

**IMPLEMENTASI DATA *MINING* DENGAN ALGORITMA *FP-GROWTH* UNTUK Mendukung Strategi Promosi Pendidikan ( Studi Kasus : Universitas Islam Kuantan Singingi )**

**SKRIPSI**

**Oleh :**

**DESPIK MICE  
NPM. 1502210134**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM KUANTAN SINGINGI  
2022**

**IMPLEMENTASI DATA *MINING* DENGAN ALGORITMA *FP-GROWTH* UNTUK Mendukung Strategi Promosi Pendidikan ( Studi Kasus : Universitas Islam Kuantan Singingi )**

**SKRIPSI**

**DIAJUKAN SEBAGAI SALAH SATU SYARAT UNTUK Mencapai Gelar Sarjana Program Studi Teknik Informatika**

Oleh :

**DESPIK MICE  
NPM. 1502210134**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM KUANTAN SINGINGI  
2022**

## PERSETUJUAN SKRIPSI

NPM : 150210134  
Nama : DESPIK MICE  
Jenjang Studi : Strata Satu (S1)  
Program Studi : Teknik Informatika  
Judul Proposal : Implementasi Data *Mining* Dengan Algoritma *Fp-Growth* Untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi)

Disetujui Oleh :

Pembimbing I,

(FEBRI HASWAN, M.Kom )  
NIDN. 1009028803

Tanggal. 15 Maret 2022

Pembimbing II

( NOFRI WANDI AL-HAFIZ, M.Kom )  
NIDN. 1002118802

Tanggal. 28 Maret 2022

Mengetahui,  
Ketua prodi teknik informatika

( JASRI, S.Kom, M.kom )  
NIDN. 1001019001

Tanggal. 30 Maret 2022

Tanggal Lulus : 30 Maret 2022

## TANDA PENGESAHAN SKRIPSI

NPM : 150210134  
Nama : DESPIK MICE  
Jenjang Studi : Strata Satu (S1)  
Program Studi : Teknik Informatika  
Judul Proposal : Implementasi Data *Mining* Dengan Algoritma *Fp-Growth*  
Untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus :  
Universitas Islam Kuantan Singingi)

Dipertahankan di depan Tim Penguji Skripsi  
Program Studi Teknik Informatika, Fakultas Teknik,  
Universitas Islam Kuantan Singingi  
Pada Tanggal : 30 Maret 2022

### Dewan Penguji

No	Nama	Jabatan	Tanda Tangan
1.	Chitra Hermawan M.T	Ketua	
2.	Febri Haswan M,Kom	Pembimbing I	
3.	Nofri Wandu Al-Hafiz, M.Kom	Pembimbing II	
4.	Jasri, S.Kom, M.kom	Penguji I	
5.	Aprizal, M.Kom	Penguji II	

### Mengetahui,

Dekan,  
Fakultas Teknik

Ketua,  
Prodi Teknik Informatika

Chitra Hermawan M.T  
NIDN. 1022068901

Jasri, S.Kom, M.kom  
NIDN. 1001019001

## PERNYATAAN

Saya yang bertanda tangan di bawah ini :

NPM : 150210134

Nama : DESPIK MICE

Tempat / Tgl Lahir : Koto Cengar / 12-12-1981

Alamat : Sei. Jering Teluk Kuantan

Saya menyatakan bahwa dalam skripsi yang berjudul ” Implementasi Data *Mining* Dengan Algoritma *Fp-Growth* Untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi)” tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana komputer disuatu perguruan tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka. Atas pernyataan ini dibuat saya siap menanggung segala resiko dan sanksi apabila dikemudian hari ditemukan pelanggaran terhadap etika keilmuan dalam skripsi saya ini.

Teluk Kuantan, 30 Maret 2022

( DESPIK MICE )

**IMPLEMENTASI DATA MINING DENGAN ALGORITMA *FP-GROWTH*  
UNTUK Mendukung STRATEGI PROMOSI Pendidikan (STUDI  
KASUS : UNIVERSITAS ISLAM KUANTAN SINGINGI)**

**ABSTRAK**

Dalam melakukan promosi pendidikan tentunya harus mempunyai cara yang tepat dalam melakukan promosi, hal ini bertujuan untuk dapat memaksimalkan biaya ataupun waktu dalam melakukan promosi pendidikan. Untuk terus meningkatkan jumlah mahasiswa yang mendaftar ke Universitas Islam Kuantan Singingi, karena berdasarkan data yang di dapatkan penulis jumlah mahasiswa yang mendaftar setiap tahunnya tidak stabil, jika dibandingkan dengan data mahasiswa yang mendaftar tahun 2018, 2019 dan tahun 2020 maka dalam 3 fase terakhir ini terjadi turun naik dalam pendaftaran mahasiswa baru. Jika dibandingkan ketiga tahun ini maka sangat terlihat grafik ketidakstabilan mahasiswa baru yang mendaftar ke Universitas Islam Kuantan Singingi, padahal promosi selalu dilakukan oleh bagian akademik dengan cara membuat brosur, baliho dan media online untuk dapat menarik minat pelajar supaya menjadikan Universitas Islam Kuantan Singingi menjadi pilihan utama dalam dunia Pendidikan untuk menggapai cita – cita yang diinginkan. Maka dengan terjadinya ketidakstabilan ini penulis melakukan penelitian dengan memanfaatkan data mahasiswa yang tersimpan dalam *database* untuk diolah menggunakan metode *Fp – Growth* sehingga didapatkan prediksi untuk dijadikan acuan dalam melakukan promosi pendidikan.

**Kata Kunci :** Uniks, *Fp-Growth*, Pendaftaran.

**IMPLEMENTASI DATA MINING DENGAN ALGORITMA *FP-GROWTH*  
UNTUK Mendukung STRATEGI PROMOSI Pendidikan (STUDI  
KASUS : UNIVERSITAS ISLAM Kuantan Singingi)**

**ABSTRACT**

In carrying out promotional education, of course, you must have the right way of doing promotions, this aims to be able to maximize the cost or time in conducting education. to continue to increase the number of students who register at the Kuantan Singingi Islamic University, because based on the data obtained the number of students who register each year is unstable, when compared to the data of students who registered in 2018, 2019 and 2020, in the last 3 phases there were ups and downs in student registration. When compared to the three this year, the graph of the instability of new students who register at the Kuantan Singingi Islamic University is very visible, even though promotions are always carried out by the academic department by making brochures, billboards and online media to attract students' interest, making Kuantan Singingi Islamic University the main choice in world of Education to achieve the desired goals. So with this instability the author conducted research by utilizing student data stored in the database to be used using the Fp - Growth method so that predictions were obtained to be used as references in promoting education.

**Keywords:** *Uniques, Fp-Growth, Registration.*

## **Riwayat Hidup**

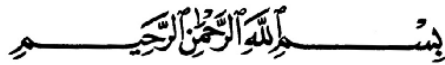
Penulis bernama Despik Mice umur 40 tahun, dilahirkan di Desa Koto Cengar Kecamatan Kuantan Mudik Tanggal 12 Desember 1981. Penulis beragama Islam, anak kedua dari tiga bersaudara dari pasangan Bapak Rusaham dan Ibu Darni ZN, S.Pd. Pendidikan formal dimulai Sekolah Dasar Negeri di SDN 021 Koto Cengar Tahun 1988-1994, Sekolah Lanjutan Tingkat Pertama di SLTP N 1 Lubuk Jambi Tahun 1995-1997, Sekolah Menengah Kejuruan di SMK N 2 Teluk Kuantan Jurusan Akuntansi Tahun 1998-2000, selanjutnya di Sekolah Tinggi Manajemen Informatika Komputer di STMIK Indonesia Padang Tahun 2000-2003, selanjutnya Keguruan di Universitas Islam Riau Tahun 2006-2007, Selanjutnya penulis melanjutkan pendidikan di S1 Teknik Informatika Fakultas Teknik Universitas Islam Kuantan Singingi. Penulis juga mengikuti Diklat Pengelolaan Barang Milik Daerah di Politeknik Keuangan Negara STAN Tangerang Selatan dan Bimtek Pengelolaan Barang Milik Daerah Barbasis Teknologi Informasi di LKKN Medan.

Teluk Kuantan, 30 Maret 2022

Despik Mice



## KATA PENGANTAR



*Assalamu 'alaikum wa rahmatullahi wa barokaatuh.*

*Alhamdulillah Robbil'aalamiin*, Puji dan syukur penulis panjatkan ke hadirat Allah subhanahu wa ta'ala yang telah melimpahkan kasih dan sayang-Nya kepada kita, sehingga penulis bisa menyelesaikan Laporan Skripsi dengan tepat waktu yang berjudul **“Implementasi Data Mining dengan Algoritma FP-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi)”** Tujuan dari penyusunan skripsi ini guna memenuhi salah satu syarat untuk bisa menempuh ujian sarjana Program Studi Teknik Informatika pada Fakultas Teknik di Universitas Islam Kuantan Singingi (UNIKS).

Didalam pengerjaan skripsi ini telah melibatkan banyak pihak yang sangat membantu dalam banyak hal. Oleh sebab itu, disini penulis sampaikan rasa terima kasih sedalam-dalamnya kepada :

1. Bapak **Prof. Dr. H. Zulfan Sa'am, M.Si** . Selaku Ketua Yayasan Pendidikan Tinggi Islam Kuantan Singingi.
2. Bapak **Dr. H. Nopriadi, S.KM, M.Kes** selaku Rektor Universitas Islam Kuantan Singingi, Riau.
3. Bapak **Chitra Hermawan M.T** selaku Dekan Fakultas Teknik Universitas Islam Kuantan Singingi, Riau.
4. Bapak **Jasri, S,Kom, M.Kom** Ketua Program Studi Teknik Informatika .
5. Bapak **Febri Haswan M,Kom** dan Bapak **Nofri Wandu Al-Hafiz, M.Kom** selaku dosen pembimbing yang telah memberikan masukan dan saran dalam penyusunan Skripsi ini..

6. Terimakasih kepada orang tua tercinta, Bapak **Rusaham** dan Ibu **Alm. Darni ZN, S.Pd** dan juga ananda tersayang **Muhammad Rayhan Purnama** serta keluarga atas semua doa dan dukungan yang selalu diberikan.
7. Terimakasih kepada teman-teman seperjuangan Teknik Informatika yang terus memberikan bantuan dan *support*.
8. Bapak / Ibu Dosen serta seluruh karyawan-karyawati Universitas Islam Kuantan Singingi yang juga telah banyak membantu penulis dalam menyelesaikan Skripsi ini.

Semoga Skripsi ini dapat bermanfaat bagi penulis khususnya maupun pembaca umumnya. Penulis sadar masih banyak kekurangan pada penulisan laporan ini. Oleh karena itu penulis berharap bisa mendapatkan masukan dari pembaca atas isi skripsi ini. Akhir kata penulis ucapkan terimakasih dan selamat membaca.

Teluk Kuantan, 30 Maret 2022

Despik Mice

## DAFTAR ISI

	Halaman
<b>HALAMAN SAMPUL</b> .....	<b>ii</b>
<b>PERSETUJUAN SKRIPSI</b> .....	<b>iii</b>
<b>TANDA PENGESAHAN SKRIPSI</b> .....	<b>iv</b>
<b>PERNYATAAN</b> .....	<b>v</b>
<b>ABSTRAK</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>vii</b>
<b>RIWAYAT HIDUP</b> .....	<b>viii</b>
<b>KATA PENGANTAR</b> .....	<b>ix</b>
<b>DAFTAR ISI</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>DAFTAR TABEL</b> .....	<b>xiv</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Identifikasi Masalah .....	3
1.3 Rumusan Masalah .....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Ruang Lingkup Penelitian.....	5
1.7 Sistematika Penulisan .....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>7</b>
2.1 Teoritis .....	7
2.2 Tinjauan Penelitian Terdahulu .....	21
<b>BAB III METODOLOGI PENELITIAN</b> .....	<b>24</b>
3.1 Tahapan – tahapan Penelitian.....	24
3.2 Waktu dan Lokasi Penelitian .....	24
3.3 Data yang Digunakan .....	25
3.4 Teknik Pengumpulan Data.....	25
<b>BAB IV ANALISA DAN PERANCANGAN APLIKASI</b> .....	<b>29</b>
4.1 Analisa Sistem.....	29
4.2 Perancangan Sistem.....	29
4.2.1. <i>Use Case Diagram</i> .....	30
4.2.2. <i>Activity Diagram</i> .....	31
4.2.2.1 <i>Activity Diagram</i> Login Admin .....	32
4.2.2.2 <i>Activity Diagram</i> Analisa Algoritma.....	32
4.2.2.3 <i>Activity Diagram</i> Laporan .....	34
4.2.2.4 <i>Activity Diagram</i> Melihat <i>History</i> .....	35
4.2.3. <i>Sequence Diagram</i> .....	35
4.3 <i>Class Diagram</i> .....	38
4.4. Desain Terinci .....	39
4.5. Struktur Tabel.....	44
<b>BAB V HASIL DAN PEMBAHASAN</b> .....	<b>47</b>
5.1 Praproses Data.....	47
5.2 Pemrosesan Data .....	47

5.3 Data Mining .....	47
5.4 <i>Transformation Data</i> .....	59
5.5 Analisis <i>Frequent Itemset</i> .....	68
5.6 Pembangunan <i>FP-Tree</i> .....	70
5.7 <i>Algoritma FP-Growth</i> .....	78
5.8 Pembentukan <i>Association Rules</i> .....	97
5.9 Implementasi Sistem.....	102
<b>BAB VI KESIMPULAN DAN SARAN.....</b>	<b>109</b>
6.1 Kesimpulan .....	109
6.2 Saran .....	109
<b>DAFTAR PUSTAKA.....</b>	<b>110</b>
<b>LAMPIRAN .....</b>	<b>111</b>

## DAFTAR GAMBAR

Gambar 2.1 Tahapan Proses KDD.....	10
Gambar 2.2 <i>Decision Tree</i> .....	12
Gambar 3.1 Alur Penelitian.....	24
Gambar 4. 1 <i>Use Case Diagram</i> .....	31
Gambar 4.2 <i>Activity Diagram Login Admin</i> .....	32
Gambar 4.3 <i>Activity Diagram Analisa Algoritma</i> .....	33
Gambar 4.4 <i>Activity Diagram Laporan</i> .....	34
Gambar 4.5 <i>Activity Diagram Melihat History</i> .....	35
Gambar 4.6 <i>Sequence Diagram Login</i> .....	356
Gambar 4.7 <i>Sequence Diagram Melihat History</i> .....	356
Gambar 4.8 <i>Sequence Diagram Analisis Algoritma</i> .....	357
Gambar 4.9 <i>Class Diagram</i> .....	38
Gambar 4.10 Desain Halaman Utama .....	39
Gambar 4.11 Desain <i>Output Data Sheet</i> .....	40
Gambar 4.12 Desain <i>Output Frequeen item</i> .....	40
Gambar 4.13 Desain <i>Output Assosiasi</i> .....	41
Gambar 4.14 Desain <i>Output History</i> .....	41
Gambar 4.15 Desain <i>Output Laporan</i> .....	42
Gambar 4.16 Desain <i>Login Admin</i> .....	43
Gambar 4.17 Desain <i>Input Data Assosiasi</i> .....	43
Gambar 4.18 Desain <i>Input Menyimpan Data</i> .....	44
Gambar 5.1 Alur Pohon <i>FP-Tree</i> .....	77
Gambar 5.2 Cabang yang berakhir D2 .....	79
Gambar 5.3 Cabang yang berakhir D1 .....	83
Gambar 5.4 Halaman <i>Login</i> .....	103
Gambar 5.5 Halaman <i>Dashboard</i> .....	103
Gambar 5.6 Halaman Data Pendaftaran .....	104
Gambar 5.7 Halaman Tambah Data .....	104
Gambar 5.8 Halaman Analisis Algoritma .....	105
Gambar 5.9 Halaman <i>Frequen Item</i> .....	105
Gambar 5.10 Halaman <i>Assosiasi</i> .....	106
Gambar 5.11 Halaman Hasil <i>Assosiasi</i> .....	107
Gambar 5.12 Halaman <i>History</i> .....	107
Gambar 5.13 Halaman Laporan.....	108
Gambar 5.14 Halaman Melihat <i>Print Laporan</i> .....	108

## DAFTAR TABEL

Tabel 4.1 Tabel <i>Users</i> .....	45
Tabel 4.2 Tabel <i>fp_growths</i> .....	45
Tabel 4.3 Tabel Dataset.....	46
Tabel 5.1 Data Uji Pendaftaran Mahasiswa Lama .....	48
Tabel 5.2 Kode pada masing - masing Item Atribut.....	60
Tabel 5.3. Data Itemset setelah dirubah dalam bentuk kode.....	62
Tabel 5.4. Nilai Support Per Item.....	68
Tabel 5.5. Data Itemset Data Mahasiswa yang memenuhi nilai support.....	70
Tabel 5.6 Kemunculan item cabang tree D2 .....	79
Tabel 5.7 Kemunculan cabang tree D1 .....	83
Tabel 5.8 Nilai Support dari Suffix D1 .....	89
Tabel 5.9 Nilai Support dari Suffix D2.....	92
Tabel 5.10 Nilai confidence dari Suffix .....	95
Tabel 5.11 Data Calon Nilai Confidence Dari Frequent Itemset .....	97
Tabel 5.12 Data Nilai Confidence Dari Frequent Itemset .....	98

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kemajuan teknologi membuat persaingan dalam mempromosikan pendidikan semakin maju, baik dengan menggunakan sosial media seperti *facebook*, *instagram*, dan situs *website* yang tentunya dapat menghemat biaya dalam melakukan promosi pendidikan. Tetapi dalam melakukan promosi pendidikan tentunya harus mempunyai cara yang tepat dalam melakukan promosi, hal ini bertujuan untuk dapat memaksimalkan biaya ataupun waktu dalam melakukan promosi pendidikan. Karena jika suatu promosi pendidikan dilakukan tanpa menggunakan cara yang tepat maka yang dilakukan akan percuma, contohnya melakukan promosi ke sekolah yang siswanya tidak pernah mendaftar ke Universitas yang melakukan promosi pendidikan.

Tujuan utama dari penelitian ini adalah untuk terus meningkatkan jumlah mahasiswa yang mendaftar ke Universitas Islam Kuantan Singingi, karena berdasarkan data yang di dapatkan penulis jumlah mahasiswa yang mendaftar setiap tahunnya tidak stabil, jika dibandingkan dengan data mahasiswa yang mendaftar tahun 2018, 2019 dan tahun 2020 maka dalam 3 fase terakhir ini terjadi turun naik dalam pendaftaran mahasiswa baru. Jika dibandingkan ketiga tahun ini maka sangat terlihat grafik ketidakstabilan mahasiswa baru yang mendaftar ke Universitas Islam Kuantan Singingi, padahal promosi selalu dilakukan oleh bagian akademik dengan cara membuat brosur, baliho dan media online untuk dapat menarik minat pelajar supaya menjadikan Universitas Islam Kuantan Singingi menjadi pilihan utama dalam dunia Pendidikan untuk menggapai cita –

cita yang diinginkan, tentunya dengan berbagai promosi yang telah dilakukan sangat perlu cara yang tepat dalam melakukan promosi pendidikan, maka dengan terjadinya ketidakstabilan ini penulis melakukan penelitian dengan memanfaatkan data mahasiswa yang tersimpan dalam *database* untuk diolah menggunakan metode *Fp – Growth* sehingga didapatkan prediksi untuk dijadikan acuan dalam melakukan promosi pendidikan.

Dalam penelitian ini, algoritma yang digunakan adalah *Frequent Pattern-Growth (FP-Growth)* yaitu pengembangan dari metode *Apriori* yang merupakan salah satu alternatif untuk menentukan himpunan data yang paling sering muncul (*frequent itemset*) dalam sebuah kumpulan data dengan membangkitkan struktur data *Tree* atau disebut dengan *Frequent Pattern Tree* (Fadlina 2014). Khususnya dalam hal ini adalah penelitian tentang bagaimana mengolah data mahasiswa yang telah mendaftar di Universitas Islam Kuantan Singingi, dengan menggunakan metode algoritma *FP-Growth* dapat membantu dalam menemukan strategi yang tepat dalam melakukan promosi pendidikan dengan melihat himpunan data yang sering muncul setelah dilakukan perhitungan.

Salah satu contoh dari pengembangan data *mining* yang juga mendorong penulis dalam mengembangkan data *mining* dengan menggunakan metode yang serupa adalah *Asosiasi Data Mining Menggunakan Algoritma Fp-Growth Untuk Market Basket Analysis* yang ditulis oleh Fathimah Fatihatul, Atje Setiawan, Rudi Rosadi Jurusan Matematika FMIPA Universitas Padjadjaran. Dengan menggunakan bantuan sistem yang dibuat ini maka dapat dilakukan strategi – strategi dalam hal ini adalah penjualan. Maka dengan demikian penulis juga ingin memanfaatkan data Mahasiswa Baru Universitas Islam Kuantan Singingi untuk



dikembangkan menjadi informasi yang bermanfaat terutama dalam promosi pendidikan Universitas Islam Kuantan Singingi.

Dengan adanya metode ini, diharapkan dapat mempercepat dan mempermudah manajemen sebuah universitas dalam melakukan promosi pendidikan untuk meningkatkan jumlah calon mahasiswa baru yang mendaftar pada Universitas untuk tahun berikutnya. Berdasarkan hal yang dikemukakan di atas, maka proposal ini diberi judul **“Implementasi Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi)”**

## **1.2 Identifikasi Masalah**

Dari beberapa uraian yang dikemukakan pada latar belakang, maka dapat diidentifikasi masalah-masalah sebagai berikut :

1. Menumpuknya data mahasiswa baru yang mendaftar ke Universitas Islam Kuantan Singingi yang tersimpan di dalam *database* pendaftaran mahasiswa tahun 2020.
2. Tidak adanya strategi baru yang muncul dalam rangka melakukan promosi pendidikan.
3. Promosi pendidikan hanya dilakukan dengan cara membuat spanduk, memberikan brosur dan memasang iklan pada media cetak dan elektronik.
4. Saat melakukan promosi pendidikan ke sekolah tidak menggunakan data sekolah mana yang paling banyak mendaftar pada jurusan yang terdapat pada Universitas Islam Kuantan Singingi.

### 1.3 Rumusan Masalah

Dari latar belakang masalah diatas maka dapat dirumuskan sebuah permasalahan yang akan dibahas yaitu : **Bagaimana algoritma *FP-Growth* dapat diimplementasikan untuk mendukung strategi dalam promosi pendidikan pada Universitas Islam Kuantan Singingi ?**

### 1.4 Tujuan Penelitian

Sesuai dengan permasalahan yang telah dikemukakan, maka tujuan utama yang ingin dicapai dalam penelitian ini adalah dengan menggunakan metode algoritma *FP-Growth* di harapkan dapat membantu dalam melakukan promosi pendidikan setiap tahunnya dengan tujuan untuk menambah mahasiswa yang mendaftar di Universitas Islam Kuantan Singingi semakin meningkat setiap tahunnya.

### 1.5 Manfaat Penelitian

Diharapkan dengan adanya penelitian ini dapat menunjang berkembangannya ilmu pengetahuan dalam dunia pendidikan. Adapun manfaat dari penelitian ini sebagai berikut:

1. Dapat menjadi solusi dalam memanfaatkan data yang tersimpan dalam *database* menjadi sebuah informasi yang dapat dimanfaatkan untuk strategi promosi pendidikan.
2. Memudahkan pihak kampus menemukan strategi atau sasaran yang lebih tepat dalam mempromosikan pendidikan.

