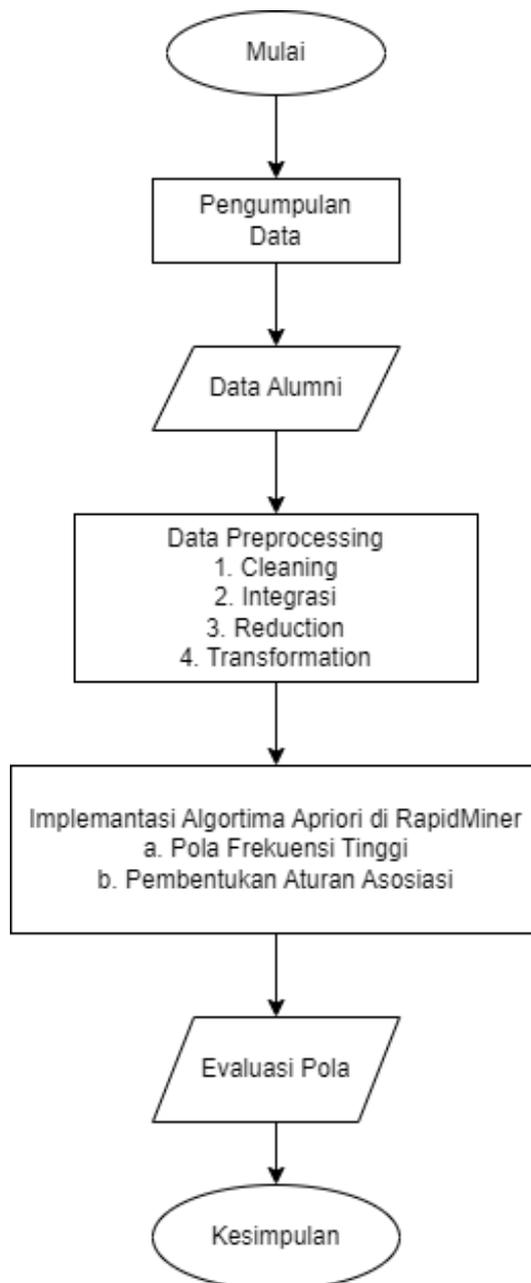


### BAB III METODOLOGI PENELITIAN

Dalam melakukan analisis penulis menggunakan *software* Rapid Miner. data yang digunakan dalam penelitian ini adalah data diri mahasiswa IIB Darmajaya jurusan Sistem Informasi. Data yang digunakan adalah data alumni Angkatan 2003 sampai 2017.

*Flowchart* tahapan penelitian dapat dilihat pada gambar 2.



Gambar 3.1 Alur Penelitian

### **3.1 Metode Pengumpulan Data**

Tugas akhir ini mengumpulkan dua jenis data, yaitu data primer dan data sekunder. Data primer adalah data yang diperoleh secara langsung dari objek penelitian, dalam hal ini adalah data alumni dari Institut Informatika dan Bisnis Darmajaya. Sedangkan data sekunder adalah data yang diperoleh melalui referensi buku, seperti skripsi, sumber dari media internet, jurnal, serta buku referensi yang terkait dengan topik penelitian.

Dalam proses pengumpulan data, digunakan metode sebagai berikut:

#### **3.1.1 Observasi**

Peneliti melakukan pengamatan langsung di Institut Informatika dan Bisnis Darmajaya guna mempelajari permasalahan yang ada. Tujuan dari observasi ini adalah untuk mendapatkan pemahaman yang lebih mendalam mengenai topik penelitian serta meningkatkan akurasi data yang diperoleh melalui wawancara. Dengan melakukan observasi langsung, penelitian dapat dengan lebih mudah mengumpulkan data dan fakta yang diinginkan dengan tingkat keakuratan yang tinggi.

#### **3.1.2 Wawancara**

Metode yang digunakan adalah dengan melakukan diskusi, tanya jawab dengan narasumber yang dianggap memiliki pengetahuan yang mendalam tentang masalah penelitian, dengan tujuan untuk mengumpulkan data primer.

