus Hofmann dari *Institute of Technologi Blanchardstown* dan *Ralf Klinkenberg* dari *rapid-i.com* dengan tampilan GUI (*Graphical User Interface*) sehingga memudahkan pengguna dalam menggunakan perangkat lunak ini. Perangkat lunak ini bersifat open source dan dibuat dengan menggunakan program Java di bawah lisensi GNU *Public Licence* dan *RapidMiner* dapat dijalankan di sistem operasi manapun. (Haryati et al., 2015)

Aplikasi ini digunakan untuk aplikasi bisnis dan komersial serta untuk penelitian, pendidikan, pelatihan, pembuatan *prototype* dengan cepat, dan pengembangan aplikasi serta mendukung semua langkah proses pembelajaran

mesin termasuk persiapan data, visualisasi hasil, validasi dan pengoptimalan. (Nofitri, 2019)

Dengan menggunakan *RapidMiner*, tidak dibutuhkan kemampuan koding khusus, karena semua fasilits sudah disediakan. *RapidMiner* dikhususkan untuk penggunaan data mining. Model yang disediakan juga cukup banyak dan lengkap, seperti Model *Bayesian, Modelling, Tree Induction, Neural Network* dan lain-lain. Banyak metode yang disediakan oleh *RapidMiner* mulai dari kllasifikasi, klustering, asosiassi dan lain-lain. (Haryati et al., 2015)

2.5. Kajian Terdahulu

Beberapa penelitian tentang data mining terdahulu yang telah berhasil diuji dengan menggunakan algoritma *Fp-Growth* diantaranya sebagai berikut :

a. Penelitian yang dilakukan oleh Icca Astrina, Muhammad Zainal Arifin, dan Utomo Pujianto pada tahun 2019 dengan judul penerapan algoritma *fp-growth* dalam penentuan pola pembelian konsumen pada kain tenun medali mas. Hasil eksperimen menunjukan indikasi yang sesuai dengan dugaan awal penelitian bahwa ada pola-pola kaidah asosiasi yang kuat dalam pembelian konsumen kain tenun Medali Mas.
b. Selanjutnya penelitian oleh Sigit Kurniawan, Windu Gata, dan Hari Wiyana dengan judul analisis algoritma *fp-growth* untuk rekomendasi produk pada data retail penjualan produk kosmetik .Dalam penelitiannya pada objek MT Shop Kelapa Gading, pengolahan data transaksi penjualan selama satu tahun pada toko tersebut yang dibentuk dalam *itemset* dan diolah menggunakan Rapidminer 8.0 didapatkan nilai *confidence* tertinggi adalah 0,899 atau sekitar 89% dengan aturan (*rule*) pada setiap pembelian produk masker beras putih dapat dipastikan akan membeli putih langsat facial foam.
c. Selanjutnya penelitian oleh Ali Ikhwan, Dicky Nofriansyah, dan Sriani dengan judul penerapan data mining dengan algoritma *fp-growth* untuk mendukung strategi promosi pendidikan. Dalam penelitian tersebut algoritma *fp-growth* dapat diterapkan untuk mendukung strategi promosi pendidikan pada Perguruan Tinggi. Adapun Informasi yang berkaitan dalam pemilihan

tempat yang akan dijadikan promosi pendidikan, sehingga pihak manajemen dari kampus sendiri dapat melakukan pengambilan keputusan dengan cepat.

d. Selanjutnya penelitian oleh Rizky Nurul Arifin dengan judul implementasi algoritma *frequent pattern growth* menentukan asosiasi antar produk dengan hasil algoritma *fp-growth* dapat diterapkan pada data transaksi untuk menentukan promosi di Minimarket Nadiamart dengan aturan asosiasi yang dihasilkan adalah hasil dari pengolahan 2020 data transaksi melalui aplikasi Market Basket Analysis dengan batasan minimum nilai *support* sebesar 7% dan *confidence* sebesar 30%, terdapat 1 pola asosiasi yang memenuhi syarat, dua pola asosiasi yaitu jika membeli snack maka membeli susu instant dengan nilai support = 8.01% dan nilai confidence = 33. 89% yang merupakan pola dengan nilai *support* dan *confidence* tertinggi.

e. Selanjutnya penelitian oleh Adi Nugroho Susanto Putro dan Richardus Indra Gunawan dengan judul Implementasi Algoritma *FP-Growth* untuk Strategi Pemasaran Ritel Hidroponik. Dalam penelitiannya menghasilkan kesimpulan bahwa Algoritma *FP-Growth* untuk strategi pemasaran ritel hidroponik telah berhasil diimplementasikan, dengan menggunakan *minimum support* 0,05 dan *minimum confidence* 0,9 menghasilkan 21 rule yang dapat digunakan sebagai strategi pemasaran PT. HAB.

f. Selanjutnya penelitian oleh Dessy Chaerunnissa1 dan Edy Mulyanto dengan judul metode association rule dalam menganalisa pola belanja konsumen pada data transaksi penjualan menggunakan algoritma *fp-growth*. Dari hasil analisa pola frekuensi yang dihasilkan dapat dijadikan sebuah rekomendasi dalam menentukan keputusan memberikan paket pembelian atau *bundling* dan bisa juga digunakan untuk tata letak *layout* atau promosi produk untuk meningkatkan penjualan di Swalayan Gelael Candi Semarang.

g. Selanjutnya penelitian oleh Dita Dwi Sapitri dan Hindriyanto Dwi Purnomo dengan judul penerapan algoritma *frequent pattern growth* untuk menganalisis pola pembelian konsumen. Dalam penelitiannya menghasilkan kesimpulan bahwa ditemukan sebanyak 6 hubungan dan keterkaitan antar barang yang dibeli oleh konsumen dengan nilai support lebih dari sama

dengan 0.01. Dari rule yang diperoleh juga didapatkan sebuah pengetahuan baru tentang pola pembelian konsumen, dari pengetahuan tersebut bisa di gunakan untuk mengatur tempat item yang saling berhubungan agar dapat memudahkan konsumen dalam membeli, bisa juga digunakan untuk menentukan keputusan persediaan barang.

BAB III
METODOLOGI PENELITIAN

3.1. Tempat dan Waktu

3.1.1. Tempat Penelitian

Dalam menyelesaikan penelitian ini tentunya membutuhkan tempat dan waktu penelitian yang mendukung penyelesaian, Penelitian ini dilakukan di PT. Cipta Niaga Alam Semesta (Mayora Group) Kota Pematangsiantar.

3.1.2. Waktu Dan Jadwal Pelaksanaan Penelitian

Waktu dan jadwal yang digunakan untuk penelitian penerapan data mining menggunakan algoritma *FP-growth* untuk menentukan pola pembelian produk makanan ini ditujukan pada Tabel 3.1

Tabel 3.1 Waktu dan Jadwal Penelitian

No	Waktu	Jadwal Penelitian				
		Desember	Januari	Februari	Maret	April
1	Perencanaan					
2	Pengumpulan Data					
3	Analisis Data Dan Perancangan Sistem					
4	Pengujian Sistem					

3.2. Bahan dan Alat Penelitian

Dalam pembuatan skripsi ini, dibutuhkan alat dan bahan yang digunakan dalam penelitian tugas akhir ini.

3.2.1. *Hardware* (Perangkat Keras)

Hardware (Perangkat Keras) yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 3.2 Perangkat Keras

No	Nama Perangkat Keras
1	*Processor core i2 Core 2,58 Ghz*
2	Kapasitas *Harddisk* 500GB
3	Besar *Memory* RAM 2GB

3.2.2. *Software* (Perangkat Lunak)

Software (Perangkat Lunak) yang digunakan dalam penelitian ini adalah sebagai berikut.

Tabel 3.3 Perangkat Lunak

No	Nama Perangkat Lunak
1	*Windows* 10 64 bit
2	*Microsoft Office Excel*
3	*RapidMiner*

3.3. Teknik Pengumpulan Data

Pada tahap pengumpulan data ini, akan ada dua langkah yang akan dilakukan yaitu :

1. Studi Observasi

   Studi observasi yaitu teknik atau pendekatan untuk mendapatkan data primer dengan cara mengamati langsung objek datanya. Pendekatan dilakukan dengan mengamati secara langsung kejadian pada sistem penjualan barang di PT. Cipta Niaga Semesta (Mayora Group).

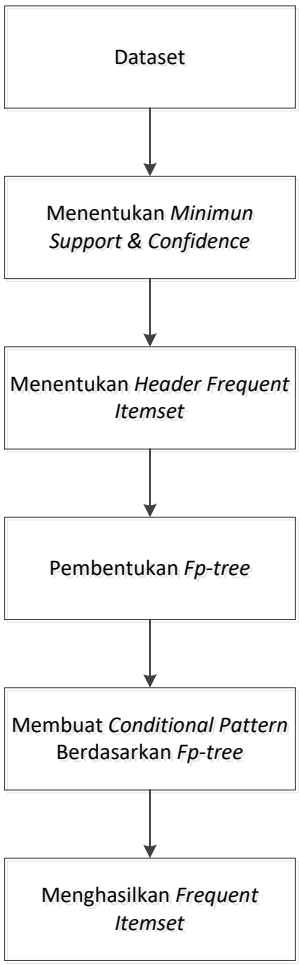
2. Studi Literatur

   Studi literatur yaitu studi yang dilakukan peneliti dalam mencari teori-teori yang bersangkutan dengan suatu penelitian yang dilakukan. Seperti teori-teori

data mining, teori metode-metode , ataupun teori-teori yang bersangkutan lainnya yang mana bisa berupa jurnal, majala ataupun buku.

3.4. Penerapan Algoritma

Pada penelitian ini akan dibahas bagaimana pencarian *frequent itemset* menggunakan algoritma *FP-Growth*. *FP-Tree* digunakan bersamaan dengan algoritma *FP-Growth* untuk menentukan *frequent itemset* (data yang paling sering muncul) dari sebuah *dataset*. Untuk menentukan *frequent itemset* pada data transaksi tersebut, dapat digambarkan dalam blok diagram sebagai berikut:



Gambar 3.1 Blok Diagram Algoritma *FP-Growth*

1. Dataset merupakan objek yang merepresentasikan data dan relasinya di *memory*, dataset berisi koleksi dari *datatable* dan *datarelation*.

2. Tentukan nilai *Minimum Support* yang diinginkan. *Minimum Support* merupakan ambang batas minimum jumlah *itemset* yang diperbolehkan, jika jumlah *item*-nya di bawah ambang batas maka *item* tersebut akan dieliminasi.
3. Menetapkan *frequent itemset* (kumpulan *item* yang muncul secara bersamaan), dengan cara mengambil *itemset* yang memiliki frekuensi *itemset* minimal sebesar *minimum support* sebelumnya.
4. *FP-tree* dibentuk oleh sebuah akar yang diberi label *null*, sekumpulan pohon yang beranggotakan item-item tertentu dan sebuah tabel *frequent header*.
5. *Conditional pattern,* pada tahap ini *support count* dari setiap item pada *conditional pattern base* dijumlahkan, lalu setiap item yang memiliki jumlah *support count* lebih besar akan dibangkitkan dengan *conditional FP-tree.*
6. Menentukan *frequent itemset,* dihasilkan dari aturan asosiasi yang memenuhi nilai *minimum support*.

3.5. Pemodelan

Pada penelitian ini akan dilakukan pemodelan menggunakan algoritma *Fp-Growth* untuk pengolahan data barang. Teknik data mining yang dipilih dalam penelitian ini adalah teknik asosiasi. Pemodelan bertujuan untuk mencarai aturan asosiasi, dimana aturan asosiasi selanjutnya dijadikan acuan untuk menentukan stok barang. Adapun langkah-langkah pembentukan model data mining dengan algoritma *Fp-Growth* adalah:

1. Menentukan data yang akan diproses. Data yang di uji adalah data transaksi penjualan produk makanan di PT. Cipta Niaga Semesta (Mayora Group) Pematangsiantar pada bulan Januari - Desember 2019.
2. Menentukan *minimum support* dan *minimum confidence*.

