

**RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU  
MENGUNAKAN ALGORITMA FP-GROWTH**

**SKRIPSI**

Oleh:  
**MOH AGUNG HADI WIJAYA**  
**NIM. 11650056**



**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2018**

**HALAMAN PENGANTAR**

**RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU  
MENGUNAKAN ALGORITMA FP-GROWTH**

**SKRIPSI**

**Diajukan Kepada:  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri (UIN)  
Maulana Malik Ibrahim Malang  
Untuk Memenuhi Salah Satu Persyaratan Dalam  
Memperoleh Gelar Sarjana Komputer (S.Kom)**

**Oleh :  
MOH AGUNG HADI WIJAYA  
NIM. 11650056**

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2018**

RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU  
MENGUNAKAN ALGORITMA FP-GROWTH

SKRIPSI

Oleh :  
Moh Agung Hadi Wijaya  
11650056

Telah Diperiksa dan Disetujui untuk Diuji:

Tanggal : 06 Juli 2018

Pembimbing I

Pembimbing II

Dr. Suhartono, M.Kom  
NIP. 19680519 200312 1 001

M. Imamuddin, Lc., MA  
NIP. 19740602 200901 1 010

Mengetahui,  
Ketua Jurusan Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Dr. Cahyo Crysdian  
NIP. 19740424 200901 1 008

**RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU  
MENGUNAKAN ALGORITMA FP-GROWTH**

**SKRIPSI**

Oleh :  
Moh Agung Hadi Wijaya  
11650056

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan  
Dinyatakan Diterima Sebagai Salah Satu Persyaratan Untuk  
Memperoleh Gelar Sarjana Komputer (S.Kom)

Pada Tanggal : 06 Juli 2018

Susunan Dewan Penguji : .....

**Penguji Utama** : Fatchurrochman, M.Kom  
NIP. 19710722 201101 1 001

**Ketua Penguji** : A'la Syaqui, M.Kom  
NIP. 19771201 200801 1 007

**Sekretaris Penguji** : Dr. Suhartono, M.Kom  
NIP. 19680519 200312 1 001

**Anggota Penguji** : M. Imamuddin, Lc., MA  
NIP. 19740602 200901 1 010

( *As.* )  
( *[Signature]* )  
( *[Signature]* )  
( *[Signature]* )

Mengetahui,

Ketua Jurusan Teknik Informatika  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Cahyo Crysdian  
NIP. 19740424 200901 1 008

**HALAMAN PERNYATAAN  
KEASLIAN PENELITIAN**

Nama : Moh Agung Hadi Wijaya  
NIM : 11650056  
Jurusan : Teknik Informatika  
Fakultas : Sains dan Teknologi  
Judul Skripsi : RANCANG BANGUN APLIKASI REKOMENDASI JUDUL  
BUKU MENGGUNAKAN ALGORITMA  
FP-GROWTH

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 06 Juli 2018

Yang membuat pernyataan,



Moh Agung Hadi Wijaya  
NIM. 11650056

## MOTTO

من شب على شيء شاب عليه

“SIAPA YANG PADA MASA MUDANYA (MEMBIASAKAN) ATAS SESUATU, NISCAYA DIA  
PADA MASA TUANYA (TERBIASA) ATASNYA.”



## HALAMAN PEREMBAHAN

Alhamdulillah segala puji bagi Allah SWT yang telah memberikan rahmat dan nikmatnya. Berkat ridlo-Nya saya dapat menyelesaikan kewajiban menuntaskan masa belajar di kampus UIN MAULANA MALIK IBRAHIM MALANG. Tanpa mengurangi rasa syukur saya persembahkan karya ini kepada :

Ibu Nurul Musdalifah serta Ayahanda Sholihudin

Atas doa dan dukungan beliaulah, pemberi dorongan utama dalam menyelesaikan tugas akhir ini. Dua insan yang senantiasa sabar serta memberikan kasih sayang kepada dua anaknya.

KH. Marzuki Mustamar

Sosok guru yang selalu memberikan keteladanan serta wawasan hati dan akal. Membimbing dengan tulus hati, “*all out dadi santri*” kata yang selalu tertancap kuat dalam sanubari para santrinya. Mendidik menjadi insan ululalbab seutuhnya.

Crew Mediasantrinu dan MajelisDiba

Teman seperjuangan dalam berkhidmah serta berorganisasi di pondok Sabilurrosyad Malang. Berproses dalam profesionalitas kerja, serta berkarya untuk bangsa. Segala bentuk pengalaman yang telah dilalui bersama.

Santri Gasek

Masa mondok yang memberikan pengalaman luas, demi meraih cita seorang santri barokah. Memperoleh banyak pengalaman berharga dari: Ahmad Nanda T, Abdullah Khoironi, M Ridwan, Aang Khunaifi, Tamami, Silva Ahmad Faizuddin, Riza Rahma, serta sahabat santri lainnya yang tak bisa saya sebutkan semua.

Siti Zaherotul Luailiyah

Adik kandung yang selalu memberikan semangat. Sosok penurut serta pemberani, mandiri, dan pintar masak.

## KATA PENGANTAR

*Assalamu'alaikum Warohmatullohi Wabarokatuh*

Segala puji bagi Allah SWT, karena atas segala rahmat dan karunia-Nya skripsi dengan judul “*Rancang Bangun Aplikasi Rekomendasi Judul Buku Menggunakan Algoritma FP-Growth*” dapat diselesaikan dengan baik dan lancar. Shalawat serta salam selalu tercurah limpah kepada tauladan terbaik Nabi Besar Muhammad SAW yang telah membimbing umatnya dari zaman kebodohan menuju Islam yang *rahmatan lil alamiin*.

Dalam penyelesaian skripsi ini, banyak rintangan serta godaan. Kepadaorang - orang yang telah memberikan bantuan baik secara moril, nasihat dan semangat maupun materiil. Atas segala bantuan yang telah diberikan, disampaikan doa dan ucapan terimakasih yang sedalam-dalamnya kepada :

1. Prof. Dr. Abdul Haris, M.Ag selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang
2. Bapak Dr. Suhartono, M.Kom selaku dosen pembimbing I yang senantiasa memberi masukan dan nasihat serta petunjuk dalam penyusunan skripsi ini.
3. Bapak M. Imamuddin, Lc., MA selaku dosen pembimbing II yang memberikan arahan dalam mengintegrasikan Sains dan Islam.
4. Bapak Dr. Cahyo Crysdiyan, selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang, yang memberikan motivasi inspiratif, serta ilmu pengetahuannya.
5. Segenap Dosen Teknik Informatika yang telah memberikan bimbingan keilmuan dan nasehat selama masa studi.

6. Semua teman – teman mahasiswa IT UIN Malang angkatan 2011 yang telah memberikan dorongan semangat dalam menyelesaikan tugas akhir.
7. Ibunda Nurul Musdalifah dan Ayahanda Sholihudin serta keluarga besar tercinta yang selalu memberi dukungan serta kepercayaannya. Memberikan tuntunan dalam setiap langkah dalam mencari ilmu di perantauan.
8. Teman-teman pondok pesantren Sabilurrosyad Gasek Malang yang selalu mengajarkan pengalaman kehidupan di tengah masyarakat. Serta selalu memberikan bantuan saat diterpa ujian/kesulitan.
9. Tim Majelis Diba yang telah bergotong royong bersama-sama berjuang “nguri-nguri warisan walisongo” berdakwah ditengah masyarakat.
10. Tim Mediasantrinu yang membuka wawasan penulis untuk ikut berjuang di lingkungan literasi, sebagai bentuk menjaga budaya kepenulisan di pesantren khususnya.

Penulis menyadari atas kekurangan dan kesalahan yang pembaca temukan dalam penulisan skripsi ini. Kritikan maupun saran terbuka luas untuk disampaikan kepada penulis guna pembenahan diri agar lebih teliti. Semoga apa yang menjadi kekurangan bisa disempurnakan oleh peneliti selanjutnya dan semoga karya ini senantiasa dapat memberi manfaat. *Amin Ya Robbal'amin*

*Wassalamualaikum Warohmatullohi Wabarokatuh*

Malang, 07 Juli 2018

Penulis

## DAFTAR ISI

HALAMAN PENGAJUAN .....	ii
HALAMAN PERSETUJUAN .....	iii
HALAMAN PENGESAHAN .....	iv
HALAMAN PERNYATAAN KEASLIAN PENELITIAN .....	v
MOTTO .....	vi
HALAMAN PEREMBAHAN .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR GAMBAR .....	xii
DAFTAR TABEL .....	xiii
ABSTRAK .....	xiv
ABSTRACT .....	xv
المخلص .....	xvi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Batasan Masalah .....	4
1.5 Manfaat Penelitian .....	5
2.1 Penelitian Terkait .....	6
2.2 Intrans Publishing .....	7
2.3 Landasan Teori .....	10
2.3.1 Pengertian Data Mining .....	10
2.3.2 Metode Data mining .....	11
2.3.3 Metode Association Rule .....	13
2.3.4 Tahapan Data Mining .....	15
2.3.5 Cross Industry Standard Process for Data Mining .....	17
2.3.6 Metode Asosiasi .....	20
2.3.7 Lift .....	21
2.3.8 FP-Tree .....	22

2.3.9	Algoritma FP-Growth.....	22
2.4	Basis Data .....	24
2.4.1	Operasi Dasar Basis Data .....	25
2.4.2	Objektif Basis Data.....	27
2.4.3	Database Management System.....	30
2.5	Unified Modelling Language (UML) .....	31
2.6	Use case Diagram .....	31
2.7	Activity Diagram .....	34
2.8	Sequence Diagram .....	34
2.9	Class Diagram.....	35
<b>BAB III</b>	<b>METODE PENELITIAN .....</b>	<b>36</b>
3.1	Desain Penelitian .....	36
3.1.1	Objek Penelitian .....	36
3.1.2	Sumber Data .....	36
3.1.3	Bahan Penelitian.....	37
3.1.4	Alat Penelitian .....	37
3.2	Prosedur penelitian .....	38
3.2.1	Pemahaman Sistem dan Studi Literatur.....	38
3.2.2	Perancangan Sistem.....	39
3.3	Analisa Kebutuhan Sistem.....	42
3.4	Analisa Masalah.....	42
3.5	Analisa Penerapan Data Mining .....	43
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN.....</b>	<b>66</b>
<b>BAB V</b>	<b>PENUTUP.....</b>	<b>77</b>
5.1	Kesimpulan .....	77
5.2	Saran .....	77

## DAFTAR GAMBAR

Gambar 2.1 Struktur Organisasi .....	8
Gambar 2.2 CRISP-DM .....	18
Gambar 3.3 Flowchart kerja Sistem .....	39
Gambar 3.4 Konteks Diagram .....	40
Gambar 3.5 DFD Level 1 .....	41
Gambar 3.6 DFD Level 2 Persiapan data .....	41
Gambar 3.7 DFD Level 2 Penentuan Kombinasi Data .....	42
Gambar 4.1 Halaman Depan Aplikasi .....	66
Gambar 4.2 Data Kategori .....	66
Gambar 4.3 Data Buku .....	67
Gambar 4.4 Data Penjualan .....	68
Gambar 4.5 Pemasukan Nilai Support dan Confidence .....	69
Gambar 4.6 Data Item Yang Terindeks .....	71
Gambar 4.7 Data Item dengan Nilai Support .....	72
Gambar 4.8 Frequent Item Set Priority .....	73
Gambar 4.9 Rekomendasi Judul Buku .....	74
Gambar 4.8 Hasil pembentukan FP-tree setelah pembacaan TID T1 .....	53
Gambar 4.9 Hasil Pembentukan FP-Tree setelah pembacaan TID T2 .....	53
Gambar 4.10 Hasil Pembentukan FP-Tree setelah pembacaan TID T3 .....	54
Gambar 4.11 Hasil Pembentukan FP-Tree setelah pembacaan TID T4 .....	54
Gambar 4.12 Hasil Pembentukan FP-Tree setelah pembacaan TID T5 .....	55
Gambar 4.13 Hasil Pembentukan FP-Tree setelah pembacaan TID T6 .....	55
Gambar 4.14 Hasil Pembentukan FP-Tree setelah pembacaan TID T7 .....	56
Gambar 4.15 Hasil Pembentukan FP-Tree Setelah Pembacaan TID T8 .....	56
Gambar 4.16 Hasil Pembentukan FP-Tree Setelah Pembacaan TID T9 .....	57
Gambar 4.17 Hasil Pembentukan FP-Tree Setelah Pembacaan TID T10 .....	57
Gambar 4.18 Struktur Fp-Tree Beserta Pointer .....	58
Gambar 4.19 Frequent Itemsets .....	62

## DAFTAR TABEL

Tabel 4.1 Struktur Data Transaksi	44
Tabel 4.2 Contoh Data Faktur Penjualan Maret 2018	45
Tabel 4.3 Data Transaksi setelah pembersihan	46
Tabel 4.4 Data training	47
Tabel 4.5 Transformasi Kode Faktur ke TID	47
Tabel 4.6 Transformasi judul buku kode buku	48
Tabel 4.7 Sampel faktur penjualan	49
Tabel 4.8 Hasil transformasi data transaksi pengunjung	50
Tabel 4.9 Frekuensi kemunculan tiap item	51
Tabel 4.10 Item yang memenuhi support sistem	52
Tabel 4.11 Tabel Data Transaksi	52



## ABSTRAK

Hadi, Agung. 2018. *Rancang Bangun Aplikasi Rekomendasi Judul Buku Menggunakan Algoritma FP-Growth*. Skripsi. Teknik Informatika. Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang.  
Pembimbing: (I) Dr. Suhartono, M.Kom (II) M. Imamuddin, Lc., MA

---

**Kata Kunci** : Fp-Growth, Judul Buku, Data Mining, Rekomendasi, Penerbit, Association Rule.

*Kemajuan teknologi membuat para pelaku bisnis berusaha memanfaatkannya untuk mempermudah dan memajukan usahanya. Selaras dengan perkembangan daya beli konsumen, menuntut pelaku bisnis menerapkan strategi pemasaran yang lebih baik dari pesaingnya. CV. Cita Intrans Selaras (CIS) merupakan salah satu perusahaan yang bergerak di bidang bisnis penerbitan dengan nama dagang Intrans Publishing produk utamanya yaitu buku. Pada Intrans Publishing, terdapat beberapa permasalahan yang kerap muncul mengenai penjualan buku. Banyak buku dengan berbagai judul yang masih belum terjual, butuh waktu lama bahkan ada yang 1 tahun baru habis terjual. Selain itu perusahaan masih kesulitan mendapatkan informasi strategis seperti kurangnya pengetahuan untuk mengenali kondisi pasar dan selera konsumen. Berdasarkan penerapan algoritma FP-Growth kedalam sistem berbasis web, sistem ini hanya menampilkan dua kombinasi itemset yang frequent dengan menggunakan 774 data, dan terbentuk aturan asosiasi dengan nilai support terbesar 30% dan confidence sebesar 50%. Jika Membeli Pengantar Manajemen maka akan membeli Hukum Islam- Indah Purbasari*

## ABSTRACT

Hadi, Agung. 2018. *Rancang Bangun Aplikasi Rekomendasi Judul Buku Menggunakan Algoritma FP-Growth*. Undergraduate Thesis. Department of Informatics Engineering. Faculty of Science and Technology. State Islamic University of Maulana Malik Ibrahim Malang  
Pembimbing: (I) Dr. Suhartono, M.Kom (II) M. Imamuddin, Lc., MA

---

**Key words:** Fp-Growth, Judul Buku, Data Mining, Rekomendasi, Penerbit, Association Rule.

*Technological advances make business people try to use it to simplify and advance business. Aligned with the development of consumer purchasing power, demanding business actors implement better marketing strategies than their competitors. Cita Intrans Selaras (CIS) is one of the companies engaged in the business of publishing under the trade name Intrans Publishing its main product is buku. Pada Intrans Publishing, there are some problems that often arise about the sale of books. Many books with titles are still unsold, it took a long time even there is a new 1 year sold out. In addition, companies are still having difficulty obtaining strategic information such as lack of knowledge to recognize market conditions and consumer tastes. Based on the application of FP-Growth algorithm into web-based system, this system only shows two combination of frequent itemsset with 774 data, and formed association rules with the greatest support value 30% and confidence of 50%. If Buying Introduction Management will then buy Islamic Law- Beautiful Purbasari*

## المخلص

هادي ، أجونج. 2018. تصميم توصية تطبيق عنوان الكتاب باستخدام الخوارزمية FP-Growth. البحث العلمي. قسم المعلوماتية. كلية العلوم والتكنولوجيا. جامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج.  
**المشرف :** (الأول) الدكتور سوهارتونو درجة الماجستير في الهندسة (الثاني) في. إمام الدين ، درجة الماجستير في الهندسة

التقدم التكنولوجي جعل رجال الأعمال يحاولون استخدامه لتبسيط الأعمال ودفعها. من خلال تطوير القوة الشرائية للمستهلكين ، فإن الجهات الفاعلة في قطاع الأعمال تطلب تنفيذ استراتيجيات تسويقية أفضل من منافسيها. شركة Cita Intrans Selaras (CIS) هي واحدة من الشركات التي تعمل في مجال النشر تحت الاسم التجاري Intrans النشر منتجها الرئيسي هو كتاب. Intrans Publishing ، هناك بعض المشاكل التي غالباً ما تنشأ عن بيع الكتب. العديد من الكتب التي تحمل عناوين لا تزال غير مبيعة ، استغرق الأمر وقتاً طويلاً حتى لو تم بيع عام جديد. بالإضافة إلى ذلك ، لا تزال الشركات تواجه صعوبة في الحصول على معلومات إستراتيجية مثل نقص المعرفة للتعرف على ظروف السوق وأذواق المستهلكين. استناداً إلى تطبيق خوارزمية FP-Growth في نظام قائم على الويب ، فإن هذا النظام لا يعرض سوى مزيجين من itemsset المتكرر مع 774 بيانات ، وتشكل قواعد الارتباط بأكثر قيمة دعم 30% وثقة 50%. إذا اشترت إدارة الشراء بعد ذلك ، فإن الشريعة الإسلامية - جميلةPurbasari

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kemajuan teknologi membuat para pelaku bisnis berusaha memanfaatkannya untuk mempermudah dan memajukan usahanya. Selaras dengan perkembangan daya beli konsumen, menuntut pelaku bisnis menerapkan strategi pemasaran yang lebih baik dari pesaingnya.