#### **3.1.3 Studi Pustaka**

Pengumpulan data dilakukan melalui studi literatur dan membaca berbagai sumber, seperti buku, internet, dan penelitian sebelumnya, untuk mendapatkan dasar teori, pemikiran, dan data yang relevan. Metode analisis data yang digunakan dalam penelitian ini menggunakan algoritma apriori yang termasuk dalam tipe aturan asosiasi atau association rule dalam data mining. Algoritma apriori digunakan untuk mengolah data mentah dan mentransformasikannya menjadi pola data yang memberikan informasi yang dapat mendukung pengambilan keputusan di Perguruan Tinggi Institut Informatika dan Bisnis Darmajaya.

### 3.2 Metode Penelitian

Dalam penelitian ini, diperlukan data alumni dari Institut Informatika dan Bisnis Darmajaya dengan rentang periode lulus tahun 2003 hingga 2017. Data ini akan dianalisis menggunakan algoritma apriori dan merupakan data primer yang digunakan dalam penelitian ini.

### 3.3 Data Alumni

Pengumpulan data mengenai alumni Institut Informatika dan Bisnis Darmajaya dilakukan selama periode 1 bulan. Data yang diteliti meliputi alumni yang lulus dalam rentang tahun 2003 hingga 2017, dengan total durasi penelitian selama 15 tahun.

Tabel 3.1 Data asli alumni

Tahun Lulus	Jumlah
2003	830 Alumni
2004	612 Alumni
2005	469 Alumni
2006	411 Alumni
2007	498 Alumni
2008	648 Alumni
2009	584 Alumni
2010	529 Alumni
2011	535 Alumni
2012	725 Alumni
2013	699 Alumni
2014	603 Alumni
2015	721 Alumni
2016	848 Alumni
2017	262 Alumni
<b>TOTAL</b>	<b>8975 Alumni</b>

### 3.4 Tahapan Preprocessing

Data alumni yang diperoleh dari BAAK mencakup periode tahun 2000-2021. Namun, dalam data tersebut masih terdapat duplikasi dan nilai yang kosong. Oleh karena itu, diperlukan proses pembersihan data (*data cleaning*) dan penanganan nilai yang hilang (*missing value*) untuk mendapatkan data yang baik dan bersih.

#### 3.4.1 Cleaning

Berdasarkan data asli yang diperoleh, dilakukan tahapan cleaning yang didapatkan total 336 data alumni.

#### 3.4.2 Missing Value

Sebelumnya sudah dilakukan tahapan *cleaning*, selanjutnya dilakukan tahapan *missing value* yang dimana membersihkan data asli dari data yang rusak, kosong dimana didapatkan total 18 data alumni *missing value*.

Data asli yang sudah di *cleaning* dan *missing value*, sebelumnya data asli yang sudah didapat dari tahun 2000-2021 terdapat 10676 data alumni. Setelah proses pembersihan data (*cleaning*) dan penanganan nilai yang hilang (*missing value*), ditemukan bahwa jumlah data alumni pada periode tahun 2003-2017 adalah sebanyak 8975 orang/data.

### 3.5 Tahapan Reduction

Atribut data ditentukan berdasarkan kebutuhan, Terdapat 10 Atribut dari data asli yang didapatkan dari BAAK yaitu:

Tabel 3.2 Atribut data asli

NO	Atribut
1	NPM
2	Tanggal Lahir
3	Jenis kelamin
4	Tanggal yudisium
5	Prodi
6	IPK
7	Lama studi
8	Nilai Test

<b>9</b>	<b>Asal Sekolah</b>
<b>10</b>	<b>Alamat Sekolah</b>

Setelah itu Atribut data asli dilakukan reduction untuk mendapatkan Atribut Terpilih, agar dapat digunakan dengan mudah penulis membuat tabel sebagai berikut:

Tabel 3.3 Atribut data terpilih

<b>NO</b>	<b>Atribut</b>
<b>1</b>	<b>Fakultas</b>
<b>2</b>	<b>IPK</b>
<b>3</b>	<b>Jenis Kelamin</b>
<b>4</b>	<b>Lama Studi</b>
<b>5</b>	<b>Status Kelulusan</b>
<b>6</b>	<b>Usia</b>
<b>7</b>	<b>Nilai Test</b>
<b>8</b>	<b>Asal Sekolah</b>
<b>9</b>	<b>Alamat Sekolah</b>

### 3.6 Transformation

Tabel atribut Tahun Lulus, NPM, dan Tanggal lahir dilakukan transformation untuk mendapatkan atribut terpilih Masa Studi, dan Usia.