3.6. Pengujian

Pengujian dilakukan untuk mengetahui hasil perhitungan yang dianalisa dan untuk mengetahui apakah fungsi bekerja dengan baik atu tidak. Setelah data dihitung secara manual, kemudian data diuji menggunakan *tools RapidMiner* untuk menghasilkan apakah hasil perhitungan manual dengan hasil yang diperoleh *Rapidminer* sama atau tidak.

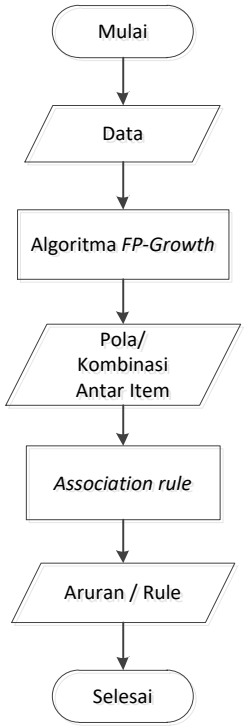
BAB IV
HASIL DAN PEMBAHASAN

4.1. Pembahasan

Dalam pambahasan ini, terdapat beberapa pembahasan yang harus selelesaikan. Adapun tahapan yang akan dibahas pada penelitian ini yaitu analisis data, representasi data, hasil analisis data , dan perancangan sebagai berikut.

4.1.1. Analisis Data

Dalam tahap ini analisis data dilakukan khusus pada penjualan produk makanan ringan di PT. Cipta Niaga Semesta (Mayora Group) Pematangsiantar dengan tujuan untuk menemukan pola penjualan produk makanan yang saling berhubungan didalam suatu transaksi. Tahapan awal yang dilakukan pada penelitian ini adalah mempersiapkan data, data yang akan diolah yaitu data penjualan dari PT. Cipta Niaga Semesta (Mayora Group) dari bulan januari sampai dngan desember 2019. Untuk mendapat analisa data maka data penjualan di *export* kedalam database *Microsoft excel* karena database *Microsoft excel* bersifat *spreadsheet* sehingga sangat mendukung dalam analisa data. *Microsoft excel* dalam analisis data digunakan sebagai database karena *Microsoft excel* sangat mendukung beberapa aplikasi data mining, dimana aplikasi (perangkat lunak) data mining digunakan sebagai tempat uji coba atau implementasi. Data mining ialah proses menambang data untuk mendapatkan ilmu pengetahuan atau informasi penting dari database khususnya database penjualan produk-produk makanan ringan.



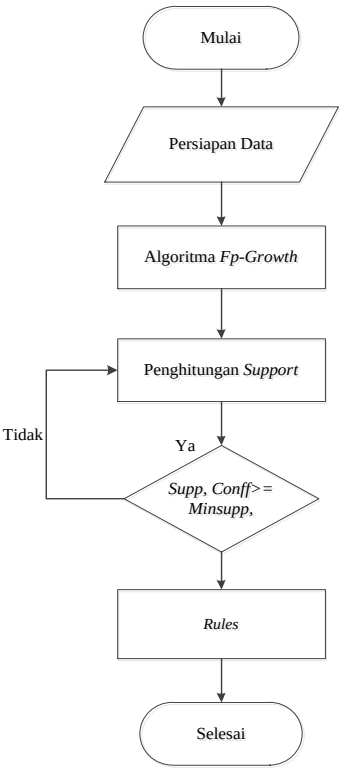
Gambar 4.1 *Flowchart* Pengolahan Data Mining

4.1.2. Representasi Data

Karakteristik algoritma *FP-Growth* adalah struktur data yang digunakan adalah *tree* yang disebut dengan *FP-Tree*. Dengan menggunakan *FP-Tree*, algoritma *FP-growth* dapat langsung mengekstrak *frequent Itemset* dari *FP-Tree.* Penggalian *itemset* yang *frequent* dengan menggunakan algoritma *FP-Growth* akan dilakukan dengan cara membangkitkan struktur data *tree* atau disebut dengan *FP-Tree*. *FP-Growth* dapat dibagi menjadi 3 tahapan utama yaitu sebagai berikut :

1. Tahap pembangkitan *conditional pattern base*
2. Tahap pembangkitan *conditional FP-Tree*, dan
3. Tahap pencarian *frequent itemset*.

Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*.



Gambar 4.2 *Flowchart* Algoritma *Fp-Growth*

Berikut ini ialah langkah-langkah perhitungan manual algoritma *Fp-Growth*. Dalam proses implementasi algoritma *Fp-Growth* terdapat beberapa tahap yang harus dilakukan. Tabel dibawah ini adalah perhitungan manual untuk mencari *frequent itemset* yang terdapat 20 sample data transaksi.

Tabel 4.1 Untuk Memberi Tanda Pada Setiap *Itemset*

ITEM
Beng-beng wafer cokelat (A)
Better Vanila Special (B)
Beng Beng Share It (C)
Better Vanila Funbites (D)
Le Minerale 600 ML (E)
Teh Pucuk 350 ML (F)
Energen All Variant (G)
Kopiko Coffee Shot Cappucino (H)
Toracafe Milky Latte (I)

Choki Choki (J)
Mie Gelas Cup Ayam Bawang (K)
Mie Gelas Kari Ayam (L)
Roma Malkist Abon (M)
Roma Malkist Crackers (N)
Roma Malkist Choco (O)
Astor Double Cokelat (P)
Astor Matcha (Q)
Super Star Chocholate Wafer (R)
Kis Cherry (S)
Thamarin Zak New (T)

Tabel 4.2 Data Transaksi

TID	Item
1	{ A,B,D,F,G,H,I,K,N,P,R,S }
2	{ A,C,D,E,F,I,M,O,Q,R }
3	{ A,C,D,E,F,G,J,K,M,N,O,Q,S }
4	{ A,B,C,D,E,F,H,I,K,N,P,Q,S }
5	{ A,B,C,E,H,I,J,M,N,P,R }
6	{ B,C,G,H,I,K,M,O,P,Q,R }
7	{ C,D,J,M,O,P,R }
8	{ A,B,D,G,H,I,L,P,S }
9	{ A,C,D,E,,G,J,Q,R,S }
10	{ B,D,F,H,I,K,M,N,O,S,T }
11	{ A,B,C,E,N,Q,S }
12	{ A,B,C,G,H,I,L,O,Q,T }
13	{ A,B,C,E,G,K,L,R,T }
14	{ B,D,F,J,M,N,O,Q }
15	{ A,B,C,D,E,G,L,Q,T }
16	{ E,M,N,O,S }
17	{ A,D,E,F,H,I,M }

18	{ C,D,L,P,T }
19	{ C,E,L,N,P,Q }
20	{ B,D,G,J,M,N,T }
21	{ D,H,I,L,P,R,S }
22	{ A,C,G,K,N,R }
23	{ A,B,C,H,L,N,Q,R,S }
24	{ A,B,C,D,K,M }
25	{ A,C,D,E,H,J,P }
26	{ A,B,C,D,G,J,P,T }
27	{ A,B,D,F,P,T }
28	{ A,B,P,S }
29	{ A,C,D,F,G,N }
30	{ A,B,H,I }

Setelah data sample di siapkan, maka data tersebut akan di ubah ke dalam bentuk format binominal yang dapat dilihat pada tabel dibawah ini :

Tabel 4.3 Transaksi Penjualan Dalam Format Binominal

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	1	0	1	0	1	1	1	1	0	0	1	0	1	0	1	1	1	1	1
1	0	1	1	1	1	0	0	1	0	0	0	1	0	1	0	1	1	0	0
1	0	1	1	1	1	1	0	1	0	1	0	1	1	1	0	1	0	1	0
1	1	1	1	1	1	0	1	1	0	1	0	0	1	0	1	1	0	1	0
1	1	1	0	1	0	0	1	1	1	0	0	1	1	0	1	0	1	0	0
0	1	1	0	0	0	1	1	1	0	1	0	1	0	1	1	1	1	0	0
0	0	1	1	0	0	0	0	0	1	0	0	1	0	1	1	0	1	0	0
1	1	0	1	0	0	1	1	1	0	0	1	0	0	0	1	0	0	1	0
1	0	1	1	1	0	1	0	0	1	0	0	0	0	0	0	1	1	1	0
0	1	0	1	0	1	0	1	1	0	1	0	1	1	1	0	0	0	1	1
1	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0
1	1	1	0	0	0	1	1	1	0	0	0	1	0	1	0	1	0	0	1
1	1	1	0	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1

0	1	0	1	0	1	0	0	0	1	0	0	1	1	1	0	1	0	0	0
1	1	1	1	1	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1
0	0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0
1	0	1	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0
0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1
0	0	1	0	1	0	0	0	0	0	0	1	0	1	0	1	1	0	0	0
0	1	0	1	0	0	1	0	0	1	0	0	1	1	0	0	0	0	0	1
0	0	0	1	0	0	0	1	1	0	0	1	0	0	0	1	0	1	1	0
1	0	1	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	1
1	1	1	0	0	0	0	1	0	0	0	1	0	1	0	0	1	1	1	0
1	1	1	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
1	0	1	1	1	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0
1	1	1	1	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1
1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
1	0	1	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0
1	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0

Data itemset pada tabel di atas menjelaskan item produk yang dijadikan atribut untuk di eksekusi, setiap produk yang terjual pada setiap transaksi akan ditandai dengan angka 1 atau bernilai *true* sedangkan untuk produk yang tidak terjual akan ditandai dengan angka 0 atau bernilai *false*.

Tabel 4.4 Frekuensi Tiap Pembeli Melakukan Transaksi

TID	Total Transaksi
1	14
2	10
3	13
4	13
5	11
6	11

7	7
8	9
9	9
10	11
11	7
12	10
13	9
14	8
15	9
16	5
17	7
18	5
19	5
20	7
21	6
22	7
23	9
24	6
25	7
26	7
27	6
28	4
29	6
30	4

Tabel 4.5 Frekuensi Kemunculan Tiap Item/Produk

Item	Produk	Frekuensi
A	Beng-beng wafer cokelat	21
B	Better Vanila Special	18
C	Beng Beng Share It	20

D	Better Vanila Funbites	19
E	Le Minerale 600 ML	12
F	Teh Pucuk 350 ML	7
G	Energen All Variant	12
H	Kopiko Coffee Shot Cappucino	12
I	Toracafe Milky Latte	11
J	Choki Choki	7
K	Mie Gelas Cup Ayam Bawang	7
L	Mie Gelas Kari Ayam	8
M	Roma Malkist Abon	11
N	Roma Malkist Crackers	13
O	Roma Malkist Choco	8
P	Astor Double Cokelat	13
Q	Astor Matcha	10
R	Super Star Chocholate Wafer	10
S	Kis Cherry	11
T	Thamarin Zak New	10

Setelah frekuensi tiap item muncul lalu produk di urutkan berdasarkan frekuensi yang paling terbesar ke terkecil, dan dapat di lihat pada tabel di bawah ini :

Tabel 4.6 Tabel Frekuensi dan Nilai *Support*

Produk	**Frekuensi**	**Nilai *Support***
Beng-beng wafer cokelat (A)	21	105 %
Beng Beng Share It (C)	20	100 %
Better Vanila Funbites (D)	19	95 %
Better Vanila Special (B)	18	90 %
Roma Malkist Crackers (N)	13	65 %
Astor Double Cokelat (P)	13	65 %
Le Minerale 600 ML (E)	12	60 %
Energen All Variant (G)	12	60 %

Kopiko Coffee Shot Cappucino (H)	12	60 %
Toracafe Milky Latte (I)	11	55 %
Roma Malkist Abon (M)	11	55 %
Kis Cherry (S)	11	55 %
Astor Matcha (Q)	10	50 %
Super Star Chocholate Wafer (R)	10	50 %
Thamarin Zak New (T)	8	40 %
Roma Malkist Choco (O)	8	40 %
Mie Gelas Kari Ayam (L)	8	40 %
Teh Pucuk 350 ML (F)	7	35 %
Choki Choki (J)	7	35 %
Mie Gelas Cup Ayam Bawang (K)	7	35 %

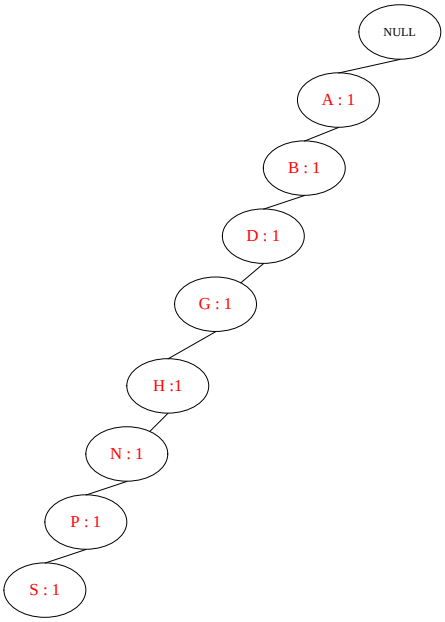
Setelah perhitungan frekuensi kemunculan setiap item diketahui barang yang berada di atas nilai support count = 55% atau item dengan frekuensi di atas 11 adalah A,C,D,B,N,P,E,G,H,I,M, dan S. Ke-12 barang inilah yang akan berpengaruh dan akan dimasukkan ke dalam *fp-tree*, selebihnya (Q,R,T,O,L,F,J, dan K) tidak digunakan karena tidak berpengaruh signifikan seperti ditunjukkan pada Tabel 4.7.