Dalam persaingan bisnis serta upaya meningkatkan pendapatan perusahaan. Data menjadi sangat berharga, untuk digali pengetahuannya. Namun pada kenyataannya, masih sedikit yang sadar akan hal itu. Data transaksi yang besar yang tersimpan di database, seringkali hanya sebagai arsip tanpa ada upaya menggali pengetahuan dari data yang terkumpul.

Para eksekutif yang menyadari pentingnya data saat ini, berharap terdapat teknologi yang mampu menghasilkan informasi yang dapat digunakan oleh mereka untuk membantu memberikan keputusan yang strategis. Sebagai pemimpin menentukan dan memberikan keputusan akan memberikan dampak kepada kegiatan atau aktivitas selanjutnya. Maka dari itu diperlukan keputusan yang tepat berdasarkan teori-teori yang ada juga berlandaskan Al-Qur'an seperti yang tertulis dalam surat AL-Ahqaaf ayat 19.

وَلِكُلِّ دَرَجَاتٍ مِمَّا عَمِلُوا وَلِيُؤْفِقَهُمْ أَعْمَالَهُمْ وَهُمْ لَا يُظْلَمُونَ

Artinya ; “Dan bagi masing-masing mereka derajat menurut apa yang telah mereka kerjakan dan agar Allah mencukupkan bagi mereka (balasan) pekerjaan-pekerjaan mereka sedang mereka tiada dirugikan”.

Percepatan teknologi memberikan kontribusi cepatnya pertumbuhan data yang ada. Sehingga ketersediaan data yang banyak dan kebutuhan akan informasi atau pengetahuan sebagai pendukung pengambil keputusan untuk membuat solusi bisnis dengan dukungan infrastruktur di bidang teknik informatika merupakan cikal-bakal dari lahirnya teknologi data mining. Data mining merupakan penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar (Enur Irdiansyah,2010)

Banyak peneliti yang mengambil topik penelitian *data mining*. Fajrin (2018) di daalam penelitiannya mengenai pola pembelian Spare part Motor menggunakan algoritma FP-Growth untuk mengetahui spare part jenis apa yang sering dibeli. Sehingga dalam penentuan pembelian stock tidak berlebihan. Demikian pula Ahmad (2012) dalam penelitiannya memanfaatkan data transaksi yang dianalisis menggunakan algoritma FP-Growth

Dengan menggunakan metode data mining diharapkan dapat membantu perusahaan dalam pengelolaan informasi yang tersimpan dalam database berupa data transaksi hingga menjadi sebuah pengetahuan (knowledge) baru. Dari pengetahuan baru yang didapatkan, perusahaan dapat meningkatkan pendapatan serta yang paling utama loyalitas pelanggan.

Foster dan Cadogan mengatakan bahwa loyalitas pelanggan akan melahirkan perilaku dan tindakan pelanggan seperti (1) perilaku pelanggan yang bersifat

memberikan rekomendasi untuk mengajak orang lain untuk melakukan pembelian atau menggunakan produk tersebut. (2) Pelanggan akan melakukan aktivitas transaksi atau mempergunakan segala bentuk layanan yang ditawarkan oleh pihak perusahaan. (3) Pelanggan akan menjadikan perusahaan tersebut sebagai pilihan pertama dalam menggunakan jasa keuangan. (4) Word of mouth yaitu perilaku pelanggan untuk membicarakan hal – hal yang bagus terhadap produk dari perusahaan tersebut kepada orang lain (Brian D. Foster, 2000).

CV. Cita Intrans Selaras (CIS) merupakan salah satu perusahaan yang bergerak di bidang bisnis penerbitan dengan nama dagang *Intrans Publishing* produk utamanya yaitu buku. Dalam rangka mengikuti perkembangan, Intrans Publishing selalu melakukan inovasi dan kreatifitas, dengan memenuhi kebutuhan produksi buku-buku bacaan untuk masyarakat umum dan perguruan tinggi.

Pada Intrans Publishing,terdapat beberapa permasalahan yang kerap muncul mengenai penjualan buku. Banyak buku dengan berbagai judul yang masih belum terjual, butuh waktu lama bahkan ada yang 1 tahun baru habis terjual. Selain itu perusahaan masih kesulitan mendapatkan informasi strategis seperti kurangnya pengetahuan untuk mengenali kondisi pasar dan selera konsumen, serta persaingan yang ketat dengan para pesaing bisnis yang sama.

Untuk itu penulis tertarik mengangkat permasalahan ini kedalam penelitian skripsi dengan judul “RANCANG BANGUN APLIKASI REKOMENDASI JUDUL BUKU MENGGUNAKAN ALGORITMA FP-GROWTH PADA CV.CITA INTRANS SELARAS”. Di harapkan dengan penelitian ini data yang selama ini tersimpan tanpa ada tindakan untuk menganalisisnya dapat diolah untuk diambil pengetahuannya.

## 1.2 Rumusan Masalah

Setelah melihat latar belakang masalah yang telah diuraikan di atas maka dapat diperoleh rumusan masalah yaitu seberapa efektif algoritma Fp-Growth dalam merekomendasikan judul buku pada CV.Cita Intrans Selaras.

## 1.3 Tujuan Penelitian

Penelitian ini dilakukan untuk memberi kontribusi dan pemahaman kepada pembaca agar dapat mengetahui hal terkait penggalian-penggalian data mentah, dengan tujuan yang lebih khusus yaitu:

1. Mengetahui bagaimana membangun aplikasi data mining berbasis web
2. Mengetahui efektifitas rekomendasi judul buku dengan menggunakan metode asosiasi algoritma FP-GROWTH.
3. Mengetahui pola pembelian produk buku yang sering dibeli secara bersamaan.
4. Membantu marketing dalam menentukan strategi yang tepat untuk penjualan buku, dengan memberikan rekomendasi pola buku yang berdekatan dengan frekuensi beli tinggi.

## 1.4 Batasan Masalah

Agar tidak melebar pembahasannya, kami tetapkan batasan masalah :

1. Tempat yang akan diteliti adalah CV. Cita Intrans Selaras
2. Data yang diteliti yaitu data transaksi penjualan berupa faktur penjualan.

### 1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah :

1. Dapat mempermudah pemimpin perusahaan mengambil keputusan.
2. Untuk mengatur display buku yang sering dibeli secara bersamaan dengan meletakkannya berdekatan.
3. Membantu marketing memasarkan produk buku secara tepat.



## BAB II

### TINJAUAN PUSTAKA

#### 2.1 Penelitian Terkait

Penelitian yang dilakukan Erwin, melakukan analisis market basket dengan algoritma Apriori dan FP-Growth menyimpulkan bahwa algoritma Apriori membutuhkan waktu komputasi yang lama untuk mendapatkan frequent itemsets, karena berulang kali melakukan pemindahan data dan membutuhkan memori yang besar sedangkan FP-Growth hanya membutuhkan dua kali scanning database dalam mencari frequent itemsets sehingga waktu yang dibutuhkan menjadi relatif singkat dan efisien (Erwin, 2009).

Menurut Rama Novta Miraldi, dkk dalam penelitiannya *Implementasi Algoritma FP-Growth untuk sistem Rekomendasi Buku di perpustakaan UKDW*. Dengan memberikan feature perkomendasi relasi buku yang sedang dicari. Feature ini membantu pengguna untuk menemukan literatur atau topik yang berhubungan. Algoritma FP-Growth dapat digunakan dalam kasus perkomendasi buku dan mendapatkan hasil keluaran berupa rekomendasi buku yang terkait, selain itu tingkat keakuratan FP-Growth dalam memberikan rekomendasi buku adalah sebesar 60,78% (Rama Novta Miraldi, 2014).

Penelitian yang lain dengan judul *Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen dengan Algoritma FP-Growth pada Data Transaksi Penjualan Spare Part Motor*. Dengan dapat mengetahui secara cepat hasil penjualan dari spare part motor sport yang paling banyak terjual pada perusahaan yang mempunyai banyak

cabang atau dealer. Akan memberikan keuntungan kepada pemilik usaha serta dengan cepat menentukan strategi penjualan (Alfannisa Annurullah Fajrin, 2018).

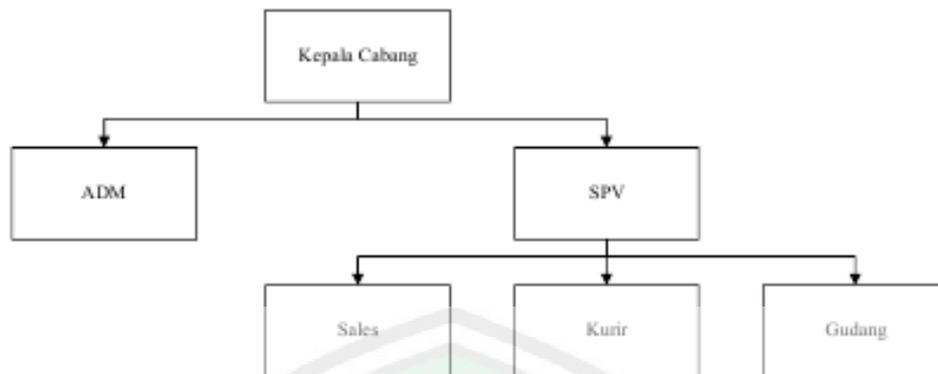
Penelitian terkait algoritma FP-Growth yang lain yaitu mengenai *Penentuan Pola Yang Sering Muncul Untuk Penerima Kartu Jaminan Kesehatan Masyarakat (Jamkesmas) Menggunakan Metode Fp-Growth*. Ketidaktepatan penilaian tim dalam memberikan penilaian terhadap pasien JAMKESMAS merupakan sumber masalah utama. Dengan menggunakan Algoritma Fp-Growth analisa data yang dihasilkan memperoleh pengetahuan pola yang sering muncul pada penerima JAMKESMAS di kelurahan Bambe kecamatan Driyorejoyang tepat sasaran masih 60% dikarenakan masih banyak penduduk yang kurang memenuhi kriteria (Meilani, 2015).

## 2.2 Intrans Publishing

Agustus 2003 Intrans Publishing didirikan dengan nama lengkap CV. Cita Intrans Selaras (CIS) dengan nama dagang Intrans Publishing yang mengusung semangat pengembangan ilmu pengetahuan dan mendorong wacana – wacana perubahan sosial, politik, ekonomi, hukum, dan budaya ke hadapan publik.

Kelompok Intrans Publishing dengan motto “*Rebut perubahan dengan Membaca*”, merupakan semangat untuk menggairahkan wacana – wacana perubahan dan pengembangan ilmu pengetahuan di Indonesia bersama dengan masyarakat (pembaca) untuk secara bersama – sama melakukannya sebagai bentuk pengabdian kepada bangsa dan negeri ini.

Perusahaan ini memproduksi buku – buku yang selalu disesuaikan dengan perkembangan ilmu pengetahuan dan disesuaikan dengan kebutuhan bahan bacaan yang baik dan berkualitas.



**Gambar 2.1** Struktur Organisasi

Deskripsi pekerjaan :

1. Kepala cabang memiliki tanggung jawab dan wewenang sebagai berikut;
  - a. Monitor proses bisnis perusahaan cabang
  - b. Menyusun dan menetapkan tujuan perusahaan cabang
  - c. Mengelola dan mengembangkan SDM
2. Administrasi memiliki tanggung jawab dan wewenang sebagai berikut;
  - a. Menerima tagihan pembayaran yang sudah dilakukan oleh sales
  - b. Memeriksa absensi karyawan untuk penentuan besarnya gaji yang harus dibayar.
  - c. Membuat laporan keuangan dan perpajakan.
3. Supervisor memiliki tanggung jawab dan wewenang sebagai berikut;
  - a. Mengatur kerjanya para staf bawahannya
  - b. Bertanggung jawab atas hasil kerja staf
  - c. Memberi motivasi kerja kepada staf bawahannya
  - d. Memberikan briefing bersama staf
  - e. Membuat planing pekerjaan harian, mingguan, bulanan, dan tahunan.

- f. Menyelesaikan masalah sebisaanya tanpa harus ditangani oleh atasan
  - g. Penghubung antara staf dan kepala cabang
  - h. Membantu tugas staf bawahan
  - i. Menampung segala keluhan dari customer yang disampaikan melalui staf untuk disampaikan ke kepala cabang.
4. Sales memiliki tugas dan fungsi sebagai berikut;
- a. Dapat mengarahkan sasaran mana dan kepada siapa produk akan ditawarkan dan dijual.
  - b. Dapat meyakinkan calon pelanggan atas manfaat dan kelebihan produk yang ditawarkan.
  - c. Dapat meyakinkan calon pelanggan yang diketahui ragu-ragu dalam mengambil keputusan atau menentukan pilihan.
5. Kurir memiliki tugas dan fungsi mengantarkan barang dengan penuh tanggung jawab dan tepat waktu.
6. Bagian gudang memiliki tugas dan fungsi sebagai berikut;
- a. Memeriksa barang – barang yang tersedia.
  - b. Menyiapkan barang - barang dengan benar dan cepat sesuai dengan Sales Order ( SO ) dan Surat Pengantar ( SP ) yang diterima.
  - c. Packing barang - barang yang akan dikirim dengan benar.
  - d. Membantu cek perlengkapan, jumlah, kondisi atas barang - barang yang akan dikirim.
  - e. Turut menjaga kebersihan, keselamatan kerja dan keutuhan barang di dalam gudang.

## 2.3 Landasan Teori

### 2.3.1 Pengertian Data Mining

*Data Mining* (DM) adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala besar yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Definisi umum dari DM itu sendiri adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data.

Menurut Abdul Kadir (2007:30), Data mining adalah kombinasi secara logis antara pengetahuan data, dan analisa statistik yang dikembangkan dalam pengetahuan bisnis atau suatu proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, tiruan dan *machine-learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat bagi pengetahuan yang terkait dari berbagai database besar.

*Data mining* meliputi tugas-tugas yang dikenal sebagai ekstraksi pengetahuan, arkeologi data, eksplorasi dalam pemrosesan pola data dan memanen informasi. Semua aktifitas ini dilakukan secara otomatis dan mengizinkan adanya penemuan cepat pengetahuan baru bahkan oleh non programmer. *Data mining* cerdas menemukan informasi di dalam data warehouse dimana laporan dan *query* tidak bisa diungkapkan secara efektif. Piranti data mining menemukan pola-pola di dalam data dan bahkan menyimpulkan aturan dari data tersebut Definisi umum dari Data Mining adalah proses pencarian pola-pola yang menarik (hidden pattern) berupa pengetahuan

(knowledge) yang tidak diketahui sebelumnya dari suatu kumpulan data dimana data tersebut dapat berada dalam database, *data warehouse*, atau media penyimpanan informasi yang lain. Data Mining merupakan proses analisis terhadap data dengan penekanan menemukan informasi yang tersembunyi pada sejumlah besar data yang disimpan ketika menjalankan bisnis perusahaan. Sumber : (Data Mining dan Web Mining) (andyku).

Pada dasarnya data mining berhubungan erat dengan analisis data dan penggunaan perangkat lunak untuk mencari pola dan kesamaan dalam sekumpulan data. Ide dasarnya adalah menggali sumber yang berharga dari suatu tempat yang sama sekali tidak diduga, seperti perangkat lunak data mining mengekstraksi pola yang sebelumnya tidak terlihat atau tidak begitu jelas sehingga tidak seorang pun yang memperhatikan sebelumnya. Analisa data mining berjalan pada data yang cenderung terus membesar dan teknik terbaik yang digunakan kemudian berorientasi kepada data berukuran sangat besar untuk mendapatkan kesimpulan dan keputusan paling layak. Data mining memiliki beberapa sebutan atau nama lain yaitu : *knowledge discovery in database* (KDD), ekstraksi pengetahuan (*knowledge extraction*), analisa data / pola (*data / pattern analysis*), kecerdasan bisnis (*business intelligence*), *data archaeology* dan *data dredging* (Larose, 2005)

### 2.3.2 Metode Data mining

Secara garis besar, Han dalam bukunya menjelaskan bahwa metode data mining dapat dilihat dari dua sudut pandang pendekatan yang berbeda, yaitu pendekatan deskriptif dan pendekatan prediktif [8]. Pendekatan deskriptif

adalah pendekatan dengan cara mendeskripsikan data inputan. Metode yang termasuk ke dalam pendekatan ini adalah :

1. Metode deskripsi konsep/kelas, yaitu data dapat diasosiasikan dengan kelas atau konsep. Ada tiga macam pendeskripsian yaitu (1) karakteristik data, dengan membuat summary karakter umum atau fitur data suatu kelas target, (2) diskriminasi data, dengan membandingkan class target dengan satu atau sekelompok kelas pembanding, (3) gabungkan antara karakterisasi dan diskriminasi.
2. Metode association rule, yaitu menemukan aturan asosiatif atau pola kombinasi dari suatu item yang sering terjadi dalam sebuah data. Pendekatan kedua adalah pendekatan prediktif, yaitu pendekatan yang dapat digunakan untuk memprediksi, dengan hasil berupa kelas atau cluster. Metode yang termasuk dalam pendekatan ini adalah :
3. Metode klasifikasi dan prediksi, yaitu metode analisis data yang digunakan untuk membentuk model yang mendeskripsikan kelas data yang penting, atau model yang memprediksikan trend data. Klasifikasi digunakan untuk memprediksi kelas data yang bersifat kategorial, sedangkan prediksi untuk memodelkan fungsi yang mempunyai nilai kontinu.
4. Metode clustering, mengelompokkan data untuk membentuk kelas-kelas baru atau sering disebut cluster. Metode clustering bertujuan untuk memaksimalkan persamaan dalam satu cluster dan meminimalkan perbedaan antar cluster.

### 2.3.3 Metode Association Rule

Association Rule atau Aturan Asosiasi adalah teknik Data mining untuk menemukan aturan asosiatif atau pola kombinasi dari suatu item. Bila kita mengambil contoh aturan asosiatif dalam suatu transaksi pembelian barang di suatu Toko Buku adalah kita dapat mengetahui berapa besar kemungkinan seorang konsumen membeli suatu item bersamaan dengan item lainnya (membeli Buku Islam bersama dengan Buku Ekonomi). Karena awalnya berasal dari studi tentang database transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama apa, maka *association rule* sering juga dinamakan *market basket analysis*.

*Association Rule* adalah bentuk jika “kejadian sebelumnya” kemudian “konsekuensinya” (If antecedent, then consequent), yang diikuti dengan perhitungan aturan support dan confidence . Bentuk umum dari *association rule* adalah Antecedent -> Consequent. Bila kita ambil contoh dalam sebuah transaksi pembelian barang di sebuah Toko Buku didapat bentuk *association rule* Buku Islam -> Buku Ekonomi. Yang artinya bahwa pelanggan yang membeli Buku Islam ada kemungkinan pelanggan tersebut juga akan membeli Buku Ekonomi, dimana tidak ada batasan dalam jumlah item-item pada bagian *antecedent* ataupun *consequent* dalam sebuah rule.