Tabel 3.4 Atribut Data

<b>NO</b>	<b>Atribut</b>	<b>Atribut Transformation</b>	
<b>1</b>	<b>NPM</b>	<b>Lama Studi</b>	
<b>2</b>	<b>Tahun Lulus</b>		<b>Usia</b>
<b>3</b>	<b>Tanggal Lahir</b>		

Pada Tabel 3.2 ditunjukkan bahwa ada 6 atribut yang digunakan untuk penentuan data seperti Fakultas, IPK, Jenis kelamin, Lama studi, Status kelulusan, Usia. Berikut adalah atribut yang digunakan untuk mencari pola aturan yang mungkin tersembunyi, dengan harapan dapat digunakan untuk pengambilan keputusan oleh pihak terkait atau berkepentingan.

Tabel 3.5 Usia

No	Usia	Ket
1	21-25 Tahun	1
2	>25 Tahun	0

Tabel Usia memiliki 2 kategori yaitu 21-25 dan >25 Tahun, untuk usia awal masuk perkuliahan semua di umur 18 Tahun dan memiliki umur yang berbeda saat mencapai kelulusan.

Tabel 3.6 Jenis Kelamin

No	Jenis Kelamin	Ket
1	Perempuan	1
2	Laki-laki	0

Tabel Jenis Kelamin memiliki 2 kategori yaitu perempuan dan laki-laki.

Tabel 3.7 Fakultas

No	Prodi	Fakultas	Ket
1	T. Informatika	Ilmu Komputer	0
2	Sistem Informasi		
3	Sistem Komputer		
4	Management	Ekonomi & Bisnis	1
5	Akuntansi		

Pada tabel 3.6 yaitu Atribut Fakultas terdapat beberapa prodi, pada fakultas ekonomi dan bisnis ada prodi Manajemen dan Akuntansi sedangkan fakultas Ilmu Komputer ada Teknik Informatika, Sistem Informasi dan Sistem Komputer.

Tabel 3.8 IPK

No	IPK	Ket
1	0 – 2.75	0
2	2.76 - 4	1

Pada tabel 3.7 dijelaskan bahwa IPK atau indeks prestasi kumulatif yaitu 0 - 2.75 kemudian ada diatas 2.76 - 4.

Tabel 3.9 Lama Studi

No	Lama Studi	Ket
1	3.5 – 4 Tahun	1
2	>4 Tahun	0

Pada tabel 3.8 terdapat lama studi yaitu, lulus tepat waktu, bila lama studi 3.5 sampai 4 Tahun dan lulus tidak tepat waktu, bila lama studi >4 tahun, untuk menentukan Lama studi penulis menggunakan data NPM dan Tahun Lulus.

Tabel 3.10 Status Kelulusan

No	Status Kelulusan	Lama Studi	Ket
1	Tepat	3.5 – 4 Tahun	1
2	Tidak Tepat	>4 Tahun	0

Tabel status kelulusan memiliki 2 yaitu tepat dan tidak tepat, untuk menentukan Status kelulusan ini menggunakan atribut Lama studi.

Tabel 3.11 Asal Sekolah

No	Asal Sekolah	Ket
1	Negeri	0
2	Swasta	1

Tabel Asal sekolah memiliki 2 keterangan yaitu Negeri & Swasta, untuk menentukan dari mana asal sekolah mahasiswa ini.

Tabel 3.12 Alamat Sekolah

No	Alamat Sekolah	Ket
1	Bandar Lampung	0
2	Luar Bandar Lampung	1

Tabel Alamat sekolah, digunakan untuk memberitahukan bahwa Alamat domisili dari sekolah tersebut.

Tabel 3.13 Nilai Test

No	Nilai Test	Ket
1	0-50	0

2	51-100	1
---	--------	---

Tabel Nilai test digunakan untuk menentukan nilai saat dia mendaftar di Universitas IIB Darmajaya.

