

**POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN  
MAHASISWA MENGGUNAKAN ALGORITMA *FOLD-  
GROWTH* (STUDI KASUS JURUSAN TEKNIK  
INFORMATIKA UIN MAULANA MALIK  
IBRAHIM MALANG)**

**SKRIPSI**

**Oleh :  
ANIS SHIVA 'ULIA DEWI  
NIM. 16650035**



**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2021**

**POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN  
MAHASISWA MENGGUNAKAN ALGORITMA *FOLD-  
GROWTH* (STUDI KASUS JURUSAN TEKNIK  
INFORMATIKA UIN MAULANA MALIK  
IBRAHIM MALANG)**

**SKRIPSI**

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

**oleh :  
ANIS SHIVA 'ULIA DEWI  
NIM. 16650035**

**JURUSAN TEKNIK INFORMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2021**

**LEMBAR PERSETUJUAN**

**POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN  
MAHASISWA MENGGUNAKAN ALGORITMA *FOLD-  
GROWTH* (STUDI KASUS JURUSAN TEKNIK  
INFORMATIKA UIN MAULANA MALIK  
IBRAHIM MALANG)**

**SKRIPSI**

**ANIS SHIVA 'ULIA DEWI  
NIM. 16650035**

Telah Diperiksa dan Disetujui untuk Diuji

Tanggal : 08 Desember 2021

Dosen Pembimbing I



Dr. Muhammad Faisal, M.T  
NIP. 19740510 200501 1 007

Dosen Pembimbing II



M. Ainul Yaqin, M.Kom  
NIP. 19761013 200604 1 004

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

Dr. Fachrul Kurniawan, M.MT  
NIP. 19771020 200912 1 001

**LEMBAR PENGESAHAN**

**POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN  
MAHASISWA MENGGUNAKAN ALGORITMA *FOLD-  
GROWTH* (STUDI KASUS JURUSAN TEKNIK  
INFORMATIKA UIN MAULANA MALIK  
IBRAHIM MALANG)**

**SKRIPSI**

**Oleh:**

**ANIS SHIVA 'ULIA DEWI  
NIM. 16650035**

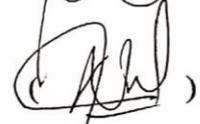
Telah Dipertahankan Di depan Dewan Penguji  
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)  
Pada Tanggal 08 Desember 2021

**Susunan Penguji**

- |                       |   |  |
|-----------------------|---|--|
| 1. Penguji Utama      | : | <u>Prof. Dr Suhartono S.Si M.Kom</u><br>NIP. 19680519 200312 1 001 |
| 2. Ketua Penguji      | : | <u>Fresy Nugroho, M.T</u><br>NIP. 19710722 201101 1 001            |
| 3. Sekretaris Penguji | : | <u>Dr. Muhammad Faisal, M.T</u><br>NIP. 19740510 200501 1 007      |
| 4. Anggota Penguji    | : | <u>M. Ainul Yaqin, M.Kom</u><br>NIP. 19761013 200604 1 004         |

**Tanda tangan**

()

()

()

()

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

Dr. Fachrul Kurniawan, M.MT  
NIP. 19771020 200912 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Anis Shiva 'Ulia Dewi  
NIM : 16650035  
Fakultas/jurusan : Sains dan Teknologi/Teknik Informatika  
Judul Skripsi : Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa Menggunakan Algoritma *FOLD-Growth* (Studi Kasus Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang)

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan pengambilalihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 08 Desember 2021  
Yang membuat pernyataan,



**Anis Shiva 'Ulia Dewi**  
**NIM. 16650035**

## **HALAMAN MOTTO**

“Maju terus pantang mundur”

## HALAMAN PERSEMBAHAN

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ

**Skripsi ini telah selesai meskipun tidak tepat waktu, namun cukup melegakan bagi diri saya. Penulis persembahkan skripsi ini kepada:**

Saya sendiri karena bisa setangguh ini dan tak ada hentinya memberikan dorongan lahir batin untuk tetap semangat dalam menyelesaikan skripsi sampai di titik pencapaian sekarang ini.

Kedua orang tua tercinta dan tersayang, Bapak Mashuri dan Ibu Kasemi Nur Mufidah, dan kedua adik saya, maupun keluarga besar lainnya yang selalu memanjatkan do'a dan memberikan motivasi untuk segera menyelesaikan skripsi ini.

Dosen pembimbing, Bapak Dr. Muhammad Faisal, M. T dan Bapak M. Ainul Yaqin, M. Kom, yang telah memberikan bimbingan dan saran sehingga penelitian ini bisa berjalan dengan lancar.

Seluruh dosen Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang, yang telah dengan ikhlas membagikan ilmu dan berbagai pengalamannya.

Support system yang telah banyak membantu dalam menyelesaikan skripsi ini.

Sahabat-sahabat saya yang begitu banyak dan tidak bisa disebutkan namanya satu per satu, saudara Andromeda Teknik Informatika 2016, keluarga besar Teknik Informatika UIN Maulana Malik Ibrahim Malang, yang tiada hentinya memberikan semangat untuk menyelesaikan skripsi ini.

## KATA PENGANTAR

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ

Puji syukur penulis panjatkan kehadiran Tuhan yang Maha Esa yang telah memberikan berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa Menggunakan Algoritma FOLD-Growth (Studi Kasus Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang)”. Tujuan dari penyusunan skripsi ini guna memenuhi salah satu syarat untuk bisa menempuh ujian sarjana komputer pada Fakultas Sains dan Teknologi Program Studi Teknik Informatika di Universitas Islam Negeri Maulana Malik Ibrahim Malang. Dalam penyusunan skripsi penulis banyak mendapatkan bantuan ataupun masukan dari berbagai pihak. Oleh sebab itu, penulis mengungkapkan rasa terima kasih kepada:

1. Bapak Prof. Dr. M. Zainuddin, MA, selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Ibu Dr. Sri Harini, M.Si, selaku Dekan Fakultas Sains dan Teknologi.
3. Bapak Dr. Fachrul Kurniawan, M.MT, selaku Ketua jurusan Teknik Informatika, Fakultas Sains dan Teknologi UIN Maliki Malang.
4. Bapak Dr. Muhammad Faisal, M. T, selaku Wali Dosen dan Pembimbing I, yang telah memberikan bimbingan dan masukan sehingga penulis dapat menyelesaikan perkuliahan dan penyusunan skripsi ini dengan baik.
5. Bapak M. Ainul Yaqin, M. Kom, selaku pembimbing II, yang telah memberikan bimbingan dan masukan sehingga penulis dapat menyelesaikan penyusunan skripsi ini dengan baik.
6. Segenap bapak dan ibu dosen yang telah mendidik dan mengajarkan ilmunya selama penulis menempuh pendidikan di Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Ibrahim Malang.
7. Mbak Citra Fidyia Atmalia, S.H dan para staff laboran, yang telah membantu dalam urusan administrasi selama perkuliahan dan penyusunan skripsi.