Tabel 4.7 Transaksi yang disesuaikan dengan *frequent list*

TID	Item
1	{ A,B,D,G,H,I,N,P,S }
2	{ A,C,D,E,I,M }
3	{ A,C,D,E,G,M,N,S }
4	{ A,B,C,D,E,H,I,N,P,S }
5	{ A,B,C,E,H,I,M,N,P }
6	{ B,C,G,H,I,M,P }
7	{ C,D,M,P }
8	{ A,B,D,G,H,I,P,S }
9	{ A,C,D,E,G,S }
10	{ B,D,H,I,M,N,S }

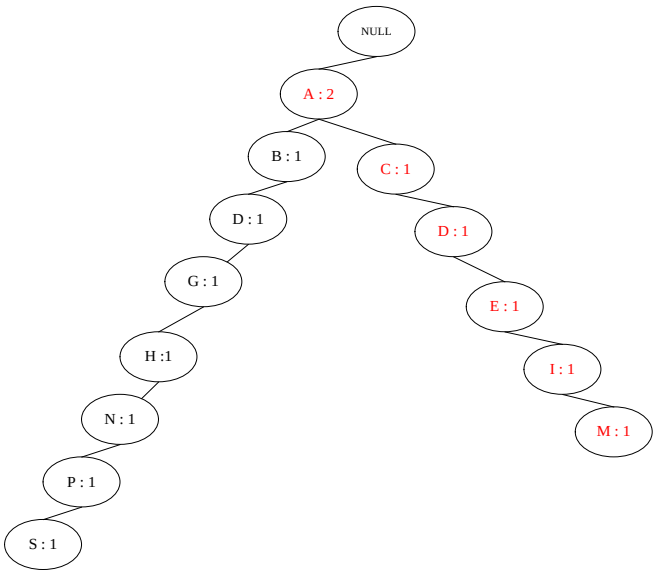
11	{ A,B,C,E,N,S }
12	{ A,B,C,G,H,I }
13	{ A,B,C,E,G }
14	{ B,D,M,N }
15	{ A,B,C,D,E,G }
16	{ E,M,N,S }
17	{ A,D,E,H,I,M }
18	{ C,D,P }
19	{ C,E,N,P }
20	{ B,D,G,M,N }
21	{ D,H,I,P,S }
22	{ A,C,G,N }
23	{ A,B,C,H,N,S }
24	{ A,B,C,D,M }
25	{ A,C,D,E,H,P }
26	{ A,B,C,D,G,P }
27	{ A,B,D,P }
28	{ A,B,P,S }
29	{ A,C,D,G,N }
30	{ A,B,H,I }

Maka langkah selanjutnya adalah membentuk tree yang merupakan langkah selanjutnya dalam proses algoritma *fp-growth* berdasarkan tabel 4.7. Pembuatan *fp-tree* yang diawali dari TID 1 yaitu { A,B,D,G,H,I,N,P,S }.

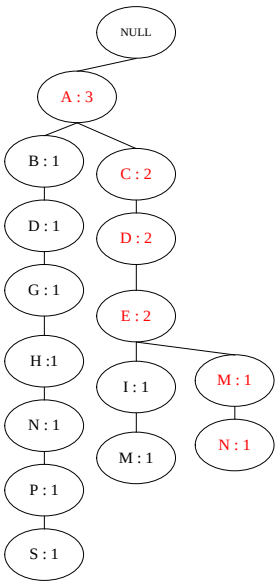


Gambar 4.3 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 1

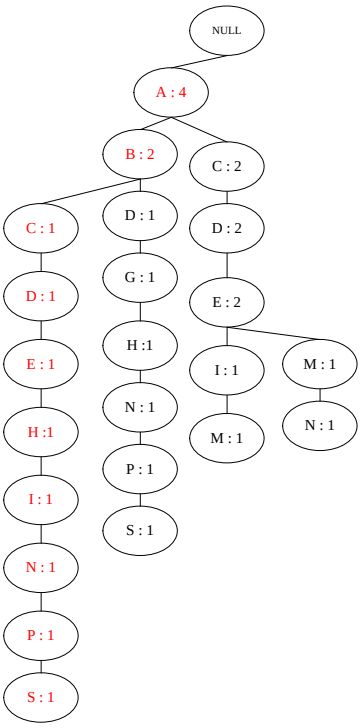
Setelah pembacaan TID 1, selanjutnya dilakukan TID 2 yaitu {A,C,D,E,I,M}. Karena item A,B,D,G,H,I,N,P,S ada pada lintasan sebelumnya yaitu TID 1, maka lintasan item {A,C,D,E,I,M} dapat menggunakan lintasan yang sama. Namun untuk lintasan yang dilintasi untuk kedua kalinya, maka akan diberi keterangan nomor seperti A : 2 yang artinya item A sudah dilewati sebanyak 2 kali.



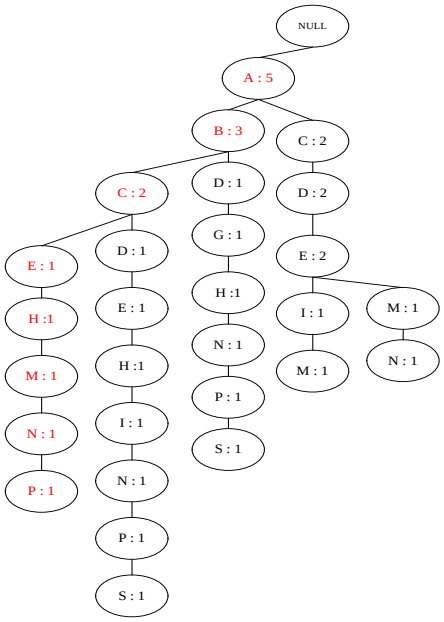
Gambar 4.4 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 2



Gambar 4.5 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 3

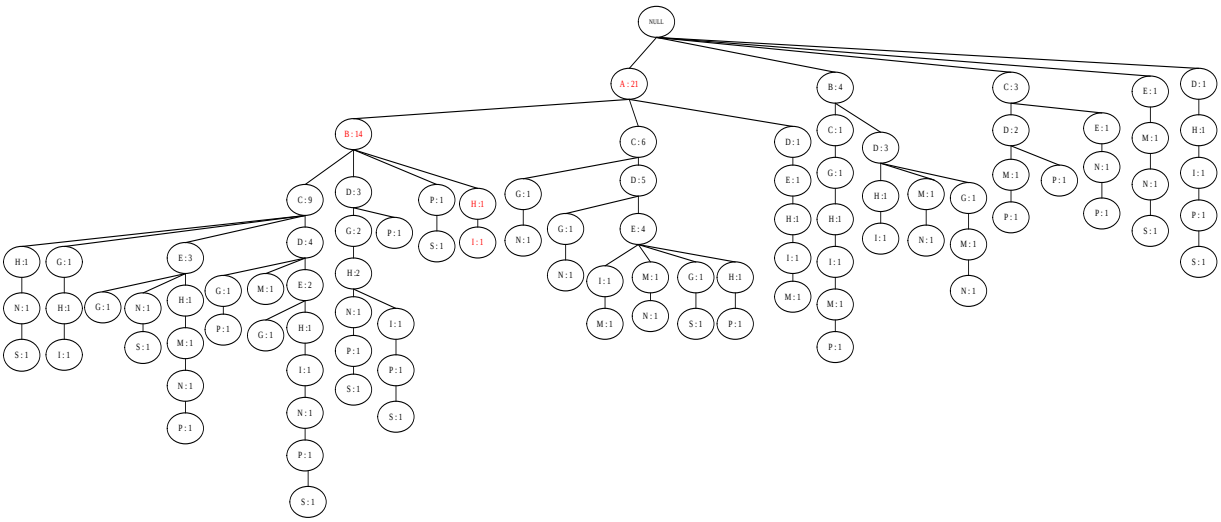


Gambar 4.6 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 4



Gambar 4.7 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 5

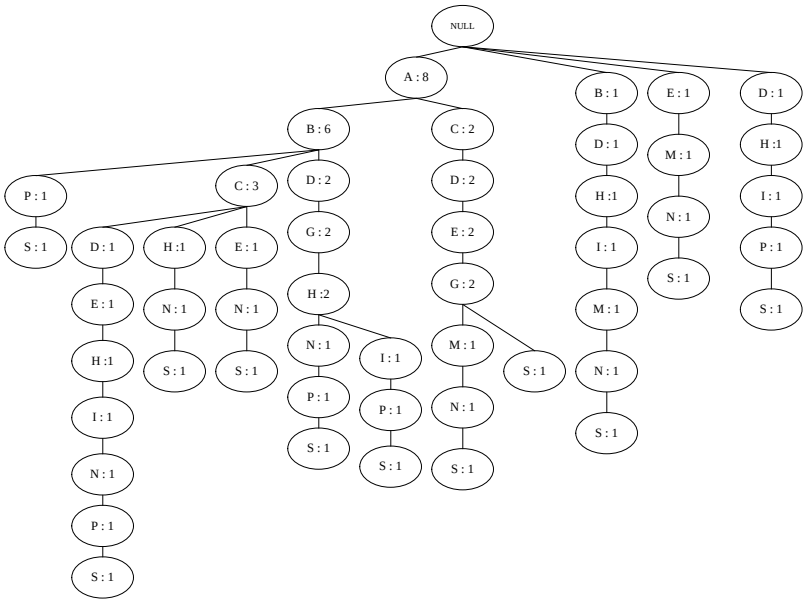
Setelah dilakukan pembacaan TID 1, TID 2, TID 3, TID 4, dan TID 5 pembacaan dilakukan dengan cara yang sama sampai dengan pembacaan TID 30 seperti Gambar dibawah ini :