*Association rule* memiliki dua tahap pengerjaan, yaitu :

1. Mencari *frequent itemset*.
2. Mendefinisikan *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* : suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item/itemset dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu item/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* : suatu ukuran yang menunjukkan hubungan antara 2 item secara conditional (misal, menghitung kemungkinan seberapa sering item B dibeli oleh pelanggan jika pelanggan tersebut membeli sebuah item A).
3. Kedua ukuran ini nantinya berguna dalam menentukan kekuatan suatu pola dengan membandingkan pola tersebut dengan nilai minimum kedua parameter tersebut yang ditentukan oleh pengguna. Bila suatu pola memenuhi kedua nilai minimum parameter yang sudah ditentukan sebelumnya, maka pola tersebut dapat disebut sebagai *interesting rule* atau *strong rule*.

### 2.3.4 Tahapan Data Mining

Menurut Syaifullah (2010:15), Dalam aplikasinya, data mining sebenarnya merupakan bagian dari proses *Knowledge Discovery in Database* (KDD), bukan sebagai teknologi yang utuh dan berdiri sendiri. Data mining merupakan suatu bagian langkah yang penting dalam proses KDD terutama berkaitan dengan ekstraksi dan dari data yang ditelaah, seperti ditunjukkan oleh gambar 1 dibawah ini



Gambar1. Knowledge discovery from data

#### 1. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai. Data hasil seleksi yang digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

## 2. Pre-processing/Cleaning

Sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi focus KDD. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu “memperkaya” data yang sudah ada dengan data informasi atau informasi lain yang relevan dan diperlukan untuk

### a. Transformasi

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

## 3. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

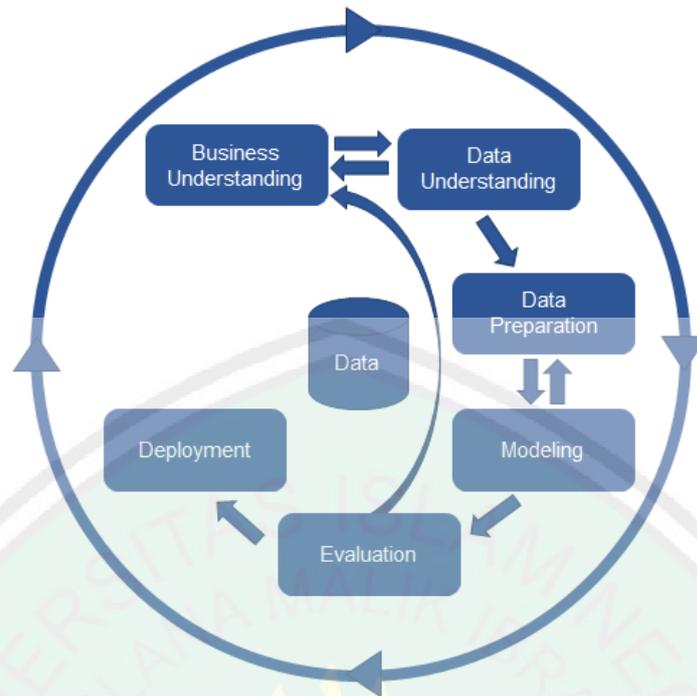
#### 4. Interpretation/Evaluation

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KKD yang disebut dengan interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

#### 2.3.5 Cross Industry Standard Process for Data Mining

Cross Industry Standard for Data Mining (CRISP – DM) yang dikembangkan tahun 1996 oleh analis dari beberapa industry seperti Daimbler Chrysler, SPSS, dan NCR. CRISP DM menyediakan standar proses data mining sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian.

Dalam CRISP – DM, sebuah proyek data mining memiliki siklus hidup yang terbagi dalam enam fase. Keseluruhan fase berurutan yang ada tersebut bersifat adaptif. Fase berikutnya dalam urutan bergantung kepada keluaran dari fase sebelumnya. Hubungan penting antarfase digambarkan dengan panah. Sebagai contoh, jika proses berada pada fase modelling. Berdasar pada perilaku dan karakteristik model, proses mungkin harus kembali kepada fase data preparation untuk perbaikan lebih lanjut terhadap data atau berpindah maju kepada fase evaluation.



Gambar 2.2 CRISP-DM

Enam fase CRISP – DM (Larose, 2005) :

1. Fase Pemahaman Bisnis (Business Understanding Phase)
  - a. Penentuan tujuan objek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
  - b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan data mining.
  - c. Menyiapkan strategi awal untuk mencapai tujuan.
2. Fase Pemahaman Data (Data Understanding Phase)
  - a. Mengumpulkan data.
  - b. Menggunakan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
  - c. Mengevaluasi kualitas data.

- d. Jika diinginkan, pilih sebagian kecil grup data yang mungkin mengandung pola dari permasalahan.

### 3. Fase Pengolahan Data (Data Preparation Phase)

- a. Siapkan dari data awal, kumpulkan data yang ingin digunakan untuk keseluruhan fase berikutnya. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif.
- b. Pilih kasus dan variabel yang ingin dianalisis dan yang sesuai analisis yang akan dilakukan.
- c. Lakukan perubahan pada beberapa variabel jika dibutuhkan.
- d. Siapkan data awal sehingga siap untuk perangkat pemodelan.

### 4. Fase Pemodelan (Modelling Phase)

- a. Pilih dan aplikasikan teknik pemodelan yang sesuai.
- b. Kalibrasi aturan model untuk mengoptimalkan hasil.
- c. Perlu diperhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan data mining yang sama.
- d. Jika diperlukan, proses dapat kembali ke fase pengolahan data untuk menjadikan data ke dalam bentuk yang sesuai dengan spesifikasi kebutuhan teknik data mining tertentu.

### 5. Fase Evaluasi (Evaluation Phase)

- a. Mengevaluasi satu atau lebih model yang digunakan dalam fase pemodelan untuk mendapatkan kualitas dan efektifitas sebelum disebarkan untuk digunakan.
- b. Menetapkan apakah terdapat model yang memenuhi tujuan pada fase awal.

- c. Menentukan apakah terdapat permasalahan penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan berkaitan dengan penggunaan hasil dari data mining.

#### 6. Fase Penyebaran (Deployment Phase)

- a. Menggunakan model yang dihasilkan. Terbentuknya model tidak menandakan telah terselesaikannya proyek.
- b. Contoh sederhana penyebaran : Pembuatan laporan.
- c. Contoh kompleks penyebaran : Penerapan proses data mining secara paralel pada departemen lain.

#### 2.3.6 Metode Asosiasi

Salah satu pengaplikasian dari *Association Rule* adalah *Market Basket Analysis*, dimana bertujuan untuk menemukan bagaimana item yang dibeli oleh pelanggan dalam supermarket atau toko saling berhubungan. Pencarian *Association Rules* dilakukan melalui dua tahap yaitu pencarian *frequent itemset* dan penyusunan rules. Penting tidaknya suatu *Association Rules* dapat diketahui dengan dua parameter, yaitu *support* (nilai penunjang) dan *confidence* (nilai kepastian). *Support* adalah ukuran yang menunjukkan tingkat dominasi itemset dari keseluruhan transaksi. Persamaan I untuk menentukan nilai suatu support

Ada 2 (dua) proses utama dalam penggalian aturan asosiasi, yaitu :

pencarian pola (frequent pattern) dari sejumlah transaksi penentuan kuatnya rule (aturan) dari pola yang dihasilkan. Dalam association rule mining ada 2 (dua) hal yang mempengaruhi :

### 1. Support

Support adalah proporsi suatu item dalam semua transaksi. Support dirumuskan sebagai berikut:

$$supp(x) = \frac{\sum x}{\sum transaksi}$$

### 2. Confidence

Confidence adalah hal yang mendasari aturan asosiasi, dengan konsep implikasi ( $x \Rightarrow y$ ), atau “if ... then ...”. Besarnya nilai confidence suatu aturan (rule) dirumuskan sebagai berikut :

$$conf(x \Rightarrow y) = \frac{\sum xy}{\sum x}$$

Nilai minimum support dan minimum confidence bisa diatur oleh pengguna ataupun ahli yang berkaitan (Han, Jiawei, 2012).

### 2.3.7 Lift

Korelasi pada suatu rule tidak hanya diukur pada nilai support dan confidence tetapi juga oleh korelasi antara kedua itemsets. Ada banyak ukuran korelasi yang dapat dipilih, salah satunya adalah lift. Lift adalah salah satu contoh ukuran korelasi sederhana. Nilai lift antara A dan B dapat diukur dengan menggunakan komputasi sebagai berikut (Tan, 2004)

### 2.3.8 FP-Tree

Frequent Pattern Tree (FP-Tree) adalah representasi pemasukan data yang dipadatkan (Kumar, 2004). FP-Tree dibentuk dengan membaca kumpulan data pada suatu transaksi dalam suatu waktu dan memetakan transaksi tersebut ke dalam lintasan FP-Tree. Transaksi yang berbeda dapat memiliki items yang sama, sehingga memungkinkan lintasannya saling menimpa. Semakin banyak lintasan yang menimpa satu sama lain, dengan menggunakan struktur FP-Tree, maka proses pemadatan akan semakin baik. Dalam, "Mining Frequent Patterns without Candidate Generation: A Frequent Pattern Tree Approach" (Han, Pei, Yin, & Mao, 2004) menjelaskan misal:  $\{dt, az, \dots, a\}$  adalah kumpulan dari item, dan database transaksi DB :  $\{T_1, T_2, \dots, T_n\}$ , dimana  $T_i$  ( $i \in [1..n]$ ) adalah sekumpulan transaksi yang mengandung item di I. Sedangkan support adalah penghitung (counter) frekuensi kemunculan transaksi yang mengandung suatu-pola. Suatu pola dikatakan sering muncul (frequent pattern) apabila support dari pola tersebut tidak kurang dari suatu konstanta minimum support (batas ambang minimum support) yang telah di definisikan sebelumnya. Permasalahan mencari pola frequent dengan batas ambang minimum support count, inilah yang dicoba untuk dipecahkan oleh FP-Growth dengan bantuan struktur FP-tree.

### 2.3.9 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. Pada algoritma Apriori diperlukan generate candidate untuk mendapatkan frequent itemsets. Akan tetapi, di algoritma FP-Growth generate candidate tidak dilakukan karena FP-Growth menggunakan konsep pembangunan tree dalam pencarian frequent itemsets. Hal tersebutlah yang menyebabkan algoritma FP-Growth lebih cepat dari algoritma Apriori.

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.

Metode FP-Growth dapat dibagi menjadi 3 tahapan utama yaitu sebagai (Erwin, 2009):

- 1) Tahap pembangkitan *conditional pattern base*

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

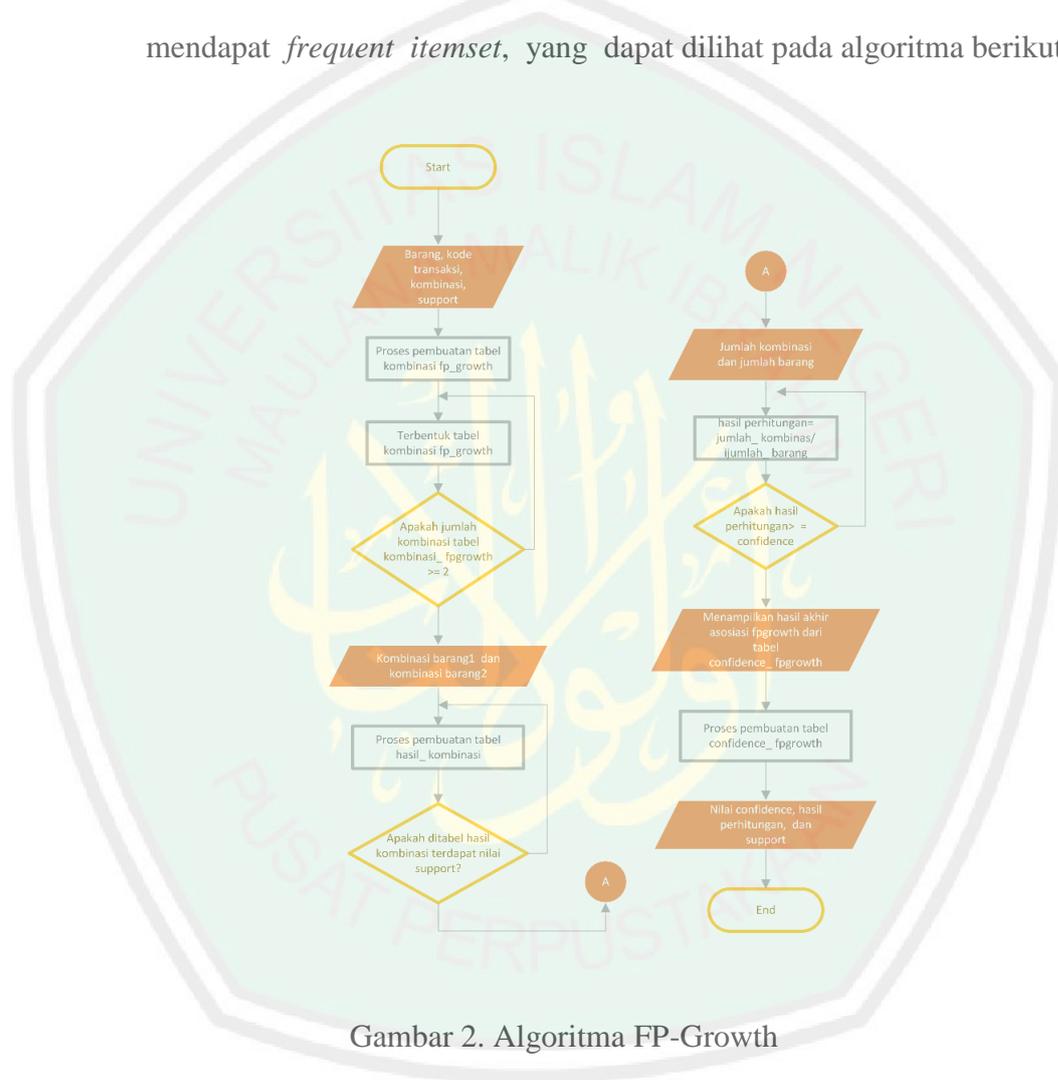
- 2) Tahap pembangkitan conditional FP-Tree

Pada tahap ini, *support count* dari setiap item pada setiap conditional pattern base dijumlahkan, lalu setiap item yang memiliki jumlah support count lebih besar sama dengan minimum *support count* akan dibangkitkan dengan Conditional FP-Tree

### 3) Tahap pencarian frequent itemset

Apabila Conditional FP-Tree merupakan lintasan tunggal (single path), maka didapatkan frequent itemset dengan melakukan kombinasi item untuk setiap Conditional FP-Tree.

Ketiga tahap tersebut merupakan langkah yang akan dilakukan untuk mendapat *frequent itemset*, yang dapat dilihat pada algoritma berikut :



Gambar 2. Algoritma FP-Growth

## 2.4 Basis Data

Basis Data terdiri atas 2 kata, yaitu Basis dan Data. Basis kurang lebih dapat diartikan sebagai markas atau gudang, tempat bersarang/berkumpul. Sedangkan Data adalah representasi fakta dunia nyata mewakili suatu objek seperti manusia

(pegawai, siswa, pembeli, pelanggan), barang, hewan, peristiwa, konsep, keadaan, dan sebagainya, yang diwujudkan dalam bentuk angka, huruf, simbol, teks, gambar, bunyi, atau kombinasinya. Sebagai satu kesatuan istilah, Basis Data (Database) sendiri dapat didefinisikan dalam sejumlah sudut pandang seperti:

- a. Himpunan kelompok data (arsip) yang saling berhubungan yang diorganisasi sedemikian rupa agar kelak dapat dimanfaatkan kembali dengan cepat dan mudah.
- b. Kumpulan data yang saling berhubungan yang disimpan secara bersamaan sedemikian rupa dan tanpa pengulangan (redudansi) yang tidak perlu, untuk memenuhi berbagai kebutuhan.
- c. Kumpulan file/ tabel/ arsip yang saling berhubungan yang disimpan dalam media penyimpan elektronik.

#### **2.4.1 Operasi Dasar Basis Data**

Di dalam sebuah disk, basis data dapat diciptakan dan dapat pula dihapuskan. Di dalam sebuah disk, kita dapat pula menempatkan beberapa (lebih dari satu) basisdata. Sementara dalam sebuah basis data, kita dapat menempatkan satu atau lebih file/tabel. Pada file/tabel inilah sesungguhnya data disimpan/ditempatkan. Setiap basis data umumnya dibuat untuk mewakili sebuah semesta data yang spesifik. Misalnya, ada basis data kepegawaian, basis data akademik, basis data inventori (Pergudangan), dan sebagainya. Sementara dalam basis data akademik, misalnya, kita dapat menempatkan file mahasiswa, file mata\_kuliah, file dosen, file jadwal, file kehadiran, file nilai, dan seterusnya. Karena itu, operasi-operasi dasar yang dapat kita lakukan berkenaan dengan basis data dapat meliputi:

1. Pembuatan basis data baru (create database), yang identik dengan pembuatan lemari arsip yang baru.
  2. Penghapusan basis data (drop database), yang identik dengan perusakan lemari arsip (sekaligus beserta isinya, jika ada).
  3. Pembuatan file/tabel dari suatu basis data (create table), yang identik dengan penambahan map arsip baru ke sebuah lemari arsip yang telah ada.
  4. Penghapusan file/tabel dari suatu basis data (drop table), yang identik dengan perusakan map arsip lama yang ada di sebuah lemari arsip.
  5. Penambahan/pengisian data baru ke sebuah file/tabel di sebuah basis data (insert), yang identik dengan penambahan ke lemari arsip ke sebuah map arsip.
  6. Pengambilan data dari sebuah file/tabel (retrieve/search) yang identik dengan pencarian lembaran arsip dari sebuah map arsip.
  7. Pengubahan data dari sebuah file/tabel (update), yang identik dengan perbaikan isi lembaran arsip yang ada di sebuah map arsip.
  8. Penghapusan data dari sebuah file/tabel (delete), yang identik dengan penghapusan sebuah lembaran arsip yang ada di sebuah map arsip.
- Operasi yang berkenaan dengan pembuatan objek (basis data dan tabel) merupakan operasi awal yang hanya dilakukan sekali dan berlaku seterusnya. Sedang operasi-operasi yang berkaitan dengan isi tabel (data) merupakan operasi rutin yang akan berlangsung berulang-ulang dan karena itu operasi-operasi inilah yang lebih tepat mewakili aktivitas pengelolaan (management) dan pengolahan (processing) data dalam basis data.