3. Mengetahui algoritma *FP-Growth* dapat memberikan alternatif pengetahuan tentang menemukan strategi dalam promosi pendidikan.

## **1.6 Ruang Lingkup Penelitian**

Untuk lebih mengarahkan penulisan proposal ini pada tujuan yang diinginkan, maka perlu ruang lingkup masalah agar tidak menyimpang dari tujuan yang telah dikemukakan. Adapun batasan masalah dalam pembuatan proposal ini adalah:

1. Data yang digunakan adalah data PMB Universitas Islam Kuantan Singingi dari tahun ajaran 2020.
2. Atribut yang digunakan adalah Program Studi, Asal Sekolah, Jenis SLTA, Jenis Klamin.

## **1.7 Sistematika Penulisan**

Pada penelitian ini direncanakan terdiri dari lima bab. Bab-bab ini saling berkaitan satu sama lain. Sistematika penulisan ini adalah :

### **BAB I : PENDAHULUAN**

Pada bab ini membahas tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

### **BAB II : TINJAUAN PUSTAKA**

Pada bab ini menjelaskan tentang teori-teori yang digunakan untuk memahami permasalahan yang dibahas pada penelitian ini seperti: *knowledge discovery in database*, *data mining* secara umum,

*association rules mining*, algoritma *FP-Growth* serta aplikasi pendukung lainnya.

### **BAB III : METODOLOGI PENELITIAN**

Bab ini merupakan penjabaran dari metode analisa yang digunakan yaitu pendekatan terstruktur dengan model prosesnya.

### **BAB IV : ANALISA DAN PEMBAHASAN**

Bab keempat akan dibahas tentang analisa yang lama kemudian menganalisa rancangan baru yang akan dibuat.

### **BAB V : IMPLEMENTASI SISTEM**

Bab kelima akan dibahas tentang menampilkan tampilan - tampilan sistem yang sudah dibuat.

### **BAB VI : KESIMPULAN DAN SARAN**

Bab ini akan menguraikan kesimpulan hasil penelitian dan juga memberikan saran-saran yang sekiranya diperlukan.

### **DAFTAR PUSTAKA**

Daftar pustaka ini membahas tentang daftar dari buku, skripsi dan jurnal yang akan dijadikan kajian untuk penelitian.

### **LAMPIRAN**

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1 Teoritis**

##### **2.1.1 Definisi Sistem**

Definisi sistem menurut Mulyadi (2016), sistem pada dasarnya merupakan sekelompok unsur yang erat dan berhubungan satu dengan yang lainnya, yang berfungsi bersama-sama untuk mencapai tujuan tertentu. Menurut Romney dan Steinbert (2015), sistem adalah serangkaian dua atau lebih komponen yang saling terkait dan berinteraksi untuk mencapai tujuan, terdiri dari subsistem yang mendukung sistem yang lebih besar. Sedangkan menurut Jogiyanto (2010), sistem adalah suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan kegiatan atau menyelesaikan suatu sasaran tertentu, sistem dapat didefinisikan dengan pendekatan prosedur dengan pendekatan komponen.

Berdasarkan dari definisi-definisi diatas, dapat disimpulkan bahwa sistem adalah sekumpulan elemen-elemen atau fungsi-fungsi tertentu yang bekerja dalam suatu proses untuk mencapai suatu tujuan tertentu.

##### **2.1.2 Definisi Informasi**

Jogiyanto (2010) menyatakan, informasi dapat didefinisikan sebagai hasil pengolahan data dalam bentuk yang berguna dan berarti bagi penerimanya yang menggambarkan kejadian-kejadian yang nyata yang digunakan untuk pengambilan keputusan.

Sedangkan menurut Romney dan Steinbart (2015), informasi adalah data yang sudah diproses dan diorganisasikan arti bagi penggunanya. Bodnar dan Hopwood (2009) informasi merupakan data yang diolah sedemikian rupa, sehingga bisa dijadikan dasar dalam mengambil keputusan yang tepat dan benar.

Berdasarkan pendapat yang dikemukakan diatas menunjukkan bahwa informasi merupakan hasil pengolahan data yang berguna sebagai dasar untuk mengambil keputusan.

### **2.1.3 Definisi Sistem Informasi**

Sistem informasi didefinisikan oleh Jogiyanto (2010), Sistem informasi adalah suatu sistem didalam suatu organisasi yang mempertemukan kebutuhan pengolahan transaksi harian, mendukung operasi, bersifat manajerial dan kegiatan strategi dari suatu organisasi dan menyediakan pihak luar tertentu dengan laporan-laporan yang diperlukan.

Disimpulkan dari definisi tersebut, menurut penulis sistem informasi adalah kumpulan elemen-elemen atau fungsi-fungsi yang membentuk sistem dan memberikan output berupa laporan yang berguna bagi pemakai sistem informasi tersebut.

### **2.1.4 Data**

Data adalah sekumpulan fakta yang nantinya jika diolah dengan baik akan memberikan informasi kepada orang lain dengan ketentuan informasi tersebut harus memiliki kualitas yang baik. Kualitas informasi dipengaruhi oleh beberapa faktor, yaitu keakuratan, kerelevanan, tepat waktu dan penyajiannya.

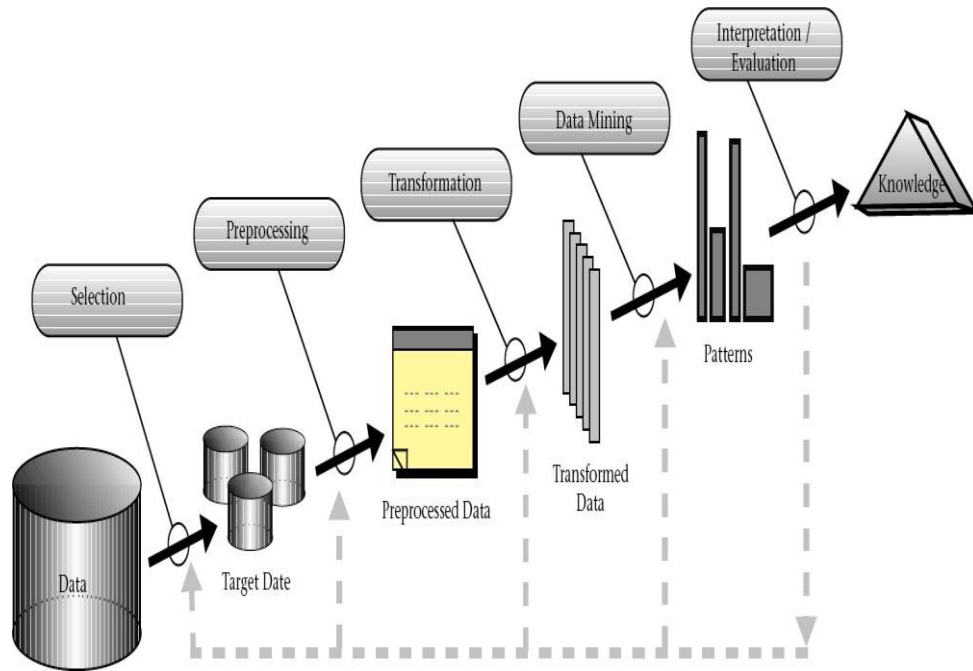
Informasi ini akan digunakan untuk penentuan keputusan dan kebijakan serta langkah apa yang akan sebaiknya dilakukan selanjutnya, jadi proses dalam mendapatkan informasi ini haruslah menggunakan proses dan teknologi yang dapat dipercaya kebenarannya.( Fathimah Fatihalu, Rudi Rosadi, 2016).

### **2.1.5 Basis Data**

Basis data (*database*) menurut Edhy Sutanta, (2004;18) dapat dipahami sebagai suatu kumpulan data terhubung (*interrelated data*) yang disimpan secara bersama-sama pada suatu media tanpa suatu kerangkapan data walaupun ada maka kerangkapan data tersebut harus seminimal mungkin dan terkontrol (*controlled redudncy*), data di disimpan dengan cara-cara tertentu, sehingga mudah untuk digunakan atau ditampilkan kembali.

### **2.1.6 Knowledge Discovery In Database (KDD)**

Data mining dan *knowledge discovery in databases* (KDD) sering digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Secara umum kedua istilah di atas mempunyai konsep yang berbeda akan tetapi selalu mempunyai keterkaitan antara satu dan yang lain. *Knowledge Discovery in Database* (KDD) adalah proses menentukan informasi yang berguna serta pola-pola yang ada dalam data, Informasi ini terkandung dalam basis data yang berukuran besar yang sebelumnya tidak diketahui dan potensial bermanfaat. (Irwan Budiman, Toni Prahasto, Yuli Cristyanto, 2012).



**Gambar 2.1** Tahapan Proses KDD

Adapun proses *Knowledge Discovery in Databases* sebagai berikut :

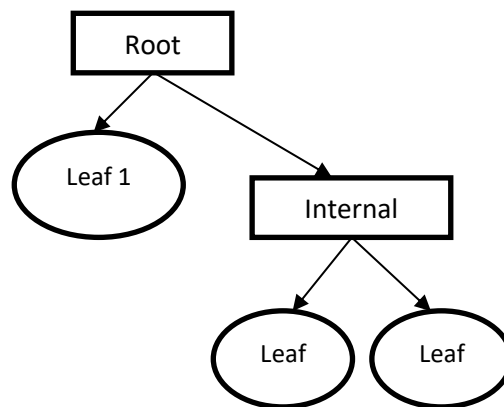
1. *Data Selection* : pemilihan data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam KDD dimulai.
2. *Preprocessing* : sebelum proses data mining dapat dilaksanakan, perlu dilakukan proses cleaning dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses memperkaya data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.



3. *Transformation* : yaitu proses coding pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses data mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam *database*.
4. *Data mining* : proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu.
5. *Interpretation / Evaluation* : pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.

#### **2.1.7 *Decision Tree/ Classification Tree***

Pembuatan model untuk memprediksikan kelas suatu objek berdasarkan atribut yang dimilikinya merupakan hal yang harus dilakukan dalam *data mining* dengan *predictive method*. *Decision tree* merupakan salah satu metode untuk melakukan pembuatan model tersebut dan cukup terkenal karena mudah untuk diinterpretasikan, tingkat akurasi yang baik, dan efisien dalam menangani atribut diskret ataupun numerik/kontinyu (Bening, 2014). *Decision tree* mempunyai konsep mengubah data menjadi pohon keputusan lalu pohon keputusan diubah menjadi aturan-aturan keputusan. Data yang terdapat pada pohon keputusan biasanya berbentuk tabel yang memiliki atribut dan record. Atribut menyatakan parameter yang digunakan sebagai kriteria dalam pembentukan tree.



**Gambar 2.2** *Decision Tree*

*Decision tree* terdiri dari Simpul akar, yang tidak memiliki cabang masukan dan berpengaruh paling besar pada suatu kelas tertentu. Simpul internal yang merepresentasikan *test* atau *subset* dari sebuah *Attribute* dan simpul daun yang merupakan sebuah sambungan *node* dari *tree* berupa *class* label.

### **2.1.8 Data Mining**

*Data mining* adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer untuk menganalisis dan mengekstrak pengetahuan secara otomatis, atau serangkaian proses untuk menggali nilai tambah dari kumpulan data berupa pengetahuan yang selama ini tidak diketahui. Dengan arti lain data mining adalah proses untuk penggalian pola – pola dari data yang tersembunyi, data mining menjadi penting karena mengubah data menjadi sebuah informasi. Hal ini sering digunakan dalam pemasaran, pengawasan dan penemuan ilmiah. (Sijabat, Alimancon, Vol. 5, No.3).

Definisi umum dari data *mining* itu sendiri adalah proses pencarian pola-pola yang tersembunyi (*hidden pattern*) berupa pengetahuan (*knowledge*) yang

tidak diketahui sebelumnya dari suatu sekumpulan data yang mana data tersebut dapat berada di dalam *database*, *data warehouse*, atau media penyimpanan informasi yang lain. Hal penting yang terkait di dalam data *mining* adalah:

1. Data *mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan diproses berupa data yang sangat besar.
3. Tujuan data mining adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat. (Kusrini dan Emha Taufiq, 2009)

#### **2.1.9 Metode Data Mining**

Dalam mengekstrak atau mengolah jumlah data yang besar tentunya mempunyai metode, berikut merupakan beberapa metode dalam data *mining* :

##### a. Deskripsi

Deskripsi merupakan cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data yang dimiliki.

##### b. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali *variable* target estimasi lebih ke arah numerik daripada ke arah kategori. Model yang dibangun menggunakan *record* lengkap yang menyediakan nilai *variable* target sebagai nilai prediksi.

##### c. Prediksi

Prediksi menerka sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang

d. Klasifikasi

Dalam klasifikasi terdapat target *variable* kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

e. Pengklasteran

Merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

f. Asosiasi

Asosiasi bertugas menemukan atribut yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja.

### 2.1.10 Tahapan – tahapan Data Mining

Data *mining* sering dipahami sebagai suatu proses yang memiliki tahapan tertentu yang bersifat interaktif dan juga ada umpan balik dari setiap tahapan sebelumnya. Adapun tahapan data *mining* tersebut adalah sebagai berikut :

#### 1. Data Cleaning

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak *valid* atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performansi dari teknik data

*mining* karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

## **2. Data Integration**

Integrasi data merupakan penggabungan data dari berbagai *database* ke dalam satu *database* baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu *database* tetapi juga berasal dari beberapa *database* atau *file* teks. Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

## **3. Data Selection**

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus *market basket analysis*, tidak perlu mengambil nama pelanggan, cukup dengan *id* pelanggan saja.

## **4. Data Transformation**

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data *mining*. Beberapa metode data *mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa

metode standar seperti analisis asosiasi dan *clustering* hanya bisa menerima *input* data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa *interval*. Proses ini sering disebut transformasi data.

#### **5. Data Mining**

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan atau informasi yang tersembunyi di data yang akan diteliti.

#### **6. Pattern Evaluation**

Merupakan suatu proses mengevaluasi data yang akan presentasikan, apakah data yang ditemukan telah mencapai standar yang ditentukan atau belum.

#### **7. Knowledge Presentation**

Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik data *mining* berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data *mining*, mencoba metode data *mining* lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

#### **2.1.11 Association Rules Mining**

*Association rules mining* digunakan untuk menemukan hubungan di antara data atau bagaimana suatu kelompok data mempengaruhi suatu keberadaan data yang (Yova & Fahrhan, 2008). *Association rules mining* meliputi dua tahap:

1. Mencari kombinasi yang paling sering terjadi dari suatu *itemset*.
  2. Mendefinisikan *Condition* dan *Result* (untuk *conditional association rule*).
- (Ulmer David, 2002)

Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran ketertarikan) yang didapatkan dari hasil pengelolaan data dengan perhitungan tertentu. Terdapat dua ukuran (Han & Kamber 2006). yaitu:

1. *Support* (nilai penunjang/pendukung): suatu ukuran yang menunjukkan seberapa besar tingkat kemunculan suatu *item/itemset* dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu *itemset* layak untuk dicari *confidence* pada tahapan selanjutnya.
2. *Confidence* (nilai kepastian/keyakinan): suatu ukuran yang menunjukkan hubungan antara 2 *item* secara *conditional*.

Kedua ukuran tersebut akan digunakan dalam menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh *user*. Batasan tersebut umumnya terdiri dari *minimum support* dan *minimum confidence*, yang digunakan pada proses pencarian *association rules*. Menurut (Hermawati, 2009), tujuan dari *association rules mining* adalah untuk menentukan semua aturan yang mempunyai  $support \geq min\_support$  dan  $confidence \geq min\_confidence$ . Sebuah *association rule* dengan *confidence* sama atau lebih besar dari *minimum confidence* dapat dikatakan sebagai *valid association rules* (Agrawal & Srikant, 1994).

Proses pencarian *association rules* terbagi menjadi dua tahap yaitu *analisis frequent itemset* dan pembentukan *association rules* (Han & Kamber 2006).

## 1. Analisis *Frequent Item*

Tahapan ini mencari kombinasi *item* yang memenuhi syarat *minimum* dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus 2.1 atau 2.2, yaitu:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total transaksi}} \dots\dots\dots (2.1)$$

Sedangkan nilai dari *support* 2 item diperoleh dari rumus berikut :

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

$$\text{Support (A, B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Total transaksi}} \dots\dots\dots (2.2)$$

## 2. Pembentukan *Asosiation Rules*

Setelah semua pola *frekuensi* tinggi ditemukan, kemudian dicari aturan asosiasi yang memenuhi syarat *minimum* untuk *confidence* dengan menghitung *confidence* aturan *asositif*  $A \rightarrow B$ . Nilai *confidence* dari aturan  $A \rightarrow B$  diperoleh dari rumus 2.3,

$$\text{Confidence (A} \rightarrow \text{B)} = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}} \times 100\% \dots\dots(2.3)$$

### 2.1.12 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. Algoritma *FP-Growth* merupakan pengembangan dari algoritma *Apriori*, sehingga kekurangan dari algoritma *Apriori* diperbaiki oleh algoritma *FP-Growth*.

*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*. *FP-tree* dibangun dengan memetakan setiap data transaksi ke dalam setiap lintasan tertentu dalam *FP-tree*. Karena dalam setiap transaksi yang dipetakan, mungkin ada transaksi yang memiliki item yang sama, maka lintasannya memungkinkan untuk saling menimpa. Semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data *FP-tree* semakin efektif.

### **2.1.13 Pencarian Association Rules Dengan Algoritma FP-Growth**

Proses pencarian *association rules* terbagi menjadi dua tahap yaitu analisis pola frekuensi tinggi dan pembentukan aturan asosiasi (Han & Kamber 2006).

1. Analisis *frequent itemset*, atau analisis pola frekuensi tinggi untuk mencari kombinasi item yang memenuhi syarat *minimum* dari nilai *support itemset*, dimana nilai *support* menunjukkan perhitungan frekuensi kemunculan pada suatu data. Proses pencarian *frequent itemset* pada sistem ini menggunakan *algoritma FP-Growth* yang melalui tiga tahapan.

- a. Pencarian *frequent item*

Melakukan pencarian item yang paling sering muncul dengan cara menghitung nilai *support* pada setiap *item*, apabila terdapat *item* yang tidak *frequent* atau kemunculan sedikit maka *item* akan dibuang. Nilai *support* dapat diperoleh dengan rumus (2.1)

b. Pembangunan *FP-Tree*

Pembangunan *FP-tree* diawali dengan pembangunan *tree* pada setiap *itemset*. Pembangunan *tree* diawali dengan *prefix* atau awalan yang sama dari setiap *itemset*. Apabila terdapat *itemset* yang memiliki *prefix* berbeda, maka *itemset* berikutnya dibangun pada lintasan berbeda (Tan, et al. 2005).

2. Pembentukan *association rules* atau aturan assosiasi untuk mencari aturan assositif  $A \rightarrow B$  yang memenuhi syarat *minimum* nilai *confidence*. Pencarian nilai *confidence* dapat dihitung dengan rumus (2.3). Dimana A adalah *antecedent* dan B adalah *consequent*. Untuk *antecedent* dapat terdiri lebih dari satu unsur, akan tetapi *consequent* hanya terdiri dari satu unsur. Ini digunakan untuk mengetahui keterhubungan antar item dalam suatu *itemset*.

#### 2.1.14 *Knowledge*

Secara konseptual, *Knowledge Management* merupakan kegiatan organisasi dalam mengelola pengetahuan sebagai aset, diperlukan upaya penyaluran pengetahuan yang tepat kepada orang yang tepat dan dalam waktu yang cepat, hingga mereka bisa saling berinteraksi, berbagi pengetahuan dan mengaplikasikannya dalam pekerjaan sehari-hari demi peningkatan kinerja organisasi. Pengertian *Knowledge Management* (KM) (Hendrik, 2003) adalah merencanakan, mengumpulkan dan mengorganisir, memimpin dan mengendalikan data dan informasi yang telah digabung dengan berbagai bentuk pemikiran dan analisis dari macam macam sumber yang kompeten.

## 2.2 Tinjauan Penelitian Terdahulu

Berikut terdapat beberapa contoh penelitian yang telah dilakukan dengan metode yang sama yaitu metode *Fp-Growth* sebagai data pembanding, yaitu :

**1. Penerapan Data Mining menggunakan Algoritma *Fp-Tree* dan *Fp-Growth* pada Data Transaksi Penjualan Obat** (Yuyun Dwi Lestari)

Penelitian ini menghasilkan: Dengan menggunakan *Fp-Tree algoritma* dan *fp-growth* maka didapatkan informasi untuk kombinasi *itemset* dalam transaksi penjualan obat sehingga pengembangan transaksi penjualan obat dapat dilakukan untuk selanjutnya, serta dapat diketahui jenis obat mana yang paling banyak terjual pada Apotek tersebut, sehingga dalam melakukan pembelian obat – obatan dapat dilakukan dengan terkontrol dengan baik.

**2. Rancang Bangun Aplikasi Rekomendasi Judul Buku Menggunakan Algoritma *FP-GROWTH* (Studi Kasus : CV. Cita Intrans Selaras)**

(Moh. Agung Hadi Wijaya). Penelitian ini menghasilkan : Algoritma *FP-Growth* dapat digunakan di dalam kasus perekomendasi judul buku dan dapat menghasilkan keluaran berupa rekomendasi judul buku yang sesuai. Berdasarkan penerapan Algoritma *FP-Growth* kedalam sistem, maka sistem ini hanya menampilkan dua kombinasi *itemset* yang *frequent* dengan menggunakan 774 data, dan terbentuk aturan asosiasi dengan nilai *support* terbesar 30% dan *confidence* 50%.

**3. Penerapan Data Mining untuk Analisis Pola Pembelian Konsumen dengan Algoritma *Fp-Growth* pada Data Transaksi Penjualan *Spare Part Motor*.** (Alfannisa Annurullah Fajrin, Algifanri Maulana). Hasil

penjualan dari *spare part* motor sport yang paling banyak terjual di TB-Damar CV.TJAJAJA BARU Padang bisa diketahui dengan menggunakan algoritma *FP-Growth*. *Spare part* yang memenuhi minimum *support* dan minimum *confidence* serta yang banyak terjual adalah *screw valve adjusting, oil seal, battery assy, axle, gasket cylinder, dan cable clutch*.

Serta Menerapkan metode Data Mining dengan algoritma *FP-Growth* kedalam aplikasi untuk analisis pola pembelian konsumen sangat bermanfaat bagi perusahaan tersebut, karena TB-Damar akan mengetahui *spare part* mana yang banyak dibeli dan membantu dalam pemesanan *spare part* pada kantor pusat.

- 4. Asosiasi Data Mining menggunakan Algoritma *Fp-Growth* untuk *Market Basket Analysis*** (Fathimah Fatihalul, Atje Setiawan, Rudi Rosadi). Dengan menggunakan algoritma *Fp-Growth* didapatkan aturan *rules* yang merupakan kumpulan *frequent itemset* dengan nilai *confidence* yang tinggi, dan dengan didapatkannya *rules* ini maka perusahaan dapat membuat strategi dalam meningkatkan penjualan.
- 5. Penentuan Pola Yang Sering Muncul Untuk Penjualan Pupuk Menggunakan Algoritma *Fp-Growth*** (Chandra Eri Firman). Hasil perhitungan nilai *support* dan nilai *confidence* dari *rules* yang dihasilkan maka diambil dua *rules* dengan nilai tertinggi yaitu :
  1. Jika dilakukan penjualan pada produk NPK MAHKOTA (KG) dan TSP (KG) maka dilakukan penjualan pada produk UREA PUTIH (KG) dengan tingkat *confidence* 92,3% dan *support* 11% dari banyaknya transaksi yang dibeli secara bersamaan.