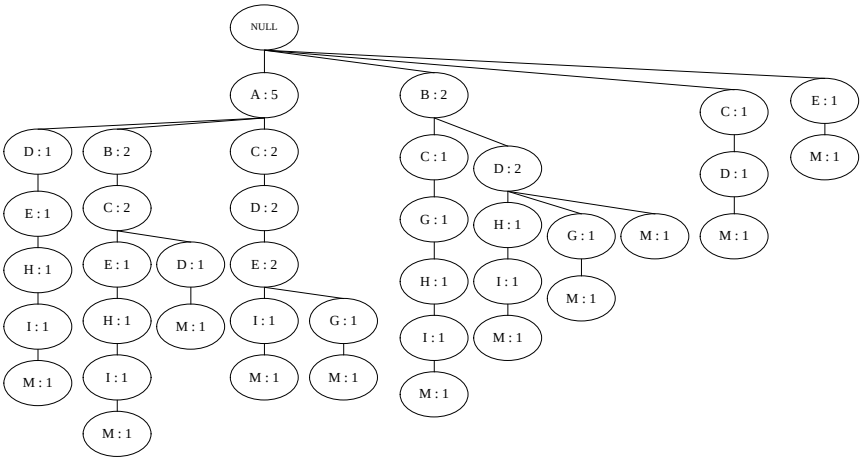
Gambar 4.8 Hasil Pembentukan *FP-tree* Setelah Pembacaan TID 30

Jika pembacaan dari TID 1 sampai TID 30 sudah dilakukan, maka langkah selanjutnya ialah mencari lintasan yang berakhiran dengan *support count* (S), (M),

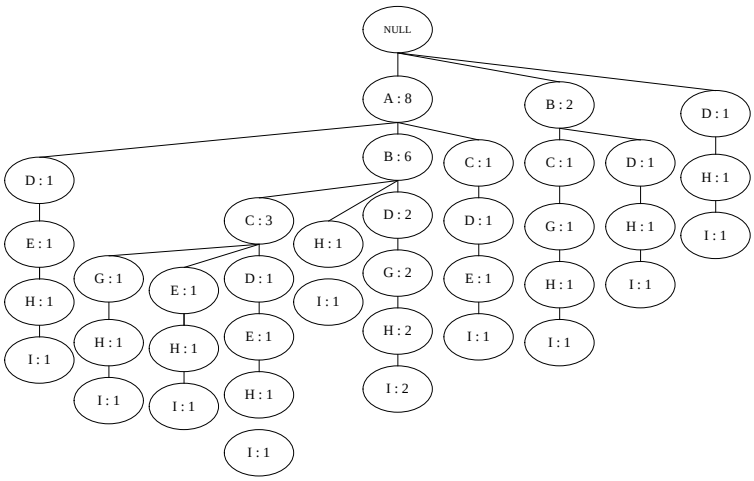
(I), ((H), (G), (E), (P), (N), (B), (D), dan (C). Contoh proses pembentukan masing-masing node dapat dilihat pada Gambar 4.9 sampai Gambar 4.19.



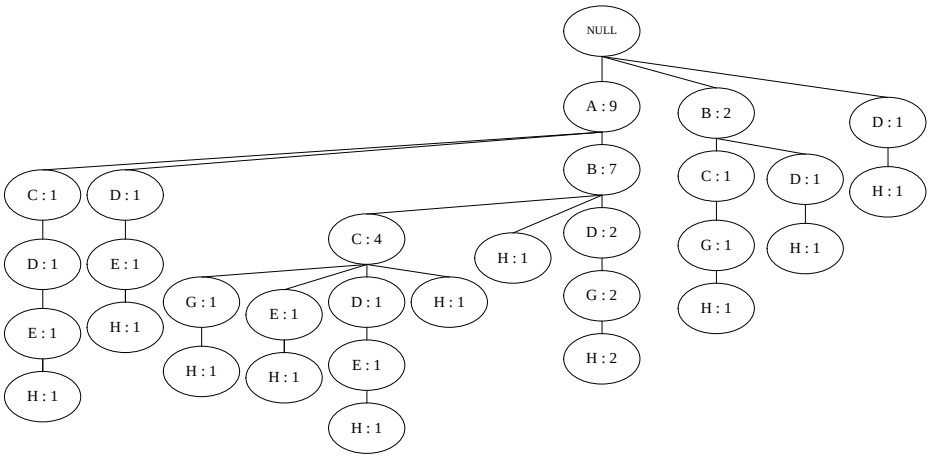
Gambar 4.9 Lintasan yang mengandung simpul S



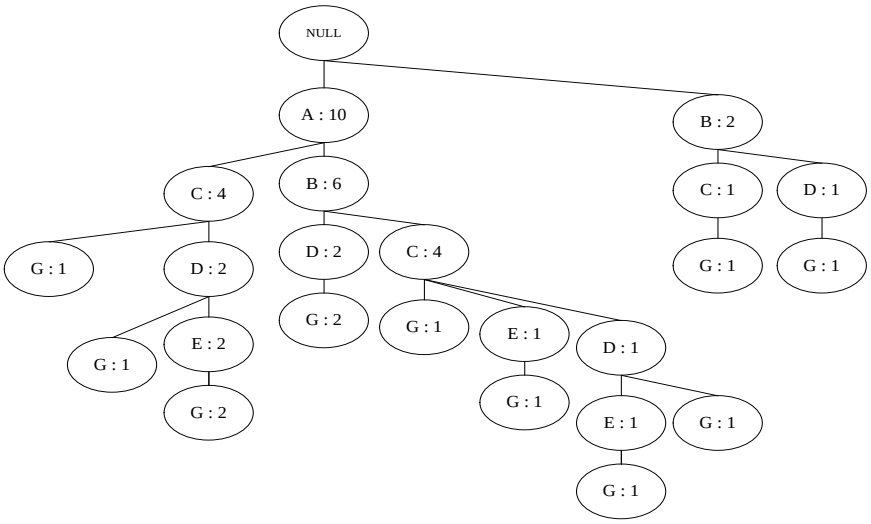
Gambar 4.10 Lintasan yang mengandung simpul M



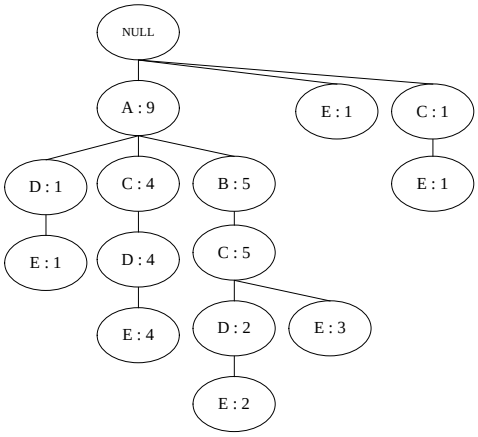
Gambar 4.11 Lintasan yang mengandung simpul I



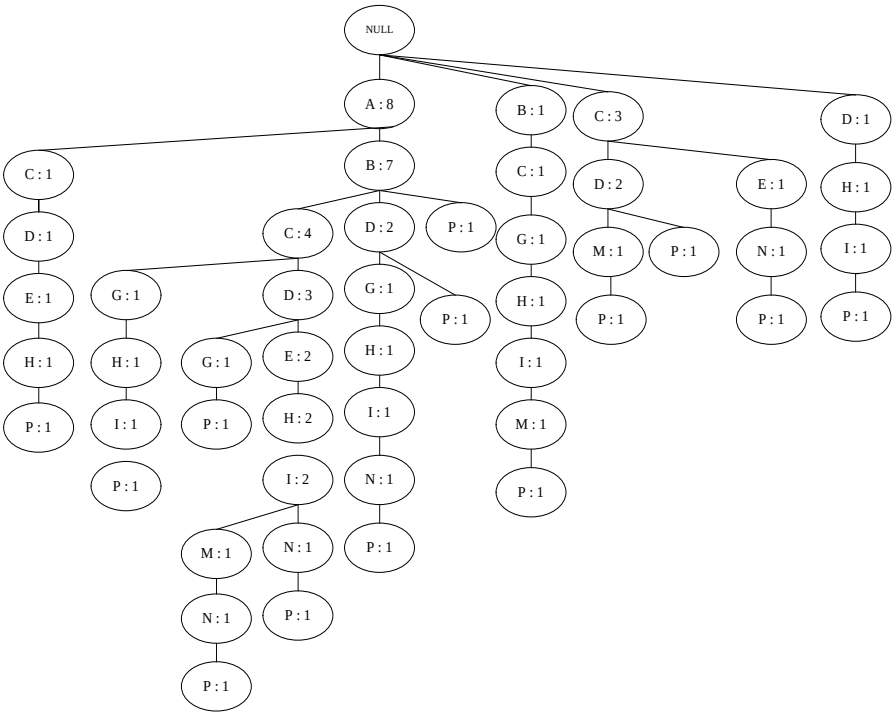
Gambar 4.12 Lintasan yang mengandung simpul H



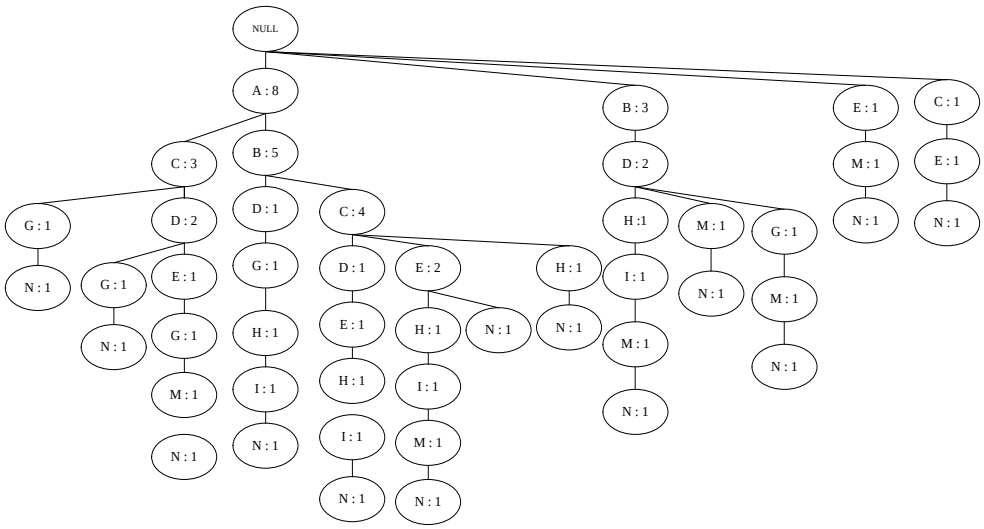
Gambar 4.13 Lintasan yang mengandung simpul G



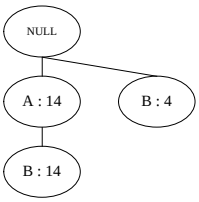
Gambar 4.14 Lintasan yang mengandung simpul E



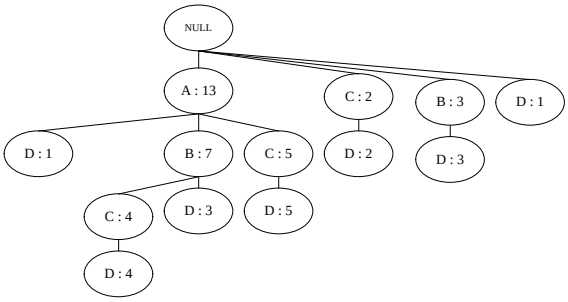
Gambar 4.15 Lintasan yang mengandung simpul P



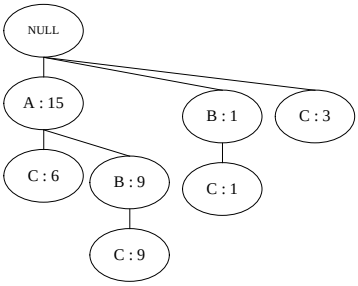
Gambar 4.16 Lintasan yang mengandung simpul N



Gambar 4.17 Lintasan yang mengandung simpul B



Gambar 4.18 Lintasan yang mengandung simpul D



Gambar 4.19 Lintasan yang mengandung simpul C

Setelah memeriksa *frequent itemset* untuk beberapa akhiran (*suffix*), hasilnya dirangkum pada Tabel 4.8.