### 2.4.2 Objektif Basis Data

Telah disebutkan di awal bahwa tujuan awal dan utama dalam pengelolaan data dalam sebuah basis data adalah agar kita dapat memperoleh/menemukan kembali data (yang kita cari) dengan mudah dan cepat. Di samping itu, pemanfaatan basis data untuk pengelolaan data, juga memiliki tujuan-tujuan lain.

Secara lebih lengkap, pemanfaatan basis data dilakukan untuk memenuhi sejumlah tujuan (objektif) seperti buku ini:

1. Kecepatan dan Kemudahan (Speed) Pemanfaatan basis data memungkinkan kita untuk dapat menyimpan data atau melakukan perubahan/manipulasi terhadap data atau menampilkan kembali data tersebut dengan lebih cepat dan mudah, daripada jika kita menyimpan data secara manual (non elektronik) atau secara elektronik (tetapi tidak dalam bentuk penerapan basis data, misalnya dalam bentuk spread sheet atau dokumen teks biasa).
2. Efisiensi Ruang Penyimpanan (Space) Karena keterkaitan yang erat antar kelompok data dalam sebuah basis data, maka redundansi (pengulangan) data pasti akan selalu ada. Banyaknya redundansi ini tentu akan memperbesar ruang penyimpanan (baik di memori utama maupun memori sekunder) yang harus disediakan. Dengan basis data, efisiensi/optimalisasi penggunaan ruang penyimpanan dapat dilakukan, karena kita dapat melakukan penekanan jumlah redundansi data, baik dengan menerapkan sejumlah pengkodean atau

dengan membuat relasi-relasi (dalam bentuk file) antar kelompok data yang saling berhubungan.

3. Keakuratan (Accuracy) Pemanfaatan pengkodean atau pembentukan relasi antar data bersama dengan penerapan aturan/batasan (constraint) tipe data, domain data, keunikan data, dan sebagainya, yang seara ketat dapat diterapkan dalam sebuah basis data, sangat berguna untuk menekan ketidakakuratan pemasukan/penyimpanan data.
4. Ketersediaan (Availability) Pertumbuhan data (baik dari sisi jumlah maupun jenisnya) sejalan dengan waktu akan semakin membutuhkan ruang penyimpanan yang besar. Padahal tidak semua data itu selalu kita gunakan/butuhkan. Karena itu kita dapat memilah adanya data utama/master/referensi, data transaksi, data histori hingga data kadaluwarsa. Data yang sudah jarang atau bahkan tidak pernah lagi kita gunakan, dapat kita atur untuk dilepaskan dari sistem basis data yang sedang aktif (menjadi off-line) baik dengan cara penghapusan atau dengan memindahkannya ke media penyimpanan off-line (seperti removable disk atau tape). Di sisi lain, karena kepentingan pemakaian data, sebuah basis data dapat memiliki data yang disebar di banyak lokasi geografis. Data nasabah sebuah bank, misalnya, dipisah-pisah dan disimpan di lokasi yang sesuai dengan keberadaan nasabah. Dengan pemanfaatan teknologi jaringan komputer, data yang berada di suatu

lokasi/cabang, dapat juga diakses (menjadi tersedia/available) bagi lokasi/cabang lain.

5. Kelengkapan (Completeness) Lengkap/tidaknya data yang kita kelola dalam sebuah basis data bersifat relatif (baik terhadap kebutuhan pemakai maupun terhadap waktu). Bila seorang pemakai sudah menganggap bahwa data yang dipelihara sudah lengkap, maka pemakai yang lain belum tentu berpendapat sama. Atau, yang sekarang dianggap sudah lengkap, belum tentu di masa yang akan datang juga demikian. Dalam sebuah basis data, di samping data kita juga harus menyimpan struktur (baik yang mendefinisikan objek-objek dalam basis data maupun definisi dari tiap objek, seperti struktur file/tabel atau indeks). Untuk mengakomodasi kebutuhan kelengkapan data yang semakin berkembang, maka kita tidak hanya dapat menambah record- record data, tetapi juga dapat melakukan perubahan struktur dalam basis data, baik dalam bentuk penambahan objek baru (tabel) atau dengan penambahan field-field baru pada suatu tabel.
6. Keamanan (Security) Memang ada sejumlah (aplikasi) pengelola basis data yang tidak menerapkan aspek keamanan dalam penggunaan basis data. Tetapi untuk sistem yang besar dan serius, aspek keamanan juga dapat diterapkan dengan ketat. Dengan begitu kita dapat menentukan siapa-siapa (pemakai) yang boleh menggunakan basis data beserta objek-objek di dalamnya dan menentukan jenis-jenis operasi apa saja yang boleh dilakukannya.

7. Kebersamaan Pemakaian (Sharability) Pemakai basis data sering kali tidak terbatas pada satu pemakai saja, atau di satu lokasi saja atau oleh satu sistem/aplikasi saja. Data pegawai dalam basis data kepegawaian, misalnya, dapat digunakan oleh banyak pemakai, dari sejumlah departemen dalam perusahaan atau oleh banyak sistem (sistem penggajian, sistem akuntansi, sistem inventori, dan sebagainya). Basis data yang dikelola oleh sistem (aplikasi) yang mendukung lingkungan multiuser, akan dapat memenuhi kebutuhan ini, tetapi tetap dengan menjaga/menghindari (karena data yang sama diubah oleh banyak pemakai pada saat yang bersamaan) atau kondisi deadlock (karena ada banyak pemakai yang saling menunggu untuk menggunakan data).

#### **2.4.3 Database Management System**

Kumpulan atau gabungan database dengan perangkat lunak aplikasi yang berbasis database tersebut dinamakan Database Management System (DBMS). DBMS merupakan koleksi terpadu dari database dan program-program komputer (utilitas) yang digunakan untuk mengakses dan memelihara database. Program-program tersebut menyediakan berbagai fasilitas operasi untuk memasukan, melacak, dan memodifikasi data ke dalam database, mendefinisikan data baru, serta mengolah data menjadi informasi yang dibutuhkan (DBMS = Database + Program Utilitas) [4]. Perangkat lunak yang termasuk DBMS seperti dBase, FoxBase, Rbase, Microsoft-Access (sering juga disingkat Ms-Access) dan Borland Pradox (untuk DBMS yang

sederhana) atau Borland-Interbase, MS-Sql, Sever, Oracle Database, IBM, DB2, Informix, Sybase, MySql, PostgreSQL (untuk DBMS yang lebih kompleks dan lengkap).

## 2.5 Unified Modelling Language (UML)

UML singkatan dari Unified Modeling Languages yang berarti bahasa pemodelan standar. Ketika kita membuat model menggunakan konsep UML ada aturan-aturan yang harus diikuti. Bagaimana elemen pada model-model yang kita buat berhubungan satu dengan yang lainnya harus mengikuti standar yang ada. UML bukan hanya sekedar diagram tetapi juga menceritakan konteksnya.

Berikut adalah beberapa model yang digunakan dalam perancangan Data mining pemaketan produk di Cv. Bukit Manikam untuk menggambarkan sistem dalam UML:

1. Diagram Use case
2. Diagram Activity
3. Diagram Sequence
4. Diagram Class

## 2.6 Use case Diagram

Diagram Use case menggambarkan apa saja aktifitas yang dilakukan oleh suatu sistem. Use case menggambarkan fungsi tertentu dalam suatu sistem berupakomponen, kejadian atau kelas. Komponen Pembentuk *Use case Diagram* adalah sebagai berikut :

1. Actor

Pada dasarnya Actor bukanlah bagian dari use case diagram, namun untuk dapat terciptanya suatu use case diagram diperlukan beberapa Actor. Actor tersebut mempresentasikan seseorang atau sesuatu (seperti perangkat, sistem lain) yang berinteraksi dengan sistem. Sebuah Actor mungkin hanya memberikan informasi inputan pada sistem, hanya menerima informasi dari sistem atau keduanya menerima, dan memberi informasi pada sistem. Actor hanya berinteraksi dengan use case, tetapi tidak memiliki kontrol atas use case. Actor digambarkan dengan stick man. Actor dapat digambarkan secara umum atau spesifik, di mana untuk membedakannya kita dapat menggunakan relationship.

## 2. Use case

Use case adalah gambaran fungsionalitas dari suatu sistem, sehingga customer atau pengguna sistem paham dan mengerti mengenai kegunaan sistem yang akan dibangun.

Catatan : Use case diagram adalah penggambaran sistem dari sudut pandang pengguna sistem tersebut (user), sehingga pembuatan use case lebih dititikberatkan pada fungsionalitas yang ada pada sistem, bukan berdasarkan alur atau urutan kejadian.

Cara menentukan Use case dalam suatu sistem:

- a. Pola perilaku perangkat lunak aplikasi.
- b. Gambaran tugas dari sebuah Actor.
- c. Sistem atau “benda” yang memberikan sesuatu yang bernilai kepada Actor.

- d. Apa yang dikerjakan oleh suatu perangkat lunak (bukan bagaimana cara mengerjakannya).

Ada beberapa relasi yang terdapat pada use case diagram:

- a. Association, menghubungkan link antar element.
- b. Generalization, disebut juga inheritance (pewarisan), sebuah elemen dapat merupakan spesialisasi dari elemen lainnya.
- c. Dependency, sebuah elemen bergantung dalam beberapa cara ke element lainnya.
- d. Aggregation, bentuk association di mana sebuah elemen berisi elemen lainnya.

Tipe relasi/ stereotype yang mungkin terjadi pada use case diagram:

- a. <<include>> , yaitu kelakuan yang harus terpenuhi agar sebuah event dapat terjadi, dimana pada kondisi ini sebuah use case adalah bagian dari use case lainnya.
- b. <<extends>>, kelakuan yang hanya berjalan di bawah kondisi tertentu seperti menggerakkan alarm.
- d. <<communicates>>, mungkin ditambahkan untuk asosiasi yang menunjukkan asosiasinya adalah *communicates association*. Ini merupakan pilihan selama asosiasi hanya tipe *relationship* yang dibolehkan antara Actor dan use case .

## 2.7 Activity Diagram

Activity diagram memiliki pengertian yaitu lebih fokus kepada menggambarkan proses bisnis dan urutan aktivitas dalam sebuah proses. Dipakai pada business modeling untuk memperlihatkan urutan aktifitas proses bisnis. Memiliki struktur diagram yang mirip flowchart atau data flow diagram pada perancangan terstruktur. Memiliki pula manfaat yaitu apabila kita membuat diagram ini terlebih dahulu dalam memodelkan sebuah proses untuk membantu memahami proses secara keseluruhan. Dan activity dibuat berdasarkan sebuah atau beberapa use case pada use case diagram.

## 2.8 Sequence Diagram

Diagram sequence menggambarkan interaksi objek pada use case dengan mendeskripsikan waktu hidup objek dan pesan yang dikirimkan dan di terima antar objek. Oleh karena itu untuk menggambarkan diagram sequence maka harus diketahui objek-objek yang terlibat dalam sebuah use case beserta metode-metode sequence juga di butuhkan untuk melihat skenario yang ada pada *use case*. *Sequence diagram* biasa digunakan untuk menggambarkan skenario atau rangkaian langkah- langkah yang dilakukan sebagai respons dari sebuah event untuk menghasilkan output tertentu. Diawali dari apa yang men-*trigger* aktivitas tersebut, proses dan perubahan apa saja yang terjadi secara internal dan output apa yang dihasilkan. Masing-masing objek, termasuk aktor, memiliki *lifeline* vertikal. *Message* digambarkan sebagai garis berpanah dari satu objek ke objek lainnya.

## 2.9 Class Diagram

Diagram kelas atau class diagram menggambarkan struktur sistem dari segi pendefinisian kelas-kelas yang akan di buat untuk membangun sistem. Kelas memiliki apa yang disebut atribut dan metode atau operasi.

- a. Atribut merupakan variabel-variabel yang di miliki oleh suatu kelas
- b. Operasi atau metode adalah fungsi-fungsi yang dimiliki oleh suatu kelas

Diagram kelas dibuat agar pembuat program atau programmer membuat kelas-kelas sesuai rancangan di dalam diagram kelas agar antara dokumentasi perancangan dan perangkat lunak sinkron. banyak berbagai kasus, perancangan kelas yang dibuat tidak sesuai dengan kelas-kelas yang dibuat pada perangkat lunak, sehingga tidaklah ada gunanya lagi sebuah perancangan karena apa yang dirancang dan hasil jadinya tidak sesuai .

## BAB III

### METODE PENELITIAN

#### 3.1 Desain Penelitian

##### 3.1.1 Objek Penelitian

Penelitian ini bertujuan merancang dan membangun aplikasi yang berguna dalam menentukan rekomendasi yang tepat. Dengan memperhatikan pola aturan yang akan terbentuk dari serangkaian proses mining data. Rekomendasi keputusan ini akan dibangun menggunakan algoritma *Association Rule FP-Growth* dengan objek data real transaksi penjualan dari perusahaan penerbitan CV. Cita Intrans Selaras. Diharapkan dengan output aplikasi rekomendasi judul buku, perusahaan CV. Cita Intrans Selaras dapat meningkatkan pelayanannya, dalam penataan *display buku* di toko online maupun pada buku Katalog penjualan. Sehingga misi perusahaan untuk memberikan pelayanan yang baik dapat terpenuhi serta meningkatkan loyalitas dari pelanggannya.

##### 3.1.2 Sumber Data

Sumber data merupakan segala sesuatu yang dapat memberikan informasi berupa data, dalam hal ini kami memiliki dua acuan data yaitu data Primer dan data Sekunder , sebagai berikut:

a. Data Primer

Studi literature dilakukan dalam proses mencari data yang sesuai yang akan digunakan dalam proses *Data Mining*. Data yang diperoleh merupakan

langsung dari rekap faktur penjualan buku yang terjadi di CV.Cita Intrans Selaras.

b. Data Sekunder

Dilakukan apabila data primer masih belum cukup untuk melakukan pencarian data penelitian. Dilakukan dengan cara tanya jawab kepada narasumber dan juga penelusuran melalui berbagai media yang lain.

### 3.1.3 Bahan Penelitian

Pada penelitian ini terdapat data faktur penjualan sebagai data uji, untuk mengetahui pola pembelian dari konsumen. Dalam terdapat lebih dari 100 judul buku yang terdeteksi dalam data transaksi penjualan bulan Maret 2018, dari 774 *record*.

### 3.1.4 Alat Penelitian

1. Kebutuhan Perangkat Keras

Perangkat Keras yang sedang berjalan yaitu sebuah laptop dengan spesifikasi berikut :

- a. Prosesor Intel Core i3-2330M CPU @2.20GHz
- b. Hardisk 500 GB
- c. RAM 4 GB
- d. Monitor LCD 14"
- e. Keyboard

Perangkat Keras yang dibutuhkan untuk mengimplementasikan sistem yang akan dibangun yaitu sebuah laptop dengan spesifikasi berikut :

- f. Prosesor Intel Core 2 Duo E7500
- g. Hardisk 160 GB

- h. RAM 1 GB
- i. Monitor LCD 14”
- j. Keyboard

## 2. Kebutuhan Perangkat Lunak

Analisis kebutuhan perangkat lunak dalam pembangunan aplikasi digunakan untuk mengetahui dan mengoptimalkan implementasi dari aplikasi yang akan dibangun.

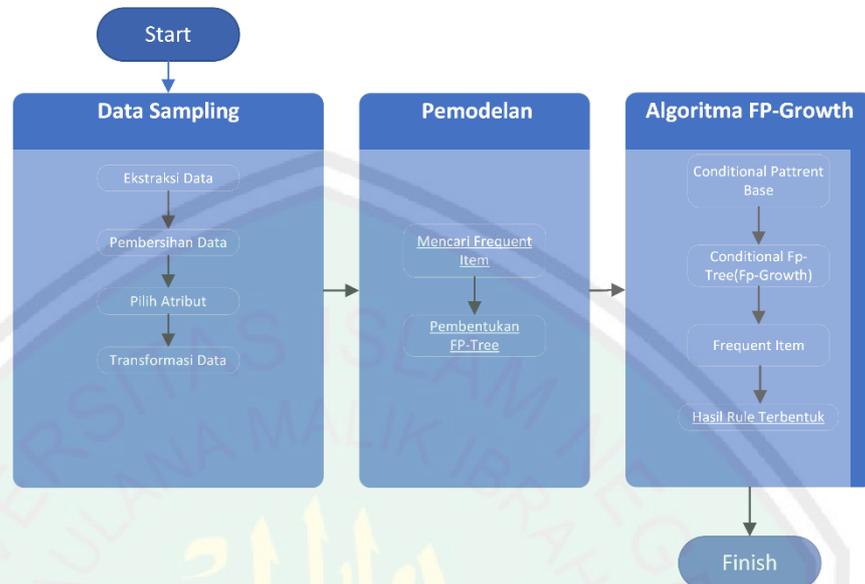
1. Analisis spesifikasi perangkat lunak pada sistem yang sedang berjalan di CV. Cita Intrans Selaras adalah:
  - a. Sistem Operasi Windows 7
  - b. Microsoft Excel 2010
  - c. Adobe Reader
2. Analisis spesifikasi kebutuhan perangkat lunak untuk menjalankan sistem yang akan dibangun adalah :
  - a. Sistem Operasi Windows 7
  - b. XAMPP
  - c. Adobe Reader
  - d. Microsoft Excel 2010

## 3.2 Prosedur penelitian

### 3.2.1 Pemahaman Sistem dan Studi Literatur

Pada tahap ini peneliti melakukan studi literatur dengan mencari literatur serta referensi buku-buku maupun jurnal penelitian sebelumnya yang berkaitan

dengan pemrosesan data dengan metode data mining menggunakan algoritma FP-Growth.