8. Bapak Imam Ahmad, M.Ag selaku Kabag. Akademik, Nuril Ma'arif, S.AP selaku Kasubbag. Informasi Akademik, dan Badrus Hidayat, S.Kom selaku Staf Informasi Akademik yang telah membantu penulis dalam urusan data penelitian dalam penyusunan skripsi.
9. Orang tua dan keluarga penulis yang telah memanjatkan do'a dan memberikan suntikan motivasi.
10. Sahabat-sahabat tercinta dan teman-teman yang sudah membantu dan memotivasi dalam segala urusan penyusunan skripsi ini.
11. Saudara Andromeda TI 16, kakak tingkat, dan adik tingkat Teknik Informatika UIN Maulana Malik Ibrahim Malang.
12. Semua pihak yang terlibat dalam penyusunan skripsi yang tidak bisa penulis sebutkan semuanya.

Penulis menyadari bahwa skripsi ini masih terdapat banyak kekurangan, namun penulis berharap skripsi ini dapat bermanfaat khususnya bagi penulis secara pribadi.

Malang, 08 Desember 2021

Penulis

## DAFTAR ISI

<b>LEMBAR PERSETUJUAN .....</b>	<b>ii</b>
<b>LEMBAR PENGESAHAN.....</b>	<b>iii</b>
<b>PERNYATAAN KEASLIAN TULISAN.....</b>	<b>iv</b>
<b>HALAMAN MOTTO .....</b>	<b>v</b>
<b>HALAMAN PERSEMBAHAN .....</b>	<b>vi</b>
<b>KATA PENGANTAR.....</b>	<b>vii</b>
<b>DAFTAR ISL.....</b>	<b>ix</b>
<b>DAFTAR GAMBAR .....</b>	<b>xi</b>
<b>DAFTAR TABEL.....</b>	<b>xiii</b>
<b>ABSTRAK.....</b>	<b>xiv</b>
<b>ABSTRACT.....</b>	<b>xv</b>
<b>المخلص.....</b>	<b>xvi</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	6
1.3 Tujuan Penelitian.....	6
1.4 Manfaat Penelitian.....	6
1.5 Batasan Masalah.....	6
<b>BAB II STUDI PUSTAKA .....</b>	<b>8</b>
2.1 <i>Data Mining</i> .....	8
2.2 Aturan Asosiasi ( <i>Association Rule</i> ).....	11
2.3 Algoritma <i>FOLDARM</i> .....	14
2.4 Algoritma <i>FP-Growth</i> .....	15
2.5 Algoritma <i>FOLD-Growth</i> .....	16
2.6 Struktur Data <i>SOTrieIT</i> .....	18
2.7 Lift Ratio.....	20
2.8 Penelitian Terkait.....	21
<b>BAB III METODE PENELITIAN.....</b>	<b>25</b>

3.1	Pengumpulan Data.....	26
3.1.1	Data Induk Mahasiswa.....	27
3.1.2	Data Kelulusan.....	27
3.2	Desain Sistem.....	29
3.2.1	Desain Sistem Bagian Input.....	30
3.2.2	Desain Sistem Bagian Proses.....	31
3.2.3	Desain Sistem Bagian Output.....	32
3.2.4	Perancangan Database.....	32
3.3	Perhitungan Manual.....	33
<b>BAB IV HASIL DAN PEMBAHASAN.....</b>		<b>61</b>
7.1	Implementasi.....	61
7.1.1	Implementasi Sistem.....	61
7.1.2	Implementasi Algoritma <i>FOLD-Growth</i> .....	62
7.1.3	Implementasi Interface.....	65
7.2	Pengujian.....	69
7.2.1	Pengujian Pengaruh <i>Minimum Support</i> dan <i>Minimum Confidence</i> .....	70
7.2.2	Pengujian <i>Lift Ratio</i> .....	74
7.3	Pembahasan.....	74
7.4	Integrasi Islam.....	81
<b>BAB V KESIMPULAN DAN SARAN.....</b>		<b>85</b>
8.1	Kesimpulan.....	85
8.2	Saran.....	86
<b>DAFTAR PUSTAKA.....</b>		<b>87</b>
<b>LAMPIRAN.....</b>		<b>90</b>

## DAFTAR GAMBAR

Gambar 2.1 Flowchart Algoritma FOLD-Growth .....	17
Gambar 2.2 Hasil SOTrieIT dari Tabel 2.1 .....	20
Gambar 3.1 Prosedur Perancangan Sistem.....	26
Gambar 3.2 Desain Sistem.....	30
Gambar 3.3 (ERD) Entity Relation Diagram.....	33
Gambar 3.4 Pembentukan 1-itemset pada SOTrieIT.....	37
Gambar 3.5 Pembentukan 2-itemset pada SOTrieIT.....	38
Gambar 3.6 Pembentukan SOTrieIT dan diurutkan dari support count terbesar..	38
Gambar 3.7 Struktur SOTrieIT yang telah dipangkas dan diurutkan .....	42
Gambar 3.8 FP-Tree Awal.....	44
Gambar 3.9 FP-Tree pada kandidat pertama .....	45
Gambar 3.10 FP-Tree pada kandidat kedua.....	45
Gambar 3.11 FP-Tree pada kandidat ketiga .....	46
Gambar 3.12 FP-Tree pada kandidat keempat.....	47
Gambar 3.13 FP-Tree pada kandidat kelima .....	47
Gambar 3.14 FP-Tree pada kandidat keenam.....	48
Gambar 3.15 FP-Tree pada kandidat ketujuh .....	49
Gambar 3.16 FP-Tree pada kandidat kedelapan .....	49
Gambar 3.17 Lintasan yang memiliki suffix {MANDIRI TERTULIS}.....	51
Gambar 3.18 Conditional FP-Tree yang terbentuk untuk suffix {MANDIRI TERTULIS} .....	51
Gambar 3.19 Lintasan yang memiliki suffix {SPAN-PTKIN}.....	52
Gambar 3.20 Conditional FP-Tree yang terbentuk untuk suffix {SPAN-PTKIN} .....	52
Gambar 3.21 Lintasan yang memiliki suffix {UM-PTKIN} .....	52
Gambar 3.22 Conditional FP-Tree yang terbentuk untuk suffix {UM-PTKIN} ..	52
Gambar 3.23 Lintasan yang memiliki suffix {SBMPTN}.....	53
Gambar 3.24 Conditional FP-Tree yang terbentuk untuk suffix {SBMPTN}.....	53
Gambar 3.25 Lintasan yang memiliki suffix {SNMPTN}.....	54