Tabel 4.8 Hasil *Frequent Itemset*

***Suffix* (Akhiran)**	***Frequent Itemset***
S	{A,B,D,G,H,I,N,P,S}, {A,C,D,E,G,M,N,S}, {A,B,C,D,E,H,I,N,P,S} {A,B,D,G,H,I,P,S}, {A,C,D,E,G,S}, {B,D,H,I,M,N,S}, {A,B,C,E,N,S}, {E,M,N,S}, {D,H,I,P,S}, {A,B,C,H,N,S}, {A,B,P,S}
M	{A,C,D,E,I,M}, {A,C,D,E,G,M}, {A,B,C,E,H,I,M}, {B,C,G,H,I,M}, {C,D,M}, {B,D,H,I,M}, {B,D,M}, {E,M}, {A,D,E,H,I,M}, {B,D,G,M}, {A,B,C,D,M}
I	{A,B,D,G,H,I}, {A,C,D,E,I}, {A,B,C,D,E,H,I}, {A,B,C,E,H,I}, {B,C,G,H,I}, {A,B,D,G,H,I}, {B,D,H,I}, {A,B,C,G,H,I}, {A,D,E,H,I}, {D,H,I}, {A,B,H,I}
H	{A,B,D,G,H}, {A,B,C,D,E,H}, {A,B,C,E,H}, {B,C,G,H}, {A,B,D,G,H}, {B,D,H}, {A,B,C,G,H}, {A,D,E,H}, {D,H}, {A,B,C,H}, {A,C,D,E,H}, {A,B,H}
G	{A,B,D,G}, {A,C,D,E,G}, {B,C,G}, {A,B,D,G}, {A,C,D,E,G}, {A,B,C,G}, {A,B,C,E,G}, {A,B,C,D,E,G}, {B,D,G}, {A,C,G}, {A,B,C,D,G}, {A,C,D,G}
E	{A,C,D,E}, {A,B,C,D,E}, {A,B,C,E}, {E}, {A,D,E}, {C,E}
P	{A,B,D,G,H,I,N,P}, {A,B,C,D,E,H,I,N,P}, {A,B,C,E,H,I,M,N,P}, {B,C,G,H,I,M,P}, {C,D,M,P}, {A,B,D,G,H,I,P}, {C,D,P}, {C,E,N,P}, {D,H,I,P}, {A,C,D,E,H,P }, {A,B,C,D,G,P}, {A,B,D,P}, {A,B,P}
N	{A,B,D,G,H,I,N}, {A,C,D,E,G,M,N}, {A,B,C,D,E,H,I,N}, {A,B,C,E,H,I,M,N}, {B,D,H,I,M,N}, {A,B,C,E,N}, {B,D,M,N}, {E,M,N}, {C,E,N}, {B,D,G,M,N}, {A,C,G,N}, {A,B,C,H,N}, {A,C,D,G,N}
B	{A,B}, {B}
D	{A,B,D}, {A,C,D}, {A,B,C,D}, {C,D}, {B,D}, {A,D}, {D}
C	{A,C}, {A,B,C}, {B,C}, {C}
A	{A}

Setelah diketahui hasil dari *frequent itemsets* pada Tabel 4.8 maka dapat dihitung nilai *confidence* nya yang dapat di lihat pada tabel di bawah ini :

Tabel 4.9 *Conditional Pattern Base*

Item	***Conditional Pattern Base***
S	{A B D G H I N P : 1}, {A C D E G M N : 1}, {A B C D E H I N P : 1} {A B D G H I P : 1}, {A C D E G : 1}, {B D H I M N : 1}, {A B C E N : 1}, {E M N : 1}, {D H I P : 1}, {A B C H N : 1}, {A B P : 1}
M	{A C D E I : 1}, {A C D E G : 1}, {A B C E H I : 1}, {B C G H I : 1}, {C D :1}, {B D H I :1}, {B D : 1}, {E :1}, {A D E H I :1}, {B D G :1}, {A B C D :1}
I	{A B D G H : 1}, {A C D E :1}, {A B C D E H : 1}, {A B C E H : 1}, {B C G H :1}, {A B D G H : 1}, {B D H : 1}, {A B C G H : 1}, {A D E H :1}, {D H : 1}, {A B H : 1}
H	{A B D G : 1}, {A B C D E : 1}, {A B C E : 1}, {B C G : 1}, {A B D G :1}, {B D :1}, {A B,C G : 1}, {A : 1}, {D : 1}, {A B C : 1}, {A C D E : 1}, {A B : 1}
G	{A B D :1}, {A C D E :1}, {B C :1}, {A B D :1}, {A C D E :1}, {A B C :1}, {A B C E :1}, {A B C D E :1}, {B D :1}, {A C :1}, {A B C D :1}, {A C D :1}
E	{A C D : 4}, {A B C D : 2}, {A B C : 3}, {A D : 1}, {C :1}
P	{A B D G H I N :1}, {A B C D E H I N : 1}, {A B C E H I M N : 1}, {B C G H I M O : 1}, {C D M O : 1}, {A B D G H I :1}, {C D :1}, {C E N :1}, {D H I :1}, {A C D E H : 1}, {A B C D G : 1}, {A B D :1}, {A B :1}
N	{A B D G H I :1}, {A C D E G M :1}, {A B C D E H I : 1}, {A B C E H I M :1}, {B D H I M N :1}, {A B C E N : 1}, {B D M :1}, {E M :1}, {C E :1}, {B D G M :1}, {A C G :1}, {A B C H :1}, {A C D G:1}
B	{A : 14}
D	{A B : 3}, {A C : 5}, {A B C :4}, {C :2}, {B :3}, {A :1}
C	{A : 6}, {A B : 9}, {B : 1}

Tabel 4.10 *Conditional Fp-Tree*

Item	***Conditional Fp-Tree***
S	<A : 8, B : 7, D : 7, G : 4, H : 6, I : 5, N : 7, P : 5, C : 5, E : 5, M : 3>
M	<A : 5, C : 6, D : 8, E : 5, I : 5, G : 3, B : 6, H : 4>
I	<A : 8, B : 8, D : 7, G : 4, H : 10, C : 4, E ; 4>
H	<A : 9, B : 9, D : 4, G : 4, C : 6, E : 3>
G	<A : 11, B : 8, D : 8, C : 9, E : 4>
E	<A : 10, C : 10, D : 7, B : 5>
P	<A : 8, B : 8, D : 9, G : 4, H : 7, I : 6, N : 4, C : 8, E : 4, M : 3>
N	A : 8, B : 8, D : 7, G : 5, H : 5, I : 4, C : 8, E : 5, M : 6, S : 2>

B	<A : 14, B : 14>
D	<A : 13, B : 10, C : 11>
C	<A : 15, B : 10>

Setelah tahap *conditional pattern base* dan *conditional fp-tree* di bentuk, maka tahap selanjutnya ialah menentukan *minimum support* dan *minimum confidence* nya.

Tabel 4.11 Kombinasi 2 *Itemsets*

Kombinasi	**Min.*Support***
A,B	43
A,C	43
A,D	37
A,E	26
A,G	18
A,H	22
A,I	13
A,M	3
A,N	9
A,P	4
A,S	8
B,C	27
B,D	35
B,E	14
B,G	26
B,H	35
B,I	24
B,M	13
B,N	15
B,P	11
B,S	7
C,D	27

C,E	29
C,G	20
C,H	20
C,I	13
C,M	12
C,N	14
C,P	8
C,S	5
D,E	22
D,G	23
D,H	27
D,I	22
D,M	13
D,N	13
D,P	13
D,S	7
E,G	7
E,H	14
E,I	11
E,M	11
E,N	12
E,P	4
E,S	4
G,H	15
G,I	11
G,M	7
G,N	6
G,P	6
G,S	3
H,I	29

H,M	9
H,N	12
H,P	11
H,S	6
I,M	10
I,N	10
I,P	10
I,S	5
M,N	10
M,P	3
M,S	8
N,P	6
N,S	6
P,S	5

Pada tahap ini digunakan untuk menentukan nilai *support* dan *confidence* pada setiap itemset dengan rumus yang sudah dijelaskan sebelumnya pada dasar teori, maka hasilnya adalah sebagai beikut :

$$Support\ A \cap B = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total Transaksi}}$$

$$Support\ (A, B) = A \cap B = \frac{43}{30} \times 100\% = 143\%$$

$$Support\ (A, C) = A \cap C = \frac{43}{30} \times 100\% = 143\%$$

$$Support\ (A, D) = A \cap D = \frac{37}{30} \times 100\% = 123\%$$

$$Support\ (A, E) = A \cap E = \frac{26}{30} \times 100\% = 87\%$$

$$Support\ (A, G) = A \cap G = \frac{18}{30} \times 100\% = 60\%$$

$$Support\ (A, H) = A \cap H = \frac{22}{30} \times 100\% = 73\%$$

$$Support\ (A, I) = A \cap I = \frac{13}{30} \times 100\% = 43\%$$

$$Support\ (A,M) = A \cap M = \frac{3}{30} \times 100\% = 10\%$$
$$Support\ (A,N) = A \cap N = \frac{9}{30} \times 100\% = 30\%$$
$$Support\ (A,P) = A \cap P = \frac{4}{30} \times 100\% = 13\%$$
$$Support\ (A,S) = A \cap S = \frac{8}{30} \times 100\% = 27\%$$
$$Support\ (B,C) = B \cap C = \frac{27}{30} \times 100\% = 90\%$$
$$Support\ (B,D) = B \cap D = \frac{35}{30} \times 100\% = 117\%$$
$$Support\ (B,E) = B \cap E = \frac{14}{30} \times 100\% = 47\%$$
$$Support\ (B,G) = B \cap G = \frac{26}{30} \times 100\% = 87\%$$
$$Support\ (B,H) = B \cap H = \frac{35}{30} \times 100\% = 117\%$$
$$Support\ (B,I) = B \cap I = \frac{24}{30} \times 100\% = 80\%$$
$$Support\ (B,M) = B \cap M = \frac{13}{30} \times 100\% = 43\%$$
$$Support\ (B,N) = B \cap N = \frac{15}{30} \times 100\% = 50\%$$
$$Support\ (B,P) = B \cap P = \frac{11}{30} \times 100\% = 37\%$$
$$Support\ (B,S) = B \cap S = \frac{7}{30} \times 100\% = 23\%$$
$$Support\ (C,D) = C \cap D = \frac{27}{30} \times 100\% = 90\%$$
$$Support\ (C,E) = C \cap E = \frac{29}{30} \times 100\% = 97\%$$
$$Support\ (C,G) = C \cap G = \frac{20}{30} \times 100\% = 67\%$$
$$Support\ (C,H) = C \cap H = \frac{20}{30} \times 100\% = 67\%$$
$$Support\ (C,I) = C \cap I = \frac{13}{30} \times 100\% = 43\%$$
$$Support\ (C,M) = C \cap M = \frac{12}{30} \times 100\% = 40\%$$
$$Support\ (C,N) = C \cap N = \frac{14}{30} \times 100\% = 47\%$$
$$Support\ (C,P) = C \cap P = \frac{8}{30} \times 100\% = 27\%$$