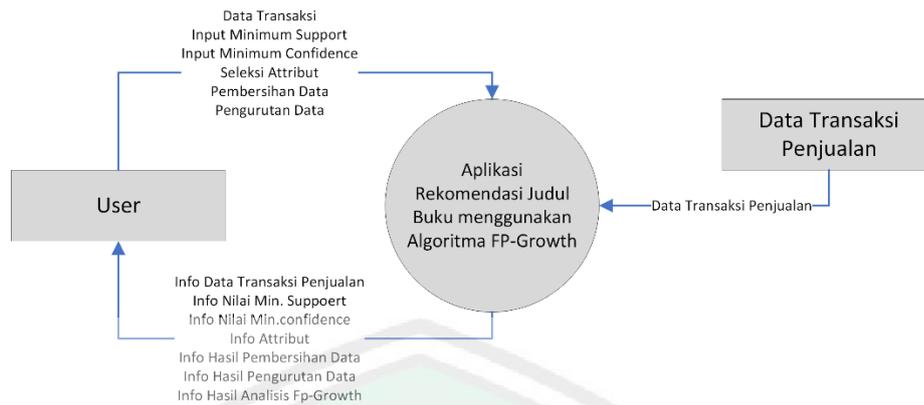


Gambar 3.3 Flowchart kerja Sistem

### 3.2.2 Perancangan Sistem

#### 3.2.2.1 Diagram Konteks

Diagram konteks atau disebut juga dengan model sistem fundamental merepresentasikan seluruh elemen sistem sebagai sebuah bubble tunggal dengan data input, output yang ditunjukkan oleh anak panah yang masuk dan keluar secara berurutan. Adapun diagram konteks perangkat lunak yang akan dibangun dapat dilihat pada gambar dibawah ini :



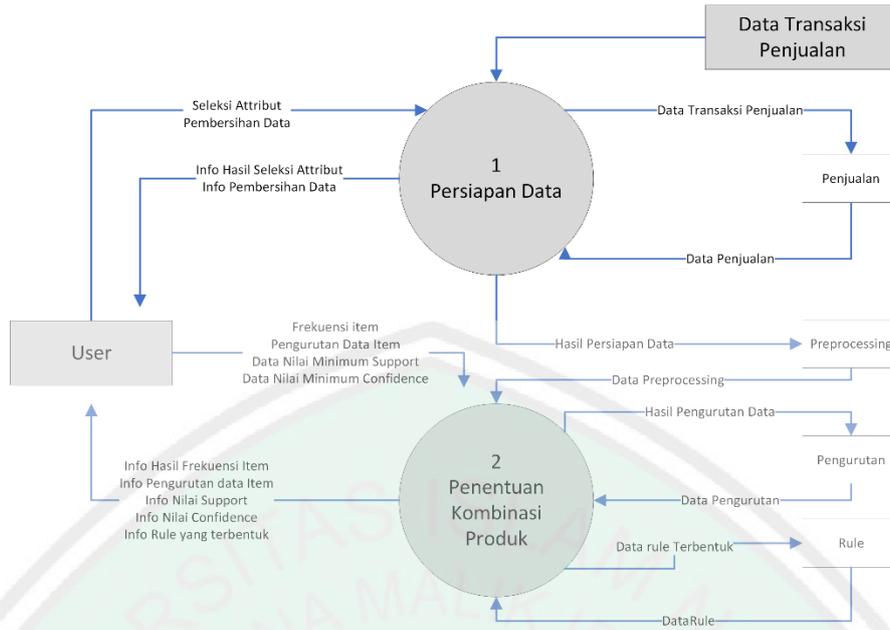
**Gambar 3.4** Konteks Diagram

### 3.2.2.2 DFD (Data Flow Diagram)

DFD adalah suatu model logika data atau proses yang dibuat untuk menggambarkan dari mana asal data dan kemana tujuan data yang keluar dari sistem, dimana data disimpan, proses apa yang menghasilkan data tersebut dan interaksi antara data yang tersimpan dan proses yang dikenakan pada data tersebut.

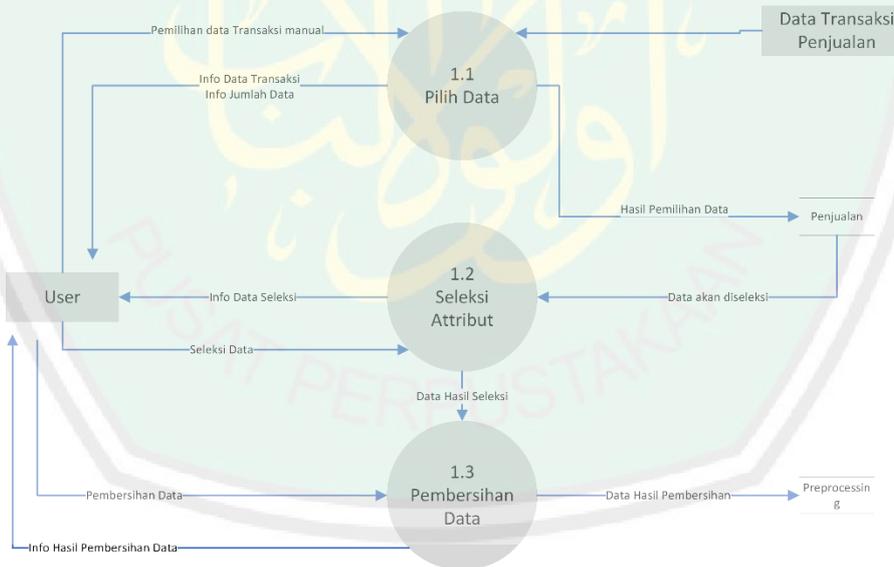
Data flow Diagram berfungsi untuk memudahkan pengguna yang kurang menguasai bidang komputer agar dapat mengerti perangkat lunak yang akan dijalankan. Berikut adalah data flow diagram untuk perangkat lunak yang akan dibangun.

1. DFD level 1



Gambar 3.5 DFD Level1

2. DFD Level 2 Proses Persiapan data



Gambar 3.6 DFD Level 2 Persiapan data

3. DFD Level 2 Penentuan Kombinasi Produk



Analisa masalah yang dilakukan adalah mengamati dan mengidentifikasi permasalahan yang terjadi dalam penelitian yang dilakukan. Melalui rumusan masalah yang telah ditetapkan sebelumnya dapat disimpulkan permasalahan yang ada pada CV.Cita Intrans Selaras yaitu tidak adanya aplikasi yang dapat membantu memberikan rekomendasi judul buku yang sesuai dengan kebutuhan pasar.

### **3.5 Analisa Penerapan Data Mining**

Metode yang digunakan dalam merancang dan membangun aplikasi Data Mining dalam penelitian ini yaitu dengan menggunakan *Cross-Industry Standart Process for Data Mining (CRISP-DM)*

#### **3.5.1 Pemahaman Bisnis**

Tahapan pemahaman bisnis merupakan tahapan pertama dilakukan dalam kerangka kerja CRISP-DM. Dalam tahapan bisnis ini terdapat beberapa tahapan lainnya, yaitu:

1. Tujuan Bisnis

Dalam proses bisnisnya CV. Cita Intrans Selaras mempunyai tujuan bisnis yaitu untuk memenuhi kebutuhan informasi pembaca dengan memberikan wacana-wacana perubahan.

Sehingga dapat meningkatkan kepuasan masyarakat sebagai pelanggan yang membutuhkan informasi.

2. Penentuan Sasaran Data Mining

Tujuan dari penerapan data mining pada rekomendasi judul buku ini adalah untuk mengetahui bagaimana pola pembelian kosumen buku sehingga terbentuk rekomendasi judul buku.

### 3.5.2 Data Sampling

Langkah selanjutnya dalam kerangka kerja CRISP-DM yaitu melakukan tahap pemahaman Data pada data *sampling*. Dalam memahami data harus melakukan beberapa tahapan yaitu:

1. Mengumpulkan data awal

Data yang digunakan dalam penelitian ini adalah sampel/ccontoh data faktur penjualan bulan maret 2018 di CV.Cita Intrans Selaras dengan jenis file (.xlsx) yang akan dimining dengan aplikasi rekomendasi judul buku.

2. Menjelaskan data

Berikut merupakan penjelasan atribut-atribut yang ada pada tabel 3.1 faktur penjualan maret 2018:

Tabel4.1 Struktur Data Transaksi

Keterangan		
Fungsi	Untuk mengetahui transaksi penjualan buku dilakukan	
Format	Microsoft Excel (.xlsx)	
Atribut	Tanggal	Tanggal penjualan buku
	Kode Faktur	Kode urut penjualan
	Customer	Nama pembeli
	Judul Buku	Judul buku yang terjual
	Kode Buku	Kode untuk membedakan judul buku
	Eks/Qty	Jumlah eksemplar buku yang terjual
	Harga	Harga satuan buku
	Brutto	Jumlah kotorpendapatan
	Disc.	Pengurangan harga berdasar prosentase

	Netto.	Jumlah bersih pendapatan
--	--------	--------------------------

### 3.5.3 Persiapan Data

Seperti tahapan sebelum-sebelumnya dalam proses *Data Mining* yang akan dilakukan selanjutnya yaitu persiapan data karena tidak semua atribut digunakan dalam me-mining data. Tahap ini harus dilakukan untuk menyeleksi data yang dibutuhkan. Adapun tahapan-tahapan *preprocessing* data dalam penelitian ini yaitu:

#### 1. Ekstraksi Data

Ekstraksi data merupakan proses pengambilan data dari sumber data dalam rangka untuk melanjutkan proses pengolahan data ketinggian selanjutnya ataupun untuk menyimpan data hasil ekstrak tersebut. Dalam penelitian ini, data yang berasal dari flat file berformat microsoft excel diekstrak, kemudian disimpan kedalam sebuah database agar memudahkan dalam proses pengolahan data. Berikut merupakan contoh sampel data transaksi bulan Maret 2018 yang diambil sebanyak 30 transaksi dari sebanyak 774 record yang terdapat dalam data tersebut adalah Tanggal, Kode Faktur, Costumer, No, Judul, Kode buku, Eks, Harga, Brutto, Disc, Netto. Data transaksi dapat dilihat pada table 4.2 di bawah ini:.

Tabel 4.2 Contoh Data Faktur Penjualan Maret 2018

Kode Faktur	Judul Buku	Harga	Eks	Jumlah
US-083418	AGAMA DAN POLITIK MORAL	\$ 57,000.00	4	\$ 228,000.00
US-083418	HUKUM ISLAM - INDAH PURBASARI	\$ 69,000.00	5	\$ 345,000.00
US-083418	HUKUM ISLAM - INDAH PURBASARI	\$ 69,000.00	1	\$ 69,000.00
US-083418	AGAMA DAN POLITIK MORAL	\$ 57,000.00	2	\$ 114,000.00
US-083618	HUKUM ISLAM - INDAH PURBASARI	\$ 69,000.00	5	\$ 345,000.00
US-083618	PENGANTAR MANAJEMEN	\$ 60,000.00	5	\$ 300,000.00
US-083618	PENGANTAR MANAJEMEN	\$ 60,000.00	2	\$ 120,000.00
US-083618	EKONOMI ISLAM - ARIF	\$ 85,000.00	12	\$ 1,020,000.00
US-083618	HUKUM ISLAM - INDAH PURBASARI	\$ 69,000.00	1	\$ 69,000.00
US-084018	EKONOMI ISLAM - ARIF	\$ 85,000.00	2	\$ 170,000.00
US-084018	IDEOLOGI PANCASILA	\$ 75,000.00	1	\$ 75,000.00
US-084018	EKONOMI ISLAM - ARIF	\$ 85,000.00	1	\$ 85,000.00

## 2. Pembersihan Data

Proses menghilangkan noise dan menghilangkan data tidak relevan atau inkonsisten disebut pembersihan data. Dalam penelitian ini data yang tidak relevan dalam rangka proses data mining yaitu data harga, Eks, Jumlah.

Tabel 4.3 Data Transaksi setelah pembersihan

Kode Faktur	Judul Buku
US-083418	AGAMA DAN POLITIK MORAL
US-083418	HUKUM ISLAM - INDAH PURBASARI
US-083418	HUKUM ISLAM - INDAH PURBASARI
US-083418	AGAMA DAN POLITIK MORAL
US-083618	HUKUM ISLAM - INDAH PURBASARI
US-083618	PENGANTAR MANAJEMEN
US-083618	PENGANTAR MANAJEMEN
US-083618	EKONOMI ISLAM - ARIF
US-083618	HUKUM ISLAM - INDAH PURBASARI
US-084018	EKONOMI ISLAM - ARIF
US-084018	IDEOLOGI PANCASILA
US-084018	EKONOMI ISLAM - ARIF
US-084018	HISTORIOGRAFI FILSAFAT ISLAM
US-084018	PENGANTAR MANAJEMEN
US-084018	PENDIDIKAN YANG MEMISKINKAN
US-084018	EKONOMI ISLAM - ARIF

## 3. Pemilihan Atribut

Setelah proses pembersihan data kemudian dilakukan proses pemilihan atribut, proses dimana atribut data akan dipilih dan diproses sesuai dengan kebutuhan data mining. Dalam

penelitian ini, atribut yang dibutuhkan hanya 2 atribut yaitu Kode Faktur dan Judul Buku. Kedua atribut ini digunakan untuk memenuhi kebutuhan proses data mining. Kode faktur digunakan untuk mengetahui dalam satu kali pembelian berapa buku yang dibeli. Sedangkan judul buku untuk mengetahui buku apa saja yang telah dibeli.

Tabel 4.4 Data training

Kode Faktur	Judul Buku
US-083418	AGAMA DAN POLITIK MORAL
US-083418	HUKUM ISLAM - INDAH PURBASARI
US-083418	HUKUM ISLAM - INDAH PURBASARI
US-083418	AGAMA DAN POLITIK MORAL
US-083618	HUKUM ISLAM - INDAH PURBASARI
US-083618	PENGANTAR MANAJEMEN
US-083618	PENGANTAR MANAJEMEN
US-083618	EKONOMI ISLAM - ARIF
US-083618	HUKUM ISLAM - INDAH PURBASARI
US-084018	EKONOMI ISLAM - ARIF
US-084018	IDEOLOGI PANCASILA
US-084018	EKONOMI ISLAM - ARIF
US-084018	HISTORIOGRAFI FILSAFAT ISLAM
US-084018	PENGANTAR MANAJEMEN
US-084018	PENDIDIKAN YANG MEMISKINKAN
US-084018	EKONOMI ISLAM - ARIF

#### 4. Transformasi data

Tahapan transformation data merupakan tahap merubah data ke dalam bentuk yang sesuai untuk di-mining. Pada tahap transformasi Kode Faktur yang terdapat pada dataset diubah menjadi kode huruf TID sedangkan judul buku diubah menjadi angka yang lebih singkat sesuai dengan kode faktur dan judul buku sebenarnya yang telah ditentukan pada saat dilakukannya proses penyeleksian data, proses transformasi adalah sebagai berikut:

Tabel 4.5 Transformasi Kode Faktur ke TID

Kode Faktur		TID
US-083418		
US-083418	→	T1
US-083418		
US-083418		
US-083618		
US-083618	→	T2
US-083618		
US-083618		
US-084018	→	T3
US-084018		
US-084018		
US-084018		
US-085318		
US-085318	→	T4
US-085318		
US-085318		

Tabel 4.6 Transformasi judul buku kode buku

Judul Buku	Kode Buku
AGAMA DAN POLITIK MORAL	3
HUKUM ISLAM - INDAH PURBASARI	89
HUKUM ISLAM - INDAH PURBASARI	89
AGAMA DAN POLITIK MORAL	3
HUKUM ISLAM - INDAH PURBASARI	89
PENGANTAR MANAJEMEN	253
PENGANTAR MANAJEMEN	253
EKONOMI ISLAM - ARIF	49
HUKUM ISLAM - INDAH PURBASARI	89
EKONOMI ISLAM - ARIF	49
IDEOLOGI PANCASILA	124
EKONOMI ISLAM - ARIF	49
HISTORIOGRAFI FILSAFAT ISLAM	72
PENGANTAR MANAJEMEN	253
PENDIDIKAN YANG MEMISKINKAN	238
EKONOMI ISLAM - ARIF	49
PENDIDIKAN YANG MEMISKINKAN	238
HISTORIOGRAFI FILSAFAT ISLAM	72
IDEOLOGI PANCASILA	124
IDEOLOGI PANCASILA	124
IDEOLOGI PANCASILA	124

### 3.5.4 Pemodelan

Pemodelan merupakan tahapan metode dalam proses data mining yang diterapkan pada data CV. Cita Intrans Selaras. Pemodelan ini dilakukan

dengan menggunakan metode FP-Growth. Metode ini memiliki dua tahapan pengerjaan yaitu:

1. Mencari *Frequent Itemset*
2. Menghasilkan *rule (generate rule)*

Dalam tahap pencarian *frequent itemset* algoritma yang akan digunakan adalah *fp-growth*. Algoritma ini memiliki dasar pengetahuan mengenai *frequent itemset* yang telah diketahui sebelumnya untuk memproses informasi lebih lanjut. Dalam algoritma *fp-growth* terhadap tiga tahapan penting, namun sebelum masuk pada tahapan untuk mempermudah pembentukan *frequent itemset* akan dilakukan pengkodean terhadap produk yang terdapat pada data yang sudah siap digunakan untuk proses *mining*, proses pengkodean ini berdasarkan nomor setiap produk yang terdapat pada transaksi, contoh “Agama dan Politik Moral” maka akan menjadi “3”.

### 3.5.5 Mencari Frequent Item

Berikut adalah perhitungan manual untuk mencari *Frequent Item* yang terdapat 10 sampel data transaksi pengunjung, dengan batas minimum support sebanyak 20%.

Tabel 4.7 Sampel faktor penjualan

Tabel 4.8 Hasil transformasi data transaksi pengunjung

TID	Itemset
T1	3, 89
T2	89, 253, 49
T3	3, 49, 72, 238, 253
T4	3, 72
T5	3, 49, 89, 119
T6	3, 49, 89, 253
T7	3, 63
T8	3, 49, 89
T9	3, 89, 253
T10	49, 72, 89

Hasil dari proses penelusuran database diketahui jumlah frekuensi kemunculan tiap item yang ada di dalam database dan mengurutkannya berdasarkan jumlah frekuensi kemunculan item yang paling terbesar.