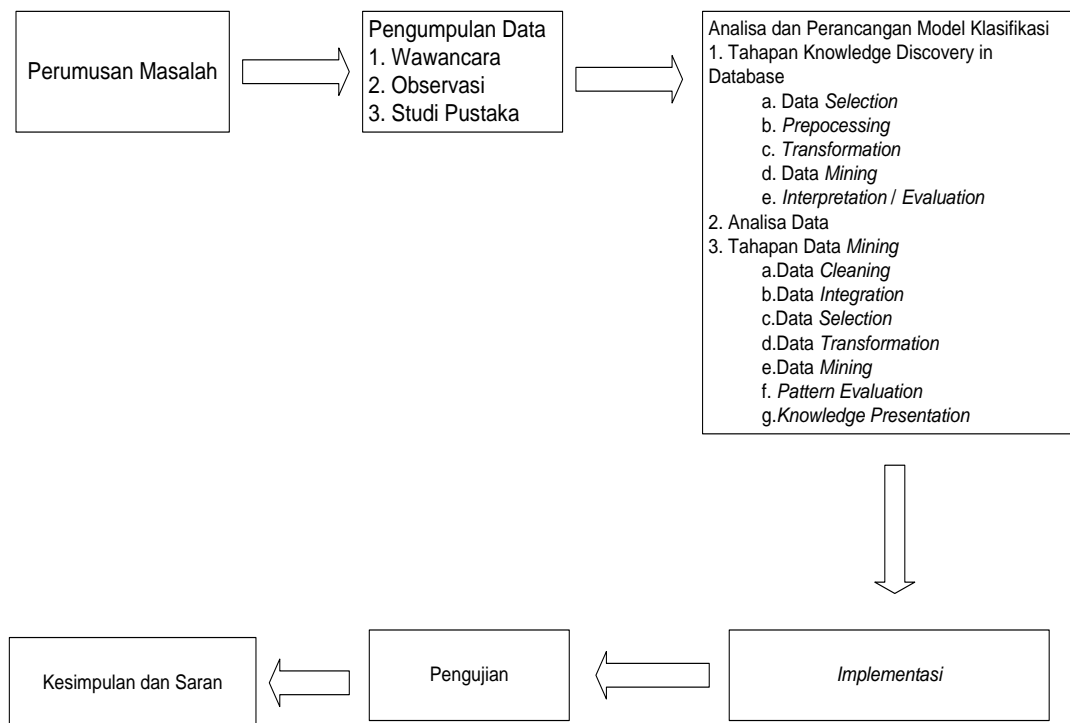
2. Jika dilakukan penjualan pada produk TSP (KG) dan DOLOMIT BR (SAK) hen dilakukan penjualan pada produk UREA PUTIH (KG) dengan tingkat *confidence* 88,9% dan *support* 8% dari banyaknya transaksi yang dibeli secara bersamaan.

# BAB III

## METODOLOGI PENELITIAN

### 3.1 Tahapan – tahapan Penelitian

Untuk mencapai keberhasilan suatu penelitian, maka harus dibuat tahapan dalam penelitian tersebut. Berikut merupakan tahapan – tahapan penelitian yang dilakukan pada Universitas Islam Kuantan Singingi:



**Gambar 3.1 : Alur Penelitian**

### 3.2 Waktu dan Lokasi Penelitian

Penelitian ini dilakukan di Universitas Islam Kuantan Singingi, Jl. Gatot Subroto KM 7, Kebun Nenas, Teluk Kuantan, Sungai Jering, Kabupaten Kuantan Singingi.

### **3.3 Data yang Digunakan**

Data yang digunakan dalam penelitian ini adalah Data Penerimaan Mahasiswa Baru pada Universitas Islam Kuantan Singingi periode 2018 – 2020.

### **3.4 Teknik Pengumpulan Data**

Untuk mendapatkan data yang terpercaya, maka dalam penelitian tersebut dibutuhkan metode demi tercapainya tujuan dalam penelitian. Ada beberapa metode yang digunakan dalam mengumpulkan data yaitu :

#### **1. Wawancara**

Wawancara merupakan salah satu cara untuk mendapatkan data, contohnya dalam hal ini adalah melakukan wawancara dengan Kepala Biro Umum Universitas Islam Kuantan Singingi dan para pegawai Biro Perencanaan Akademik, Kemahasiswaan, Alumni dan Sistem Informasi untuk mengetahui bagaimana cara mereka dalam melakukan pengolahan data mahasiswa yang mendaftar, serta dapat mengetahui berapa jumlah mahasiswa yang mendaftar ke Universitas Islam Kuantan Singingi.

#### **2. Observasi**

Observasi merupakan pengamatan langsung dilapangan untuk memperoleh informasi tentang penerimaan mahasiswa baru, serta mengamati apa saja permasalahan yang ditemukan saat melakukan promosi untuk Penerimaan Mahasiswa Baru.

### 3. Studi Pustaka

Studi Pustaka dilakukan untuk mengetahui metode apa yang paling terbaik digunakan dalam menyelesaikan masalah yang sedang diteliti, serta mendapatkan dasar-dasar referensi yang kuat dalam menerapkan suatu metode yang akan digunakan dalam skripsi ini, yaitu dengan mempelajari buku-buku, artikel-artikel dan jurnal-jurnal maupun data dari internet yang berhubungan dengan penelitian untuk dijadikan acuan dalam penelitian ini.

### **3.5 Metode/ Pemodelan yang Digunakan**

Dalam penelitian ini penulis menjelaskan metode serta pemodelan yang digunakan dalam penelitian ini. Adapun prosenya dapat di jelaskan sebagai berikut :

#### 1. Perumusan Masalah

Dalam hal ini penulis merumuskan permasalahan permasalahan yang terjadi dalam penerimaan mahasiswa baru, seperti menurunnya jumlah mahasiswa yang mendaftar ke Universitas Islam Kuantan Singingi, setelah itu penulis melakukan pengumpulan data untuk menunjang penelitian yang akan dilaksanakan.

#### 2. Pengumpulan Data

Dalam melakukan pengumpulan data ini, penulis mengambil data mahasiswa yang telah mendaftar ke Universitas Islam Kuantan Singingi Periode 2018 – 2020 serta melakukan wawancara mengenai proses penerimaan mahasiswa baru dengan pihak kampus bagian kemahasiswaan.



3. Setelah mendapatkan data penulis melakukan proses *Discovery in Database* yaitu metode dalam menggali informasi yang terdapat dalam data mahasiswa, membuang duplikasi data yang dapat membuat kesalahan dalam penelitian serta menentukan metode yang terbaik untuk tercapainya hasil atau informasi yang mudah dimengerti bagi orang yang membutuhkannya.
4. Setelah data tersebut di analisa dan melewati proses *Discovery in Databases* maka tahap selanjutnya adalah proses data *mining*, Adapun tahapannya sebagai berikut :
  - a. Proses memilih atribut – atribut mana saja yang akan digunakan dalam proses selanjutnya, karena dari data yang didapatkan banyak atribut – atribut yang tidak menjadi pilihan dalam proses data mining, maka harus dihilangkan terlebih dahulu.
  - b. Melakukan proses pencarian item – item mana saja yang sering muncul dalam database yang di dapatkan, seperti siswa dari sekolah mana yang paling banyak masuk Universitas Islam Kuantan Singingi.
5. Setelah data itemset didapatkan, maka tahap selanjutnya adalah proses kombinasi item yang memenuhi syarat minimum dari nilai *support* dalam *database* dengan rumus 2.1, yaitu jumlah transaksi yang mengandung A dibagi total transaksi. Sedangkan untuk mencari nilai *support* 2 item menggunakan rumus 2.2, yaitu jumlah transaksi yang mengandung A dan B dibagi total transaksi.

6. Setelah kedua data didapatkan menggunakan rumus 2.1 dan 2.2 maka proses selanjutnya menggunakan rumus 2.3, yaitu jumlah transaksi mengandung A dan B dibagi transaksi yang mengandung A kemudian di kali seratus persen (100%).
7. Setelah itu barulah dibuat pohon keputusan atau *Fp - Tree* untuk memudahkan dalam melihat item – item yang sering muncul, serta untuk memudahkan dalam melihat item yang sama maka lintasannya memungkinkan untuk saling menimpa, semakin banyak data transaksi yang memiliki item yang sama, maka proses pemampatan dengan struktur data *Fp - Tree* semakin efektif.
8. Setelah pembangunan proses *Fp - Tree* selesai dikerjakan, maka perlu dilakukan proses pengujian data, dalam hal ini adalah RapidMiner, karena RapidMiner dapat menemukan algoritma yang akan berguna untuk klasifikasi, prediksi dan teknik lainnya dalam data *mining*.

## **BAB IV**

### **ANALISA DAN PERANCANGAN APLIKASI**

#### **4.1 Analisa Sistem**

Analisis Sistem atau *System Analysis* adalah suatu teknik atau metode pemecahan masalah dengan cara menguraikan sistem ke dalam komponen-komponen pembentuknya untuk mengetahui bagaimana komponen-komponen tersebut bekerja dan saling berinteraksi satu sama lain untuk mencapai tujuan sistem. Analisa sistem perlu dilakukan sebagai dasar pembangunan sistem yang baru, sistem yang sedang berjalan menjadi dasar pembangunan sistem yang diusulkan pada pembangunan aplikasi yang dapat mengoptimalkan data.

Hasil analisa dari permasalahan yang ada adalah tidak adanya kestabilan dalam tingkat pendaftaran mahasiswa di Universitas Islam Kuantan Singingi maka dengan ini peneliti menganalisa sistem yang akan dibuat adalah sebuah sistem yang dapat menghasilkan analisa sebuah data menggunakan algoritma *Fp-Growth* untuk data setiap mahasiswa yang sudah melakukan pendaftaran berdasarkan prodi, asal sekolah, jenis sekolah dan jenis kelamin mahasiswa yang mana hasilnya akan jadi pedoman admin dalam menyusun rencana promosi untuk meningkatkan pendaftaran.

#### **4.2 Perancangan Sistem**

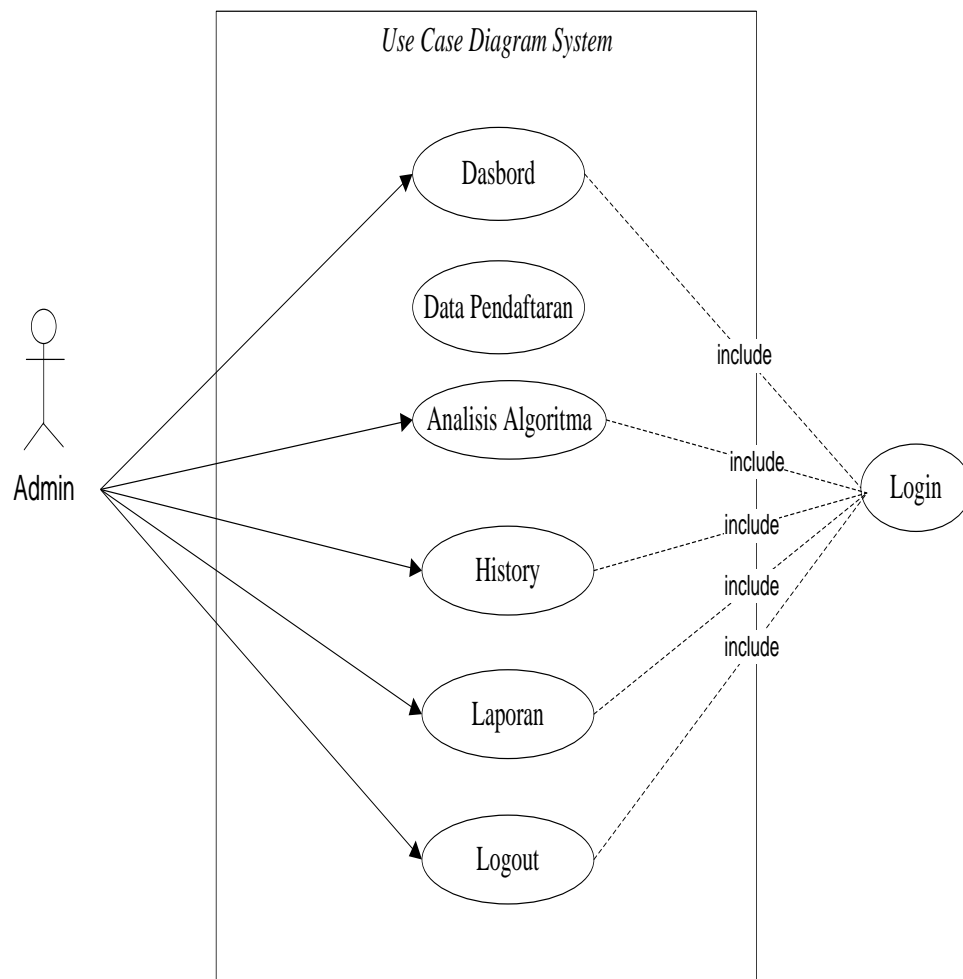
Setelah tahap analisa sistem selesai dilakukan, analisa ini akan memberikan suatu pandangan terhadap pembangunan sistem yang baru. Berdasarkan permasalahan yang ditemukan yang ada di UNIKS , maka diperlukan perancangan aplikasi untuk menunjang proses promosi untuk meningkat kan

minat memilih universitas kuantan singingi dengan sistem yang dapat mempermudah universitas untuk memilih tempat dan kebijakan yang akan diambil untuk melakukan promosi peningkat minat memilih Universitas ini.

Adapun rancangan aplikasi yang akan peneliti buat berupa desain global menjelaskan tentang *Use Case Diagram*, *Activity Diagram*, *Sequence Diagram*, desain *Output* dan desain *Input* sebagai berikut.

#### **4.2.1. Use Case Diagram**

*Use case* diagram ini menggambarkan bagaimana *Actor* yaitu Pengguna berinteraksi dengan sistem. Untuk sistem ini *actornya* hanya ada admin saja. Dikarenakan sistem dibuat untuk membuat sebuah data yang berbentuk laporan strategi untuk meningkatkan promosi pendaftaran ke UNIKS ini berupa data yang mesti dijaga kerahasiannya dan aksesnya pun terbatas untuk admin saja. Hanya admin yang bisa memiliki akses penuh pada sistem ini. Untuk lebih jelas dapat dilihat pada gambar *use case* diagram perancangan **Implementasi Data Mining dengan Algoritma FP-Growth untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi)** sebagai berikut.



**Gambar 4. 1 Use Case Diagram**

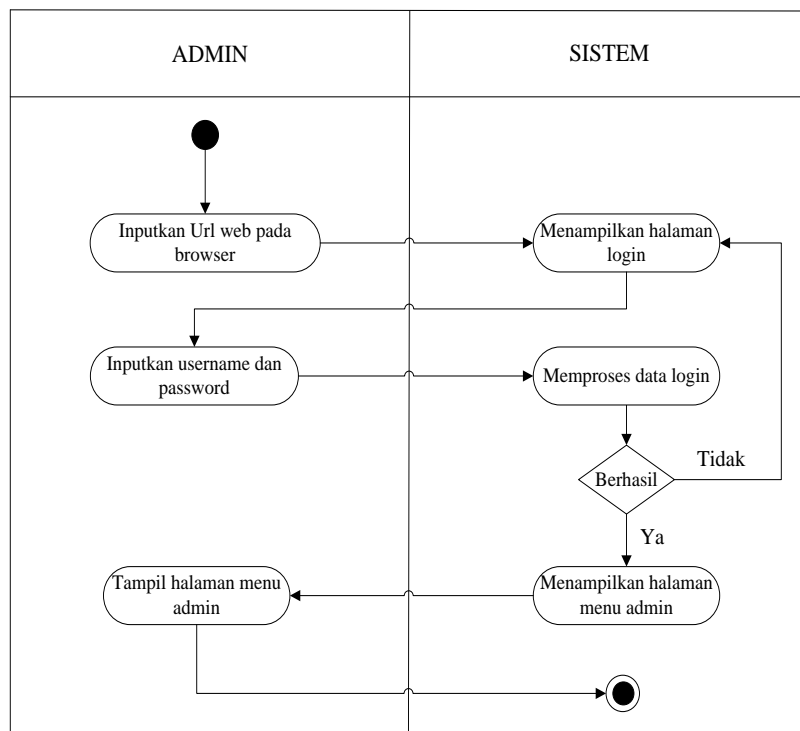
#### **4.2.2. Activity Diagram**

Menurut Sukamto dan Shalahuddin (2014:161) diagram aktivitas atau *activity diagram* adalah menggambarkan aliran kerja atau aktifitas dari sebuah sistem atau proses bisnis atau menu yang ada pada perangkat lunak. Diagram aktifitas menggambarkan aktifitas sistem bukan apa yang dilakukan oleh aktor.

Activity diagram yang merupakan alat aktifitas sistem yang sedang dirancang, bagaimana masing-masing alir berawal dan alir berakhir. Berikut akan digambarkan *activity diagram* sistem.

#### 4.2.2.1 Activity Diagram Login Admin

*Activity diagram login admin* menggambarkan bagaimana proses admin login ke dalam sistem ini menggunakan web guna meningkatkan promosi minat mendaftar ke UNIKS sehingga mudah untuk dipahami. Untuk lebih jelasnya dapat dilihat gambaran *activity diagram login admin* sebagai berikut.

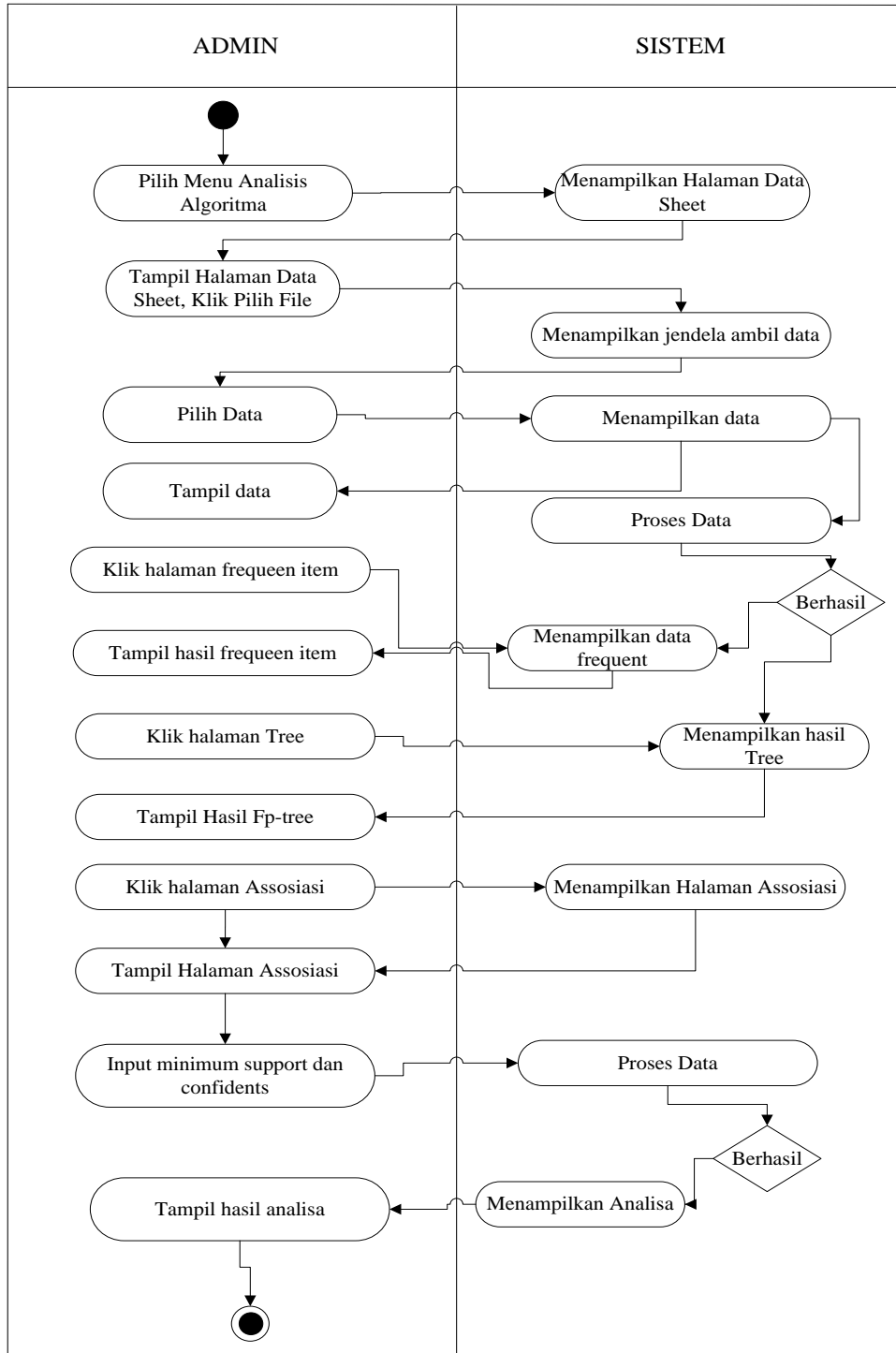


**Gambar 4.2 Activity Diagram Login Admin**

#### 4.2.2.2 Activity Diagram Analisa Algoritma

*Activity diagram* analisa algoritma fp-growth dimulai dari mengklik menu analisa algoritma. Setelah itu data yang sudah dimasukkan di data pendaftaran sebelumnya akan tampil kembali di halaman ini, setelah itu sistem akan memprosesnya lalu mencari dan menampilkan hasil dari frequent itemnya di halaman frequent item maka selanjutnya untuk menampilkan hasil analisa algoritma dalam bentuk asosiasi dengan minimum size yang telah ditentukan

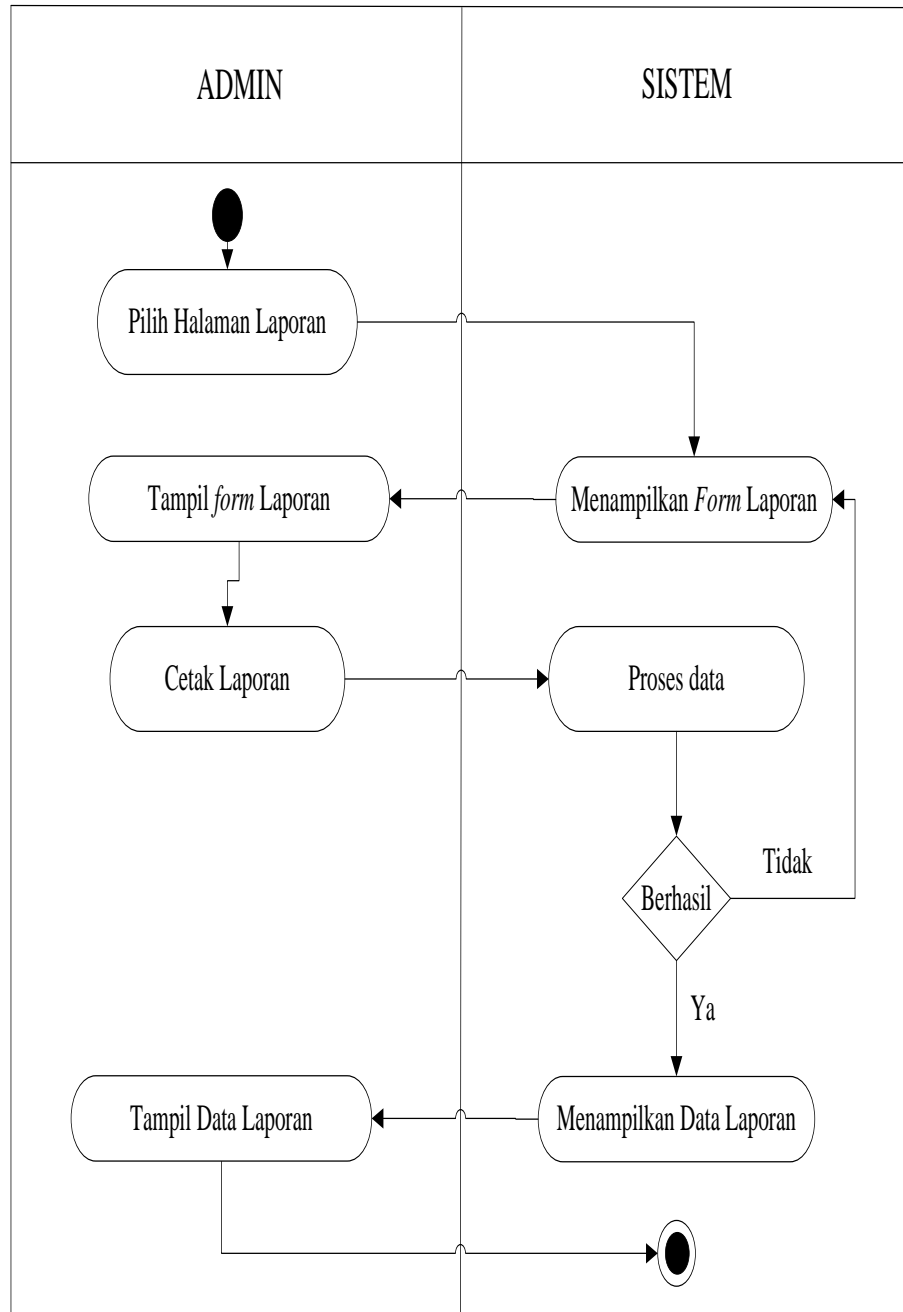
maka sistem akan menampilkannya di halaman *asosiasi*. Berikut adalah gambaran *activity diagram admin* mengolah data yang ada pada penelitian ini.