$$Support\ (C,S) = C \cap S = \frac{22}{30} \times 100\% = 17\%$$
$$Support\ (D,E) = D \cap E = \frac{22}{30} \times 100\% = 17\%$$
$$Support\ (D,G) = D \cap G = \frac{23}{30} \times 100\% = 77\%$$
$$Support\ (D,H) = D \cap H = \frac{27}{30} \times 100\% = 90\%$$
$$Support\ (D,I) = D \cap I = \frac{22}{30} \times 100\% = 73\%$$
$$Support\ (D,M) = D \cap M = \frac{13}{30} \times 100\% = 43\%$$
$$Support\ (D,N) = D \cap N = \frac{13}{30} \times 100\% = 43\%$$
$$Support\ (D,P) = D \cap P = \frac{13}{30} \times 100\% = 43\%$$
$$Support\ (D,S) = D \cap S = \frac{7}{30} \times 100\% = 23\%$$
$$Support\ (E,G) = E \cap G = \frac{7}{30} \times 100\% = 23\%$$
$$Support\ (E,H) = E \cap H = \frac{14}{30} \times 100\% = 47\%$$
$$Support\ (E,I) = E \cap I = \frac{11}{30} \times 100\% = 37\%$$
$$Support\ (E,M) = E \cap M = \frac{11}{30} \times 100\% = 37\%$$
$$Support\ (E,N) = E \cap N = \frac{12}{30} \times 100\% = 40\%$$
$$Support\ (E,P) = E \cap P = \frac{4}{30} \times 100\% = 13\%$$
$$Support\ (E,S) = E \cap S = \frac{4}{30} \times 100\% = 13\%$$
$$Support\ (G,H) = G \cap H = \frac{15}{30} \times 100\% = 50\%$$
$$Support\ (G,I) = G \cap I = \frac{11}{30} \times 100\% = 37\%$$
$$Support\ (G,M) = G \cap M = \frac{7}{30} \times 100\% = 23\%$$
$$Support\ (G,N) = G \cap N = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (G,P) = G \cap P = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (G,S) = G \cap S = \frac{3}{30} \times 100\% = 10\%$$

$$Support\ (H,I) = H \cap I = \frac{29}{30} \times 100\% = 97\%$$
$$Support\ (H,M) = H \cap M = \frac{9}{30} \times 100\% = 30\%$$
$$Support\ (H,N) = H \cap N = \frac{12}{30} \times 100\% = 40\%$$
$$Support\ (H,P) = H \cap P = \frac{11}{30} \times 100\% = 37\%$$
$$Support\ (H,S) = H \cap S = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (I,M) = I \cap M = \frac{10}{30} \times 100\% = 33\%$$
$$Support\ (I,N) = I \cap N = \frac{10}{30} \times 100\% = 33\%$$
$$Support\ (I,P) = I \cap P = \frac{10}{30} \times 100\% = 33\%$$
$$Support\ (I,S) = I \cap S = \frac{5}{30} \times 100\% = 17\%$$
$$Support\ (M,N) = M \cap N = \frac{10}{30} \times 100\% = 33\%$$
$$Support\ (M,P) = M \cap P = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (M,S) = M \cap S = \frac{8}{30} \times 100\% = 27\%$$
$$Support\ (N,P) = N \cap P = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (N,S) = N \cap S = \frac{6}{30} \times 100\% = 20\%$$
$$Support\ (P,S) = P \cap S = \frac{5}{30} \times 100\% = 17\%$$

$$Confidence\ (\mathrm{A} \rightarrow \mathrm{B}) = P\ (A|B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}}$$
$$Confidence\ (A \rightarrow B) = \frac{43}{21} \times 100\% = 205\%$$
$$Confidence\ (A \rightarrow C) = \frac{43}{21} \times 100\% = 205\%$$
$$Confidence\ (A \rightarrow D) = \frac{37}{21} \times 100\% = 176\%$$
$$Confidence\ (A \rightarrow E) = \frac{26}{21} \times 100\% = 124\%$$
$$Confidence\ (A \rightarrow G) = \frac{18}{21} \times 100\% = 86\%$$
$$Confidence\ (A \rightarrow H) = \frac{22}{21} \times 100\% = 105\%$$

$$Confidence\ (A \rightarrow I) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (A \rightarrow M) = \frac{3}{21} \times 100\% = 14\%$$
$$Confidence\ (A \rightarrow N) = \frac{9}{21} \times 100\% = 43\%$$
$$Confidence\ (A \rightarrow P) = \frac{4}{21} \times 100\% = 19\%$$
$$Confidence\ (A \rightarrow S) = \frac{8}{21} \times 100\% = 38\%$$
$$Confidence\ (B \rightarrow C) = \frac{27}{21} \times 100\% = 129\%$$
$$Confidence\ (B \rightarrow D) = \frac{35}{21} \times 100\% = 167\%$$
$$Confidence\ (B \rightarrow E) = \frac{14}{21} \times 100\% = 67\%$$
$$Confidence\ (B \rightarrow G) = \frac{26}{21} \times 100\% = 124\%$$
$$Confidence\ (B \rightarrow H) = \frac{35}{21} \times 100\% = 167\%$$
$$Confidence\ (B \rightarrow I) = \frac{24}{21} \times 100\% = 114\%$$
$$Confidence\ (B \rightarrow M) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (B \rightarrow N) = \frac{15}{21} \times 100\% = 71\%$$
$$Confidence\ (B \rightarrow P) = \frac{11}{21} \times 100\% = 52\%$$
$$Confidence\ (B \rightarrow S) = \frac{7}{21} \times 100\% = 33\%$$
$$Confidence\ (C \rightarrow D) = \frac{27}{21} \times 100\% = 129\%$$
$$Confidence\ (C \rightarrow E) = \frac{29}{21} \times 100\% = 138\%$$
$$Confidence\ (C \rightarrow G) = \frac{20}{21} \times 100\% = 95\%$$
$$Confidence\ (C \rightarrow H) = \frac{20}{21} \times 100\% = 95\%$$
$$Confidence\ (C \rightarrow I) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (C \rightarrow M) = \frac{12}{21} \times 100\% = 57\%$$
$$Confidence\ (C \rightarrow N) = \frac{14}{21} \times 100\% = 67\%$$

$$Confidence\ (C \rightarrow P) = \frac{8}{21} \times 100\% = 38\%$$
$$Confidence\ (C \rightarrow S) = \frac{5}{21} \times 100\% = 24\%$$
$$Confidence\ (D \rightarrow E) = \frac{22}{21} \times 100\% = 105\%$$
$$Confidence\ (D \rightarrow G) = \frac{23}{21} \times 100\% = 110\%$$
$$Confidence\ (D \rightarrow H) = \frac{27}{21} \times 100\% = 129\%$$
$$Confidence\ (D \rightarrow I) = \frac{22}{21} \times 100\% = 105\%$$
$$Confidence\ (D \rightarrow M) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (D \rightarrow N) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (D \rightarrow P) = \frac{13}{21} \times 100\% = 62\%$$
$$Confidence\ (D \rightarrow S) = \frac{7}{21} \times 100\% = 33\%$$
$$Confidence\ (E \rightarrow G) = \frac{7}{21} \times 100\% = 33\%$$
$$Confidence\ (E \rightarrow H) = \frac{14}{21} \times 100\% = 67\%$$
$$Confidence\ (E \rightarrow I) = \frac{11}{21} \times 100\% = 52\%$$
$$Confidence\ (E \rightarrow M) = \frac{11}{21} \times 100\% = 52\%$$
$$Confidence\ (E \rightarrow N) = \frac{12}{21} \times 100\% = 57\%$$
$$Confidence\ (E \rightarrow P) = \frac{4}{21} \times 100\% = 19\%$$
$$Confidence\ (E \rightarrow S) = \frac{4}{21} \times 100\% = 19\%$$
$$Confidence\ (G \rightarrow H) = \frac{15}{21} \times 100\% = 71\%$$
$$Confidence\ (G \rightarrow I) = \frac{11}{21} \times 100\% = 52\%$$
$$Confidence\ (G \rightarrow M) = \frac{7}{21} \times 100\% = 33\%$$
$$Confidence\ (G \rightarrow N) = \frac{6}{21} \times 100\% = 29\%$$
$$Confidence\ (G \rightarrow P) = \frac{6}{21} \times 100\% = 29\%$$

$$Confidence\ (G \rightarrow S) = \frac{3}{21} \times 100\% = 14\%$$
$$Confidence\ (H \rightarrow I) = \frac{29}{21} \times 100\% = 138\%$$
$$Confidence\ (H \rightarrow M) = \frac{9}{21} \times 100\% = 43\%$$
$$Confidence\ (H \rightarrow N) = \frac{12}{21} \times 100\% = 57\%$$
$$Confidence\ (H \rightarrow P) = \frac{11}{21} \times 100\% = 52\%$$
$$Confidence\ (H \rightarrow S) = \frac{6}{21} \times 100\% = 29\%$$
$$Confidence\ (I \rightarrow M) = \frac{10}{21} \times 100\% = 48\%$$
$$Confidence\ (I \rightarrow N) = \frac{10}{21} \times 100\% = 48\%$$
$$Confidence\ (I \rightarrow P) = \frac{10}{21} \times 100\% = 48\%$$
$$Confidence\ (I \rightarrow S) = \frac{5}{21} \times 100\% = 24\%$$
$$Confidence\ (M \rightarrow N) = \frac{10}{21} \times 100\% = 48\%$$
$$Confidence\ (M \rightarrow P) = \frac{3}{21} \times 100\% = 14\%$$
$$Confidence\ (M \rightarrow S) = \frac{8}{21} \times 100\% = 38\%$$
$$Confidence\ (N \rightarrow P) = \frac{6}{21} \times 100\% = 29\%$$
$$Confidence\ (N \rightarrow S) = \frac{6}{21} \times 100\% = 29\%$$
$$Confidence\ (P \rightarrow S) = \frac{5}{21} \times 100\% = 24\%$$

Tabel 4.12 *Rule* Yang Terbentuk

Rule	***Support***	***Confidence***
A,B	143%	205%
A,C	143%	205%
A,D	123%	176%
B,D	117%	167%
B,H	117%	167%
C,E	97%	138%
H,I	97%	138%
B,C	90%	129%
C,D	90%	129%

D,H	90%	129%
A,E	87%	124%
B,I	80%	114%
D,G	77%	110%
A,H	73%	105%
D,E	73%	105%
D,I	73%	105%
C,G	67%	95%
C,H	67%	95%
A,G	60%	86%
B,N	50%	71%
G,H	50%	71%
C,N	47%	67%
EH	47%	67%
A,I	43%	62%
B,M	43%	62%
C,I	43%	62%
D,M	43%	62%
D,N	43%	62%
D,P	43%	62%
C,M	40%	57%
E,N	40%	57%
H,N	40%	57%
B,P	37%	52%
E,I	37%	52%
E,M	37%	52%
G,I	37%	52%
H,P	37%	52%
I,M	33%	48%
I,N	33%	48%
I,P	33%	48%
M,N	33%	48%
A,N	30%	43%
H,M	30%	43%
A,S	27%	38%
C,P	27%	38%
M,S	27%	38%
B,S	23%	33%
D,S	23%	33%
E,G	23%	33%
G,M	23%	33%
G,N	20%	29%
G,P	20%	29%
N,P	20%	29%

N,S	20%	29%
C,S	17%	24%
I,S	17%	24%
P,S	17%	24%
A,P	13%	19%
E,P	135	19%
E,S	13%	19%
A,M	10%	14%
G,S	10%	14%
M,P	10%	14%

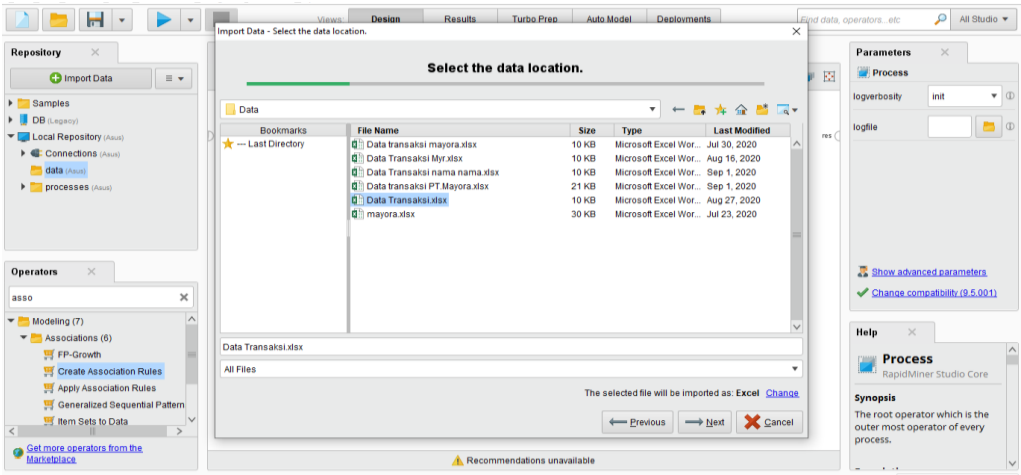
Dari perhitungan *confidence* terhadap pola yang terbentuk diatas, maka *Association Rule* yang memenuhi syarat *confidence* ≥ 0,75 adalah: H→ A = 0,75 (jika konsumen membeli kopiko coffee shot cappucino, maka membeli beng-beng wafer cokelat), G → C = 0,75 (jika konsumen membeli energen all variant, maka membeli beng beng share it), H → B = 0,75 (jika konsumen membeli kopiko coffee shot cappucino, maka membeli better vanila special).