Frekuensi kemunculan tiap item berdasar Tabel 5 dapat dilihat

Kode Faktur	TID	Judul Buku	Kode Buku
US-083418	T1	AGAMA DAN POLITIK MORAL	3
US-083418		HUKUM ISLAM - INDAH PURBASARI	89
US-083418		HUKUM ISLAM - INDAH PURBASARI	89
US-083418	T2	AGAMA DAN POLITIK MORAL	3
US-083618		HUKUM ISLAM - INDAH PURBASARI	89
US-083618		PENGANTAR MANAJEMEN	253
US-083618		PENGANTAR MANAJEMEN	253
US-083618	T3	EKONOMI ISLAM - ARIF	49
US-083618		HUKUM ISLAM - INDAH PURBASARI	89
US-084018		EKONOMI ISLAM - ARIF	49
US-084018	T4	AGAMA DAN POLITIK MORAL	3
US-084018		EKONOMI ISLAM - ARIF	49
US-084018		HISTORIOGRAFI FILSAFAT ISLAM	72
US-084018		PENGANTAR MANAJEMEN	253
US-084018		PENDIDIKAN YANG MEMISKINKAN	238
US-084018		EKONOMI ISLAM - ARIF	49
US-084018	T5	PENDIDIKAN YANG MEMISKINKAN	238
US-085318		HISTORIOGRAFI FILSAFAT ISLAM	72
US-085318		AGAMA DAN POLITIK MORAL	3
US-085318	T6	AGAMA DAN POLITIK MORAL	3
US-085318		AGAMA DAN POLITIK MORAL	3

abel berikut:

Tabel 4.9 Frekuensi kemunculan tiap item

Item	Frekuensi
3	8
89	7
49	6
253	4
72	3
63	1
119	1
238	1

Proses pembentukan 1 itemset dengan jumlah minimum support = 20% dengan rumus sebagai berikut

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

$$\text{Support}(3) = \frac{8}{10} \times 100\% = 80\%$$

$$\text{Support}(89) = \frac{7}{10} \times 100\% = 70\%$$

$$\text{Support}(49) = \frac{6}{10} \times 100\% = 60\%$$

$$\text{Support}(253) = \frac{4}{10} \times 100\% = 40\%$$

$$\text{Support}(72) = \frac{3}{10} \times 100\% = 30\%$$

$$\text{Support}(63) = \frac{1}{10} \times 100\% = 10\%$$

$$\text{Support}(119) = \frac{1}{10} \times 100\% = 10\%$$

$$\text{Support}(238) = \frac{1}{10} \times 100\% = 10\%$$

Tabel 4.10 Item yang memenuhi support sistem

Item	Frekuensi	Support
3	8	80%
89	7	70%
49	6	60%
253	4	40%
72	3	30%
63	1	10%
119	1	10%
238	1	10%

Dari proses pembentukan itemset pada tabel 7 dengan minimum support 20 % dapat diketahui yang memenuhi standar minimum support adalah N,SB,BI,E,M.

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

Tabel 4.11 Tabel Data Transaksi

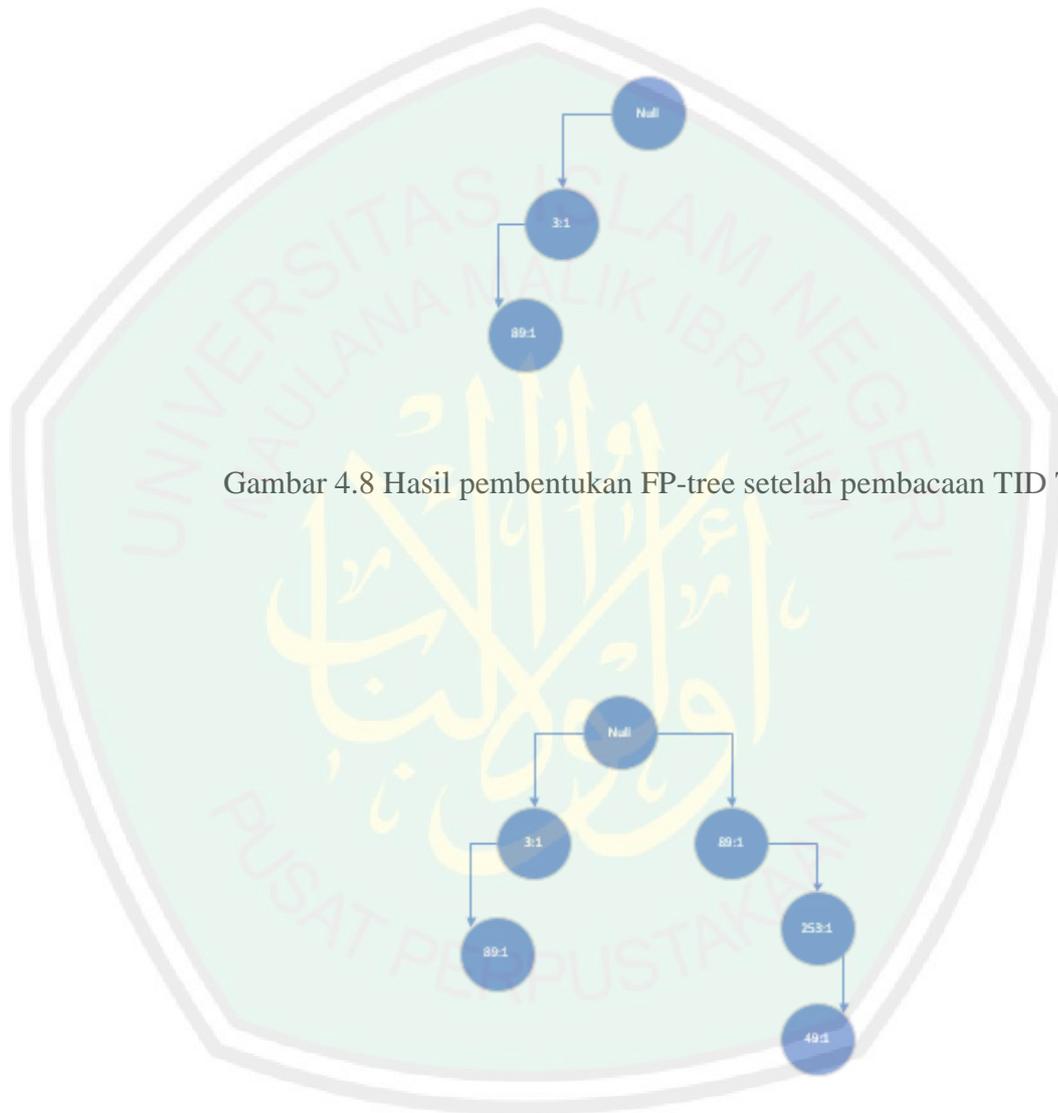
TID	Itemset
T1	3, 89
T2	89, 253, 49
T3	3, 49, 72, 253
T4	3, 72
T5	3, 49, 89
T6	3, 49, 89, 253
T7	3
T8	3, 49, 89
T9	3, 89, 253
T10	49, 72, 89

### 3.5.6 Pembentukan FP-Tree

Setelah data sudah diurutkan berdasarkan prioritas dan memenuhi minimum support yang telah di tentukan, langkah selanjutnya adalah

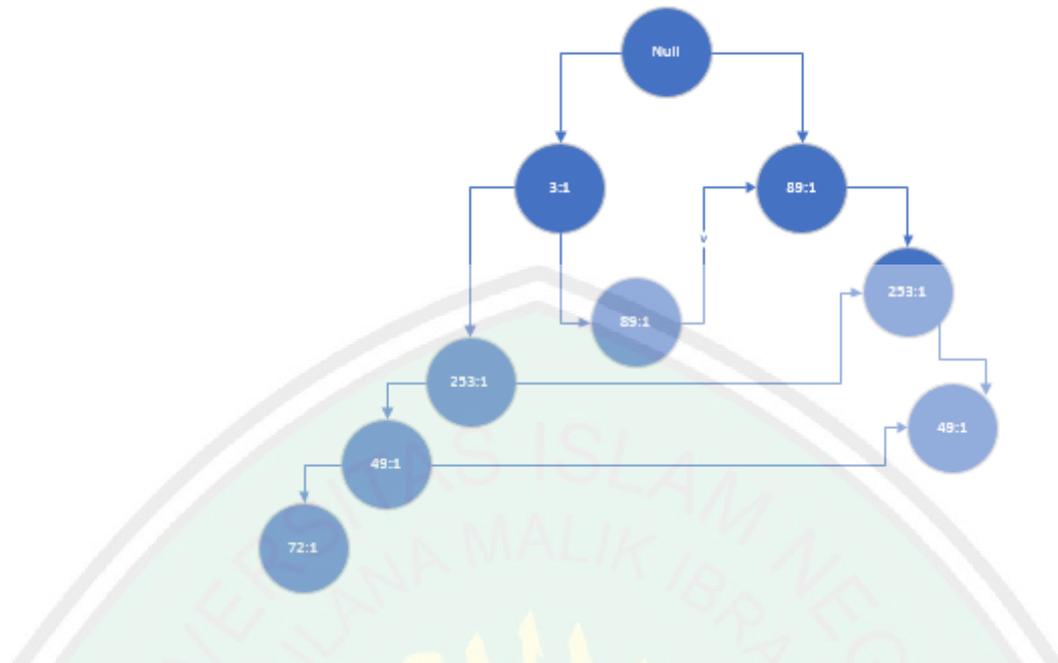
membuat struktur pohon FP-Tree untuk setiap transaksi. Berikut adalah struktur pohon dari semua transaksi yang ada.

Gambar di bawah ini memberikan ilustrasi mengenai pembentukan FP-tree dari tiap-tiap TID.



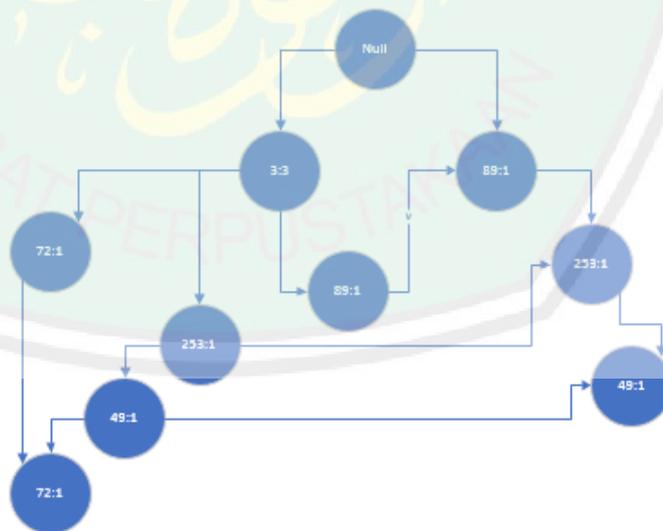
Gambar 4.8 Hasil pembentukan FP-tree setelah pembacaan TID T1

Gambar 4.9 Hasil Pembentukan FP-Tree setelah pembacaan TID T2

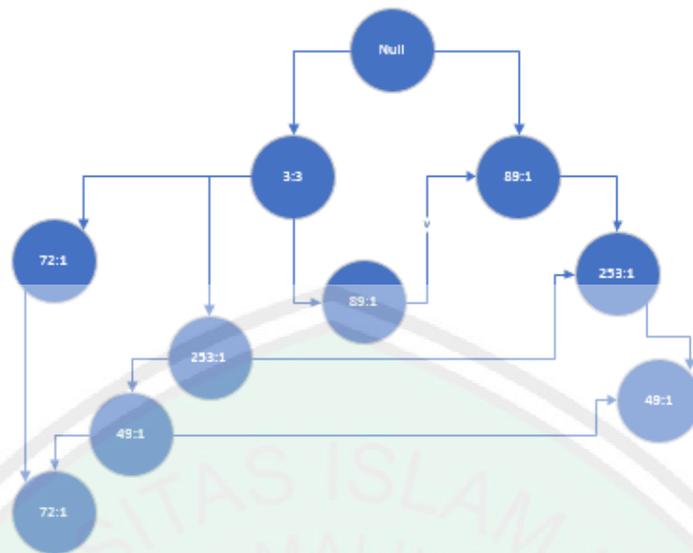


Gambar 4.10 Hasil Pembentukan FP-Tree setelah pembacaan TID T3

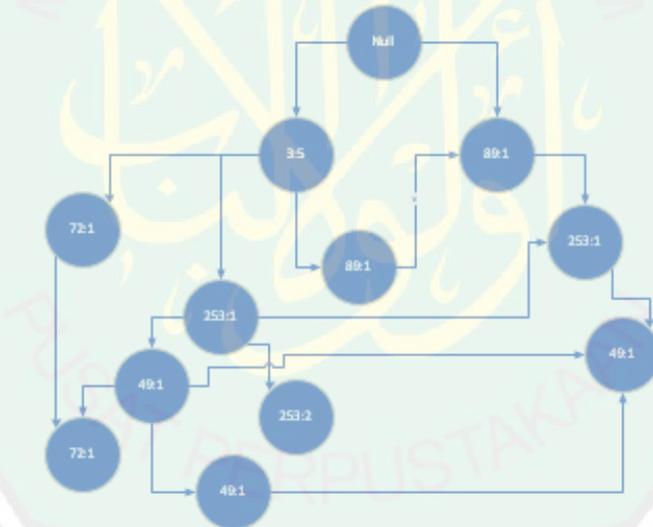
Pembacaan data sampai TID T3 menunjukkan frequent dari 3 bertambah karena TID baru melalui lintasan yang sudah ada.



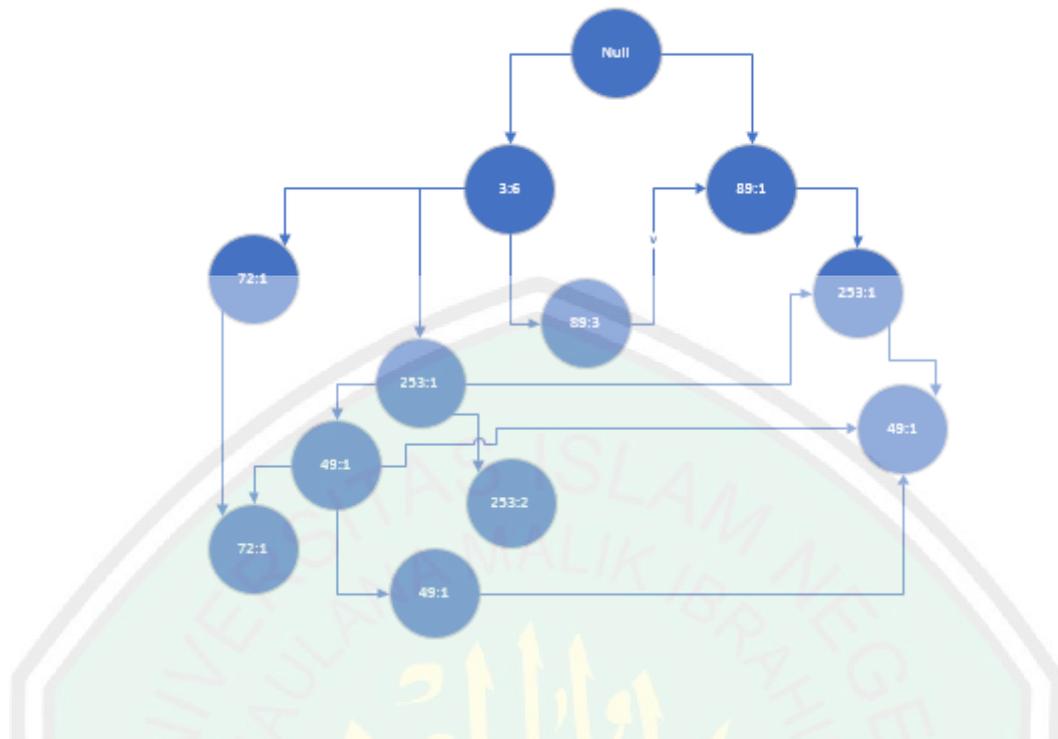
Gambar 4.11 Hasil Pembentukan FP-Tree setelah pembacaan TID T4



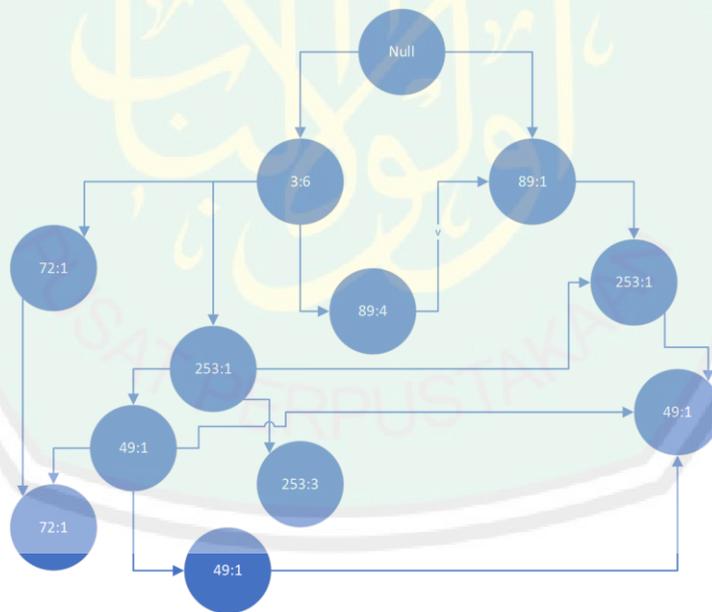
Gambar 4.12 Hasil Pembentukan FP-Tree setelah pembacaan TID T5



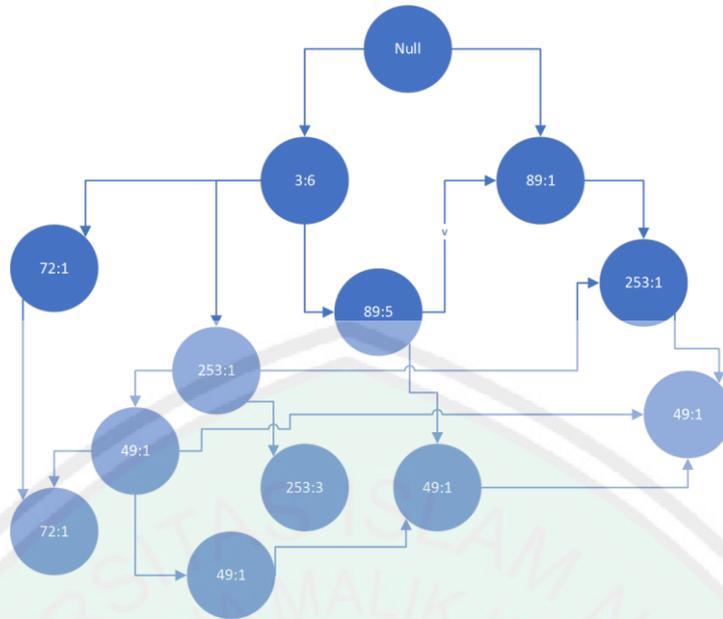
Gambar 4.13 Hasil Pembentukan FP-Tree setelah pembacaan TID T6