**Gambar 4.3 Activity Diagram Analisa Algoritma**

#### 4.2.2.3 Activity Diagram Laporan

Gambar di bawah ini merupakan gambaran aktivitas yang dilakukan Admin ketika mengambil laporan.

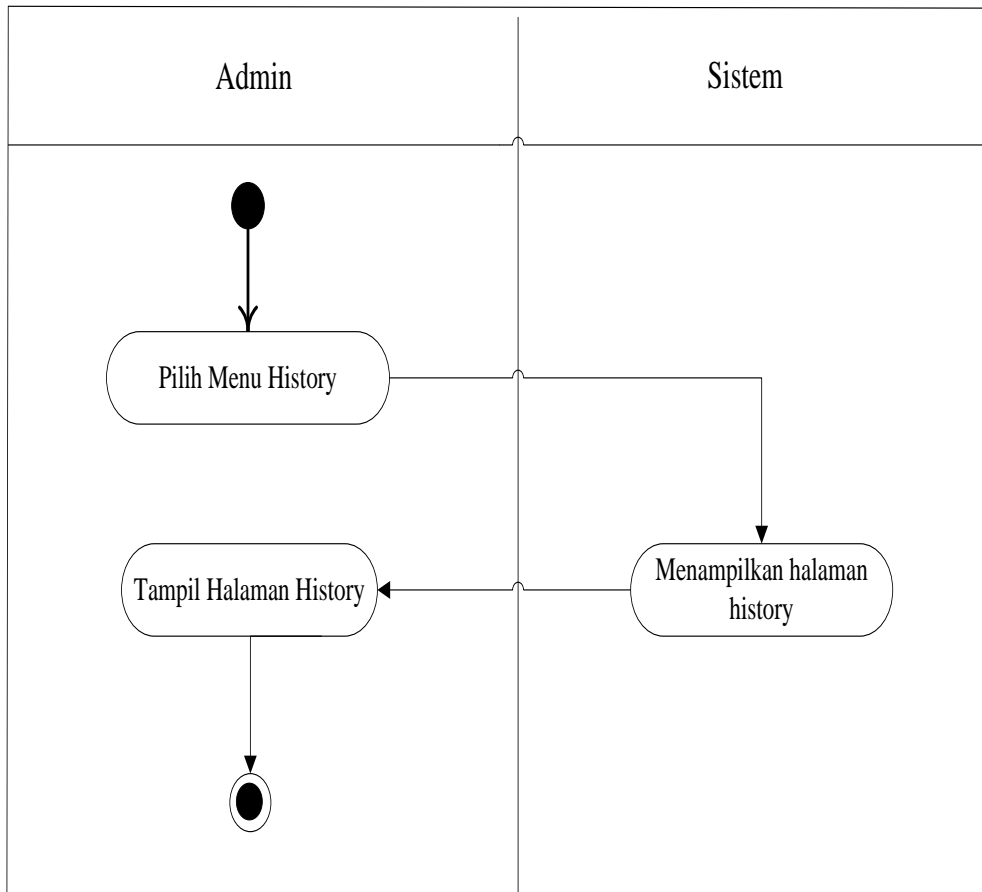


**Gambar 4.4 Activity Diagram Laporan**



#### 4.2.2.4 Activity Diagram Melihat History

Gambar di bawah ini merupakan gambaran aktivitas yang dilakukan *Admin* ketika melihat *history*.



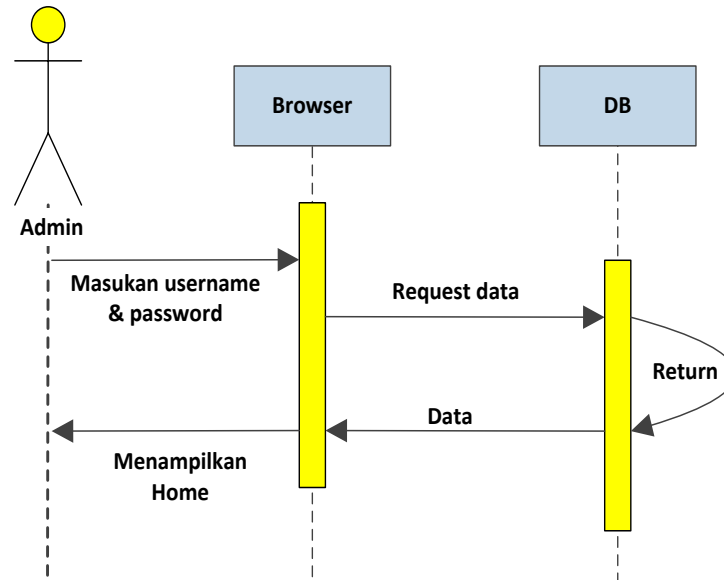
**Gambar 4.5 Activity Diagram Melihat History**

#### 4.2.3. Sequence Diagram

*Diagram - diagram* ini mendeskripsikan bagaimana *entitas* dalam sistem berinteraksi, termasuk pesan yang digunakan saat interaksi.

### 1. *Sequence Diagram Login*

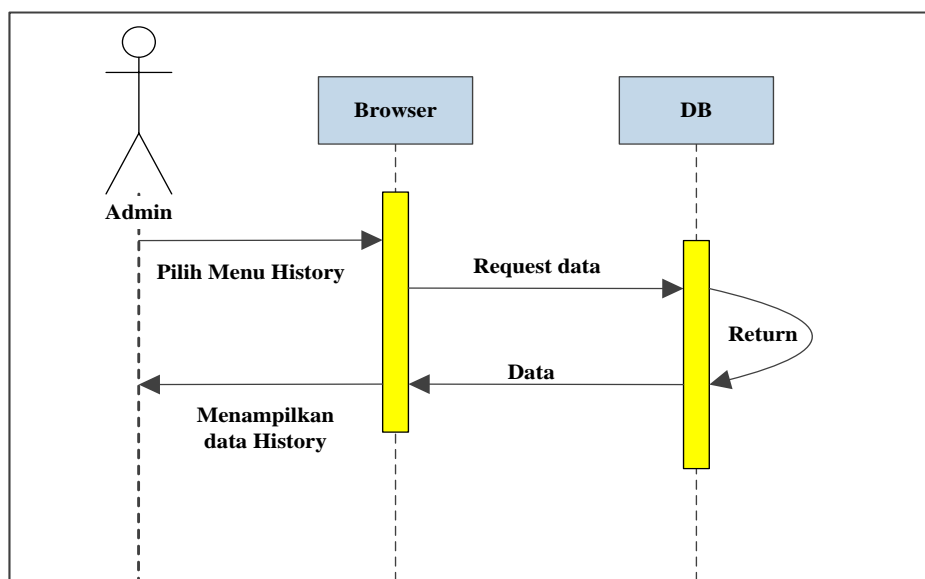
Gambar di bawah ini merupakan *sequence diagram* ketika admin melakukan *login*.



Gambar 4.6 *Sequence Diagram Login*

### 2. *Sequence Diagram Melihat History*

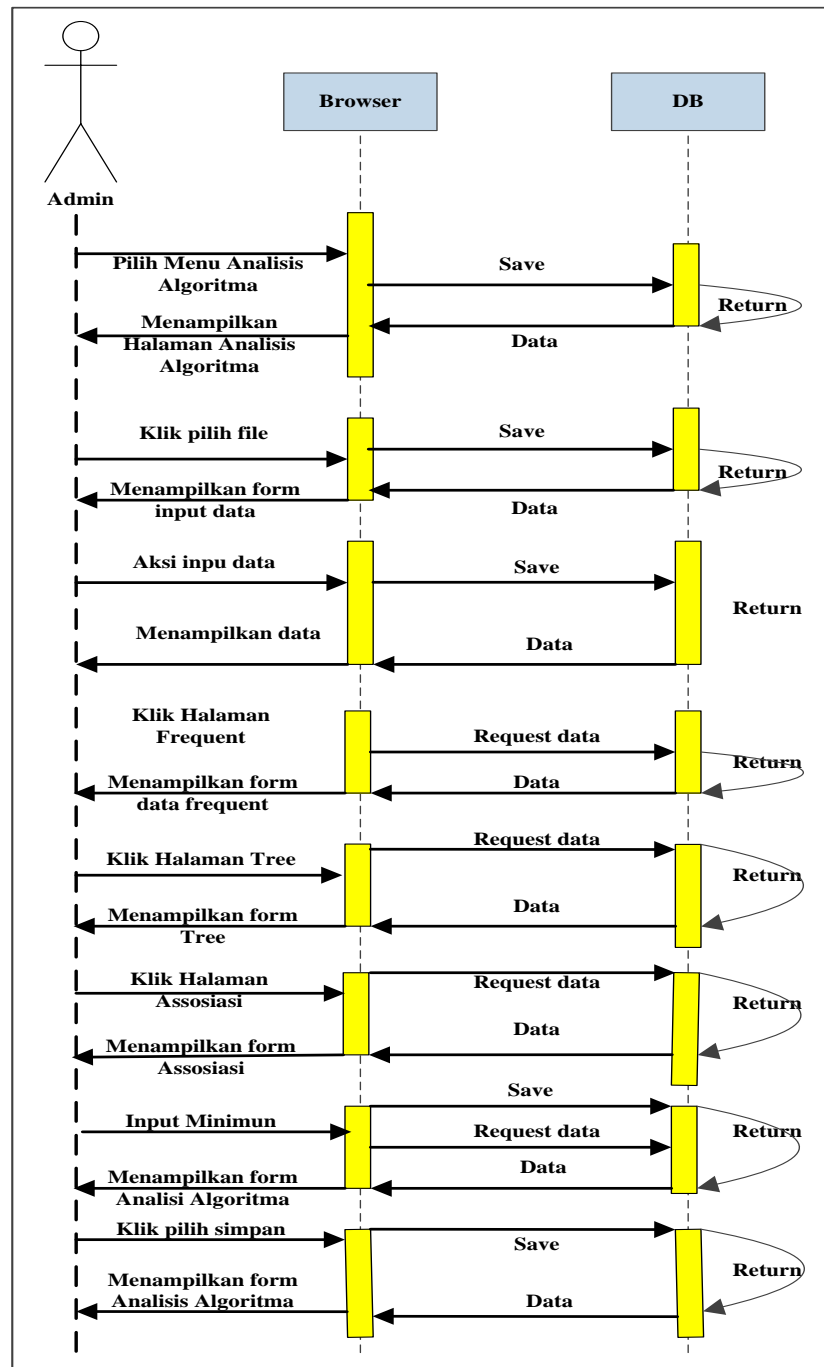
Gambar di bawah ini merupakan *sequence diagram* ketika admin melihat history.



Gambar 4.7 *Sequence Diagram Melihat History*

### 3. Sequence Diagram Admin Di Menu Analisis Algoritma

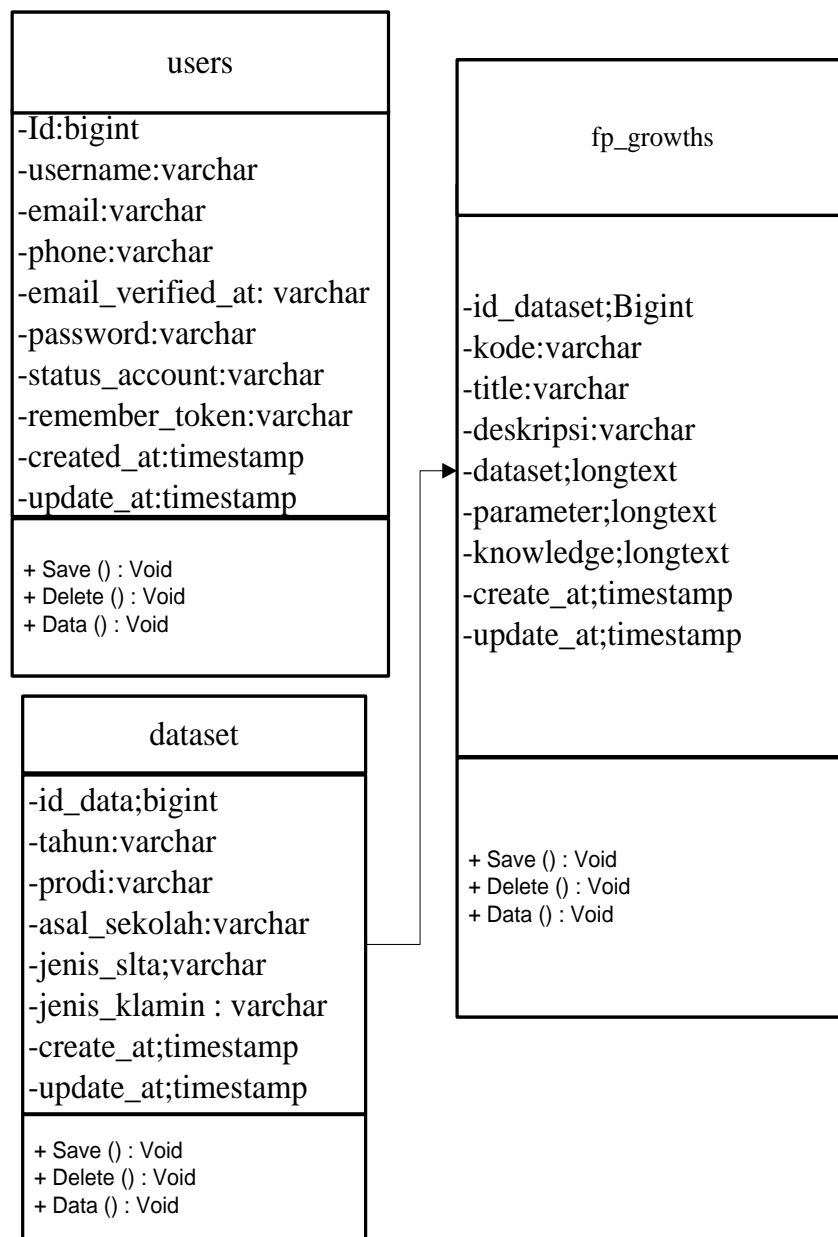
Gambar di bawah ini merupakan *sequence diagram* ketika *admin* melihat data hasil dari *fp-growth*, memasukkan data dan melihat hasil data strategi yang akan digunakan.



Gambar 4.8 Sequence Diagram Analisis Algoritma

### 4.3 Class Diagram

*Class diagram* ini memberikan gambaran tentang atribut-atribut yang digunakan pada *database* Sistem Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi) dan juga untuk mengetahui *relasi* antar *tabel* yang ada. Untuk lebih jelasnya dapat dilihat pada *class diagram* sistem ini sebagai berikut.



**Gambar 4.9 Class Diagram**

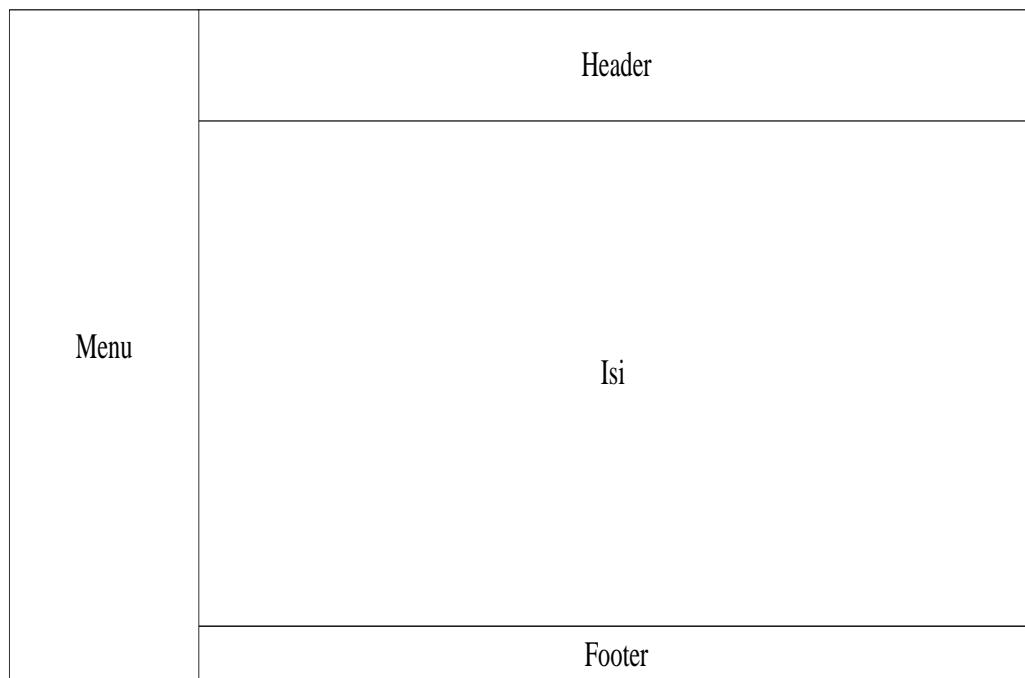
#### 4.4. Desain Terinci

Desain terinci adalah gambaran dari keseluruhan mengenai Sistem Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi). Desain terinci ini menjelaskan tentang *desain output*, *desain input*, struktur tabel, dan relasi tabel.

##### 4.4.1 Desain Output

Desain *output* merupakan suatu bentuk rancangan tampilan keluaran yang dihasilkan oleh suatu sistem. desain *output* atau keluaran merupakan hal yang tidak dapat diabaikan, dikarenakan laporan atau keluaran yang dihasilkan harus memudahkan bagi setiap unsur manusia yang membutuhkannya.

Adapun *output* perancangan Sistem Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi) berikut ini:



**Gambar 4.10 Desain Halaman Utama**

## 1. *Desain Output Data Sheet*

*Desain output data sheet* ini berfungsi untuk menghasilkan data mahasiswa pendaftar setiap tahunnya yang akan menjadi acuan untuk mengolah data selanjutnya. Untuk lebih jelasnya dapat dilihat pada gambar sebagai berikut.

Program Studi	Jenis Kelamin	Alamat	Tahun Ijazah	Asal Sekolah	Pekerjaan
Xxx/999	Xxx/999	Xxx/999	Xxx/999	Xxx/999	Xxx/999
↓	↓	↓	↓	↓	↓
Xxx/999	Xxx/999	Xxx/999	Xxx/999	Xxx/999	Xxx/999

**Gambar 4.11 Desain Output Data Sheet**

## 2. *Desain Output Frequeen Item*

*Desain output frequeen item* ini berisi data sheet yang sudah di ubah ke frequeen item untuk selanjutnya di proses menggunakan rumus *Fp-Growth*. Untuk lebih jelasnya dapat dilihat pada gambar sebagai berikut.

NO	Keyword	Frequeen
123	Xxx/999	Xxx/999
↓	↓	↓
123	Xxx/999	Xxx/999

**Gambar 4.12 Desain Output Frequeen item**

### 3. *Desain Output Asosiasi*

*Desain output Asosiasi* ini berisi data yang sudah menjadi hasil dari pengelolaan rumus *Fp-Growth* yang mana hasil dari data ini bisa menjadi acuan bagi pihak universitas memikirkan strategi promosi pendidikan untuk meningkatkan minat pendaftaran ke Universitas Islam Kuantan Singingi. Untuk lebih jelasnya dapat dilihat pada gambar sebagai berikut.

Data sheet				
No	accident	consequent	confident	support
123	Xxx/999	Xxx/999	123	123
↓	↓	↓	↓	↓
123	Xxx/999	Xxx/999	123	123

**Gambar 4.13 Desain Output Asosiasi**

### 4. *Desain Output History*

*Desain output History* ini berisi catatan data setiap melakukan perumusan di *Fp-Growth*. Untuk lebih jelasnya dapat dilihat pada gambar sebagai berikut.

History				
No	Judul	Deskripsi	Tanggal	Actions
123	Xxx/999	Xxx/999	date	
↓	↓	↓	↓	
123	Xxx/999	Xxx/999	date	

**Gambar 4.14 Desain Output History**

## 5. *Desain Output Laporan*

*Desain output* Laporan. Untuk lebih jelasnya dapat dilihat pada gambar sebagai berikut.

Laporan			
No	Judul	Tanggal	Print
123	Xxx/999	date	
↓	↓	↓	
123	Xxx/999	date	

**Gambar 4.15** *Desain Output Laporan*

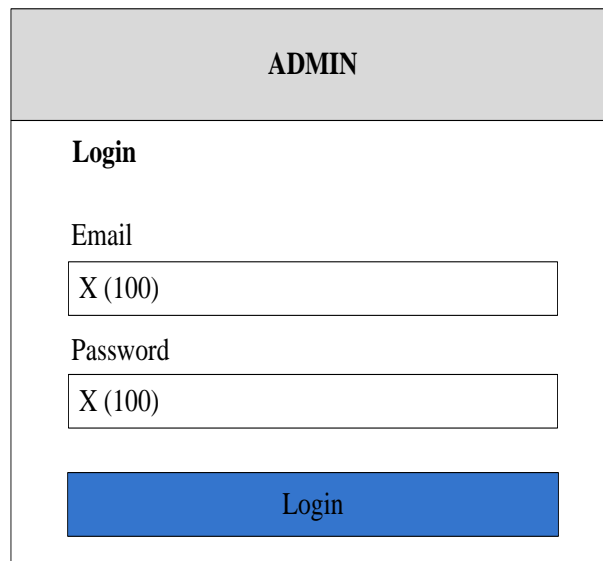
### 4.4.2 *Desain Input*

Masukan (*input*) merupakan awal dimulainya proses informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi. Sampah yang masuk sampah yang keluar (*garbage in garbage out*). Oleh karena itu desain input harus benar-benar menerima *input* bukan sampah. Desain *input* terinci dimulai dari desain dokumen dasar sebagai penangkap *input* yang pertama kali. Jika dokumen dasar tidak didesain dengan baik, kemungkinan *input* yang tercatat dapat salah bahkan kurang atau berlebihan. Untuk lebih jelasnya berikut beberapa desain *input* pada Sistem Data Mining dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi).