4.2. Hasil

4.2.1. Implementasi Sistem

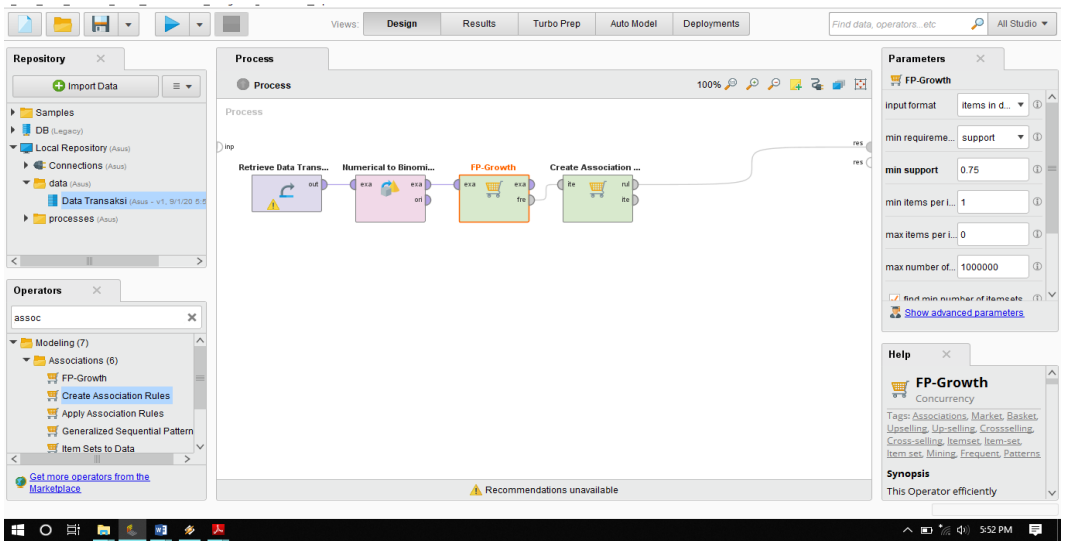
Untuk membuktikan kebenaran hasil analisa diperlukan sebuah proses pengujian untuk menguji kebenaran dari hasil pengolahan data yang telah dikerjakan secara manual sebelumnya, untuk proses pengujian tersebut kita dapat menggunakan salah satu *software* aplikasi seperti *Rapidminer*, dengan langkah-langkah sebagai berikut.

Dataset yang terdiri 30 record data yang di simpan dalam apliksi *microsoft excel* dengan nama file Data Transaksi.xls dan akan dicoba menggunakan *software rapidminer* 9.3.1 untuk melihat sama tidaknya hasil pencarian *frequent itemset*nya.



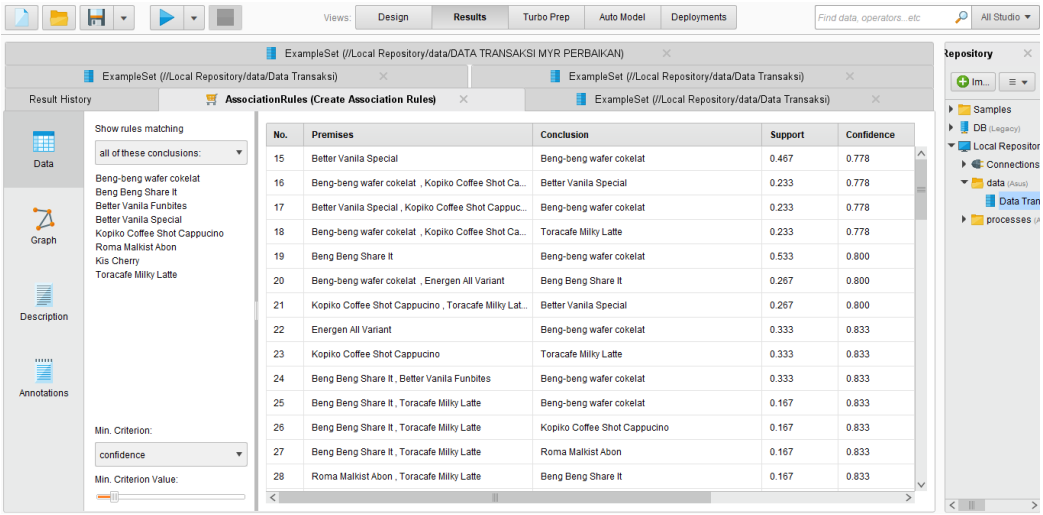
Gambar 4.20 Tampilan *Import Excel Sheet*

Gambar 4.20 adalah proses dari import Data Transaksi.xls ke dalam *rapidminer*.



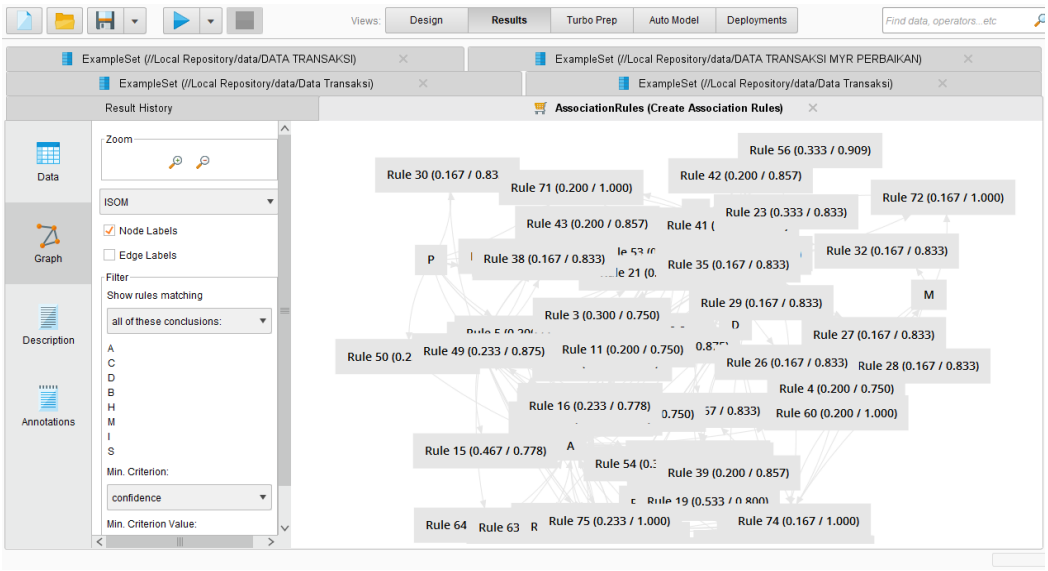
Gambar 4.21 Fp-*Growth* Pada *Rapidminer*

Pada gambar 4.21 merupakan proses menghubungkan fp-*growth*, dan pada proses tersebut telah diatur *minimum support* dan *confidence*nya.

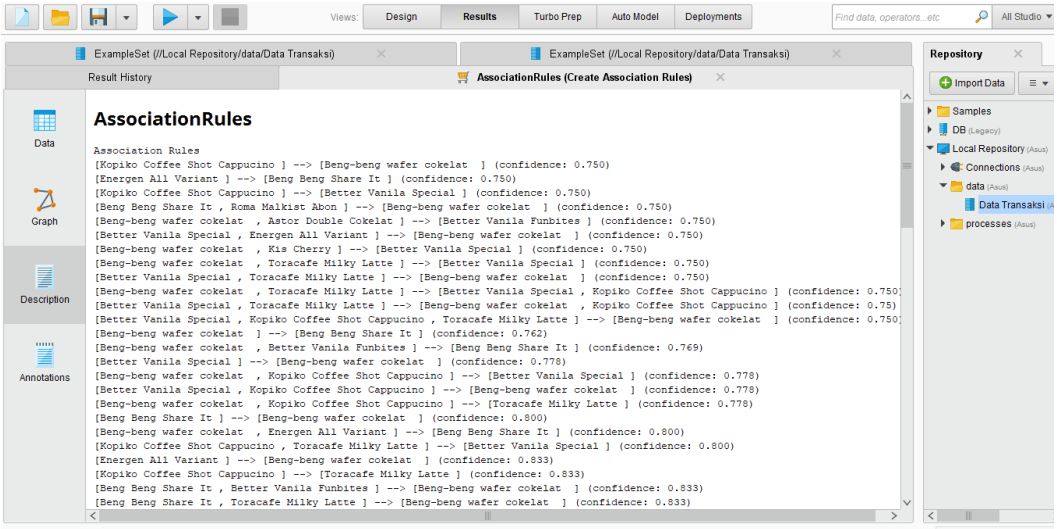


No.	Premises	Conclusion	Support	Confidence
15	Better Vanila Special	Beng-beng wafer cokelat	0.467	0.778
16	Beng-beng wafer cokelat , Kopiko Coffee Shot Ca...	Better Vanila Special	0.233	0.778
17	Better Vanila Special , Kopiko Coffee Shot Cappuc...	Beng-beng wafer cokelat	0.233	0.778
18	Beng-beng wafer cokelat , Kopiko Coffee Shot Ca...	Toracafe Milky Latte	0.233	0.778
19	Beng Beng Share It	Beng-beng wafer cokelat	0.533	0.800
20	Beng-beng wafer cokelat , Energen All Variant	Beng Beng Share It	0.267	0.800
21	Kopiko Coffee Shot Cappucino , Toracafe Milky Lat...	Better Vanila Special	0.267	0.800
22	Energen All Variant	Beng-beng wafer cokelat	0.333	0.833
23	Kopiko Coffee Shot Cappucino	Toracafe Milky Latte	0.333	0.833
24	Beng Beng Share It , Better Vanila Funbites	Beng-beng wafer cokelat	0.333	0.833
25	Beng Beng Share It , Toracafe Milky Latte	Beng-beng wafer cokelat	0.167	0.833
26	Beng Beng Share It , Toracafe Milky Latte	Kopiko Coffee Shot Cappucino	0.167	0.833
27	Beng Beng Share It , Toracafe Milky Latte	Roma Malkist Abon	0.167	0.833
28	Roma Malkist Abon , Toracafe Milky Latte	Beng Beng Share It	0.167	0.833

Gambar 4.22 *Support* Dan *Confidence* Menggunakan *Rapidminer* Dari Sample Data Penjualan



Gambar 4.23 Grafis *Rule* Sample Data Penjualan Menggunakan *Rapidminer*



Gambar 4.24 *Association Rules* Dari *Rapidminer*

Setelah diuraikan dari *rule* di atas, maka dapat diambil kesimpulan bahwasannya konsumen cenderung membeli item yang saling berhubungan seperti pada : jika konsumen membeli kopiko coffe shot cappuccino maka juga membeli beng-beng wafer cokelat, dengan nilai *confidence* 0,75.

Kemudian hasil akhir yang diperoleh dari *rule* tersebut adalah sebuah pengetahuan baru tentang pola pembelian konsumen yang selama ini jarang diketahui. Hal itu dapat dimanfaatkan untuk membantu perusahaan agar dapat menjaga kualitasnya dengan baik kepada konsumen, dengan cara produk makanan harus tetap tersedia saat konsumen membutuhkannya. Untuk itu perusahaan harus selalu menyediakan produk yang diinginkan oleh konsumen.