Gambar 3.26 Conditional FP-Tree yang terbentuk untuk suffix {SNMPTN} .....	54
Gambar 3.27 Lintasan yang memiliki suffix {SPAN-PTKIN} .....	54
Gambar 3.28 Conditional FP-Tree yang terbentuk untuk suffix {SPAN-PTKIN} .....	55
Gambar 3.29 Lintasan yang memiliki suffix {UM-PTKIN} .....	55
Gambar 3.30 Conditional FP-Tree yang terbentuk untuk suffix {UM-PTKIN} ..	55
Gambar 3.31 Lintasan yang memiliki <i>suffix</i> {SBMPTN} .....	56
Gambar 3.32 <i>Conditional FP-Tree</i> yang terbentuk untuk <i>suffix</i> {SBMPTN}.....	56
Gambar 4.1 Tampilan Login .....	66
Gambar 4.2 Tampilan Home .....	66
Gambar 4.3 Tampilan Data Mahasiswa .....	67
Gambar 4.4 Tampilan Proses Hasil.....	68
Gambar 4.5 Tampilan Hasil Proses FOLD-Growth.....	68
Gambar 4.6 Tampilan Hasil Rule .....	69
Gambar 4.7 Tampilan Hasil Grafik Rule dan Nilai Lift Ratio .....	69
Gambar 4.8 Grafik Pengaruh Nilai Minimum Support dan Nilai Minimum Confidence terhadap jumlah rule.....	75
Gambar 4.9 Grafik Nilai Support dan Nilai Confidence dari Association Rule yang terbentuk.....	79
Gambar 4.10 Grafik Lift Ratio dari Association Rule yang terbentuk .....	80

## DAFTAR TABEL

Tabel 2.1 Basis Data Transaksi .....	19
Tabel 2.2 Penelitian Terkait Pola Asosiasi .....	23
Tabel 3.1 Atribut Data Mahasiswa .....	27
Tabel 3.2 Atribut Data Kelulusan .....	28
Tabel 3.3 Indeks Prestasi Kumulatif (IPK).....	29
Tabel 3.4 Sample Data Mahasiswa Pertama Masuk dan Data Kelulusan hasil Preprocessing.....	34
Tabel 3.5 Kategori Tingkat Kelulusan .....	35
Tabel 3.6 Sample Data Mahasiswa dan Kategori Kelulusan .....	36
Tabel 3.7 Hasil Perhitungan Support Count 1-itemset .....	39
Tabel 3.8 Hasil Perhitungan Support 2-itemset .....	39
Tabel 3.9 Pemangkasan Support 1-itemset.....	41
Tabel 3.10 Pemangkasan Support 2-itemset.....	42
Tabel 3.11 Kandidat 1-itemset.....	43
Tabel 3.12 Kandidat 2-itemset.....	43
Tabel 3.13 Conditional Pattern Base yang terbentuk .....	50
Tabel 3.14 <i>Frequent Itemset</i> dari data.....	56
Tabel 3.15 Nilai <i>Confidence</i> dari <i>frequent itemset</i> .....	58
Tabel 3.16 Rule yang bisa dibangkitkan .....	58
Tabel 3.17 Lift Ratio Rules.....	60
Tabel 4.1 Hasil pengujian pengaruh minimum support dan minimum confidence .....	71
Tabel 4.2 Hasil rule dari masing-masing minimum support dan minimum confidence .....	71
Tabel 4.3 Hasil Pengujian Lift Ratio.....	74
Tabel 4.4 Rule yang dihasilkan .....	76
Tabel 4.5 Hasil Rule yang dihasilkan.....	81

## ABSTRAK

Dewi, Anis Shiva ‘Ulia. 2021. **Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa Menggunakan Algoritma *FOLD-Growth* (Studi Kasus Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang)**. Skripsi. Jurusan Teknik Informatika Fakultas Sains Dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (I) Dr. Muhammad Faisal, M.T. (II) M. AinulYaqin, M.Kom.

---

Kata Kunci : Penggalan Data, Pola Asosiasi, *Fold-Growth*, *Lift Ratio*.

Penerimaan mahasiswa tiap tahunnya di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengakibatkan terjadinya akumulasi data mahasiswa dalam jumlah besar. Perlunya pemanfaatan data dalam jumlah besar dengan banyaknya variasi pada basis data mahasiswa di Universitas Islam Negeri Maulana Malik Ibrahim Malang yang optimal. Pengolahan data mahasiswa pada saat pertama kali masuk universitas dan data kelulusan mahasiswa dapat dianalisis dan diketahui pola-pola dari kedua data tersebut menjadi sebuah informasi baru. Untuk mendapatkan informasi tersebut diperlukan sebuah sistem yang mengetahui pola asosiasi tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa menggunakan algoritma *FOLD-Growth*. Proses penggalan data algoritma *FOLD-Growth* ini lebih cepat dibandingkan algoritma yang sebelumnya seperti algoritma apriori. Algoritma *FOLD-Growth* memiliki 4 tahapan dalam proses penggalan data yaitu penggalan  $L_1$  dan  $L_2$  dengan menggunakan *SOTrieIT*, pemangkasan item-item yang tidak *frequent*, membangun *FP-Tree* dan penggalan *itemset frequent*. Apabila pola asosiasi sudah terbentuk maka akan dilakukan perhitungan pengujian *lift ratio* untuk mengetahui kekuatan dari pola asosiasi. Penelitian ini menggunakan 90 data mahasiswa jurusan Teknik Informatika angkatan 2013. Hasil dari penelitian dengan menggunakan nilai *minimum support* sebesar 4% dan 8% dan nilai *minimum confidence* sebesar 30% dan 60% yaitu 7 pola asosiasi dengan nilai *lift ratio* tertinggi sebesar 7,31 dan terendah sebesar 3,28.