## 1. Desain *Login Admin*

Berikut adalah desain *input login* yang ada pada sistem ini :



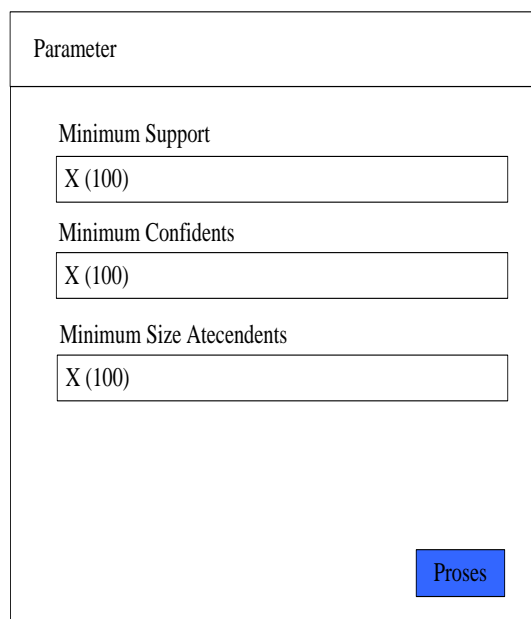
ADMIN
<b>Login</b>
Email X (100)
Password X (100)
<b>Login</b>

**Gambar 4.16 Desain *Login Admin***

## 2. Desain *Input Data Asosiasi*

Desain *input* data asosiasi ini digunakan oleh admin untuk mengatur minimum support, accidents dan size antecedents data sheet yang ingin di proses.

Berikut adalah desain input data asosiasi yang ada pada aplikasi.



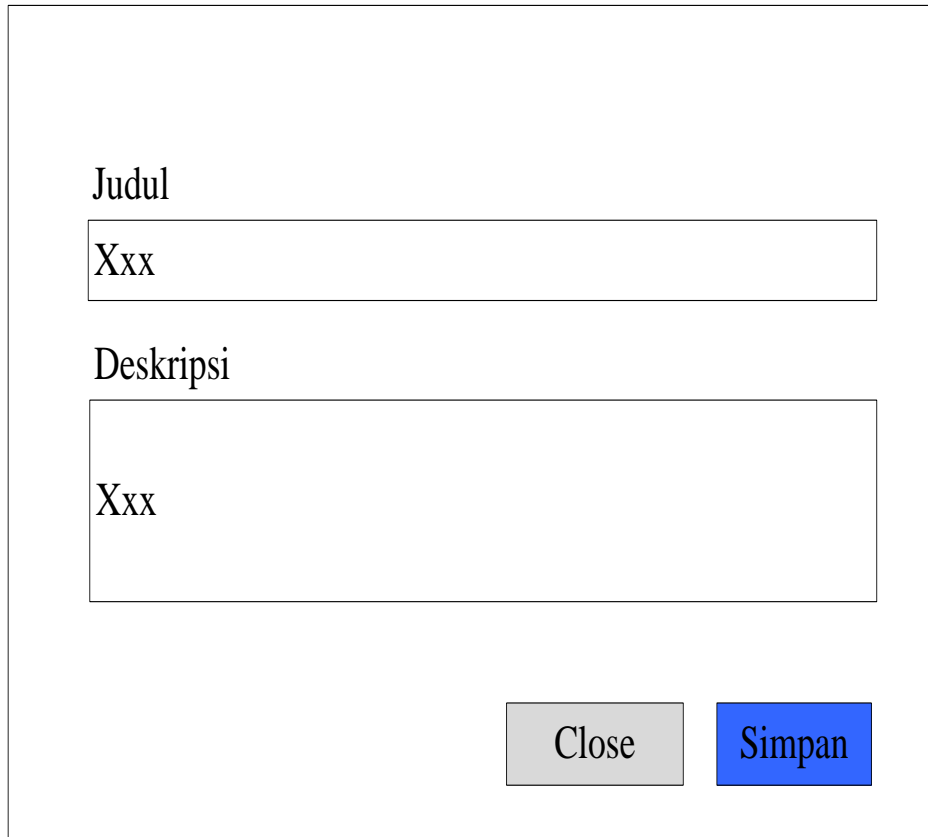
Parameter
Minimum Support X (100)
Minimum Confidants X (100)
Minimum Size Atecedents X (100)
<b>Proses</b>

**Gambar 4.17 Desain *Input Data Asosiasi***

### 3. Desain *Input* Menyimpan Data

Berikut ini adalah desain *input* menyimpan data yang sudah di asosiasi.

Data ini berguna sebagai data laporan nantinya.



The image shows a user interface for saving data. It features two text input fields. The first field is labeled 'Judul' and contains the placeholder text 'Xxx'. The second field is labeled 'Deskripsi' and also contains the placeholder text 'Xxx'. Below the input fields, there are two buttons: a grey button labeled 'Close' and a blue button labeled 'Simpan'.

**Gambar 4.18 Desain *Input* Menyimpan Data**

### 4.5. Struktur Tabel

Berikut Struktur tabel yang digunakan dalam perancangan sistem ini, sehingga dapat menentukan struktur fisik *database* yang menunjukkan struktur dari elemen data yang menyatakan panjang elemen data dan jenis datanya. Struktur *file* dari *tabel* dalam database yang akan dirancang yaitu sebagai berikut :

## 1. Tabel *Users*

Nama Tabel : *users*

Jumlah Field : 9

Primary Key : id

Foreign Key : -

**Tabel 4.1 Tabel *Users***

No	Field	Tipe Data	Size	Keterangan
1	id	Bigint	20	Id User
2	username	Varchar	191	Nama User
3	email	Varchar	191	Alamat Email
4	phone	Varchar	191	No Hp
5	email_verified_at	timestamp		Waktu Email verifikasi
6	password	Varchar	100	Password User
7	status_account	Varchar	15	Status akun
8	created_at	Timestamp		Waktu Buat
9	update_at	Timestamp		Waktu Update

## 2. Tabel *fp\_growth*

Nama Tabel : *fp\_growth*

Jumlah Field : 9

Primary Key : id\_dataset

Foreign Key :kode

**Tabel 4.2 Tabel *fp\_growths***

No	Field	Tipe Data	Size	Keterangan
1	id_dataset	Bigint	20	Id

2	kode	Varchar	50	Kode
3	title	Varchar	191	Judul
4	deskripsi	Varchar	191	Deskripsi
5	dataset	longtext		
6	parameter	longtext		
7	knowledge	longtext		
8	create_at	timestamp		Waktu Buat
9	update_at	timestamp	-	Waktu Update

### 3. Tabel Dataset

Nama Tabel : dataset

Jumlah Field : 8

Primary Key : kode

Foreign Key : -

**Tabel 4.3 Tabel Dataset**

No	Field	Tipe Data	Size	Keterangan
1	id_data	varchar	20	kode
2	tahun	Varchar	20	Data path
3	prodi	Varchar	191	keterangan
4	asal_sekolah	Varchar	191	Asal Sekolah
5	jenis_slta	Varchar	191	Jenis Slta
6	jenis_kelamin	Varchar	25	Jenis Kelamin
7	Created_at	timestamp		Waktu Buat
8	Updated_at	timestamp	-	Waktu Update

## **BAB V**

### **HASIL DAN PEMBAHASAN**

Pada Bab V ini penulis akan menjelaskan tentang bagaimana proses data dapat dirubah atau diterapkan ke dalam bentuk *Fp – Tree* sehingga dapat dimengerti dan setelah itu data yang telah disusun dapat di uji menggunakan aplikasi yang telah ditentukan. Berikut merupakan alur hasil dari penelitian ini.

#### **5.1 Praproses Data**

Tahap praproses data merupakan tahap menseleksi data yang didapatkan dengan tujuan untuk mendapatkan data yang memenuhi kriteria dan membuang data yang di anggap tidak memenuhi kriteria dalam penelitian serta penetapan sample penelitian. data calon mahasiswa yang akan dianalisis adalah program studi, asal sekolah, kota sekolah dan jenis sekolah.

Tahapan selanjutnya adalah melakukan *change* data pada beberapa tipe data *atribut* dengan tujuan untuk mempermudah dalam memahami dan mengelolah data.

#### **5.2 Pemrosesan Data**

Pemrosesan data merupakan tahapan dalam pengolahan data yang dilakukan untuk menghasilkan data yang diinginkan, yaitu dengan menggunakan *association rules* dari data mahasiswa yang telah didapatkan.

#### **5.3 Data Mining**

Tahapan ini adalah tahap memilih dataset yang digunakan untuk proses data mining. Berikut ini merupakan data pandaftar mahasiswa tahun 2020 yang mewakili keseluruhan data dari Universitas Islam Kuantan Singingi yang akan

dilakukan proses pencarian *association rules*, untuk lebih jelasnya dapat dilihat pada tabel dibawah ini.

**Tabel 5.1** Data Uji Pendaftaran Mahasiswa Lama

<b>NO</b>	<b>PRODI</b>	<b>ASAL SEKOLAH</b>	<b>JENIS SLTA</b>	<b>JENIS KELAMIN</b>
1	AGROTEKNOLOGI	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI
2	AGROTEKNOLOGI	SMA N 1 SINGINGI HILIR	SMA	LAKI-LAKI
3	AGROTEKNOLOGI	SMA N 3 SIAK HULU	SMA	PEREMPUAN
4	AGROTEKNOLOGI	SMA N 1 GUNUNG TOAR	SMA	PEREMPUAN
5	AGROTEKNOLOGI	SMA N 1 CERENTI	SMA	LAKI-LAKI
6	AGROTEKNOLOGI	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI
7	AGROTEKNOLOGI	SMA N 1 KUANTAN HILIR SEBERANG	SMA	PEREMPUAN
8	AGROTEKNOLOGI	SMK MUHAMMADIYAH CERENTI	SMK	PEREMPUAN
9	AGROTEKNOLOGI	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI

10	AGROTEKNOLOGI	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI
11	AGROTEKNOLOGI	SMK N 1 KUANTAN HILIR	SMK	PEREMPUAN
12	AGROTEKNOLOGI	MAS PONPES NURUL ISLAM	MA	PEREMPUAN
13	AGROTEKNOLOGI	SMA N 1 KUANTAN HILIR	SMA	PEREMPUAN
14	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
15	AGROTEKNOLOGI	SMK N 1 BENAI	SMK	LAKI-LAKI
16	AGROTEKNOLOGI	SMA N 1 PANGEAN	SMA	LAKI-LAKI
17	AGROTEKNOLOGI	SMA N 1 GUNUNG TOAR	SMA	LAKI-LAKI
18	AGROTEKNOLOGI	SMA N 2 SENTAJO RAYA	SMA	LAKI-LAKI
19	AGROTEKNOLOGI	SMA N 1 SINGINGI	SMA	PEREMPUAN
20	AGROTEKNOLOGI	SMA N 2 SENTAJO RAYA	SMA	LAKI-LAKI
21	AGROTEKNOLOGI	SMK N 1 TALUK KUANTAN	SMK	LAKI-LAKI
22	AGROTEKNOLOGI	MAS BAHRUL ULUM	MA	LAKI-LAKI

23	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	PEREMPUAN
24	AGROTEKNOLOGI	SMK N 1 LOGAS TANAH DARAT	SMK	PEREMPUAN
25	AGROTEKNOLOGI	SMA N 1 KUANTAN HILIR	SMA	PEREMPUAN
26	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
27	AGROTEKNOLOGI	SMK N 1 KEMPAS	SMK	LAKI-LAKI
28	AGROTEKNOLOGI	SMA N 2 SINGINGI HILIR	SMA	PEREMPUAN
29	AGROTEKNOLOGI	SMK N 1 KUANTAN MUDIK	SMK	LAKI-LAKI
30	AGROTEKNOLOGI	SMK N 1 LOGAS TANAH DARAT	SMK	LAKI-LAKI
31	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
32	AGROTEKNOLOGI	SMA N 1 GUNUNG TOAR	SMA	LAKI-LAKI
33	AGROTEKNOLOGI	SMK N 1 PANGEAN	SMK	LAKI-LAKI
34	AGROTEKNOLOGI	SMA N 1 CERENTI	SMA	LAKI-LAKI
35	AGROTEKNOLOGI	SMK N 1 TELUK KUANTAN	SMK	LAKI-LAKI



36	AGROTEKNOLOGI	SMK N 1 LOGAS TANAH DARAT	SMK	LAKI-LAKI
37	AGROTEKNOLOGI	MAS BAHRUL ULUM	MA	LAKI-LAKI
38	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
39	AGROTEKNOLOGI	SMK N 1 BATANG PERANAP	SMK	LAKI-LAKI
40	AGROTEKNOLOGI	SMA N 1 INUMAN	SMA	LAKI-LAKI
41	AGROTEKNOLOGI	SMA N 2 KUANTAN MUDIK	SMA	LAKI-LAKI
42	AGROTEKNOLOGI	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
43	AGROTEKNOLOGI	SMA N 1 BENAI	SMA	PEREMPUAN
44	AGROTEKNOLOGI	SMK N 1 PERTANIAN PASIR PENYU	SMK	LAKI-LAKI
45	AGROTEKNOLOGI	SMA N 1 PANGEAN	SMA	LAKI-LAKI
46	PETERNAKAN	SMA N 1 SENTAJO RAYA	SMA	LAKI-LAKI
47	PETERNAKAN	SMA N 1 PANGEAN	SMA	LAKI-LAKI
48	PETERNAKAN	SMA N 1 SENTAJO	SMA	LAKI-LAKI

		RAYA		
49	PETERNAKAN	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI
50	PETERNAKAN	SMK 3 MODEL TELUK KUANTAN	SMK	PEREMPUAN
51	PETERNAKAN	SMA N 1 GUNUNG TOAR	SMA	PEREMPUAN
52	PETERNAKAN	SMA N 1 KUANTAN HILIR SEBERANG	SMA	PEREMPUAN
53	PETERNAKAN	SMA N 1 KUANTAN HILIR	SMA	PEREMPUAN
54	PETERNAKAN	SMA N 1 BENAI	SMA	LAKI-LAKI
55	PETERNAKAN	SMA N 1 GUNUNG TOAR	SMA	PEREMPUAN
56	PETERNAKAN	SMK N 3 TELUK KUANTAN	SMK	LAKI-LAKI
57	PETERNAKAN	MAN 1 KUANTAN SINGINGI	MA	PEREMPUAN
58	PETERNAKAN	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
59	PETERNAKAN	SMK N 1 MARANCAR	SMK	PEREMPUAN
60	PETERNAKAN	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN

61	PETERNAKAN	SMA N 1 TALUK KUANTAN	SMA	LAKI-LAKI
62	PETERNAKAN	SMA N 1 KUANTAN MUDIK	SMA	PEREMPUAN
63	AGRIBISNIS	SMA N 2 TELUK KUANTAN	SMA	LAKI-LAKI
64	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
65	AGRIBISNIS	SMK N 1 BENAI	SMK	PEREMPUAN
66	AGRIBISNIS	SMA N 1 PANGEAN	SMA	PEREMPUAN
67	AGRIBISNIS	SMK MUHAMMADIYAH CERENTI	SMK	PEREMPUAN
68	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
69	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	LAKI-LAKI
70	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
71	AGRIBISNIS	SMK N 1 KUANTAN MUDIK	SMK	PEREMPUAN
72	AGRIBISNIS	SMK N 1 BENAI	SMK	LAKI-LAKI
73	AGRIBISNIS	SMA N 1	SMA	LAKI-LAKI

		PANGEAN		
74	AGRIBISNIS	SMK N 1 PUCUK RANTAU	SMK	LAKI-LAKI
75	AGRIBISNIS	SMK N 1 BENAI	SMK	LAKI-LAKI
76	AGRIBISNIS	SMA N 2 SENTAJO RAYA	SMA	LAKI-LAKI
77	AGRIBISNIS	SMK N 1 BENAI	SMA	LAKI-LAKI
78	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
79	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
80	AGRIBISNIS	SMA N 1 SENTAJO RAYA	SMA	LAKI-LAKI
81	AGRIBISNIS	SMA N 1 LOGAS TANAH DARAT	SMA	LAKI-LAKI
82	AGRIBISNIS	MAS PONPES NURUL ISLAM	MA	LAKI-LAKI
83	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
84	AGRIBISNIS	MAS PP SYAFA'ATURRAS UL	MA	LAKI-LAKI
85	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	LAKI-LAKI

86	AGRIBISNIS	MAS BABUSSALAM SAPI	MA	PEREMPUAN
87	AGRIBISNIS	SMA N 1 BENAI	SMA	PEREMPUAN
88	AGRIBISNIS	SMA N 1 TELUK KUANTAN	SMA	LAKI-LAKI
89	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
90	AGRIBISNIS	SMK N 1 SENTAJO RAYA	SMK	LAKI-LAKI
91	AGRIBISNIS	SMK N 1 TELUK KUANTAN	SMK	LAKI-LAKI
92	AGRIBISNIS	SMK TEKNOLOGI YPL LIRIK	SMK	LAKI-LAKI
93	AGRIBISNIS	SMA N 1 KUANTAN HILIR	SMA	PEREMPUAN
94	AGRIBISNIS	SMK N 1 PUCUK RANTAU	SMK	PEREMPUAN
95	AGRIBISNIS	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
96	AGRIBISNIS	SMA N 1 CONCONG	SMA	PEREMPUAN
97	AGRIBISNIS	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN

98	AGRIBISNIS	SMA N 1 KUANTAN HILIR	SMA	PEREMPUAN
99	AGRIBISNIS	SMA N 1 PANGEAN	SMA	PEREMPUAN
100	TEKNIK SIPIL	SMA N 1 BESITANG	SMA	PEREMPUAN
101	TEKNIK SIPIL	SMK N 1 KUANTAN MUDIK	SMK	LAKI-LAKI
102	TEKNIK SIPIL	MA PP NURUL ISLAM KAMPUNG BARU	MA	LAKI-LAKI
103	TEKNIK SIPIL	SMK N 1 TELUK KUANTAN	SMK	LAKI-LAKI
104	TEKNIK SIPIL	SMK N 1 BATANG PERANAP	SMK	PEREMPUAN
105	TEKNIK SIPIL	SMA N 1 GUNUNG TOAR	SMA	PEREMPUAN
106	TEKNIK SIPIL	SMA N 1 TELUK KUANTAN	SMA	PEREMPUAN
107	TEKNIK SIPIL	SMA N 1 PANGEAN	SMA	LAKI-LAKI
108	TEKNIK SIPIL	SMK N 1 PERANAP	SMK	LAKI-LAKI
109	TEKNIK SIPIL	SMK N 1 BATANG PERANAP	SMK	LAKI-LAKI

110	TEKNIK SIPIL	SMK N 2 TELUK KUANTAN	SMK	LAKI-LAKI
111	TEKNIK SIPIL	SMK N 1 KUANTAN HILIR	SMK	PEREMPUAN
112	TEKNIK SIPIL	SMA QUEEN ALFALAH	SMA	LAKI-LAKI
113	TEKNIK SIPIL	SMA N 1 BENAI	SMA	LAKI-LAKI
114	PERENCANAAN WILAYAH DAN KOTA	SMA N 1 GUNUNG TOAR	SMA	LAKI-LAKI
115	PERENCANAAN WILAYAH DAN KOTA	SMA N 1 GUNUNG TOAR	SMA	PEREMPUAN
116	PERENCANAAN WILAYAH DAN KOTA	SMA N 1 KUANTAN MUDIK	SMA	PEREMPUAN
117	PERENCANAAN WILAYAH DAN KOTA	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
118	PERENCANAAN WILAYAH DAN KOTA	SMA N 1 PANGEAN	SMA	PEREMPUAN
119	PERENCANAAN WILAYAH DAN KOTA	MAS PPNI KAMPUNG BARU	MA	LAKI-LAKI

120	PERENCANAAN WILAYAH DAN KOTA	SMA N 1 KUANTAN MUDIK	SMA	LAKI-LAKI
121	PERENCANAAN WILAYAH DAN KOTA	SMK N 2 TELUK KUANTAN	SMK	PEREMPUAN
122	PERENCANAAN WILAYAH DAN KOTA	SMK N 3 TANJUNG PINANG	SMK	LAKI-LAKI
123	PERENCANAAN WILAYAH DAN KOTA	MAN TEMBILAHAN	MA	LAKI-LAKI
124	PERENCANAAN WILAYAH DAN KOTA	SMK N 1 KUANTAN MUDIK	SMK	LAKI-LAKI
125	PERENCANAAN WILAYAH DAN KOTA	SMA N 2 SINGINGI	SMA	PEREMPUAN
126	PERENCANAAN WILAYAH DAN KOTA	SMK N 1 TELUK KUANTAN	SMK	LAKI-LAKI
127	TEKNIK INFORMATIKA	SMK N 1 BENAI	SMK	LAKI-LAKI
128	TEKNIK INFORMATIKA	SMK N 1 PUCUK RANTAU	SMK	LAKI-LAKI
129	TEKNIK	SMK N 1	SMK	LAKI-LAKI



	INFORMATIKA	KUANTAN MUDIK		
130	TEKNIK INFORMATIKA	SMA N 1 BENAI	SMA	PEREMPUAN
131	TEKNIK INFORMATIKA	SMA N 1 BENAI	SMA	PEREMPUAN
132	TEKNIK INFORMATIKA	SMK N 1 BENAI	SMK	LAKI-LAKI
133	TEKNIK INFORMATIKA	MA BAHRU ULUM	MA	LAKI-LAKI
134	TEKNIK INFORMATIKA	SMK N 2 TELUK KUATAN	SMK	PEREMPUAN
135	TEKNIK INFORMATIKA	SMA N 1 SENTAJO RAYA	SMA	PEREMPUAN
136	TEKNIK INFORMATIKA	SMA N 2 TELUK KUANTAN	SMA	PEREMPUAN
137	TEKNIK INFORMATIKA	SMA N 1 BENAI	SMA	LAKI-LAKI
138	TEKNIK INFORMATIKA	SMK N 2 TELUK KUANTAN	SMK	LAKI-LAKI
139	TEKNIK INFORMATIKA	SMA N 1 BENAI	SMA	LAKI-LAKI

#### 5.4 *Transformation Data*

Tahapan *transformation* data merupakan tahap merubah data ke dalam bentuk yang sesuai untuk dilakukan proses *assosiasi*. Dari data tabel di atas

terdapat 139 mahasiswa yang masing-masing memiliki kelompok *atribut* yang berbeda-beda. Maka diberikan kode pada setiap *item atribut* agar memudahkan proses pencarian rules. Untuk lebih jelasnya dapat dilihat pada tabel dibawah ini.