BAB V
KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan data dan hasil pembahasan maka dapat diambil kesimpulan sebagai berikut:

1. Secara keseluruhan dari data sample penjualan diperoleh 88 *rule* yang terdiri dari 53 *rule* asosiasi yang memenuhi *support* dengan ambang batas 75% dan 4 *rule* yang memenuhi *confidence* 90%.
2. Dari *rule* yang diperoleh didapatkan sebuah pengetahuan baru tentang pola pembelian konsumen. Dari pengetahuan tersebut bisa digunakan untuk menentukan keputusan persediaan barang dan memberikan paket diskon terhadap pola pembelian produk yang memiliki nilai *confidence* tertinggi.
3. Algoritma *fp-growth* dapat membantu perusahaan untuk memantau stok barang yang sering dibeli oleh konsumen sehingga tidak akan terjadi kelangkaan pasokan.

5.2. Saran

Berdasarkan hasil penelitian, ada beberapa saran yang dapat dilakukan agar pengembangan sistem ini menjadi lebih baik, diantaranya sebagai berikut :

1. Dalam penerapan *association rules* menggunakan algoritma *fp-growth* sebaiknya dilakukan pengecekan secara berkala untuk evaluasi secara terus menerus.
2. Pada penelitian selanjutnya dapat menggunakan algoritma yang berbeda dari penelitian yang telah dilakukan saat ini. Dapat juga menggunakan algoritma *k-nearest neighbor* dan *decision tree*.

DAFTAR PUSTAKA

Anggraeni, R. M. 2014. Perbandingan Algoritma *Apriori* dan Algoritma *FP-Growth* untuk Perekomendasi Pada Transaksi Peminjaman Buku di Perpustakaan Universitas Dian Nuswantoro. Jurnal Teknik Informatika. Vol.1, No.6.

Anto, A., Susilo, T., Informatika, P. T., Selatan, S., & Selatan, S. 2018. Penerapan Algoritma Apriori pada Pengolahan Data Transaksi Penjualan di Minimarket Priyo Kota Lubuklinggau. Jurnal Teknik Informatika.Vol.1, No.03.

Amalia, Y. rizki. 2018. Penerapan data mining untuk prediksi penjualan produk elektronik terlaris menggunakan metode *k-nearest neighbor*. Jurnal Ilmiah Komputer dan Informatika. Vol. 12, No. 3.

A., Marisa, F., & Purnomo, D. 2016. Penerapan Algoritma Apriori Terhadap Data Penjualan di Toko Gudang BM. *JOINTECS (Journal of Information Technology and Computer Science)*, Vol.1, No.1.

Astrina. I, Zainal. A. M & Pujianto. U. 2019. Penerapan Algoritma FP-Growth Dalam Penentuan Pola Pembelian Konsumen Pada Kain Tenun Medali Mas. Jurnal Matrix. Vol 9, No 1.

Eska, J. 2016. Penerapan Data Mining Untuk Prekdiksi Penjualan Wallpaper Menggunakan Algoritma C4.5 STMIK Royal Ksiaran. Jurnal Teknologi Dan Sistem Informasi. Vol.2, No. 9.

Fahrin. A. A. dan Maulana. A. 2018. Penerapan Data Mining Untuk Analisa Pola Pembelian Konsumen Dengan Algoritma *FP-Growth* Pada Data Transaksi Penjualan *Spare Part* Motor. Kumpulan Jurnal Ilmu Komputer, Vol. 05, pp.27-36.

Furqan. M, Sajadin, Embong. A, Azwan. M. 2009. Meningkatkan Prestasi Akademik Mahasiswa dengan Penerapan Teknik Data Mining.

Gunadi, G., & Sensuse, D. I. 2012. Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan

Menggunakan Algoritma *Apriori* Dan *Frequent Pattern Growth* (*Fp-Growth*). Jurnal Telematika. Vol.4, No.1.

Haryati, S., Sudarsono, A., & Suryana, E. 2015. Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu). Jurnal Media Infotama. Vol.11, No.2.

Herasmus.H. 2017. Analisa Customer Service System Menggunakan Metode Data Mining Dengan Algoritma Fp-Growth. Jurnal Teknik Ibnu Sina. Vol. 2, No. 2.

Ikhwan, A., Nofriansyah, D., & Sriani. 2015. Penerapan Data Mining dengan Algoritma *Fp-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus Kampus STMIK Triguna Dharma). Jurnal Saintikom. Vol.14, No.3.

Junaidi. A. 2019. Implementasi Algoritma Apriori dan FP-Growth Untuk Menentukan Persediaan Barang. Jurnal Sisfokom. Vol 8, No 1.

Kurniawan, S., Gata, W., & Wiyana, H. 2018. Analisis Algoritma *Fp-Growth* Untuk Rekomendasi Produk Pada Data Retail Penjualan Produk Kosmetik (Studi Kasus : Mt Shop Kelapa Gading). Jurnal Teknologi Informasi. Vol.7, No.5.

Lestari, Y. D. 2017. Penerapan Data Mining Menggunakan Algoritma *Fp-Tree* Dan *Fp-Growth* Pada Data Transaksi Penjualan Obat. Jurnal Snastikom.

Nofitri, R., & Irawati, N. 2019. Analisis Data Hasil Keuntungan Menggunakan Software Rapidminer. Jurnal Teknologi dan Sistem Informasi. Vol 5, No 2.

Nurul . A . R. 2015. Implementasi Algoritma Frequent Pattern Growth (*Fp-Growth)* Menentukan Asosiasi Antar Produk (Study KAsus Nadia Mart). Jurnal Teknik Informatika. Vol 5, No 11.

R. M. Miraldi, A. Rachmat and B. Santoso. 2014. Implementasi Algoritma Fp-Growth Untuk Sistem Rekomendasi Buku Di Perpustakaan UKDW. Jurnal Informatika, Vol 10, No 1.

Samuel, D. 2008. Penerapan Stuktur *FP-Tree* dan Algoritma *FP-Growth* dalam Optimasi Penentuan Frequent Itemset. Jurnal Teknik Informatika. Vol.6, No.2.

Sriani, Ikhwan. A, Danti. R. A. 2015. Penerapan Data Mining Untuk Analisa Resiko Premi Asuransi Jiwa Dengan Menggunakan Algoritma C4.5.

Triyanto, W. A. 2014. Association Rule Mining Untuk Penentuan Rekomendasi Promosi Produk. *Journal Simetris*. Vol.5, No.2.

Wahyuni, S. 2018. Implementasi Data Mining dalam Memprediksi Stok Barang Menggunakan Algoritma Apriori. Jurnal Ilmiah Teknologi Terapan. Vol.5, No.1.

LAMPIRAN 1

SURAT KETERANGAN PENELITIAN

KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA MEDAN
FAKULTAS SAINS DAN TEKNOLOGI
Jl. IAIN Nomor 1 Medan. Kode Pos 20235
Url: www.saintek.uinsu.ac.id, E-mail: saintek@uinsu.ac.id

Nomor : B.301/ST.I/ST.V.2/PP.00.9/03/2020 09 Maret 2020
Lampiran :
Hal : Izin Penelitian

Yth.
Pimpinan PT. Cipta Niaga Semesta Mayora Group
Jl. Karang Sari Pematangsiantar

Assalamu'alaikum Wr. Wb.
Dengan hormat, sehubungan dengan penyelesaian Skripsi (karya Ilmiah) mahasiswa Fakultas Sains dan Teknologi UIN Sumatera Utara Medan, maka kami mohon menerima dan membimbing mahasiswa di bawah ini:

Nama	: **Uli Audina Nursyani**
Nim	: 71153028
Tempat/ Tgl. Lahir	: Pematangsiantar, 11 Agustus 1997
Sem./Prodi	: X (Sepuluh) / Ilmu Komputer
Alamat	: Jl. M. Yakub No. 118 A/7 Medan

Untuk melaksanakan penelitian skripsi dengan judul "*Penerapan Data Mining Menggunakan Algoritma FP- Growth Untuk Strategi Penjualan Produk Makanan*".

Demikian disampaikan, atas perhatian dan kerjasama yang baik kami ucapkan terima kasih.

Wassalam,
an. Dekan
Wakil Dekan Bidang Akademik
dan Kelembagaan

Dr. Rina Filia Sari, M. Si
NIP. 197703012005012006

Tembusan :
1. Dekan Fakultas Sains dan Teknologi UIN Sumatera Utara Medan;
2. Ka. Prodi Ilmu Komputer Fakultas Sains dan Teknologi UIN Sumatera Utara Medan.

LAMPIRAN 2

DAFTAR RIWAYAT HIDUP

I. Data Pribadi

Nama	: Uli Audina Nursyani
Tempat/ Tanggal Lahir	: Pematangsiantar, 11 Agustus 1997
Alamat	: Jl. M. Yakub Nomor 118 A/7, Medan
Agama	: Islam
Jenis Kelamin	: Perempuan
Status	: Belum Menikah
No. Hp	: 0823-7090-0899
Email	: uliaudina1108@gmail.com

II. Pendidikan formal

1. 2002 – 2003 TK YPHI – Pematangsiantar
2. 2003 – 2009 Madrasah Taman Pendidikan Islam Pematangsiantar
3. 2009 – 2012 SMP Negeri 2 Pematangsiantar
4. 2012 – 2015 SMK Negeri 1 Pematangsiantar
5. 2015 – 2020 Universitas Islam Negeri Sumatera Utara, Fakultas Sains dan Teknologi, Program Studi Ilmu Komputer S1

LAMPIRAN 3

KARTU BIMBINGAN SKRIPSI

KARTU BIMBINGAN SKRIPSI

Semester Gasal/Genap Tahun Akademik ...../.....

Nama	: Uli Audina Nursyani	Pembimbing I	: Dr. Mhd Furqan, S.Si., M.Comp. Sc
NIM	: 71153028	Pembimbing II	: Griani, S.Kom., M.Kom
Prog. Studi	: ILMU KOMPUTER	SK Pembimbing	:

Judul Skripsi :
PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA FP-GROWTH UNTUK STRATEGI PENJUALAN PRODUK MAKANAN

PERT	PEMBIMBING I			PEMBIMBING II		
	Tgl.	Materi Bimbingan	Tanda Tangan	Tgl.	Materi Bimbingan	Tanda Tangan
I	7/2/2020	- Revisi judul proposal - Revisi bab I	Furqan	10/10/2019	- Revisi judul proposal - Revisi penulisan latar belakang, rumusan, batasan, tujuan dan manfaat penelitian Bab I	
II	12/2/2020	Revisi bab I - III	Furqan	9/12/2019	ACC bab I Lanjut bab II	
III	14/2/2020	Acc keseluruhan proposal	Furqan	18/12/2019	Perbaiki penulisan bab II, Perbaiki kutipan bab II, Perjelas tahapan metode, sesuaikan teori pemodelan sistem & tools yg digunakan.	
IV	15/10/2020	Bimbingan Bab IV dan revisi	Furqan	9/1/2020	Acc bab II Lanjut bab III	
V	20/10/2020	- Revisi rumusan masalah dan Tujuan masalah	Furqan	16/1/2020	Jelaskan diagram kerja penelitian pada bab III	

VI	24/10/2020	Acc keseluruhan Proposal	Furqan	23/1/2020	Perbaiki daftar Pustaka, kata Pengantar dan daftar isi	
VII				5/2/2020	Acc keseluruhan Proposal	
VIII				5/9/2020	Bimbingan Bab IV dan Revisi	
IX				8/9/2020	Acc Bab IV lanjut bab V	
X				16/9/2020	Acc keseluruhan Proposal	

Medan, 26 Oktober 2020
An. Dekan
Ketua Jurusan/Program Studi

Furqan

Dr. Mhd. Furqan, S.Si, M.Comp.Sc.
NIP. 198008062006041003

Catatan: *Pada saat bimbingan, kartu ini harus diisi dan ditandatangani oleh pembimbing*