## ABSTRACT

Dewi, Anis Shiva 'Ulia. 2021. **Association Pattern of Pathways to Student Graduation Using FOLD-Growth Algorithm (Case Study department of Informatics Engineering Department of Maulana Malik Ibrahim State University Malang)**. Essay. Department of Informatics Engineering, Faculty of Science and Technology, Islamic State University of Maulana Malik Ibrahim of Malang. Supervisor: (I) Dr. Muhammad Faisal, M.T. (II) M. AinulYaqin, M.Kom.

---

Keyword : Data Mining, Association Rule, Fold-Growth, Lift Ratio.

Admission of students every year at Maulana Malik Ibrahim State Islamic University Malang resulted in the accumulation of student data in large quantities. The need for the utilization of large amounts of data with the many variations in the student database at Maulana Malik Ibrahim Malang State Islamic University is optimal. Processing student data at the first time of university entry and student graduation data can be analyzed and known patterns of both data into a new information. To get this information, a system is needed that knows the pattern of association of student graduation rates with student entrance using the FOLD-Growth algorithm. The process of extracting FOLD-growth algorithm data is faster than previous algorithms such as a priori algorithms. FOLD -Growth algorithm has 4 stages in the process of extracting data, namely L1 and L2 excavation using SOTrieIT, trimming items that are not frequent, building FP-Tree and digging frequent itemset. If the association pattern has been formed, it will be done the calculation of lift ratio testing to find out the strength of the association pattern. This study used 90 student data majoring in Informatics Engineering class of 2013. The results of the study using minimum support values of 4% and 8% and minimum confidence values of 30% and 60% are 7 association patterns with the highest lift ratio value of 7.31 and the lowest of 3.28.

## الملخص

دويو, أنيس شيفا عليا. ٢٠٢١. نمط ارتباط المسار نحو تخرج الطلاب باستخدام خوارزمية النمو الديناميكي السريع عبر الإنترنت. (قسم دراسة الحالة لهندسة المعلوماتية مولانا مالك إبراهيم مالانج). أطروحة. قسم علوم و الطقنية, كلية العلوم و الهندسة, الجامعة الإسلامية الحكومية مولانا مالك إبراهيم مالانج. المشرف: (١) الدكتور محمد فيصل الماجستير (٢) محمد عين اليقين الماجستير.

---

الكلمات الرئيسية: تنقيب في البيانات, قاعدة الاقتران, FOLD- النمو, نسبة الرفع.

أدى القبول السنوي للطلاب في جامعة مولانا مالك إبراهيم الإسلامية في مالانج إلى تراكم كميات كبيرة من بيانات الطلاب. الحاجة إلى استخدام كميات كبيرة من البيانات مع العديد من الاختلافات في قاعدة بيانات الطلاب في جامعة مولانا مالك إبراهيم مالانج الإسلامية الدولية الأمثل. معالجة بيانات الطلاب في المرة الأولى من دخول الجامعة وبيانات تخرج الطلاب يمكن تحليلها وأنماط معروفة من كل من البيانات في معلومات جديدة. للحصول على هذه المعلومات، هناك حاجة إلى نظام يعرف نمط ارتباط معدلات تخرج الطلاب مع دخول الطلاب باستخدام خوارزمية FOLD-Growth. عملية استخراج بيانات خوارزمية FOLD-Growth أسرع من الخوارزميات السابقة مثل خوارزميات سابقة خوارزمية FOLD-Growth لديها 4 مراحل في عملية استخراج البيانات، وهي L1 و L2 الحفر باستخدام دعم أمرت تري مجموعة العناصر، وتقليم العناصر التي ليست متكررة، وبناء نمط شجرة متكررة وحفر مجموعة العناصر متكررة. إذا تم تشكيل نمط الارتباط، سيتم إجراء حساب اختبار نسبة الرفع لمعرفة قوة نمط الاقتران. استخدمت هذه الدراسة 90 بيانات الطلاب تخصص في فئة هندسة المعلوماتية من عام 2013. نتائج الدراسة باستخدام الحد الأدنى لقيم الدعم من 4% و 8% وقيم الثقة الدنيا من 30% و 60% هي 7 أنماط الارتباط مع أعلى قيمة نسبة رفع من 7.31 وأدنى من 3.28.

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Pada saat ini perguruan tinggi dituntut memiliki keunggulan untuk bersaing dengan sumber daya yang ada. Perlunya proses evaluasi, perencanaan, dan pengelolaan yang baik dalam perguruan tinggi untuk dapat memenangkan persaingan pada era ini. Sumber daya manusia maupun sarana prasana yang ada diperlukan juga suatu sistem informasi yang dapat digunakan untuk meningkatkan keunggulan dari suatu perguruan tinggi. Sistem informasi ini dapat digunakan untuk mengolah dan memberikan informasi yang dapat dijadikan pen gambil keputusan yang terbaik.

Penerimaan mahasiswa tiap tahunnya di Universitas Islam Negeri Maulana Malik Ibrahim Malang mengakibatkan terjadinya akumulasi data mahasiswa dalam jumlah besar. Perlunya pemanfaatan data dalam jumlah besar dengan banyaknya variasi pada basis data mahasiswa di Universitas Islam Negeri Maulana Malik Ibrahim Malang yang optimal. Pengolahan data mahasiswa pada saat pertama kali masuk universitas dan data kelulusan mahasiswa dapat dianalisis dan diketahui pola-pola dari kedua data tersebut menjadi sebuah informasi baru yang dapat bermanfaat.

Dalam Statistik Pendidikan Tinggi 2019 jumlah lembaga Perguruan Tinggi di Indonesia ada 3.129 Perguruan Tinggi Swasta (PTS) dan 122 Perguruan Tinggi Negeri (PTN) berdasarkan kelompok Perguruan Tinggi (Attamimi, Nirmala, &

Putri, 2018). Banyaknya jumlah lembaga Perguruan Tinggi di Indonesia, tidak semua Perguruan Tinggi diakui dan dipercaya oleh masyarakat atau bisa disebut tidak semua Pendidikan Tinggi dikatakan bagus.