**Tabel 5.2** Kode pada masing - masing Item Atribut

<b>ATRIBUT</b>	<b>ITEM</b>	<b>KODE</b>
JURUSAN	AGRIBISNIS	A1
	AGROTEKNOLOGI	A2
	PERENCANAAN WILAYAH DAN KOTA	A3
	PETERNAKAN	A4
	TEKNIK INFORMATIKA	A5
	TEKNIK SIPIL	A6
ASAL SEKOLAH	MAS BAHRUL ULUM	B6
	MAS PONPES NURUL ISLAM	B7
	SMA N 1 BENAI	B10
	SMA N 1 CERENTI	B12
	SMA N 1 GUNUNG TOAR	B14
	SMA N 1 KUANTAN HILIR	B16
	SMA N 1 KUANTAN HILIR SEBERANG	B17
	SMA N 1 KUANTAN MUDIKA	B18
	SMA N 1 PANGEAN	B20

	SMA N 1 SENTAJO RAYA	B21
	SMA N 1 TELUK KUANTAN	B25
	SMA N 2 SENTAJO RAYA	B27
	SMA N 2 TELUK KUANTAN	B30
	SMK MUHAMMADIYAH CERENTI	B34
	SMK N 1 BATANG PERANAP	B35
	SMK N 1 BENAI	B36
	SMK N 1 KUANTAN HILIR	B38
	SMK N 1 KUANTAN MUDIK	B39
	SMK N 1 LOGAS TANAH DARAT	B40
	SMK N 1 PUCUK RANTAU	B45
	SMK N 1 TELUK KUANTAN	B48
	SMK N 2 TELUK KUANTAN	B49
JENIS SEKOLAH	MA	C1
	SMA	C2
	SMK	C3
JENIS KELAMIN	LAKI-LAKI	D1
	PEREMPUAN	D2

Tabel 5.3, dibawah merupakan *itemset* dari data penerimaan mahasiswa pendaftar tahun 2020 Berikut merupakan data yang sudah diubah berdasarkan kode pada Tabel 5.2

**Tabel 5.3.** Data Itemset setelah dirubah dalam bentuk kode

No	Item
1	A2,B30,C2,D1
2	A2,B5,C2,D1
3	A2,B3,C2,D2
4	A2,B14,C2,D2
5	A2,B14,C2,D1
6	A2,B30,C2,D1
7	A2,B49,C2,D2
8	A2,B17,C3,D2
9	A2,B30,C2,D1
10	A2,B30,C2,D1
11	A2,B30,C3,D2
12	A2,B38,C1,D2
13	A2,B7,C2,D2
14	A2,B18,C2,D1
15	A2,B36,C3,D1
16	A2,B20,C2,D1
17	A2,B14,C2,D1
18	A2,B14,C2,D1
19	A2,B27,C2,D2

20	A2,B4,C2,D1
21	A2,B27,C3,D1
22	A2,B25,C1,D1
23	A2,B18,C2,D2
24	A2,B18,C3,D2
25	A2,B40,C2,D2
26	A2,B18,C2,D1
27	A2,B18,C3,D1
28	A2,B2,C2,D2
29	A2,B3,C3,D1
30	A2,B39,C3,D1
31	A2,B18,C2,D1
32	A2,B14,C2,D1
33	A2,B37,C3,D1
34	A2,B40,C2,D1
35	A2,B32,C3,D1
36	A2,B48,C3,D1
37	A2,B40,C1,D1
38	A2,B18,C2,D1
39	A2,B39,C3,D1
40	A2,B35,C2,D1
41	A2,B42,C2,D1
42	A2,B18,C2,D1
43	A2,B10,C2,D2

44	A2,B10,C3,D1
45	A2,B20,C2,D1
46	A4,B20,C2,D1
47	A4,B20,C2,D1
48	A4,B20,C2,D1
49	A4,B30,C2,D1
50	A4,B30,C3,D2
51	A4,B14,C2,D2
52	A4,C2,D2
53	A4,C2,D2
54	A4,B10,C2,D1
55	A4,B14,C2,D2
56	A4,C3,D1
57	A4,C1,D2
58	A4,B18,C2,D1
59	A4,C3,D2
60	A4,B49,C3,D2
61	A4,C2,D1
62	A4,B18,C2,D2
63	A1,B30,C2,D1
64	A1,B49,C3,D2
65	A1,B36,C3,D2
66	A1,B20,C2,D2
67	A1,C3,D2

68	A1,B49,C3,D2
69	A1,B49,C3,D1
70	A1,B49,C3,D2
71	A1,C3,D2
72	A1,B36,C3,D1
73	A1,B20,C2,D1
74	A1,C3,D1
75	A1,B36,C3,D1
76	A1,B4,C2,D1
77	A1,B36,C2,D1
78	A1,B49,C3,D2
79	A1,B49,C3,D2
80	A1,B43,C2,D1
81	A1,B2,C2,D1
82	A1,B43,C1,D1
83	A1,B49,C3,D2
84	A1,B37,C1,D1
85	A1,B49,C3,D1
86	A1,B5,C1,D2
87	A1,B10,C2,D2
88	A1,B5,C2,D1
89	A1,B49,C3,D2
90	A1,B44,C3,D1
91	A1,B45,C3,D1

92	A1,B44,C3,D1
93	A1,B4,C2,D2
94	A1,B5,C3,D2
95	A1,B18,C2,D1
96	A1,B5,C2,D2
97	A1,B49,C3,D2
98	A1,B5,C2,D2
99	A1,B20,C2,D2
100	A6,B19,C2,D2
101	A6,B19,C3,D1
102	A6,B21,C1,D1
103	A6,B5,C3,D1
104	A6,B9,C3,D2
105	A6,B14,C2,D2
106	A6,B5,C2,D2
107	A6,B20,C2,D1
108	A6,B45,C3,D1
109	A6,B23,C3,D1
110	A6,B49,C3,D1
111	A6,B21,C3,D2
112	A6,B6,C2,D1
113	A6,B10,C2,D1
114	A3,B14,C2,D1
115	A3,B14,C2,D2



116	A3,B18,C2,D2
117	A3,B49,C3,D2
118	A3,B20,C2,D2
119	A3,B9,C1,D1
120	A3,B18,C2,D1
121	A3,B49,C3,D2
122	A3,B7,C3,D1
123	A3,B4,C1,D1
124	A3,B3,C3,D1
125	A3,B43,C2,D2
126	A3,B49,C3,D1
127	A5,B36,C3,D1
128	A5,B2,C3,D1
129	A5,B9,C3,D1
130	A5,B10,C2,D2
131	A5,B10,C2,D2
132	A5,B36,C3,D1
133	A5,B11,C1,D1
134	A5,B32,C3,D2
135	A5,B3,C2,D2
136	A5,B30,C2,D2
137	A5,B10,C2,D1
138	A5,B49,C3,D1
139	A5,B10,C2,D1

### 5.5 Analisis Frequent Itemset

Frequent itemset adalah algoritma yang digunakan untuk mencari frequent itemset (Tjandra. S, 2019). Pembentukan frequent terdiri dari beberapa item atau field data yang ada pada dataset, kode atau pelabelan yang dibuat untuk melambangkan data tersebut, frekwency item yaitu jumlah kemunculan setiap item yang muncul di dataset. Dan yang terakhir adalah support itemset. Support itemset yang diperoleh dari frekwency yang dibagi dengan total dataset yang ada.

Tahapan ini memberikan batasan pada *dataset* menggunakan *support count* yang telah ditentukan, pada penelitian ini peneliti menentukan nilai *minimum support* = 0,05 atau 5%, jika item yang tidak memenuhi *minimum support* maka tidak akan muncul kedalam tabel. Setelah selesai dari tahap penyeleksian minimum support nantinya data akan dibangun menjadi sebuah *Tree*. Adapun rumus yang digunakan untuk mencari nilai support yaitu :

$$Support (A) = \frac{\text{Jumlah Transaksi Mengandung A}}{\text{Total transaksi}} \dots\dots\dots (5.1)$$

Berikut ini adalah tabel dengan semua *item* yang telah digabungkan serta dirumuskan.

**Tabel 5.4.** Nilai Support Per Item

No	Item	Kode	Frequency	Total	Support
1	AGRIBISNIS	A1	37	139	27%
2	AGROTEKNOLOGI	A2	45	139	32%
3	PERENCANAAN WILAYAH DAN KOTA	A3	13	139	9%
4	PETERNAKAN	A4	17	139	12%
5	TEKNIK INFORMATIKA	A5	13	139	9%

6	TEKNIK SIPIL	A6	14	139	1%
7	MAS BAHRUL ULUM	B6	2	139	15%
8	MAS PONPES NURUL ISLAM	B7	2	139	1%
9	SMA N 1 BENAI	B10	8	139	6%
10	SMA N 1 CERENTI	B12	2	139	1%
11	SMA N 1 GUNUNG TOAR	B14	8	139	6%
12	SMA N 1 KUANTAN HILIR	B16	5	139	4%
13	SMA N 1 KUANTAN HILIR SEBERANG	B17	2	139	1%
14	SMA N 1 KUANTAN MUDIK	B18	11	139	8%
15	SMA N 1 PANGEAN	B20	8	139	6%
16	SMA N 1 SENTAJO RAYA	B21	4	139	3%
17	SMA N 1 TELUK KUANTAN	B25	2	139	1%
18	SMA N 2 SENTAJO RAYA	B27	3	139	2%
19	SMA N 2 TELUK KUANTAN	B30	7	139	5%
20	SMK MUHAMMADIYAH CERENTI	B34	2	139	1%
21	SMK N 1 BATANG PERANAP	B35	3	139	2%
22	SMK N 1 BENAI	B36	7	139	5%
23	SMK N 1 KUANTAN	B38	2	139	1%

	HILIR				
24	SMK N 1 KUANTAN MUDI	B39	5	139	4%
25	SMK N 1 LOGAS TANAH DARAT	B40	3	139	2%
26	SMK N 1 PUCUK RANTAU	B45	3	139	2%
27	SMK N 1 TELUK KUANTAN	B48	4	139	3%
28	SMK N 2 TELUK KUANTAN	B49	15	139	11%
29	MA	C1	11	139	8%
30	SMA	C2	73	139	53%
31	SMK	C3	55	139	40%
32	LAKI-LAKI	D1	82	139	59%
33	PEREMPUAN	D2	57	139	41%

## 5.6 Pembangunan *FP-Tree*

Fp-tree yang dibangun berdasarkan *frequent item* pada setiap mahasiswa, aturannya jika *item* tidak *support* akan dihilangkan yaitu minimum support 5%. Berikut data mahasiswa lama yang memenuhi nilai *support* dapat dilihat pada tabel di bawah ini.

**Tabel 5.5.** Data Itemset Data Mahasiswa yang memenuhi nilai support

No	Item
1	A2,B30,C2,D1

2	A2,C2,D1
3	A2,C2,D2
4	A2,B14,C2,D2
5	A2,C2,D1
6	A2,B30,C2,D1
7	A2,C2,D2
8	A2,C3,D2
9	A2,B30,C2,D1
10	A2,B30,C2,D1
11	A2,C3,D2
12	A2,C1,D2
13	A2,C2,D2
14	A2,B18,C2,D1
15	A2,B36,C3,D1
16	A2,B20,C2,D1
17	A2,B14,C2,D1
18	A2,C2,D1
19	A2,C2,D2
20	A2,C2,D1
21	A2,C3,D1
22	A2,C1,D1
23	A2,B18,C2,D2
24	A2,C3,D2
25	A2,C2,D2

26	A2,B18,C2,D1
27	A2,C3,D1
28	A2,C2,D2
29	A2,C3,D1
30	A2,C3,D1
31	A2,B18,C2,D1
32	A2,B14,C2,D1
33	A2,C3,D1
34	A2,C2,D1
35	A2,C3,D1
36	A2,C3,D1
37	A2,C1,D1
38	A2,B18,C2,D1
39	A2,C3,D1
40	A2,C2,D1
41	A2,C2,D1
42	A2,B18,C2,D1
43	A2,B10,C2,D2
44	A2,C3,D1
45	A2,B20,C2,D1
46	A4,C2,D1
47	A4,B20,C2,D1
48	A4,C2,D1
49	A4,B30,C2,D1

50	A4,C3,D2
51	A4,B14,C2,D2
52	A4,C2,D2
53	A4,C2,D2
54	A4,B10,C2,D1
55	A4,B14,C2,D2
56	A4,C3,D1
57	A4,C1,D2
58	A4,B18,C2,D1
59	A4,C3,D2
60	A4,B49,C3,D2
61	A4,C2,D1
62	A4,B18,C2,D2
63	A1,B30,C2,D1
64	A1,B49,C3,D2
65	A1,B36,C3,D2
66	A1,B20,C2,D2
67	A1,C3,D2
68	A1,B49,C3,D2
69	A1,B49,C3,D1
70	A1,B49,C3,D2
71	A1,C3,D2
72	A1,B36,C3,D1
73	A1,B20,C2,D1

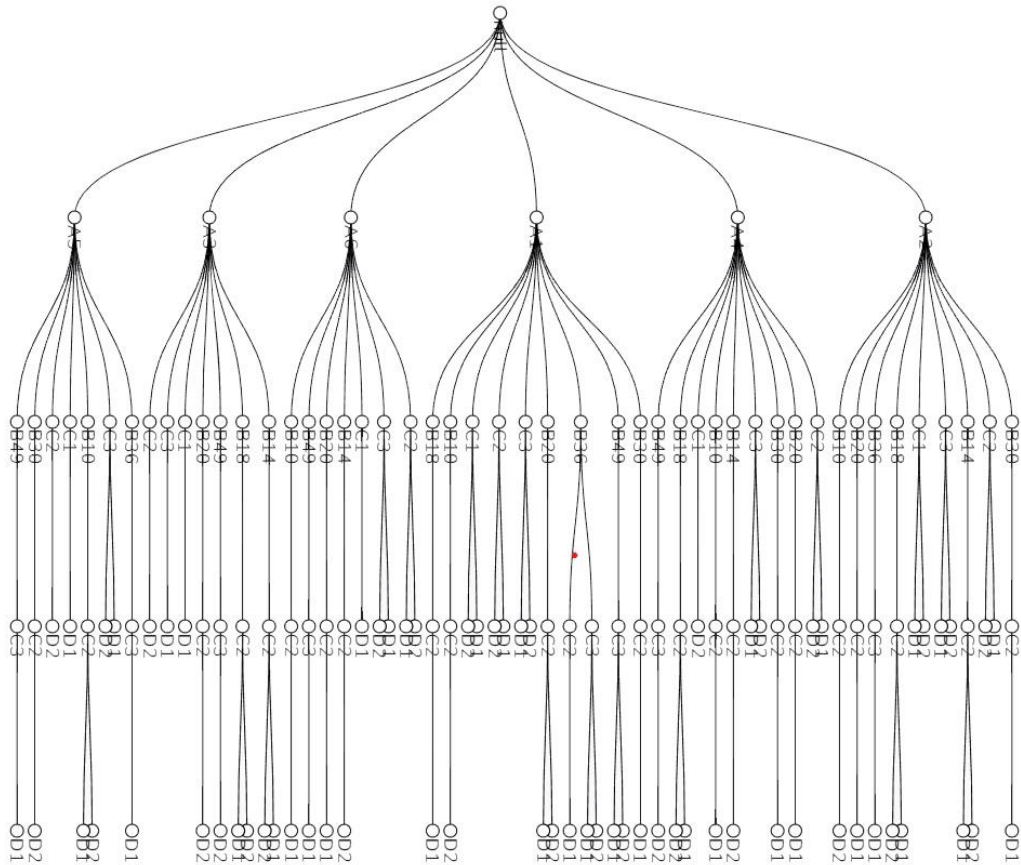
74	A1,C3,D1
75	A1,B36,C3,D1
76	A1,C2,D1
77	A1,B36,C2,D1
78	A1,B49,C3,D2
79	A1,B49,C3,D2
80	A1,C2,D1
81	A1,C2,D1
82	A1,C1,D1
83	A1,B49,C3,D2
84	A1,C1,D1
85	A1,B49,C3,D1
86	A1,C1,D2
87	A1,B10,C2,D2
88	A1,C2,D1
89	A1,B49,C3,D2
90	A1,C3,D1
91	A1,C3,D1
92	A1,C3,D1
93	A1,C2,D2
94	A1,C3,D2
95	A1,B18,C2,D1
96	A1,C2,D2
97	A1,B49,C3,D2



98	A1,C2,D2
99	A1,B20,C2,D2
100	A6,C2,D2
101	A6,C3,D1
102	A6,C1,D1
103	A6,C3,D1
104	A6,C3,D2
105	A6,B14,C2,D2
106	A6,C2,D2
107	A6,B20,C2,D1
108	A6,C3,D1
109	A6,C3,D1
110	A6,B49,C3,D1
111	A6,C3,D2
112	A6,C2,D1
113	A6,B10,C2,D1
114	A3,B14,C2,D1
115	A3,B14,C2,D2
116	A3,B18,C2,D2
117	A3,B49,C3,D2
118	A3,B20,C2,D2
119	A3,C1,D1
120	A3,B18,C2,D1
121	A3,B49,C3,D2

122	A3,C3,D1
123	A3,C1,D1
124	A3,C3,D1
125	A3,C2,D2
126	A3,C3,D1
127	A5,B36,C3,D1
128	A5,C3,D1
129	A5,C3,D1
130	A5,B10,C2,D2
131	A5,B10,C2,D2
132	A5,B36,C3,D1
133	A5,C1,D1
134	A5,C3,D2
135	A5,C2,D2
136	A5,B30,C2,D2
137	A5,B10,C2,D1
138	A5,B49,C3,D1
139	A5,B10,C2,D1

Adapun hasil dari tabel diatas dapat dihasilkan pembentukan sebuah FP-Tree. Berikut hasil pembentukan *FP-Tree* :



**Gambar 5.1** Alur pohon *FP-Tree*

Dari *itemset* di atas, menunjukkan terbentuknya *FP-Tree* setiap *TID* dibaca. Setiap simpul pada *FP-Tree* nama sebuah *item* dan *counter support* yang berfungsi untuk menghitung frekuensi kemunculan *item* tersebut dalam tiap cabang *itemset*.

*FP-Tree* yang mempresentasikan data *itemset* dibentuk dengan cara sebagai berikut :

1. Kumpulan data akan dipindahkan untuk menentukan support count dari setiap *item*. *Item* yang tidak terpengaruh frekuensi akan dibuang, sedangkan *itemset* yang memenuhi frekuensi akan diambil dimulai dari urutan kode terbesar sampai urutan terkecil.
2. Setelah pemindaian kedua, yaitu pembacaan *TID* pertama (A1, C1,D1) maka akan membuat sebuah simpul A1, C1 dan D1, sehingga terbentuk lintasan *itemset* Null→A5→C1→D1. *Support count* dari setiap cabang bernilai awal 1.

3. Setelah pembacaan *itemset* kedua (A1,C2,D2) terbentuk cabang kedua yaitu  $\text{Null} \rightarrow A1 \rightarrow C2 \rightarrow D2$ . *Support count* masing-masing dari setiap simpul juga bernilai awal 1. Walaupun A4 ada pada *itemset* kedua, namun karena prefix *itemset*nya tidak sama, maka *itemset* kedua ini tidak bisa dimasukkan dalam satu lintasan.

### 5.7 Algoritma FP-Growth

Penggalian *itemset* yang *frequent* dengan menggunakan *FP-Growth* akan dilakukan dengan cara membangkitkan struktur data *tree* atau disebut dengan *FP-Tree*. Metode *FP-Growth* dapat dibagi menjadi 3 tahapan utama yaitu sebagai berikut :

1. Tahap Pembangkitan *Conditional Pattern Base*.

Pembangkit *Conditional Pattern Base* didapat melalui *FP-Tree* yang telah dibangun sebelumnya.

2. Tahap Pembangkitan *Conditional FP-Tree*

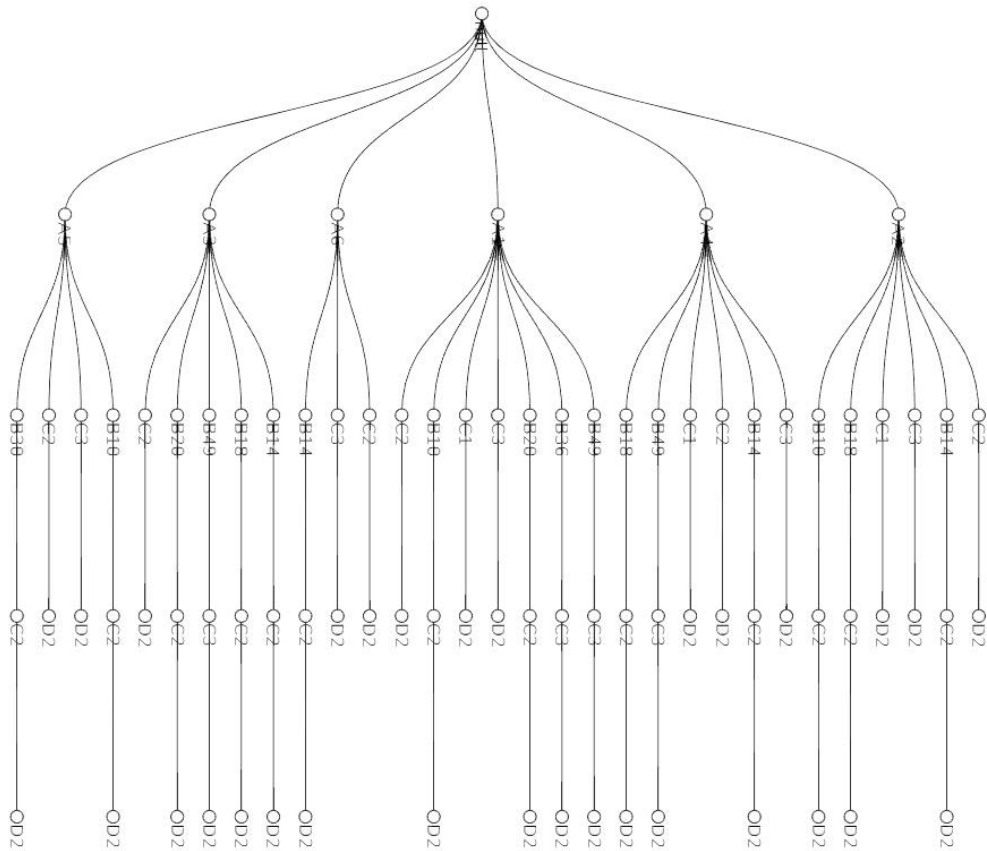
Pada tahap ini, *support count* dari setiap item pada setiap *Conditional Pattern Base* dikumpulkan, kemudian setiap item yang memiliki jumlah *support count* lebih besar sama dengan *minimum support count* akan dibandingkan dengan *Conditional FP-Tree*.

3. Tahap Pencarian frequent itemset

Tahapan ini bertujuan untuk pencarian frequent itemset dari Penerimaan Mahasiswa Baru.

Untuk menemukan frequent itemset dari contoh 1, maka perlu ditentukan cabang pohon dengan lintasan yang berakhiran support count paling kecil yaitu D2. Proses pembentukan adalah sebagai berikut :

1. Gambar dibawah ini merupakan gambar yang hanya menampilkan cabang yang berakhiran D2 yang berjumlah 43 cabang, berikut gambarannya.