### 3.7 Tabulasi Data

1	usia mahasiswa	jenis kelamin	Fakultas	IPK	Lama studi (Tahun)	status kelulusan	Asal sekolah	Alamat Sekolah	Nilai Test
2	0	0	1	1	0	0	1	0	0
3	0	0	1	0	0	0	1	0	0
4	0	0	1	1	0	0	0	0	1
5	0	0	1	0	0	0	0	0	0
6	0	0	1	0	0	0	0	1	0
7	0	0	1	0	0	0	0	1	0
8	0	0	1	0	0	0	0	0	1
9	0	0	1	0	0	0	1	0	1
10	0	0	1	1	1	1	0	0	0
11	0	0	1	1	0	0	0	0	1
12	0	0	1	0	0	0	0	0	0
13	0	0	1	0	0	0	0	0	0
14	0	1	1	1	1	1	1	0	1
15	0	1	0	0	1	1	0	0	0
16	0	0	1	1	0	0	0	0	1
17	0	0	1	1	0	0	0	0	0
18	0	1	0	0	1	1	0	0	0
19	0	0	1	1	0	0	0	0	0
20	0	0	1	1	0	0	0	0	0
21	0	0	1	1	0	0	0	0	0
22	0	0	1	0	0	0	0	0	0
23	0	0	1	0	1	1	0	0	0
24	0	0	0	0	1	1	0	0	0
25	0	0	1	1	0	0	0	0	0

Gambar 3.2 Format Tabular Data

Gambar tabel 3.2 menampilkan 25 contoh tabular data yang terdiri dari 9 Atribut.

### 3.8 Penentuan Nilai Support dan Confidence

hasil perhitungan nilai *Support* diubah menjadi persentase untuk menentukan nilai *Confidence* dari masing-masing item. Dalam penelitian ini, nilai minimum support untuk pola frekuensi tertinggi adalah 30%.

#### 3.8.1 Kombinasi *itemset* 1

Pembentukan itemset C1 atau 1 itemset, penyelesaian ini berdasarkan data pada tabel diatas

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

### 3.8.2 Kombinasi 2 *itemset*

Berikut adalah proses pembentukan 2-*itemset* berdasarkan data yang terdapat pada Tabel 3.10, dengan menggunakan rumus sebagai berikut:

$$\text{Support (A, B)} = \sum \frac{\text{jumlah transaksi mengandung A dan B}}{\text{Total transaksi}}$$

### 3.8.3 Kombinasi 3 *itemset*

Proses pembentukan 3 *itemset* dengan jumlah *minimum support* = 30% dapat diselesaikan dengan rumus berikut:

$$\text{Support (A, B)} = \sum \frac{\text{jumlah transaksi mengandung A,B dan C}}{\text{Total transaksi}}$$

### 3.8.4 Pembentukan aturan *assosiasi*

Setelah semua pola frekuensi tinggi ditemukan, langkah selanjutnya adalah mencari aturan asosiasi yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* dari aturan asosiasi A>B.

Dalam penelitian ini, syarat minimum untuk *confidence* adalah sebesar 70%.

Nilai *confidence* dari aturan A>B dapat dihitung sebagai berikut:

$$\text{Support P(B|A)} = \sum \frac{\text{transaksi mengandung A dan B}}{\text{Transaksi Mengandung A}}$$

## 3.9 Penjelasan hasil Data

Sebelum diimplementasikan dengan bantuan alat bantu RapidMiner, terdapat beberapa langkah-langkah dalam metode Apriori yang perlu dilakukan, antara lain:

a. Ekstraksi Data

Variabel data-data yang digunakan ada 9 seperti

Tabel 3.14 Variabel Terpilih

NO	Atribut
1	Fakultas
2	IPK
3	Jenis Kelamin
4	Lama Studi

<b>5</b>	<b>Status Kelulusan</b>
<b>6</b>	<b>Usia</b>
<b>7</b>	<b>Nilai Test</b>
<b>8</b>	<b>Asal Sekolah</b>
<b>9</b>	<b>Alamat Sekolah</b>

b. Tahapan Preprocessing

Dimana terdapat 4 tahapan:

1. *Cleaning* yaitu melakukan pembersihan terhadap data alumni yang digunakan
2. *Missing value* yaitu memilah data alumni yang kosong, rusak yang akan digunakan
3. *Reduction* yang dimana menentukan atribut data asli yang sudah di dapatkan dari BAAK
4. *Transformation* yaitu merubah beberapa data atribut yang akan digunakan untuk menentukan data atribut terpilih.

c. Menentukan Nilai *Minimum Support dan Confidence*

Dalam menentukan nilai *minimum support* dan *minimum confidence*, prosesnya dilakukan secara manual pada operator *FP-growth* dan *Create Association Rules* dengan menggunakan pecahan desimal.