Indikator untuk mengetahui Pendidikan Tinggi bisa dikatakan bagus dapat dilihat dari akreditasi dari masing-masing Pendidikan Tinggi. Akreditasi adalah suatu kegiatan penilaian untuk menentukan kelayakan Program Studi dan Perguruan Tinggi. Tujuan dilakukannya akreditasi adalah untuk menentukan kelayakan Program Studi dan Perguruan Tinggi berdasarkan kriteria yang mengacu pada Standar Nasional Pendidikan Tinggi dan menjamin mutu Program Studi dan Perguruan Tinggi secara eksternal baik di bidang akademik maupun non-akademik untuk melindungi kepentingan mahasiswa dan masyarakat. Kriteria titik berat penilaian Akreditasi Perguruan Tinggi salah satunya adalah kebijakan sistem penerimaan mahasiswa baru (Naskah Akademik Instrumen Akreditasi Perguruan Tinggi 3.0, 2019).

Sementara indikator untuk mengukur kualitas lulusan perguruan tinggi salah satunya dilihat dari masa studi. Dalam pedoman akreditasi perguruan tinggi, masa studi merupakan salah satu indikator untuk menilai efektifitas dan produktivitas pendidikan. Perguruan tinggi dituntut untuk mengevaluasi masa studi mahasiswa sehingga dapat menjadi pertimbangan dalam menyusun berbagai program peningkatan efisiensi pendidikan (Rahmi & Yozza, 2020).

Banyak faktor yang dapat mempengaruhi masa studi diantaranya adalah Indeks Prestasi Kumulatif (IPK) mahasiswa ketika menyelesaikan studinya.

Semakin tinggi nilai Indeks Prestasi Kumulatif (IPK) mahasiswa ketika menyelesaikan studinya, hal ini disebabkan karena mahasiswa dengan nilai Indeks Prestasi Kumulatif (IPK) lulus yang tinggi, biasanya memiliki Indeks Prestasi (IP) semester yang juga tinggi. Indeks Prestasi (IP) semester akan menentukan jumlah mata kuliah yang dapat diambil mahasiswa yang diukur dengan Sistem Kredit Semester (SKS). Sehingga mahasiswa dengan Indeks Prestasi Kumulatif (IPK) tinggi mempunyai kecenderungan untuk menyelesaikan studi lebih cepat (Febrilia, Yani, & Anwar, 2020).

Data yang ada pada sistem informasi dapat digunakan untuk mendapatkan, mengolah, dan menyebarkan informasi baru untuk menunjang dalam membantu beberapa kegiatan dan mengambil keputusan yang penting. Penggunaan data pada sistem informasi untuk menunjang kegiatan maupun dalam mengambil keputusan tidak hanya menggantungkan pada data saja, dibutuhkan analisa data untuk menggali potensi-potensi informasi pada data tersebut.

Pada data mahasiswa dan data kelulusan di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang ini belum diketahui pola asosiasi data keduanya. Dalam proses mining tidak semua atribut dalam data mahasiswa akan diambil, hanya beberapa atribut yang berguna dan tidak terlalu acak saja. Penggunaan data yang terlalu acak bisa membutuhkan waktu yang lama dalam proses mining dan tingkat hubungannya pun relatif rendah. Variabel data mahasiswa yang terdiri atas jalur masuk universitas dengan tingkat kelulusan mahasiswa yang dilihat dari lama studi dan Indeks Prestasi Kumulatif (IPK) apakah

dari variabel data tersebut terdapat pola asosiasi yang menunjukkan keterkaitan antar kedua data tersebut.

Adapun dalil Al-Qur'an menjelaskan mengenai peraturan hukum manusia yang terdapat dalam nash Qur'an Surat Ar-Ra'du ayat 37, Allah ﷻ berfirman :

وَكَذَلِكَ أَنْزَلْنَاهُ حُكْمًا عَرَبِيًّا وَلَئِنْ اتَّبَعْتَ أَهْوَاءَهُمْ بَعْدَمَا جَاءَكَ مِنَ الْعِلْمِ مَا لَكَ مِنَ اللَّهِ مِنْ وَلِيٍّ وَلَا وَاقٍ

*“Demikianlah Kami telah menurunkannya (Al-Qur'an) sebagai penentu hukum yang berbahasa Arab. Sungguh, jika engkau mengikuti keinginan mereka setelah datang pengetahuan kepadamu, niscaya engkau sekali-kali tidak mempunyai pelindung dan tidak (pula) pemelihara dari (siksa) Allah” (QS Ar-Ra'du : 37).*

Maksud dari ayat diatas menurut Tafsir Jalalayn dijelaskan bahwa tafsiran dari Surat Ar-Ra'du ayat 37 adalah (Dan demikianlah) penurunan itu (Kami telah menurunkannya) Al-Qur'an itu (sebagai peraturan dalam bahasa Arab) yaitu dengan memakai bahasa Arab, yang dengannya engkau putuskan hukum-hukum di antara manusia. (Dan seandainya kamu mengikuti hawa nafsu mereka) hawa nafsu orang-orang kafir, dalam hal apa yang mereka inginkan, supaya kamu melakukannya menurut peraturan agama mereka. Ini hanyalah merupakan perumpamaan (setelah datang pengetahuan kepadamu) tentang tauhid (maka sekali-kali tidak ada bagimu terhadap Allah) huruf min di sini adalah zaidah (seorang penolong pun) penolong yang mau membantu menyelamatkanmu (dan tidak pula seorang pemelihara) yang dapat mencegah azab Allah yang menimpa dirimu.

Relasi dari ayat diatas adalah sistem yang dibuat pada penelitian ini bahwa ketika manusia mengambil keputusan yang terbaik hendaknya mengikuti sesuai dengan ajaran yang telah dijelaskan pada Al-Qur'an.

Pencarian pola asosiasi data induk mahasiswa terhadap data kelulusan pernah diteliti oleh (Noranita & Bahtiar, 2010). Penelitian tersebut menggunakan inputan data induk mahasiswa pada angkatan 2000, 2002, dan 2003 serta data kelulusan dari tahun 2004 sampai 2008. Transformasi data pada proses data mining menghasilkan 6 kategori untuk mengukur tingkat kelulusan mahasiswa. Penelitian ini menggunakan algoritma *apriori*, dan proses *mining* yang digunakan hanya dibatasi untuk mengukur tingkat keberhasilan PSSB dan SPMB terhadap tingkat kelulusan mahasiswa. Algoritma *apriori* ini mudah untuk dipahami dan diimplementasikan akan tetapi algoritma ini juga memiliki kelemahan. Kelemahannya adalah saat melakukan pencarian *frequent itemset*, harus melakukan *scanning database* berulang kali untuk setiap kombinasi item sehingga membutuhkan waktu yang lama untuk melakukan *scanning database*. Selain itu juga dibutuhkan *generate candidate* yang besar untuk mendapatkan kombinasi item dari *database* (Erwin, 2009).