**Gambar 5.2** Cabang yang berakhir D2

Maka hasil dari setiap cabang dapat kita lihat dalam tabel berikut :

**Tabel 5.6** Kemunculan item cabang tree D2

No	Tree	Kemunculan
1	A2	45
	C2	13
	D2	6
2	A2	45
	B14	3
	C2	3
	D2	1
3	A2	45
	C3	12
	D2	3

4	A2	45
	C1	3
	D2	1
5	A2	45
	B18	6
	C2	6
	D2	1
6	A2	45
	B10	1
	C2	1
	D2	1
7	A4	17
	C2	5
	D2	2
8	A4	17
	C3	3
	D2	2
9	A4	17
	B14	2
	C2	2
	D2	2
10	A4	17
	C1	1
	D2	1
11	A4	17
	B18	2
	C2	2
	D2	1
12	A4	17
	B49	1

	C3	1
	D2	1
13	A1	37
	B49	10
	C3	10
	D2	8
14	A1	37
	B36	4
	C3	3
	D2	1
15	A1	37
	B20	3
	C2	3
	D2	2
16	A1	37
	C3	7
	D2	3
17	A1	37
	C2	7
	D2	3
18	A1	37
	C1	3
	D2	1
19	A1	37
	B10	1
	C2	1
	D2	1
20	A6	14
	C2	3

	D2	2
21	A6	14
	C3	6
	D2	2
22	A6	14
	B14	1
	C2	1
	D2	1
23	A3	13
	B14	2
	C2	2
	D2	1
24	A3	13
	B18	2
	C2	2
	D2	1
25	A3	13
	B49	2
	C3	2
	D2	2
26	A3	13
	B20	1
	C2	1
	D2	1
27	A3	13
	C2	1
	D2	1
28	A5	13





2	A2	45
	C2	13
	D1	7
3	A2	45
	B14	3
	C2	3
	D1	2
4	A2	45
	C3	12
	D1	9
5	A2	45
	C1	3
	D1	2
6	A2	45
	B18	6
	C2	6
	D1	5
7	A2	45
	B36	1
	C3	1
	D1	1
8	A2	45
	B20	2
	C2	2
	D1	2

9	A4	17
	C2	5
	D1	3
10	A4	17
	B20	1
	C2	1
	D1	1
11	A4	17
	B30	1
	C2	1
	D1	1
12	A4	17
	C3	3
	D1	1
13	A4	17
	B10	1
	C2	1
	D1	1
13	A4	17
	B18	2
	C2	2
	D1	1
14	A1	37
	B30	1
	C2	1
	D1	1

15	A1	37
	B49	10
	C3	10
	D1	2
16	A1	37
	B36	4
	C3	3
	D1	2
17	A1	37
	B36	4
	C2	1
	D1	1
18	A1	37
	B36	4
	B20	3
	C2	3
	D1	1
19	A1	37
	B36	4
	C3	7
	D1	4
20	A1	37
	B36	4
	C2	7
	D1	4

21	A1	37
	B36	4
	C1	3
	D1	2
22	A1	37
	B36	4
	B18	1
	C2	1
	D1	1
23	A6	14
	C2	3
	D1	1
24	A6	14
	C3	6
	D1	4
25	A6	14
	C1	1
	D1	1
26	A6	14
	B20	1
	C2	1
	D1	1
27	A6	14
	B49	1
	C3	1

	D1	1
28	A6	14
	B10	1
	C2	1
	D1	1
29	A3	13
	B14	2
	C2	2
	D1	1
30	A3	13
	B18	2
	C2	2
	D1	1
31	A3	13
	C1	2
	D1	2
32	A3	13
	C3	3
	D1	3
33	A5	13
	B36	2
	C3	2
	D1	2
34	A5	13
	C3	3

	D1	2
35	A5	13
	B10	4
	C2	4
	D1	2
36	A5	13
	C1	1
	D1	1
37	A5	13
	B49	1
	C3	1
	D1	1

Dari tabel yang ada diatas maka tahapan selanjutnya adalah menghitung dan menentukan frequency dan minimum support dari setiap item. Untuk Perhitungan frequencynya didapat dari penjumlahan dari setiap item yang muncul pada cabang tree sedangkan untuk minimum size didapat dengan rumus *Support*

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

Pengecekan nilai *support* pada *suffix* D1 dan

D2 dapat dilihat pada tabel di bawah ini.

**Tabel 5.8** Nilai Support dari Suffix D1

No	Item	Frequency	Support	Total data
1	D1	82	59%	139
2	C2,D1	106	76%	139
3	C3,D1	84	60%	139

4	C1,D1	18	13%	139
5	B30,C2,D1	18	13%	139
6	A2,C2,D1	65	47%	139
7	B14,C2,D1	13	9%	139
8	A2,C3,D1	66	47%	139
9	A2,C1,D1	50	36%	139
10	B18,C2,D1	30	22%	139
11	B36,C3,D1	33	24%	139
12	B20,C2,D1	19	14%	139
13	A4,C2,D1	25	18%	139
14	A4,C3,D1	21	15%	139
15	B10,C2,D1	16	12%	139
16	B49,C3,D1	28	20%	139
17	B36,C2,D1	21	15%	139
18	B36,C1,D1	9	6%	139
19	A6,C2,D1	18	13%	139
20	A6,C3,D1	24	17%	139
21	A6,C1,D1	16	12%	139
22	A3,C1,D1	17	12%	139
23	A3,C3,D1	19	14%	139
24	A5,C3,D1	18	13%	139
25	A5,C1,D1	15	11%	139
26	A2,B30,C2,D1	57	41%	139
27	A2,B14,C2,D1	53	38%	139



28	A2,B18,C2,D1	62	45%	139
29	A2,B36,C3,D1	48	35%	139
30	A2,B20,C2,D1	51	37%	139
31	A4,B20,C2,D1	20	14%	139
32	A4,B30,C2,D1	20	14%	139
33	A4,B10,C2,D1	20	14%	139
34	A4,B18,C2,D1	22	16%	139
35	A1,B30,C2,D1	40	29%	139
36	A1,B49,C3,D1	59	42%	139
37	A1,B36,C3,D1	98	71%	139
38	A1,B36,C2,D1	95	68%	139
39	B36,B20,C2,D1	11	8%	139
40	A1,B36,C1,D1	46	33%	139
41	B36,B18,C2,D1	7	5%	139
42	A6,B20,C2,D1	17	12%	139
43	A6,B49,C3,D1	17	12%	139
44	A6,B10,C2,D1	17	12%	139
45	A3,B14,C2,D1	18	13%	139
46	A3,B18,C2,D1	18	13%	139
47	A5,B36,C3,D1	19	14%	139
48	A5,B10,C2,D1	23	17%	139
49	A5,B49,C3,D1	16	12%	139
50	A1,B36,B20,C2,D1	48	35%	139
51	A1,B36,B18,C2,D1	44	32%	139

**Tabel 5.9** Nilai Support dari *Suffix* D2

<b>No</b>	<b>Suffix</b>	<b>Frequency</b>	<b>Support</b>	<b>Total Data</b>
1	D2	57	41%	139
2	C2,D2	90	65%	139
3	C3,D2	70	50%	139
4	C1,D2	10	7%	139
5	A2,C2,D2	64	46%	139
6	B14,C2,D2	21	15%	139
7	A2,C3,D2	60	43%	139
8	A2,C1,D2	49	35%	139
9	B18,C2,D2	23	17%	139
10	B10,C2,D2	16	12%	139
11	A4,C2,D2	24	17%	139
12	A4,C3,D2	22	16%	139
13	A4,C1,D2	19	14%	139
14	B49,C3,D2	37	27%	139
15	B36,C3,D2	8	6%	139
16	B20,C2,D2	11	8%	139
17	A1,C3,D2	47	34%	139
18	A1,C2,D2	47	34%	139
19	A1,C1,D2	41	29%	139
20	A6,C2,D2	19	14%	139
21	A6,C3,D2	22	16%	139
22	A3,C2,D2	15	11%	139

23	A5,C3,D2	17	12%	139
24	A5,C2,D2	15	11%	139
25	B30,C2,D2	3	2%	139
26	A2,B14,C2,D2	52	37%	139
27	A2,B18,C2,D2	58	42%	139
28	A2,B10,C2,D2	48	35%	139
29	A4,B14,C2,D2	23	17%	139
30	A4,B18,C2,D2	22	16%	139
31	A4,B49,C3,D2	20	14%	139
32	A1,B49,C3,D2	65	47%	139
33	A1,B36,C3,D2	45	32%	139
34	A1,B20,C2,D2	45	32%	139
35	A1,B10,C2,D2	40	29%	139
36	A6,B14,C2,D2	17	12%	139
37	A3,B14,C2,D2	18	13%	139
38	A3,B18,C2,D2	18	13%	139
39	A3,B49,C3,D2	19	14%	139
40	A3,B20,C2,D2	16	12%	139
41	A5,B10,C2,D2	23	17%	139
42	A5,B30,C2,D2	16	12%	139

Setelah perhitungan frequency dan support didapatkan maka tahapan selanjutnya yaitu Mengecek kembali nilai support dari *suffix* itemset lebih besar atau tidak dari *minimum support* yang ditentukan pada saat penelitian. Untuk

penelitian ini, peneliti mengambil keputusan minimum support yang diambil yaitu sebesar 10% dan untuk minimum size adalah 4.

Selanjutnya pengecekan dapat dilihat dari tabel 5.8 dengan melakukan pengecekan awal terhadap D1 (*suffix*), nilai *support* D1 memenuhi *minimum support* namun tidak memenuhi minimum size. Kemudian dilakukan pengecekan kembali terhadap dua *suffix* yaitu C2,D1 dan C3,D1 masih belum memenuhi *minimum support* dan *minimum size*. Dari pengecekan nilai *support* A1,B30,C2,D1, A2,B30,C2,D1, A2,B14,C2,D1, A2,B18,C2,D1, A2,B36,C3,D1, A2,B20,C2,D1, A1,B49,C3,D1, A1,B36,C1,D1, A5,B49,C3,D1 dari semuanya hampir semua dari itemset memenuhi *minimum support*, hanya saja nilai support A5,B49,C3,D1 yang tidak memenuhi *minimum support*. Kemudian pengecekan terhadap tiga *suffix* terakhir yang mengandung nilai *support* A1,B36,B20,C2,D1 dan B36,B18,C2,D1 juga memenuhi *minimum support*, maka A1,B30,C2,D1, A2,B30,C2,D1, A2,B14,C2,D1, A2,B18,C2,D1, A2,B36,C3,D1, A2,B20,C2,D1, A1,B49,C3,D1, A1,B36,C1,D1, A5,B49,C3,D1 merupakan *Frequent* itemset.

Dari tabel 5.9 di atas dilakukan pengecekan awal terhadap D2 (*suffix*), nilai *support* D2 tidak memenuhi *minimum support*, kemudian dilakukan pengecekan kembali terhadap dua *suffix* yaitu A1,C1,D2 dan A5,C3,D2 belum memenuhi *minimum support*. Kemudian pengecekan nilai *support* A2,B14,C2,D2, A2,B18,C2,D2, A2,B10,C2,D2, A1,B49,C3,D2, A1,B36,C3,D2 dari semuanya memenuhi *minimum support*, Kemudian pengecekan terakhir terhadap dua *suffix* terakhir yang mengandung nilai *support* A5,B10,C2,D2 dan A5,B30,C2,D2 tidak memenuhi *minimum support*, maka

A2,B14,C2,D2 , A2,B18,C2,D2 , A2,B10,C2,D2 , A1,B49,C3,D2 ,  
A1,B36,C3,D2 merupakan *Frequent* itemset.

Setelah selesai pengecekan terhadap *suffix* pertama dan kedua. Maka tahapan selanjutnya membuat sebuah tabel yang sudah memnuhi saringan minimum support dan minimum size yang sudah ditentukan dan dengan menambah pencarian nilai confidence yang dapat dihitung dengan rumus *Confidence* (A(item paling awal) → B(item yang berdampinan dengan A) =

$$\frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}} \times 100\% .$$

Berikut adalah tabel hasil dari

perhitungan tersebut :

**Tabel 5.10** Nilai confidence dari Suffix

No	Item	jumlah	frequency	support	confidence
1	A2,B14,C2,D2	45	52	37%	87%
2	A2,B18,C2,D2	45	58	42%	78%
3	A2,B10,C2,D2	45	48	35%	9%
4	A4,B14,C2,D2	17	23	17%	74%
5	A4,B18,C2,D2	17	22	16%	77%
6	A4,B49,C3,D2	17	20	14%	85%
7	A1,B49,C3,D2	37	65	47%	57%
8	A1,B36,C3,D2	37	45	32%	82%
9	A1,B20,C2,D2	37	45	32%	82%
10	A1,B10,C2,D2	37	40	29%	93%
11	A6,B14,C2,D2	14	17	12%	82%
12	A3,B14,C2,D2	13	18	13%	72%
13	A3,B18,C2,D2	13	18	13%	72%
14	A3,B49,C3,D2	13	19	14%	68%
15	A3,B20,C2,D2	13	16	12%	81%

16	A5,B10,C2,D2	13	23	17%	57%
17	A5,B30,C2,D2	13	16	12%	81%
18	A2,B30,C2,D1	45	57	41%	79%
19	A2,B14,C2,D1	45	53	38% %	85%
20	A2,B18,C2,D1	45	62	45%	73%
21	A2,B36,C3,D1	45	48	35%	95%
22	A2,B20,C2,D1	45	51	37%	88%
23	A4,B20,C2,D1	17	20	14%	85%
24	A4,B30,C2,D1	17	20	14%	85%
25	A4,B10,C2,D1	17	20	14%	85%
26	A4,B18,C2,D1	17	22	16%	77%
27	A1,B30,C2,D1	37	40	29%	92%
28	A1,B49,C3,D1	37	59	42%	63%
29	A1,B36,C3,D1	37	98	71%	38%
30	A1,B36,C2,D1	37	95	68%	39%
31	A1,B36,C1,D1	37	46	33%	80%
32	A6,B20,C2,D1	14	17	12%	82%
33	A6,B49,C3,D1	14	17	12%	82%
34	A6,B10,C2,D1	14	17	12%	82%
35	A3,B14,C2,D1	13	18	13%	72%
36	A3,B18,C2,D1	13	18	13%	72%
37	A5,B36,C3,D1	13	19	14%	68%
38	A5,B10,C2,D1	13	23	17%	57%
39	A5,B49,C3,D1	13	16	12%	81%
40	A1,B36,B20,C2,D1	37	48	35%	77%
41	A1,B36,B18,C2,D1	37	44	32%	84%

## 5.8 Pembentukan *Association Rules*

Analisis asosiasi dikenal juga sebagai salah satu teknik Data Mining yang menjadi dasar dari berbagai teknik Data Mining lainnya. Khususnya salah satu tahap dari analisis asosiasi yang disebut analisis pola frekuensi tinggi (*frequent pattern mining*) menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* (nilai penunjang) yaitu persentase kombinasi item tersebut dalam database dan *confidence* (Budianto.H & Jeje.R, 2020).

Pembentukan *association rules* atau aturan *assosiasi* dilakukan untuk mencari aturan asosiatif  $A \rightarrow B$  yang memenuhi syarat minimum nilai *confidence*, yaitu untuk mencari keterhubungan antar item dalam suatu itemset, pada kasus ini mencari seberapa keterhubungan jenis sekolah terhadap asal sekolah dan program studi. Di dalam penelitian ini peneliti mengambil *sample* dengan nilai minimum *support* sebesar 30 % atau 0.30 , minimum *size*nya 4 dengan nilai *confidence* 50% atau 0.50%, pencarian nilai *confidence* dapat dihitung dengan rumus

$$Confidence (A \rightarrow B) = \frac{\text{Jumlah Transaksi Mengandung A dan B}}{\text{Jumlah Transaksi Mengandung A}} \times 100\% \dots \dots \dots (5.2)$$

dan dapat dilihat pada tabel di bawah ini.

**Tabel 5.11** Data Calon Nilai *Confidence* Dari *Frequent Itemset*

No	path	jumlah	frequency	support	confidence	Total data
1	A2,B14,C2,D2	45	52	0.37	87 %	139
2	A2,B18,C2,D2	45	58	0.42	78 %	139
3	A2,B10,C2,D2	45	48	0.35	94 %	139
4	A1,B49,C3,D2	37	65	0.47	57 %	139
5	A1,B36,C3,D2	37	45	0.32	83 %	139
6	A1,B20,C2,D2	37	45	0.32	83 %	139
7	A2,B30,C2,D1	45	57	0.41	79 %	139
8	A2,B14,C2,D1	45	53	0.38	85 %	139

9	A2,B18,C2,D1	45	62	0.45	73 %	139
10	A2,B36,C3,D1	45	48	0.35	94 %	139
11	A2,B20,C2,D1	45	51	0.37	89 %	139
12	A1,B49,C3,D1	37	59	0.42	63 %	139
13	A1,B36,C1,D1	37	46	0.33	81 %	139
14	A1,B36,B20,C2,D1	37	48	0.35	78 %	139
15	A1,B36,B18,C2,D1	37	44	0.32	85 %	139

Berdasarkan tabel dan diagram di atas, rules memiliki nilai confidence yang memenuhi nilai minimum support, size dan confidence ada 15 yaitu semua data dari tabel diatas. Data nilai *confidence* yang memenuhi minimum *confidence* dapat dilihat pada tabel dibawah ini :

**Tabel 5.12** Data Nilai Confidence Dari Frequent Itemset

<b>N o</b>	<b>Itemset</b>	<b>Jml</b>	<b>Frekuensi</b>	<b>Support</b>	<b>Confidence</b>
1	AGROTEKNOLOGI SMA N 1 GUNUNG TOAR SMA PEREMPUAN	45	52	37%	87%
2	AGROTEKNOLOGI SMA N 1 KUANTAN MUDIK SMA PEREMPUAN	45	58	42%	78%
3	AGROTEKNOLOGI SMA N 1 BENAI SMA PEREMPUAN	45	48	35%	94%
4	AGRIBISNIS SMK N 2 TELUK KUANTAN	37	65	47%	57%



	SMK PEREMPUAN				
5	AGRIBISNIS SMK N 1 BENAI SMK PEREMPUAN	37	45	32%	83%
6	AGRIBISNIS SMA N 1 PANGEAN SMA PEREMPUAN	37	45	32%	83%
7	AGROTEKNOLOGI SMA N 2 TELUK KUANTAN SMA LAKI-LAKI	45	57	41%	79%
8	AGROTEKNOLOGI SMA N 1 GUNUNG TOAR SMA LAKI-LAKI	45	53	38%	85%
9	AGROTEKNOLOGI SMA N 1 KUANTAN MUDIK SMA LAKI-LAKI	45	62	45%	73%
10	AGROTEKNOLOGI SMK N 1 BENAI SMK LAKI-LAKI	45	48	35%	94%
11	AGROTEKNOLOGI SMA N 1 PANGEAN SMA LAKI-LAKI	45	51	37%	89%
12	AGRIBISNIS	37	59	42%	63%

	SMK N 2 TELUK KUANTAN SMK LAKI-LAKI				
13	AGRIBISNIS SMK N 1 BENAI MA LAKI-LAKI	37	46	33%	81%
14	AGRIBISNIS SMK N 1 BENAI SMA N 1 PANGEAN SMA LAKI-LAKI	37	48	35%	78%
15	AGRIBISNIS SMK N 1 BENAI SMA N 1 KUANTAN MUDIK SMA LAKI-LAKI	37	44	32%	85%

Selanjutnya karena nilai tabel diatas sudah memenuhi nilai minimum *confidence* maka dapat dijadikan *rules*, dimana memiliki tingkat kemunculan tinggi dan keterhubungan yang tinggi pada suatu data calon mahasiswa. *Association rules* yang dapat diperoleh yaitu :

1. Nilai A2,B14,C2,D2 mempunyai nilai frequent 52 dengan *support* 26 % Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMAN 1 Gunung Toar dengan jenis sekolah SMA dan Berjenis kelamin perempuan serta dengan tingkat *confidence* sebesar 87%.
2. Nilai A2,B18,C2,D2 mempunyai nilai frequent 58 dengan *support* 62% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMA Negeri 1 Kuantan Mudik dengan jenis sekolah SMA serta dengan tingkat *confidence* sebesar 78%.

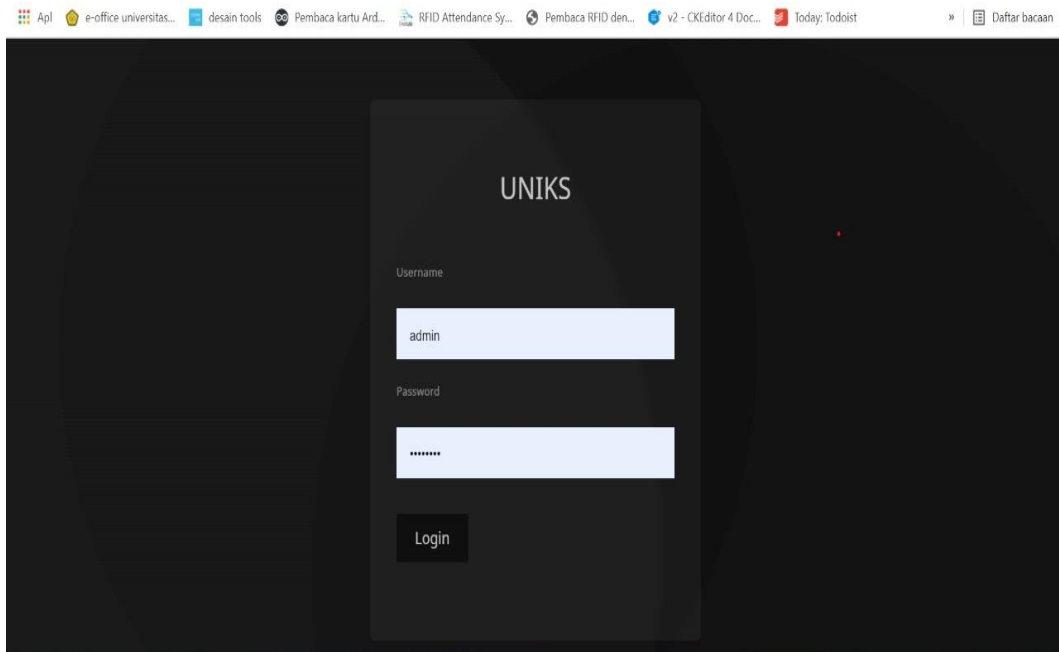
3. Nilai A2,B10,C2,D2 mempunyai nilai frequent 48 dengan *support* 26 % Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMAN 1 Gunung Toar dengan jenis sekolah SMA dan Berjenis kelamin perempuan serta dengan tingkat *confidence* sebesar 94%.
4. Nilai A1,B49,C3,D2 mempunyai nilai frequent 65 dengan *support* 47% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK Negeri 2 Telukkuantan dengan jenis sekolah SMK dan berjenis kelamin perempuan serta dengan tingkat *confidence* sebesar 57%.
5. Nilai A1,B36,C3,D2 mempunyai nilai frequent 45 dengan *support* 32% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK Negeri 1 Benai dengan jenis sekolah SMK dan berjenis kelamin perempuan serta dengan tingkat *confidence* sebesar 83%.
6. Nilai A1,B20,C2,D2 mempunyai nilai frequent 45 dengan *support* 32% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMA N 1 Pangean dengan jenis sekolah SMA dan Berjenis kelamin perempuan serta dengan tingkat *confidence* sebesar 83%.
7. Nilai A2,B30,C2,D1 mempunyai nilai frequent 57 dengan *support* 41% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMA N 2 Telukkuantan dengan jenis sekolah SMA serta dengan tingkat *confidence* sebesar 79%.
8. Nilai A2,B14,C2,D1 mempunyai nilai frequent 53 dengan *support* 38% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMAN 1 Gunung Toar dengan jenis sekolah SMA dan Berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 85%.
9. Nilai A2,B18,C2,D1 mempunyai nilai frequent 62 dengan *support* 45% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMA Negeri 1 Kuantan Mudik dengan jenis sekolah SMA dan berjenis kelamin Laki - laki serta dengan tingkat *confidence* sebesar 73%.
10. Nilai A2,B36,C3,D1 mempunyai nilai frequent 48 dengan *support* 35% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMKN 1 Benai dengan jenis sekolah SMk dan Berjenis kelamin Laki - laki serta dengan tingkat *confidence* sebesar 94%.

11. Nilai A2,B20,C2,D1 mempunyai nilai frequent 51 dengan *support* 37% Artinya Argoteknologi akan dipilih oleh mahasiswa yang berasal dari SMA N 1 Pangean dengan jenis sekolah SMA dan berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 89%.
12. Nilai A1,B49,C3,D1 mempunyai nilai frequent 59 dengan *support* 42% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK Negeri 2 Teluk Kuantan dengan jenis sekolah SMK dan berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 63%.
13. Nilai A1,B36,C1,D1 mempunyai nilai frequent 46 dengan *support* 33% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK N 1 Benai dengan jenis sekolah MA dan berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 81%.
14. Nilai A1,B36,B20,C2,D1 mempunyai nilai frequent 48 dengan *support* 35% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK N 1 Benai dan SMA N 1 Pangeandengan jenis sekolah SMA dan berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 78%.
15. Nilai A1,B36,B18,C2,D1 mempunyai nilai frequent 44 dengan *support* 32% Artinya Argibisnis akan dipilih oleh mahasiswa yang berasal dari SMK N 1 Benai dan SMA N 1 KuantanMudik dengan jenis sekolah SMA dan berjenis kelamin laki - laki serta dengan tingkat *confidence* sebesar 85%.