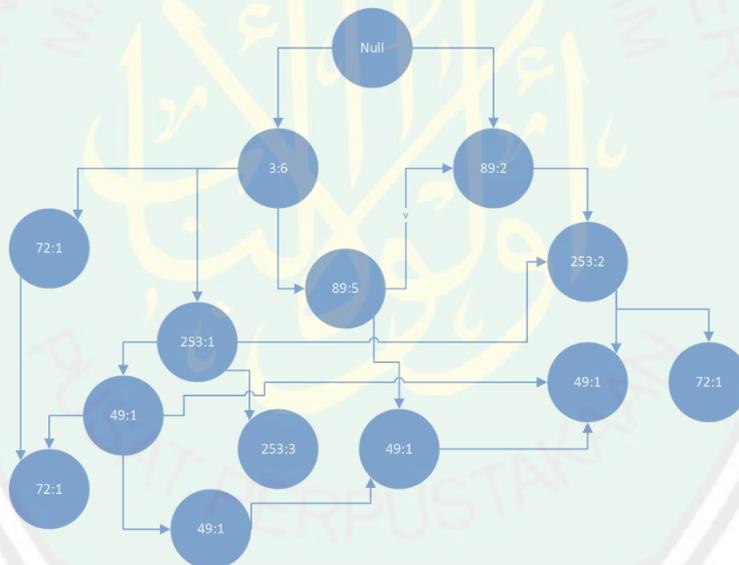
Gambar 4.14 Hasil Pembentukan FP-Tree setelah pembacaan TID T7



Gambar 4.15 Hasil Pembentukan FP-Tree Setelah Pembacaan TID T8



**Gambar 4.16** Hasil Pembentukan FP-Tree Setelah Pembacaan TID T9

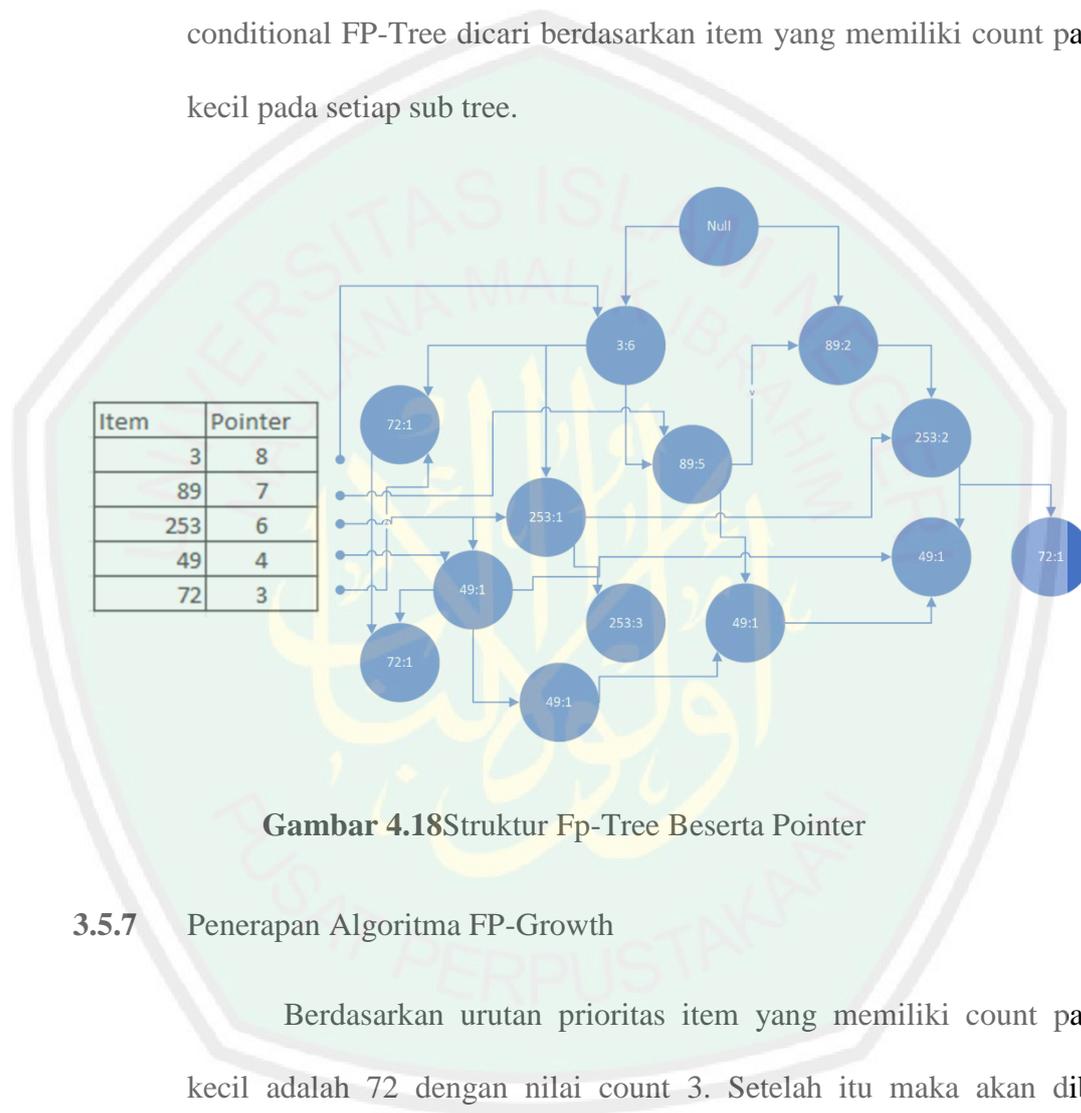


**Gambar 4.17** Hasil Pembentukan FP-Tree Setelah Pembacaan TID T10

Struktur pohon FP-Tree untuk setiap baris data didapatkan setelah data telah diurutkan berdasarkan tingkat prioritas dan telah memenuhi minimum support yang ditentukan. Gambar 8 – 17 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.

Setelah pohon FP-Tree terbentuk, conditional pattern base dan conditional FP-Tree dicari berdasarkan item yang memiliki count paling kecil pada setiap sub tree.



**Gambar 4.18** Struktur Fp-Tree Beserta Pointer

### 3.5.7 Penerapan Algoritma FP-Growth

Berdasarkan urutan prioritas item yang memiliki count paling kecil adalah 72 dengan nilai count 3. Setelah itu maka akan dibuat subtree yang berakhir node 72. Conditional pattern base, conditional FP-tree dan frequent itemset akan ditemukan dari subtree tersebut.

- a. Conditional Pattern Base :

72 ➔  $\{\{3:1\},\{3,253,49:1\},\{89,253:1\}\}$  Conditional pattern base diperoleh dengan cara membaca setiap lintasan berakhir node 72.

b. Conditional FP-Tree:

Setelah conditional pattern base diperoleh, maka conditional FP-Tree terbentuk dengan mengabaikan item tunggal. Pada lintasan berakhir node 72 ini, item 49 dan 89 diabaikan Karena merupakan item tunggal. 72  $\{3:2, 253:2\}$

c. Frequent Item Sets:

Selanjutnya untuk mendapatkan frequent itemsets, lakukan pengkombinasian dari item yang akan dibuat conditional FP-tree dengan syarat count dari item tersebut memenuhi minimum support. Frequent Item Sets 72  $\{3,72:2\},\{253,72:2\},\{3,253,72:2\}$

Urutan prioritas item yang memiliki count paling kecil kedua adalah 49 dengan nilai count 4. Maka akan dibuat subtree yang berakhir node 49. Conditional pattern base, conditional FP-tree dan frequent itemset akan ditemukan dari subtree tersebut.

a. Conditional Pattern Base :

49  $\{\{3,253:1\},\{3,89,253:1\},\{3:1\},\{89,253:1\}\}$  Conditional pattern base diperoleh dengan cara membaca setiap lintasan berakhir node 49

b. Conditional FP-Tree:

Setelah conditional pattern base diperoleh, maka conditional FP-Tree terbentuk dengan mengabaikan item tunggal. Pada lintasan berakhiran node E ini, tidak ada item yang diabaikan 49  
 $\{3:3,253:3,89:2\}$

c. Frequent Item Sets:

Selanjutnya untuk mendapatkan frequent itemsets, lakukan pengkombinasian dari item yang akan dibuat conditional FP-tree dengan syarat count dari item tersebut memenuhi minimum support.

Frequent Item Sets 49  $\{3,49:3\},\{253,49:3\},\{89,49:2\}$

Urutan prioritas item yang memiliki count paling kecil kedua adalah 253 dengan nilai count 6. Maka akan dibuat subtree yang berakhiran node 253. Conditional pattern base, conditional FP-tree dan frequent itemset akan ditemukan dari subtree tersebut.

a. Conditional Pattern Base :

253  $\{\{3:1\},\{3,89:2\},\{89:2\}\}$  Conditional pattern base diperoleh dengan cara membaca setiap lintasan berakhiran node 253.

b. Conditional FP-Tree:

Setelah conditional pattern base diperoleh, maka conditional FP-Tree terbentuk dengan mengabaikan item tunggal. Pada lintasan berakhiran node BI ini, tidak ada item yang diabaikan

253  $\{3:2,89:4\}$

c. Frequent Item Sets:

Selanjutnya untuk mendapatkan frequent itemsets, lakukan pengkombinasian dari item yang akan dibuat conditional FP-tree dengan syarat count dari item tersebut memenuhi minimum support.

Frequent Item Sets 253 {3, 253:2}, {89, 253:4}

Urutan prioritas item yang memiliki count paling kecil kedua adalah 89 dengan nilai count 7. Maka akan dibuat subtree yang berakhiran node 89. Conditional pattern base, conditional FP-tree dan frequent itemset akan ditemukan dari subtree tersebut.

a. Conditional Pattern Base:

89{{3:4}} Conditional pattern base diperoleh dengan cara membaca setiap lintasan berakhiran node 89

b. Conditional FP-Tree:

Setelah conditional pattern base diperoleh, maka conditional FP-Tree terbentuk dengan mengabaikan item tunggal. Pada lintasan berakhiran node 89 ini, tidak ada item yang diabaikan

Conditional FP-Tree 89 {3:4}

c. Conditional FP-Tree:

Selanjutnya untuk mendapatkan frequent itemsets, lakukan pengkombinasian dari item yang akan dibuat conditional FP-tree dengan syarat count dari item tersebut memenuhi minimum support. Frequent Item Sets 89 {3,89:4}

Berikut ini adalah daftar frequent itemsets yang didapatkan. Untuk pembuatan aturan asosiasi yang dipakai adalah frequent itemset dengan jumlah item terbanyak.

Tabel 4.12 Frequent Itemsets

Item	Frekuensi Item Set
72	{3,72:2}, {253,72:2}, {3,253,72:2}
49	{3,49:3}, {253,49:3}, {89,49:2}
253	{3,253:2}, {89,253:4}
89	{3,89:4}

Berdasarkan 9 frequent itemsets yang telah terbentuk di atas, semua akan dihitung dalam proses selanjutnya karena memenuhi syarat frequent itemsets dalam menghasilkan association rule yaitu minimal memiliki 2 item dimana jika membuka kategori A maka akan membuka kategori B. Maka terdapat 8 subsets yang layak untuk dihitung tingkat confidence-nya.

Tabel 4.13 Frequent item set

Item	Frekuensi Item Set
72	{3,72:2}, {253,72:2}, {3,253,72:2}
49	{3,49:3}, {253,49:3}, {89,49:2}
253	{3,253:2}, {89,253:4}
89	{3,89:4}

Setelah mendapatkan subsets yang memenuhi syarat, kemudian nilai confidence dihitung berdasarkan nilai minimum confidence yang telah ditentukanyaitu 20% untuk mengukur seberapa besar valid tidaknya aturan asosiasi tersebut.

Tabel 4.14 Frequent Pattern

Frequent Pattern	Frekuensi
3,72	2
3,49	3
3,253	2
3,89	4
253,72	2
253,89	3
89,49	2
89,253	4

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:

$$\text{Support}(A, B) = P(A \cap B)$$

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

$$\text{Support}(3,72) = (3 \cap 72) = \frac{2}{10} \times 100\% = 20\%$$

$$\text{Support}(3,49) = (3 \cap 49) = \frac{3}{10} \times 100\% = 30\%$$

$$\text{Support}(3,253) = (3 \cap 253) = \frac{2}{10} \times 100\% = 20\%$$

$$\text{Support}(3,89) = (3 \cap 89) = \frac{4}{10} \times 100\% = 40\%$$

$$\text{Support}(253,72) = (253 \cap 72) = \frac{2}{10} \times 100\% = 20\%$$

$$\text{Support}(253,89) = (253 \cap 89) = \frac{3}{10} \times 100\% = 30\%$$

$$\text{Support}(89,49) = (89 \cap 49) = \frac{2}{10} \times 100\% = 20\%$$

$$\text{Support}(89,253) = (89 \cap 253) = \frac{4}{10} \times 100\% = 40\%$$

$$\text{Confidence } (A \rightarrow B) = \frac{\text{Jumlah Transaksi Mengandung } A \cap B}{\text{Total Transaksi } A}$$

$$\text{Confidence } (3 \rightarrow 72) = \frac{2}{8} \times 100\% = 25\%$$

$$\text{Confidence } (3 \rightarrow 49) = \frac{3}{8} \times 100\% = 38\%$$

$$\text{Confidence } (3 \rightarrow 253) = \frac{2}{8} \times 100\% = 25\%$$

$$\text{Confidence } (3 \rightarrow 89) = \frac{4}{8} \times 100\% = 50\%$$

$$\text{Confidence } (253 \rightarrow 72) = \frac{2}{6} \times 100\% = 25\%$$

$$\text{Confidence } (253 \rightarrow 89) = \frac{3}{6} \times 100\% = 75\%$$

$$\text{Confidence } (89 \rightarrow 49) = \frac{2}{7} \times 100\% = 29\%$$

$$\text{Confidence } (89 \rightarrow 253) = \frac{4}{7} \times 100\% = 57\%$$

Tabel 4.15 Rule yang terbentuk

Rule	Support	Confidence
253,89	30%	75%
89,253	40%	57%
3,89	40%	50%
3,49	30%	38%
89,49	20%	29%
3,72	20%	25%
3,253	20%	25%
253,72	20%	25%

### Hasil Rekomendasi buku

Berikut merupakan hasil dari asosiasi rule yang digunakan sebagai rekomendasi judul buku, kode yang tercantum pada tabel di atas seperti 89 - > Hukum Indah Purbasari

yang telah dibeli kemungkinan juga akan membeli 49-> Ekonomi Islam Arif, dibawah merupakan keterangan dari tabel 4.15.

<p>Jika Membeli Pengantar Manajemen maka akan membeli Hukum Islam- Indah Purbasari Nilai Support 30% dan Confidence 75%</p>
<p>Jika Membeli Hukum Islam-Indah Purbasari maka akan membeli Pengantar Manajemen Nilai Support 40% dan Confidence 57%</p>
<p>Jika Membeli Agama dan Politik Sosial maka akan membeli Hukum Islam-Indah purbasari Nilai Support 40% dan Confidence 50%</p>
<p>Jika Membeli Agama dan Politik Sosial maka akan membeli Ekonomi Islam - Arif Nilai Support 30% dan Confidence 38%</p>
<p>Jika Membeli Hukum Islam-Indah Purbasari maka akan membeli Ekonomi Islam-Arif Nilai Support 20% dan Confidence 29%</p>
<p>Jika Membeli Agama dan Politik Sosial maka akan membeli Historiografi Filsafat Islam Nilai Support 20% dan Confidence 25%</p>
<p>Jika Membeli Agama dan Politik Sosial maka akan membeli Pengantar Manajemen Nilai Support 20% dan Confidence 25%</p>
<p>Jika Membeli Pengantar Manajemen maka akan membeli Historiografi Filsafat Islam Nilai Support 20% dan Confidence 25%</p>

## BAB IV HASIL DAN PEMBAHASAN

### 4.1 Implementasi Program

#### 4.1.1 Menampilkan Beranda Dashboard

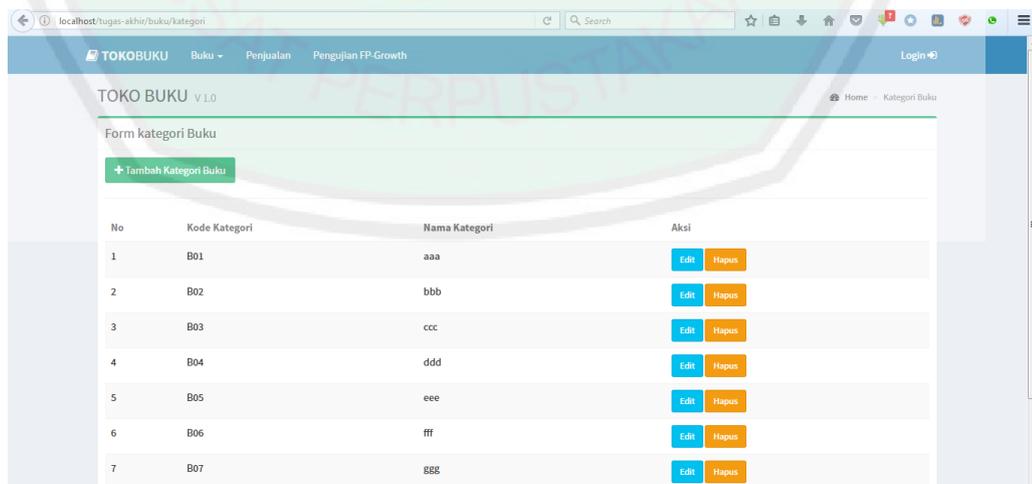
Tampilan halaman depan dari aplikasi rekomendasi judul buku menggunakan algoritma Fp-Growth.



**Gambar 4.1** Halaman Depan Aplikasi

#### 4.1.2 Proses Input Data Kategori

Data awal yang dibutuhkan yaitu menginputkan kategori buku sebagai pelengkap data yang di simpan dalam data base. Data ini berkaitan dengan data buku yang akan di inputkan setelah data kategori telah ada.