Pada penelitian ini menggunakan algoritma *FOLD-Growth* untuk memperbaiki kekurangan dari algoritma *apriori* yang telah digunakan pada penelitian sebelumnya. Penggabungan dua algoritma *FP-Growth* dengan *FOLDARM* (*Fast Online Dynamic Association Rule*) pada metode asosiasi menghasilkan sebuah algoritma baru yaitu *FOLD-Growth* yang memiliki kinerja lebih baik. Algoritma *FOLD-Growth* ini menggunakan struktur data *SOTrieIT* sehingga menghasilkan proses kinerja yang lebih fleksibel dan efisien apabila dibandingkan dengan algoritma-algoritma aturan asosiasi yang lain seperti *Apriori* dan *FP-Growth* (Woon et al, 2004). Hasil informasi yang akan ditampilkan adalah

nilai *support* (nilai penunjang), nilai *confidence* (nilai kepastian), dan kombinasi item hubungan antara tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang diatas, penelitian ini membahas bagaimana pola asosiasi tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang?

## **1.3 Tujuan Penelitian**

Tujuan dari penelitian ini untuk mengetahui pola asosiasi tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang dapat memberikan informasi dan dijadikan pengambil keputusan yang terbaik.

## **1.4 Manfaat Penelitian**

Penelitian ini menghasilkan sebuah sistem untuk mendapatkan informasi melalui pola-pola asosiasi hubungan tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang yang dapat meningkatkan keunggulan dari suatu perguruan tinggi.

## **1.5 Batasan Masalah**

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Informasi yang akan ditampilkan berupa nilai *support*, nilai *confident*, dan kombinasi item hubungan antara tingkat kelulusan mahasiswa dengan jalur masuk mahasiswa di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Data yang dikelola merupakan data mahasiswa dan data kelulusan di Jurusan Teknik Informatika Universitas Islam Negeri Maulana Maliki Ibrahim Malang angkatan 2013 yang menempuh Pendidikan dari tahun 2013 sampai tahun 2016.

## BAB II

### STUDI PUSTAKA

#### 2.1 *Data Mining*

*Data mining* menurut (Han et al, 2006) dalam buku *Data Mining Concepts and Techniques* merupakan proses mengekstraksi pola-pola yang menarik (implisit, tidak diketahui sebelumnya, dan berpotensi untuk dimanfaatkan) dari data yang berukuran besar. *Data mining* adalah penggalian atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data yang sangat besar.

*Data mining* lebih tentang mengekstraks atau menggali informasi dari data dalam jumlah besar. Diharapkan *data mining* akan menghasilkan informasi yang sangat akurat dan mudah dipahami. *Data mining* terkait dengan bidang keilmuan lainnya, seperti sistem *database*, *data warehousing*, statistik, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi. *Data mining* juga didukung oleh ilmu lain, seperti *neural network*, pengenalan pola, analisis data spasial, *database gambar*, pemrosesan sinyal (Han et al, 2006).

Tahap-tahap data mining yaitu (Han et al, 2012):

##### 1. Pembersihan Data (*Data Cleaning*)

Pembersihan data adalah proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang di peroleh, 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. Integrasi Data (*Data Integration*)

Integrasi data adalah penggabungan data dari berbagai *database* ke dalam satu *database* baru. Pada tahap ini data yang dibutuhkan untuk data *mining* tidak hanya berasal dari satu *database* melainkan juga berasal dari beberapa *database*. 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. Seleksi Data (*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. Transformasi Data (*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. Proses *Mining*

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

#### 6. Evaluasi pola (*Pattern Evaluation*)

Untuk mengidentifikasi pola-pola menarik ke dalam *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.

#### 7. Presentasi pengetahuan (*Knowledge Presentation*)

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses *data mining* adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami *data mining*. Karenanya presentasi hasil *data mining* dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses *data mining*. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining* (Sutradana & Wahyudi, 2017).

#### 2.2 Aturan Asosiasi (*Association Rule*)

*Association rule* adalah salah satu teknik mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Contoh dari aturan asosiatif dari analisa pembelian di suatu pasar swalayan bisa diketahui berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Adanya aturan asosiasi ini pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* yaitu persentase kombinasi item tersebut dalam *database* dan

*confidence* yaitu kuatnya hubungan antar item dalam aturan asosiatif (Pramudiono, 2003).

Misalkan himpunan literal (item) adalah  $I = \{i_1 ; i_2 ; \dots ; i_m\}$ , sedangkan himpunan transaksi adalah  $T$ , dan setiap transaksi  $T$  adalah himpunan item sehingga  $T \subseteq I$ . Setiap transaksi memiliki pengenal yang unik. Apabila transaksi  $T$  mengandung  $X$ , himpunan beberapa item dalam  $I$ , jika  $X \subseteq T$ . *Association rule* merupakan suatu implikasi dengan bentuk umum pada rumus 2.1

$$X \Rightarrow Y \quad (2.1)$$

dimana  $X \subset I$ ,  $Y \subset I$ , dan  $X \cap Y = \emptyset$ .  $X$  disebut *antecedent* dan  $Y$  disebut *consequent*. (Agrawal & Srikant, 1994)

Parameter yang biasanya digunakan untuk menentukan suatu *association rule* ada 2 yaitu :

- **Support** merupakan suatu ukuran yang menunjukkan seberapa besar tingkat dominasi suatu item/itemset dari keseluruhan transaksi. Ukuran ini menentukan apakah suatu item/itemset layak untuk dicari *confidence factor*-nya
- **Confidence** merupakan suatu ukuran yang menunjukkan hubungan antara dua atau lebih item secara *conditional*. (contohnya, seberapa sering *item B* dibeli jika orang membeli *item A*).

Kedua parameter ini nantinya akan digunakan dalam menentukan interesting association rule, yaitu untuk dibandingkan dengan batasan (*threshold*)

yang ditentukan oleh pengguna (*user*). Batasan tersebut umumnya terdiri dari *min\_support* dan *min\_confidence*. Apabila memenuhi kedua batasan maka sebuah *rule* dapat disebut *interesting rule* (Budhi et al, 2006).