## **5.9 Implementasi Sistem**

### **1. Halaman Login**

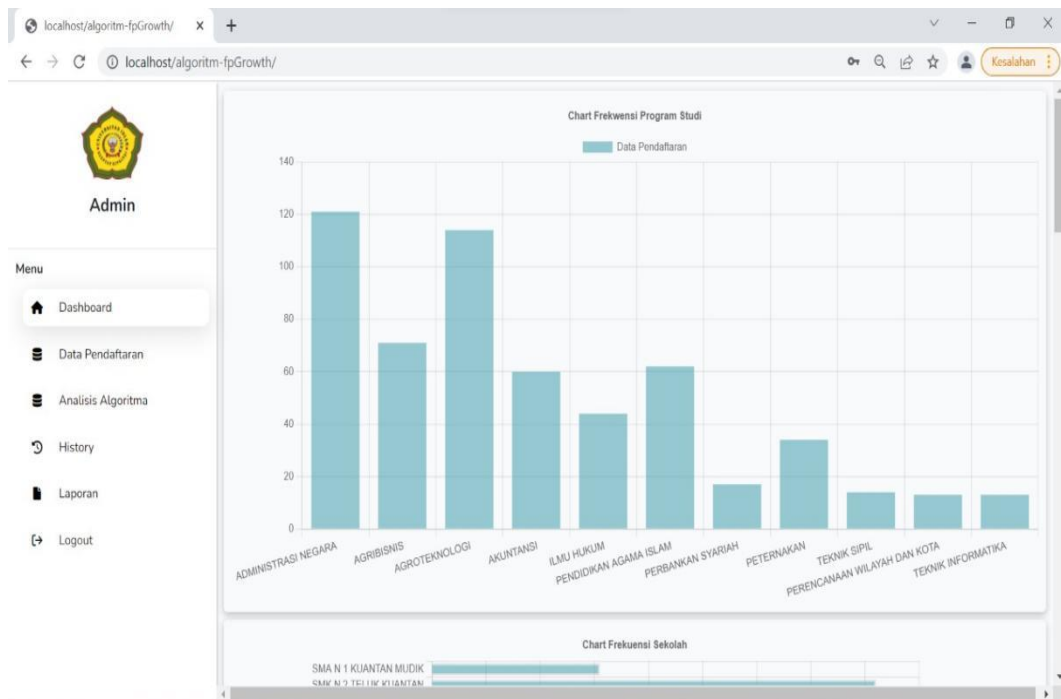
Halaman ini digunakan untuk menjaga agar *sistem* yang dibuat terjaga kerahasiaannya yang hanya dapat diakses admin yang memiliki hak akses pada sistem tersebut.



**Gambar 5.4 Halaman *Login***

## 2. Halaman *Dashboard*

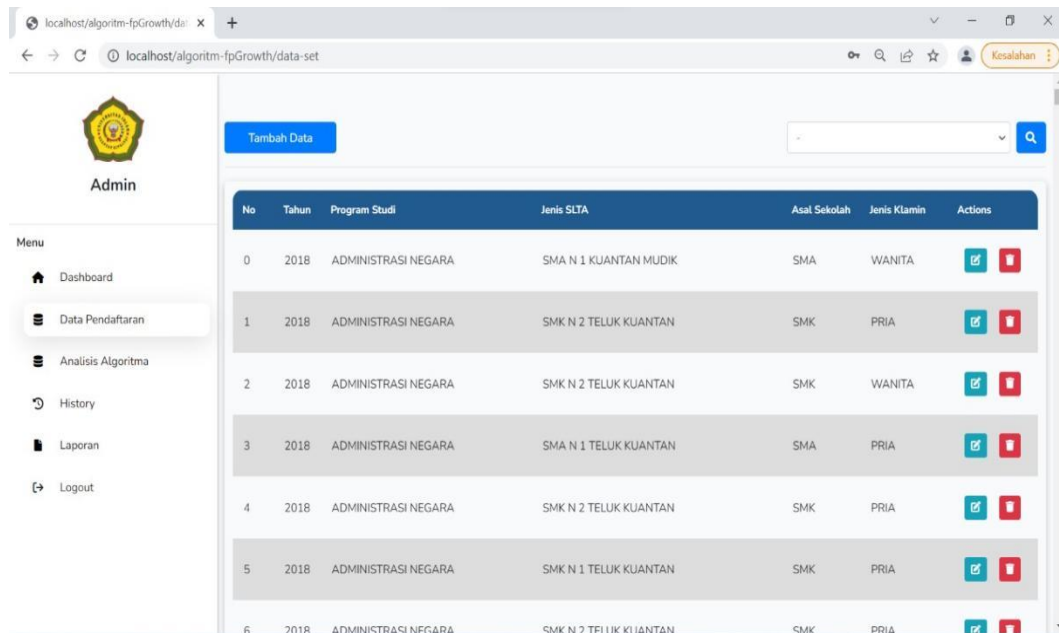
Halaman ini akan muncul untuk pertama kali setelah admin login ke sistem yaitu halaman *dashboard*. Berikut tampilan *dashboard* :



**Gambar 5.5 Halaman *Dashboard***

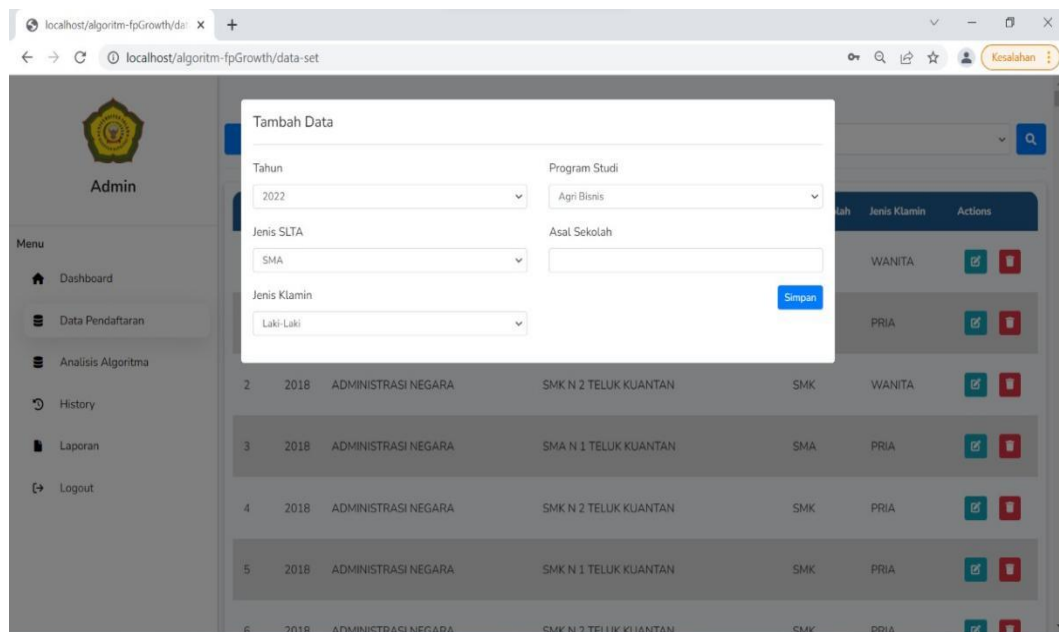
### 3. Halaman Data Pendaftaran

Berikut ini adalah halaman data pendaftaran, dimana halaman ini admin dapat menambahkan data baru.



**Gambar 5.6 Halaman Data Pendaftaran**

Nantinya jika admin mengklik tambah data untuk memasukkan data pendaftaran baru untuk di asosiasikan maka tampilan yang akan muncul



**Gambar 5.7 Halaman Tambah Data**

#### 4. Halaman Analisis Algoritma

Berikut ini adalah halaman Analisis Algoritma

Option Data  
Tahun

No	Tahun	Program Studi	Jenis SLTA	Asal Sekolah	Jenis Klamir
0		ADMINISTRASI NEGARA	SMA N 1 KUANTAN MUDIK	SMA	WANITA
1		ADMINISTRASI NEGARA	SMK N 2 TELUK KUANTAN	SMK	PRIA
2		ADMINISTRASI NEGARA	SMK N 2 TELUK KUANTAN	SMK	WANITA
3		ADMINISTRASI NEGARA	SMA N 1 TELUK KUANTAN	SMA	PRIA
4		ADMINISTRASI NEGARA	SMK N 2 TELUK KUANTAN	SMK	PRIA
5		ADMINISTRASI NEGARA	SMK N 1 TELUK KUANTAN	SMK	PRIA
6		ADMINISTRASI NEGARA	SMK N 2 TELUK KUANTAN	SMK	PRIA

Gambar 5.8 Halaman Analisis Algoritma

#### 5. Halaman Frequeen Item

Berikut ini adalah halaman *Frequeen item* di menu Analisis Algoritma

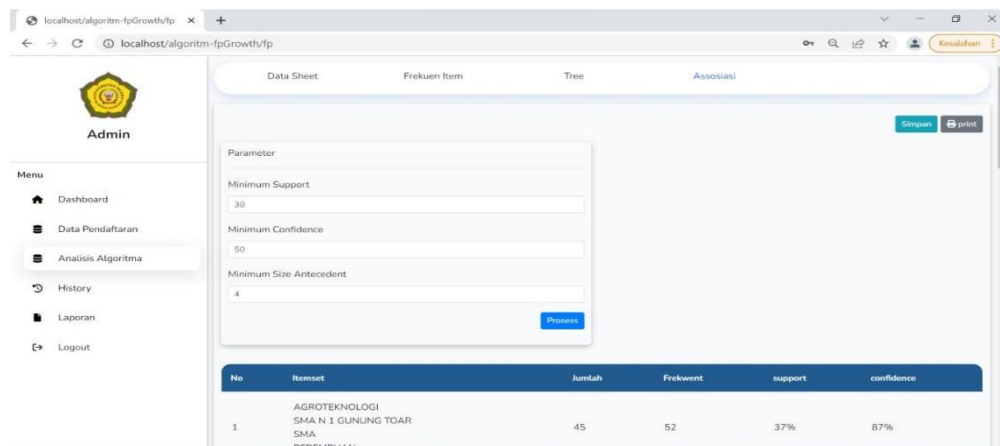
Minimal Frekwensi kemunculan 2 kali

No	keyword	kode	Frekwensi	Total	Support
0	ADMINISTRASI NEGARA	A1	73	257	29 %
1	AGRIBISNIS	A2	23	257	9 %
2	AGROTEKNOLOGI	A3	42	257	16 %
3	AKUNTANSI	A4	34	257	13 %
4	ILMU HUKUM	A5	29	257	11 %
5	PENDIDIKAN AGAMA ISLAM	A6	56	257	22 %
6	MAN 2 KUANSING	B15	3	257	1 %

Gambar 5.9 Halaman *Frequen Item*

## 6. Halaman Assosiasi

Berikut ini adalah halaman assosiasi dimana di halaman ini terdapat perhitungan menggunakan metode *Fp-Growth* sesuai kebutuhan dan di halaman ini juga tempat untuk melihat acuan dalam membuat strategi.



**Gambar 5.10 Halaman Assosiasi**

Setelah gambar 5.10 selesai kemudian klik proses maka akan muncul halaman hasil dari proses metode *assosiasi* seperti dibawah ini. Di halaman ini lah admin dapat membaca data yang sudah diolah dengan metode *fp-Growth*. Dimana hasil yang di dapat dari analisa ini dapat mudah dibaca oleh admin dengan hanya melihat bobotnya saja. Dimana adapun nilai bobot yang ditentukan adalah :

- 0 - 30 : Tidak Rekomendasi
- 31 - 60 : Sedang
- 61 - 80 : Baik
- 81 - 100 : Sangat Baik







No	Itemset	Jumlah	Frekwent	support	confidence	Bobot
1	AGROTEKNOLOGI SMA N 1 GUNUNG TOAR SMA PEREMPUAN	45	52	37%	87%	Sangat Baik
2	AGROTEKNOLOGI SMA N 1 KUANTAN MUDIK SMA PEREMPUAN	45	58	42%	78%	Baik
3	AGROTEKNOLOGI SMA N 1 BENAI SMA PEREMPUAN	45	48	35%	94%	Sangat Baik
4	AGRIBISNIS SMK N 2 TELUK KUANTAN SMK PEREMPUAN	37	65	47%	57%	Sedang

**Gambar 5.11 Halaman Hasil Assosiasi**

## 7. Halaman History

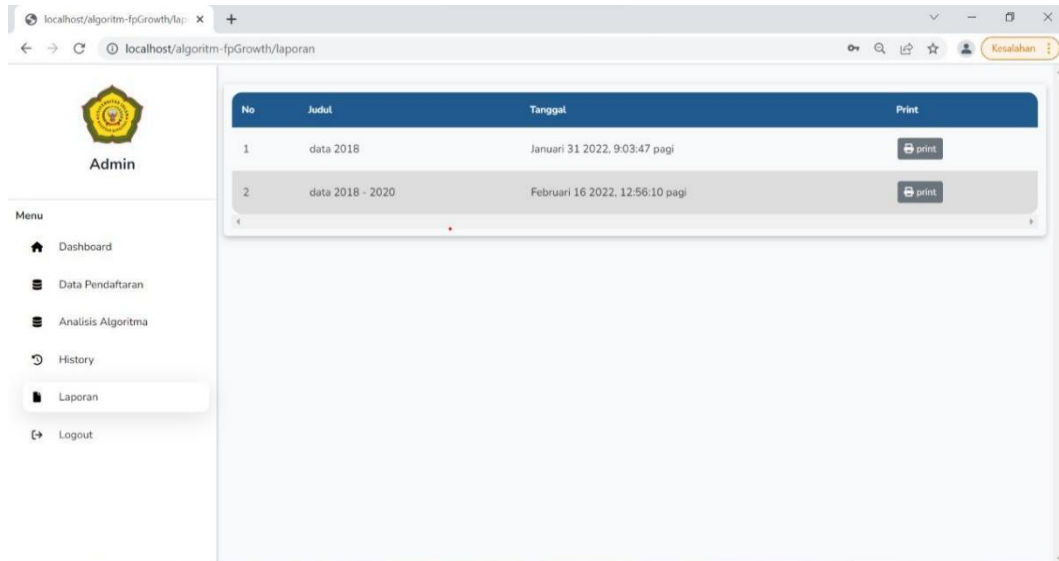
Berikut ini adalah halaman *history* yang mencatat dan menyimpan semua data yang sudah melakukan perumusan *Fp-Growth*.

No	Judul	Deskripsi	Tanggal	Actions
1	data 2018	analisis data 2018	Januari 31 2022, 9:03:47 pagi	 
2	data 2018 - 2020	data analisis 2018 sampai dengan 2020	Februari 16 2022, 12:56:10 pagi	 

**Gambar 5.12 Halaman History**

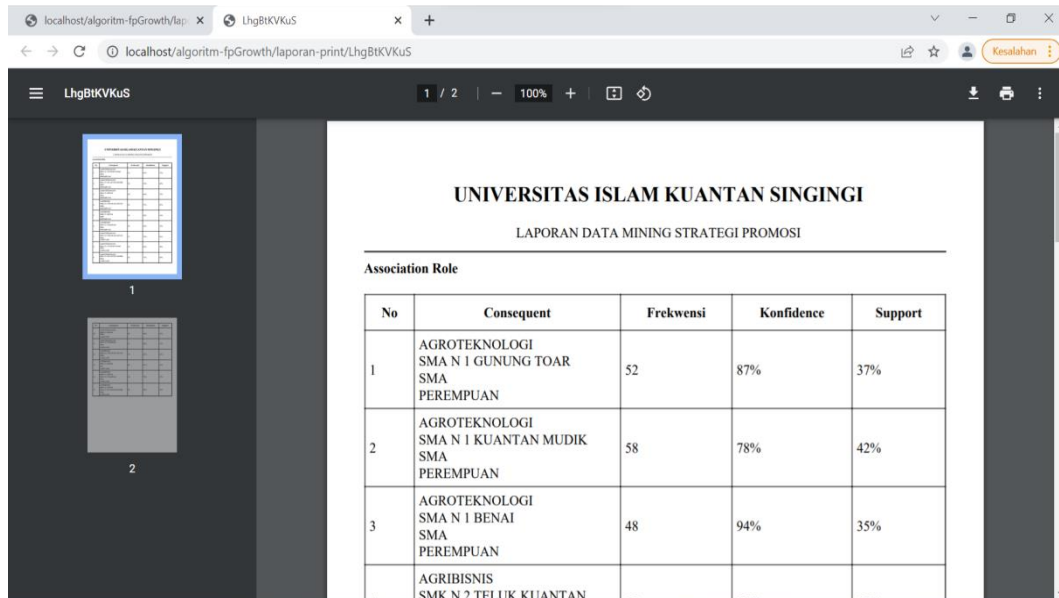
## 8. Halaman Laporan

Berikut ini adalah Halaman untuk membuat laporan.



Gambar 5.13 Halaman Laporan

Kemudian jika mengklik tombol print maka akan muncul tampilan seperti ini



Gambar 5.14 Halaman Melihat *Print* Laporan

## BAB VI

### KESIMPULAN DAN SARAN

#### 6.1 Kesimpulan

Dari hasil penelitian pembuatan sistem Data *Mining* dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi) maka dapat di ambil beberapa kesimpulan sebagai berikut :

1. Sistem ini dapat mempermudah memahami strategi dalam meningkatkan minat pendaftaran ke Universitas Islam Kuantan Singingi ini.
2. Dengan adanya sistem ini akan dapat membantu mempercepat pembuatan laporan yang dibutuhkan oleh instansi, dan termasuk setiap laporan tersimpan secara terkomputerisasi.

#### 6.2 Saran

Berdasarkan evaluasi terhadap proses dan hasil Sistem Data *Mining* dengan Algoritma *FP-Growth* untuk Mendukung Strategi Promosi Pendidikan (Studi Kasus : Universitas Islam Kuantan Singingi) Penulis Mengemukakan beberapa saran yang diharapkan dapat menjadikan bahan pertimbangan :

1. Diharapkan pengembangan selanjutnya dapat membuat sistem yang lebih dari yang penulis buat.
2. Dengan adanya penggantian sistem diharapkan UNIKS mampu mengoperasikan UNIKS yang diusulkan dalam pembuatan laporan.

## DAFTAR PUSTAKA

- Al-Hafiz, N. W., Mesran, & Suginam. (2017). Sistem Pendukung Keputusan Penentuan Kredit Pemilikan Rumah Menerapkan Multi-Objective Optimization on the Basis of Ratio Analysis ( Moora ). *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 1(1), 306–309. <http://www.stmik-budidarma.ac.id/ejurnal/index.php/komik/article/viewFile/513/455>
- Budianto, H., & Riana, J. (2020). Penerapan Data Mining menggunakan algoritma Fp-Growth Untuk Penentuan Strategi Promosi Fakultas Ilmu Komputer Universitas Kuningan. *Jurnal Cloud Information*, 5(1), 22–29.
- Ikhwan, A., Nofriansyah, D., & Sriani. (2015). Penerapan Data Mining dengan Algoritma Fp-Growth untuk Mendukung Strategi Promosi Pendidikan ( Studi Kasus Kampus STMIK Triguna Dharma ). *Saintikom*, 14(3), 211–226.
- Irwan. (2019). Strategi Promosi Dengan Algoritma Fp-Growth (Studi Kasus Di Universitas Harapan Medan). *Information Technology and Computer Science (INTECOMS)*, 3, 76–81.
- Mayasari, N., Niska, D. Y., & Samosir, E. R. (2019). Implementasi Data Mining untuk Memprediksi Itemset Promosi Penjualan Pada CV. Sumber Segar Utama. *Jurnal Teknik Dan Informatika Volume 6 | Nomor 1 | Januari 2019*, 6, 31–36.
- Parmonangan Hutabarat, C. (2021). Penerapan Data Mining Association Rule Menggunakan Algoritma FP-Growth Untuk Persediaan Sparepart pada Bengkel. 5(2), 112–121. <https://doi.org/10.31603/komtika.v5i2.6251>
- Sepri, D., Afdal, M., & Riau, S. (2017). Analisa Dan Perbandingan Metode Algoritma Apriori Dan Fp-Growth Untuk Mencari Pola Daerah Strategis Pengenalan Kampus Studi Kasus Di Stkip Adzkie Padang. *Jurnal Sistem Informasi Kaputama (JSIK)*, 1(1). <https://www.jurnal.kaputama.ac.id/index.php/JSIK/article/view/27>
- Tjandra, S. (2019). Analisa Pencarian Frequent Itemsets Menggunakan Algoritma Fp-Max. *Jurnal Teknik Informatika Sekolah Tinggi Teknik Surabaya*. 1-7
- Rosnelly, R. (2012). Buku Sistem Pakar. *Buku Sistem Pakar*.
- Winarti, D., Studi, P., Informatika, T., Komputer, F. I., Indonesia, U. D., Barat, S., & Asosiasi, R. (2018). Analisis Data Mining Dengan Algoritma Fp-Growth Dalam Mendukung Strategi Promosi. *Simtika*, 1(1), 27–31.

## DOKUMENTASI PENELITIAN

---



Wawancara dengan staf di lingkungan Universitas Islam Kuantan Singingi



Proses pengambilan data kepada salah seorang staf di Universitas Islam Kuantan Singingi



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM KUANTAN SINGINGI**  
ot Subroto KM. 7 Kebun Nenas, Desa Jake, Kec. Kuantan Tengah

**KARTU BIMBINGAN SKRIPSI**

N P M : 150210134  
Nama : DESPIK MICE  
Pembimbing 1 : FEBRI HASWAN, M.Kom  
Pembimbing 2 : NOFRI WANDI AL-HAFIZ, M.Kom  
Judul: : **Implementasi Data Mining Dengan Algoritma *Fp-Growth*  
Untuk Mendukung Strategi Promosi Pendidikan (Studi  
Kasus : Universitas Islam Kuantan Singingi)**



NO	TANGGAL	KOMENTAR PEMBIMBING	PARAF

....., .....

Pembimbing 1

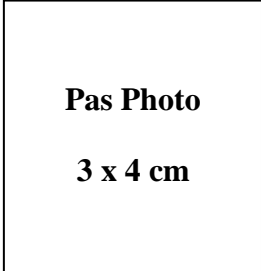
**FEBRI HASWAN, S.Kom., M.Kom**  
**NIDN. 1009028803**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS ISLAM KUANTAN SINGINGI**  
ot Subroto KM. 7 Kebun Nenas, Desa Jake, Kec. Kuantan Tengah

**KARTU BIMBINGAN SKRIPSI**

N P M : 150210134  
Nama : DESPIK MICE  
Pembimbing 1 : FEBRI HASWAN, M.Kom  
Pembimbing 2 : NOFRI WANDI AL-HAFIZ, M.Kom  
Judul: : **Implementasi Data Mining Dengan Algoritma *Fp-Growth*  
Untuk Mendukung Strategi Promosi Pendidikan (Studi  
Kasus : Universitas Islam Kuantan Singingi)**



NO	TANGGAL	KOMENTAR PEMBIMBING	PARAF

.....  
Pembimbing 2

**NOFRI WANDI AL-HAFIZ, S.Kom., M.Kom  
NIDN. 1002118802**