**Gambar 4.2** Data Kategori

```

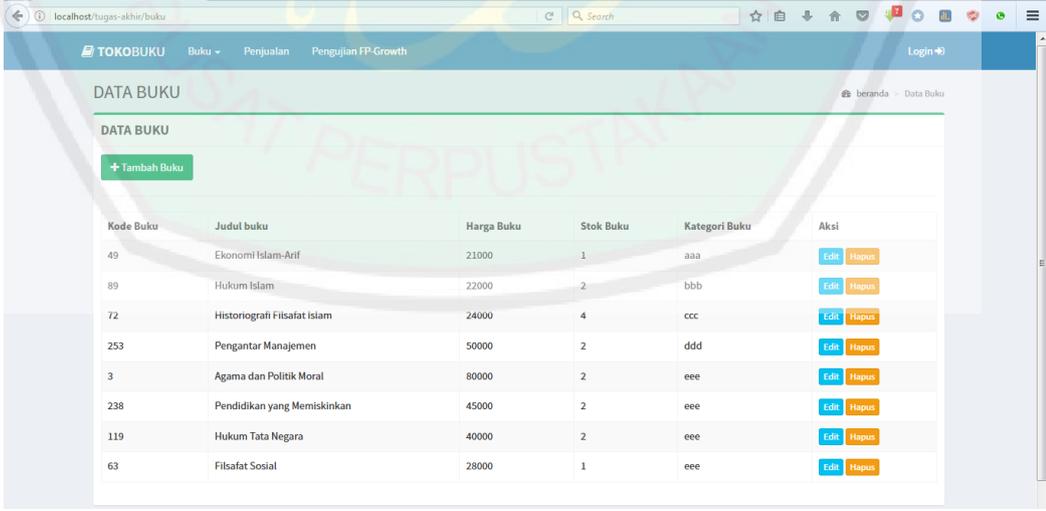
public function tambah()
{
    $data['kategori'] = $this->M_buku->data_kategori();
    $this->load->view('v_tambah_buku', $data);
}

public function edit()
{
    $data['buku'] = $this->M_buku->
    >edit_buku($_GET['kode']);
    $data['kategori'] = $this->M_buku->data_kategori();
    $this->load->view('v_edit_buku', $data);
}

```

#### 4.1.3 Proses Input Data Buku

Proses penginputan data kategori yang telah ada sebelumnya. Data kategori akan muncul pada list kategori yang ada pada form tambah buku. Jumlah buku koleksi yang ada di dalam data base, nantinya akan digunakan sebagai data transaksi penjualan.



Kode Buku	Judul buku	Harga Buku	Stok Buku	Kategori Buku	Aksi
49	Ekonomi Islam-Arif	21000	1	aaa	Edit Hapus
89	Hukum Islam	22000	2	bbb	Edit Hapus
72	Historiografi Filsafat Islam	24000	4	ccc	Edit Hapus
253	Pengantar Manajemen	50000	2	ddd	Edit Hapus
3	Agama dan Politik Moral	80000	2	eee	Edit Hapus
238	Pendidikan yang Memiskinkan	45000	2	eee	Edit Hapus
119	Hukum Tata Negara	40000	2	eee	Edit Hapus
63	Filsafat Sosial	28000	1	eee	Edit Hapus

Gambar 4.3 Data Buku

```

public function tambah_buku()

```

```

{
$params = $this->input->post();

$data['kode_buku'] = $params['kode_buku'];

$data['judul_buku'] = $params['judul_buku'];

$data['pengarang_buku'] = $params['pengarang_buku'];

        $data['id_kategori_buku'] =
$params['kategori_buku'];

        $data['stok_buku'] = $params['stok_buku'];

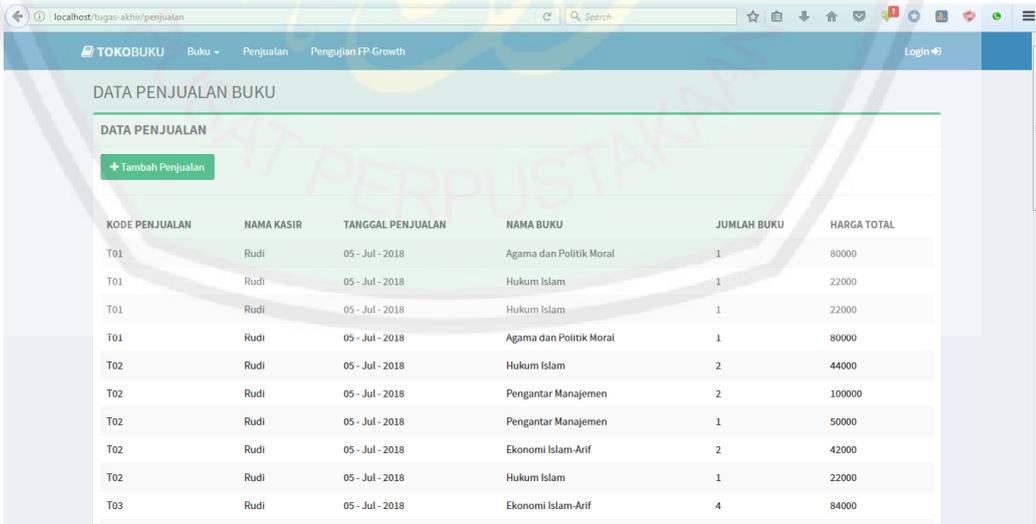
        $data['harga_buku'] = $params['harga_buku'];

$result = $this->M_buku->tambah_buku($data);
if ($result) {    redirect('buku');}
}

```

#### 4.1.4 Transaksi Data Penjualan

Data transaksi penjualan yang terjadi akan terekam pada database. Data transaksi inilah yang akan digunakan sebagai data pengujian sistem menggunakan algoritma Fp-Growth.



KODE PENJUALAN	NAMA KASIR	TANGGAL PENJUALAN	NAMA BUKU	JUMLAH BUKU	HARGA TOTAL
T01	Rudi	05 - Jul - 2018	Agama dan Politik Moral	1	80000
T01	Rudi	05 - Jul - 2018	Hukum Islam	1	22000
T01	Rudi	05 - Jul - 2018	Hukum Islam	1	22000
T01	Rudi	05 - Jul - 2018	Agama dan Politik Moral	1	80000
T02	Rudi	05 - Jul - 2018	Hukum Islam	2	44000
T02	Rudi	05 - Jul - 2018	Pengantar Manajemen	2	100000
T02	Rudi	05 - Jul - 2018	Pengantar Manajemen	1	50000
T02	Rudi	05 - Jul - 2018	Ekonomi Islam-Arif	2	42000
T02	Rudi	05 - Jul - 2018	Hukum Islam	1	22000
T03	Rudi	05 - Jul - 2018	Ekonomi Islam-Arif	4	84000

**Gambar 4.4** Data Penjualan

```
public function tambah_penjualan()
```

```

{
    $params = $this->input->post();
    $data['kode_penjualan'] = $params['kode_penjualan'];
    $data['nama_kasir'] = $params['nama_kasir'];
    $data['kode_buku'] = $params['kode_buku'];
    $data['tanggal_transaksi'] = date('Y-m-d H:i:s');
    $data['jumlah_buku'] = $params['jumlah_buku'];
    $res = $this->M_buku->tambah_penjualan($data);
    if ($res) {
        redirect('penjualan');
    }
}

```

#### 4.1.5 Pengujian Metode

##### 4.1.5.1 Input nilai Minimum Support dan Confidence

Nilai Input dari minimum support ini yang akan digunakan sebagai acuan perhitungan dari frequent item set dengan menghitung nilai setiap item dibagi dengan keseluruhan transaksi yang telah dilakukan.

**Gambar 4.5** Pemasukan Nilai Support dan Confidence

```

<form action="<?php echo site_url('metode/pengujian') ?>"
method='POST' class="form-horizontal">

    <div class="form-group">

        <label class="control-label col-
md-2">MIN SUPPORT</label>

        <div class="col-md-8">

            <div class="input-group">

                <input type="text"
name="min_support" class="form-control">

                <span class="input-
group-addon"><i>%</i></span>

            </div>

        </div>

    </div>

    <div class="form-group">

        <label class="control-label col-
md-2">MIN CONFIDENCE</label>

        <div class="col-md-8">

            <div class="input-group">

                <input type="text"
name="min_confidence" class="form-control">

                <span class="input-
group-addon"><i>%</i></span>

            </div>

        </div>

    </div>

```

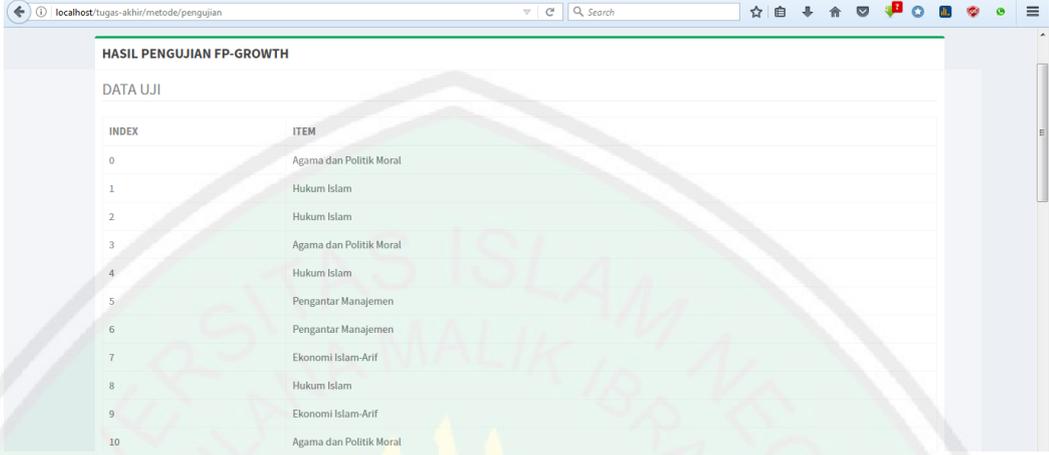
```

class Metode extends CI_Controller
{
    public $frequentItem = array();
    public $minimumSupportCount;
    public $minConfidence;
    public $supportCount = array();
    public $orderedFrequentItem = array();
}

```

#### 4.1.5.2 Menampilkan Data Uji (Item Set)

Data awal yang ditampilkan yaitu data item yang diambil berdasarkan judul buku.



INDEX	ITEM
0	Agama dan Politik Moral
1	Hukum Islam
2	Hukum Islam
3	Agama dan Politik Moral
4	Hukum Islam
5	Pengantar Manajemen
6	Pengantar Manajemen
7	Ekonomi Islam-Arif
8	Hukum Islam
9	Ekonomi Islam-Arif
10	Agama dan Politik Moral

**Gambar 4.6** Data Item Yang Terindeks

```

public function index()
{
    $this->load->view('v_metode');
}

public function set($t)
{
    if (is_array($t)) {
        $this->frequentItem[] = $t;
    }
}

public function get()
{
    // echo "<pre>";
    return $this->frequentItem;
    // echo "</pre>";
}

```

```

public function getFrequentItem()
{
    // echo "<pre>";
    // print_r($this->frequentItem);
    // echo "</pre>";
    return $this->frequentItem;
}

```

### 4.1.5.3 Frequent Item Set

Data item yang telah dihitung dengan nilai support per item dibagi dengan seluruh transaksi yang terjadi. Kemudian data ditampilkan menggunakan Array



ITEM	ORDER
Agama dan Politik Moral	5
Hukum Islam	4
Pengantar Manajemen	3
Ekonomi Islam-Arif	4
Historiografi Filsafat Islam	2
Pendidikan yang Memiskinkan	2

**Gambar 4.7** Data Item dengan Nilai Support

```

public function countSupportCount($data)
{
    if(is_array($data))
    {
        foreach ($data as $key => $value) {
            if(is_array($value))
            {
                foreach ($value as $k => $v)
                {
                    if (empty($this->supportCount[$v])) {
                        $this->supportCount[$v] = 1;
                    }else{

```



#### 4.1.5.5 Tampilan Rekomendasi Judul Buku

Hasil yang diperoleh setelah melalui serangkaian perhitungan berdasarkan Minimum Support sistem dan Minimum Confidence yang telah diinputkan di awal pada **Gambar 4.5**. Hasil rule yang terbentuk pada sistem yang berjalan yaitu Jika Membeli Buku RINGKASAN HUKUM PERNIKAHAN maka akan membeli Buku ANATOMI DAN PERKEMBANGAN SOSIAL dengan nilai Lifratio 0.198.

JUDUL BUKU	LIFRATIO
Jika Membeli Buku RINGKASAN HUKUM PERNIKAHAN maka akan membeli Buku ANATOMI DAN PERKEMBANGAN SOSIAL	0.198
Jika Membeli Buku HUKUM PENGADAAN TANAH maka akan membeli Buku KONSEP DAN TEORI GERAKAN SOSIAL	0
Jika Membeli Buku EKONOMI ISLAM - HAVIS maka akan membeli Buku SOSIOLOGI PENDIDIKAN	0
Jika Membeli Buku KONSEP DAN TEORI GERAKAN SOSIAL maka akan membeli Buku PENGANTAR HUKUM INDONESIA - REVISI	0.178

**Gambar 4.9** Rekomendasi Judul Buku

```
$benchmark = $itemconsequent/$jml;
$lifratio = $confidence/$benchmark;
$lifratio = number_format($lifratio,2);
```

## 4.2 Integrasi Sains dalam Islam

فَإِنَّ مَعَ الْعُسْرِ يُسْرًا . إِنَّ مَعَ الْعُسْرِ يُسْرًا

“Karena sesungguhnya sesudah kesulitan itu ada kemudahan. Sesungguhnya sesudah kesulitan itu ada kemudahan.”

(QS. Asy-Syarh: 5-6)

Allah SWT memberitahukan bahwa bersama kesulitan terdapat kemudahan. Kemudian dia mempertegas berita tersebut Ibnu Jarir meriwayatkan dari al-Hasan, Dia

Berkata”Nabi Muhammad pernah keluar rumah pada suatu hari dalam keadaan senang dan gembira dan beliau juga dalam keadaan tertawa seraya bersabda.

*“Satu kesulitan itu tidak akan pernah mengalahkan dua kemudahan. Satu kesulitan itu tidak akan pernah mengalahkan dua kemudahan, karena bersama kesulitan itu terdapat kemudahan sesungguhnya bersama kesulitan itu terdapat kemudahan.”*

Di era teknologi yang maju seperti saat ini, memberikan kemudahan dengan memanfaatkan sistem informasi sangat membantu mempermudah pekerjaan. Sehingga lebih cepat dan efektif dalam menyelesaikan pekerjaan.

Dalam suatu pekerjaan pasti terdapat permasalahan terkadang membuat depresi ketika tidak dapat mencari solusi dari permasalahan tersebut. Keputusan, keraguan, kadang selalu menghampiri untuk menekankan bahwa disetiap permasalahan/kesulitan terdapat solusi/kemudahan maka pernyataan Allah SWT, dikuatkan dengan ﴿قُلْ﴾ atau sesungguhnya. Keputusan yang dapat menjadikan pengingkaraan atas

kemahakuasaan Allah SWT maka Allah SWT mengulangi lagi pernyataan tersebut dengan menggunakan kalimat yang sama, *إِنَّ مَعَ الْعُسْرِ يُسْرًا* sesungguhnya bersama dengan kesulitan itu pasti ada kemudahan. Sudah barang tentu, makna yang tercakup dalam pernyataan yang kedua ini, lebih menguatkan daripada yang pertama.

Ibnu Katsir menegaskan, bahwa kesulitan itu dapat diketahui pada dua keadaan dimana kalimatnya dalam bentuk *mufrad* Sedangkan kemudahan (*al-yusr*) dalam bentuk *nakirah* (tidak ada ketentuannya) sehingga bilangannya bertambah banyak. Hal ini menunjukkan bahwa kedua ayat tersebut mengandung makna “setiap satu kesulitan akan dibarengi dengan dua kemudahan atau banyak kemudahan.

*Data mining* metode yang digunakan untuk mempermudah pengelolaan data yang jumlahnya sangat banyak. Menganalisa data yang jumlahnya sangat banyak sangat menguras banyak waktu. Dalam dunia serba teknologi, waktu begitu sangat berharga. Sebagaimana penjelasan ayat Al-Quran :

وَالْعَصْرِ (1) إِنَّ الْإِنْسَانَ لَفِي خُسْرٍ (2) إِلَّا الَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ وَتَوَّصَوْا بِالْحَقِّ وَتَوَّصَوْا بِالصَّبْرِ

”Demi masa. Sesungguhnya manusia itu benar-benar berada dalam kerugian. Kecuali orang-orang yang beriman dan mengerjakan amal sholeh dan saling menasihati supaya menaati kebenaran dan saling menasihati supaya menepati kesabaran” (QS. Al ‘Ashr).

CV. Cita Intrans Selaras merupakan salah satu penerbit yang memiliki data penjualan yang besar. Untuk mengurangi SDM dan biaya maka di butuhkanlah aplikasi yang dapat membantu memudahkan tugas dan kewajiban di setiap devisi. Terutama direktur yang berperan sebagai pemberi keputusan untuk kemajuan perusahaannya.

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Dalam penelitian ini, dapat disimpulkan bahwa Algoritma FP-Growth dapat digunakan didalam kasus rekomendasi judul buku dan mendapatkan hasil keluaran berupa rekomendasi judul bukuyang sesuai.

Berdasarkan penerapan algoritma FP-Growth kedalam sistem berbasis web, sistem ini hanya menampilkan dua kombinasi itemset yang frequent dengan menggunakan 774 data, dan terbentuk aturan asosiasi dengan nilai support terbesar 30% dan confidence sebesar 50% ..

#### **5.2 Saran**

Berdasarkan hasil penelitian, ada beberapa saran yang dapat dilakukan guna 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 aturan asosiasi lainnya dengan menggabungkan algoritma FP-Growth dengan metode lainnya untuk menghasilkan suatu pengetahuan yang baru.

## DAFTAR PUSTAKA

- Alfannisa Annurullah Fajrin, A. M. (2018). Penerapan Data Mining Untuk Analisis Pola Pembelian Konsumen dengan Algoritma FP-Growth pada Data Transaksi Penjualan Spare Part Motor. *KLIK*, 05, 27-36.
- andyku. (n.d.). Retrieved mei 14, 2018, from [http://andyku.wordpress.com/2008/04/17/data-mining-dan web-mining](http://andyku.wordpress.com/2008/04/17/data-mining-dan-web-mining)
- Brian D. Foster, J. W. (2000). Relationship selling and customer loyalty :an Empirical Investigation. *emeraldinsight*, 18(4), 185-199.
- Erwin. (2009). Analisis Market Basket dengan Algoritma Apriori dan Algoritma FP-Growth. *Jurnal Generik*. hal. 26-30.
- Larose, D. T. (2005). *Discovering Knowledge in Data : An Introduction*.
- Meilani, A. (2015). PENENTUAN POLA YANG SERING MUNCUL UNTUK PENERIMA KARTU JAMINAN KESEHATAN MASYARAKAT (JAMKESMAS) MENGGUNAKAN METODE FP-GROWTH. *IDeaTech*, 424-431.
- Rama Novta Miraldi, A. R. (2014). Implementasi Algoritma FP-GROWTH untuk Sistem Rekomendasi Buku di Perpustakaan UKDW. *INFORMATIKA*, 10, 29-39.
- Tan, P. K. (2004). Selecting the Right Objective Measure for Association Analysis. *Information Systems*, 293-313.