Suatu contoh informasi bahawa pelanggan yang membeli computer yang bersamaan dengan membeli *antivirus computer* dipresentasikan dalam *association rule* sebagai berikut :

***Computer* → *antivirus\_software* → [*support* = 2%, *confidence* = 60%]**

Aturan asosiasi diatas terdapat *support* = 2% yang berarti bahwa 2% dari keseluruhan transaksi yang dianalisis menunjukkan bahwa *computer* dan *antivirus software* dibeli secara bersamaan. Dan *confidence* = 60% yang berarti bahwa 60% pelanggan yang membeli *computer* juga membeli *antivirus software* (Han et al, 2006).

Untuk mendapatkan nilai *support* dan *confidence*, diperlukan rumus untuk kedua parameter. Untuk mencari nilai *support* dari 1 item maka menggunakan rumus 2.2. Sementara untuk mencari nilai *support* dari 2 item maka digunakan rumus 2.3.

$$Support(A) = \frac{Jumlah\ Transaksi\ yang\ Mengandung\ A}{Total\ Transaksi} \quad (2.2)$$

$$Support(A, B) = \frac{Jumlah\ Transaksi\ yang\ Mengandung\ A\ dan\ B}{Total\ Transaksi} \quad (2.3)$$

Untuk mencari nilai *confidence* dari aturan  $A \rightarrow B$  maka digunakan rumus

$$Confidence(A \rightarrow B) = \frac{Support(A, B)}{Support(A)} \quad (2.4)$$

$$\text{Confidence } (A \rightarrow B) = P(B|A) = \frac{\text{Jumlah Transaksi yang Mengandung A dan B}}{\text{Jumlah Transaksi yang Mengandung A}} \quad (2.5)$$

Dimana *support* A,B adalah hasil perhitungan dari jumlah transaksi yang mengandung A dan B dibagi dengan total transaksi seperti pada rumus 2.3 (Zhao & Bhowmick, 2003).

### 2.3 Algoritma *FOLDARM*

Algoritma *Fast Online Dynamic Association Rule Mining (FOLDARM)* diusulkan sebagai perpanjangan dari *RARM (Rapid Association Rule Mining)* sehingga penambahan aturan asosiasi dapat dilakukan lebih efektif dalam *electronic commerce*. Seperti *RARM*, *FOLDARM* membangun struktur data baru yang disebut *Support-Ordered Trie Itemset (SOTrieIT)*. Struktur *trie-like tree* ini menyimpan *support counts* dari semua *1-itemsets* dan *2-itemsets* dalam *database*. Semua transaksi yang diproses sebelumnya; semua *1-itemsets* dan *2-itemsets* diekstraksi dari setiap transaksi. Informasi yang diekstrak akan digunakan untuk memperbarui *SOTrieIT* tanpa perlu pengetahuan sebelumnya tentang *support threshold*. Struktur ini diurutkan berdasarkan *support counts* dari setiap node dalam *descending order*. *FOLDARM* menggunakan *SOTrieIT* untuk menemukan *large 1-itemsets* dan *2-itemsets* dengan cepat tanpa *scanning database* (Woon et al, 2002).

Algoritma *FOLDARM (Fast Online Dynamic Association Rule Mining)* adalah algoritma *data mining* yang menggunakan struktur data *SOTrieIT*. Ketika ukuran *frequent itemset* maksimum atau  $k_{\max} \leq 10$  maka struktur tersebut memiliki kinerja yang cepat, sedangkan algoritma *FP-Growth* memiliki kinerja yang cepat saat *frequent itemset* maksimum atau  $k_{\max} \geq 10$  (Soelaiman & Arini WP, 2006).

## 2.4 Algoritma *FP-Growth*

Algoritma *FP-Growth* adalah pengembangan dari algoritma *Apriori*, sehingga kekurangan dari algoritma *Apriori* diperbaiki di algoritma *FP-Growth*. Algoritma ini menentukan *frequent itemset* yang berakhiran *suffix* tertentu dengan menggunakan metode *divide* dan *conquer* untuk memecah *problem* menjadi *subproblem* yang lebih kecil. *FP-Growth* menggunakan konsep pembangunan *tree* dalam pencarian *frequent itemset*, tidak menggunakan *generate candidate* seperti algoritma *Apriori*. Hal ini yang menyebabkan algoritma *FP-Growth* lebih cepat dari algoritma *Apriori* (Han et al, 2011). Adapun *pseudocode* dari algoritma *FP-Growth* sebagai berikut :

**Input** : the *FP-Tree Tree*  
**Output** :  $R_t$  Complete set of frequent patterns  
**Method** : Call *FP-Growth* (*Tree*, null)  
**procedure** *FP\_growth* (*Tree*  $\alpha$ )  
 (1) **if** *Tree* contains a single path *P* then  
 (2)   **for each** combination (denoted as  $\beta$ ) of the nodes in the path *P*  
 (3)     generate pattern  $\beta \cup \alpha$  with *support\_count* = *minimum support count of nodes in  $\beta$* :  
 (4) **else for each**  $a_i$  in the header of *Tree* {  
 (5)   generate pattern  $\beta = a_i \cup \alpha$  with *support\_count* =  $a_i$  *support\_count*;  
 (6)   construct  $\beta$ 's conditional pattern base and then  $\beta$ 's conditional *FP-tree Tree*  $\beta$ ;  
 (7)   **if**  $Tree_\beta \neq \emptyset$  **then**  
 (8)     call *FP\_growth*( $Tree_\beta, \beta$ ); }

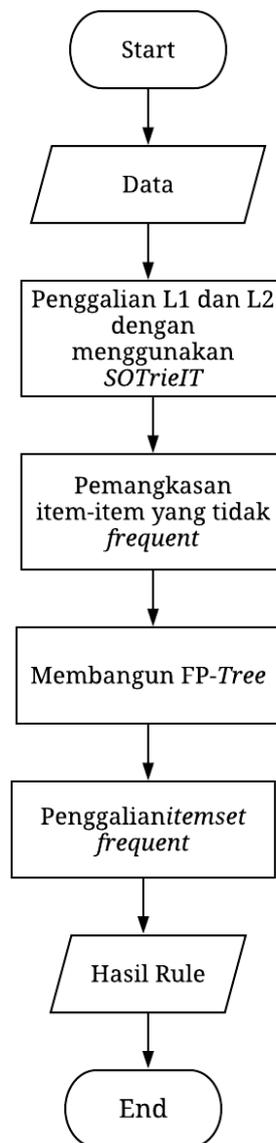
Algoritma *FP-Growth* mencari himpunan *item* yang sering muncul dengan membangun *FP-Tree* secara rekursif dan menggabungkan *itemset* yang ada pada *conditional FP-Tree* secara berturut-turut. Selain lebih padat, juga tetap lengkap merupakan keuntungan dari penggunaan struktur data *FP-Tree*. Informasi yang

tidak relevan dikurangi, item-item yang tidak sering muncul (*infrequent*) dihilangkan, tetapi tetap menjaga kelengkapan dari informasi yang dibutuhkan untuk *frequent pattern mining* (Han et al, 2006).

## 2.5 Algoritma *FOLD-Growth*

Algoritma *FOLD-Growth* merupakan hasil gabungan antara algoritma *FOLDARM* dan *FP-Growth*. Diharapkan dapat menggabungkan keuntungan dari dua algoritma antara *FOLDARM* dan *FP-Growth*. Algoritma *FOLDARM* yang memiliki kinerja cepat dalam pengukuran *itemset frequent* maksimum ( $K_{\max}$ ) adalah kecil atau  $K_{\max} \leq 10$ , sedangkan algoritma *FP-Growth* memiliki kinerja yang cepat pada saat  $K_{\max} > 10$  (Woon et al, 2004).

Berikut empat tahapan utama dalam algoritma *FOLD-Growth* (Soelaiman & WP, 2006):



**Gambar 4.1** Flowchart Algoritma FOLD-Growth

1. Penggalian  $L_1$  dan  $L_2$  dengan menggunakan *SOTrieIT*

Di tahap ini dilakukan scan basis data sebanyak satu kali untuk membaca transaksi-transaksi yang ada dalam basis data. Setiap transaksi yang akan dibangkitkan semua kemungkinan-kemungkinan *1-itemset* dan *2-itemset* lalu dicatat dalam *SOTrieIT*.

2. Pemangkasan item-item yang tidak *frequent*

Di tahap ini akan dilakukan pemangkasan pada setiap transaksi yang ada dalam basis data dengan memakai  $L_1$  dan  $L_2$ . Untuk setiap transaksi  $T$ , pada *itemset*  $L_k$  yang terdapat dalam transaksi tersebut dimana Panjang  $k$  lebih dari 2, maka akan dilakukan pengecekan dengan memakai  $L_1$  dan  $L_2$ . Sehingga untuk item-item yang dianggap tidak *frequent* akan dilakukan pemangkasan. Apabila nilai *support count*-nya kurang dari batas *minimum support* yang telah ditentukan oleh pengguna maka item tersebut dikatakan tidak *frequent*. Setelah dilakukan pemangkasan terhadap transaksi  $T$ , maka item-item yang terdapat dalam transaksi tersebut diurutkan berdasarkan nilai *support count* yang paling besar. Dengan menggunakannya  $L_1$  dan  $L_2$  yang didapatkan melalui *SOTrieIT*, maka akan dihasilkan *Ordered Frequent Items* yang telah dipangkas.

### 3. Membangun *FP-Tree*

Pada tahapan ini, pembangunan *FP-Tree* dengan menggunakan transaksi  $T$  yang telah dipangkas dan diurutkan berdasarkan nilai *support count*. Dibangunnya *FP-Tree* dengan perolehan *frequent items* setelah dipangkas dan diurutkan.

### 4. Penggalian *itemset frequent*

Tahap ini penggalian *itemset frequent* dengan algoritma menggunakan algoritma *FP-Growth* pada *FP-Tree* tersebut.

## 2.6 Struktur Data *SOTrieIT*

*SOTrieIT* (Support Ordered-Trie Itemset) adalah suatu *tree* yang dibangun dengan mengekstrak 1-2 *itemset* dari setiap transaksi. Penggalian 1-2 *itemset*

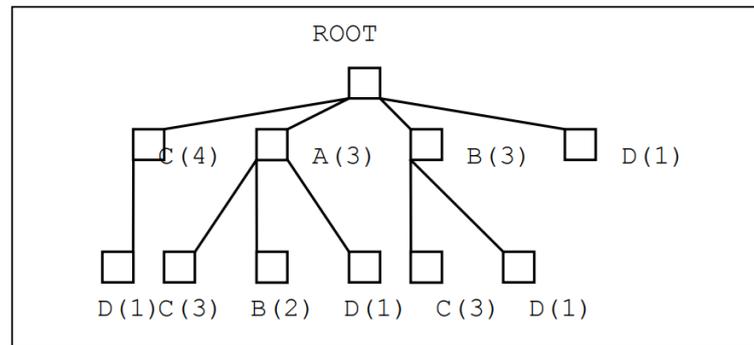
tersebut dilakukan dengan membaca *database* satu kali untuk membaca semua transaksi yang ada dalam *database* (Zhao & Bhowmick, 2003).

*SOTrieIT* memiliki dua tingkatan dari *node-node tree* yang menyatakan bahwa setiap *node*  $w$  memiliki sebuah label  $l$  yang dinyatakan sebagai *item* dan sebuah notasi  $j$  yang menyimpan nilai *support count* yang berhubungan. Setiap *node tree* terhubung pada beberapa *item* yang terdapat dalam *itemset*  $I$  (dinotasikan  $a_i \in I$ ) maka untuk  $w_i$  mengacu pada *node* yang memiliki hubungan dengan  $a_i \in I$ . Himpunan *SOTrieIT* dimungkinkan memiliki *parent node* yang berbeda-beda seperti  $w_1, w_2, \dots, w_n$ , yang dibangun dari sebuah basis data yang menyimpan *support count* dari semua *1-itemset* dan *2-itemset*, maka digunakan *node* khusus yang dinamakan *ROOT* untuk menghubungkan semua *SOTrieIT* bersama-sama (Soelaiman & Arini WP, 2006).

**Tabel 4.1** Basis Data Transaksi

TID	Item
1	AC
2	BC
3	ABC
4	ABCD

Dari data transaksi pada tabel 2.1 maka hasil dari *SOTrieIT* nya adalah pada gambar 2.1.



**Gambar 4.2** Hasil SOTrieIT dari Tabel 2.1

## 2.7 Lift Ratio

*Lift ratio* adalah suatu ukuran untuk mengetahui kekuatan sebuah aturan asosiasi yang didapatkan dari perbandingan nilai *confidence* dengan nilai *benchmark confidence*. Jika nilai *lift ratio* lebih besar dari 1, maka menunjukkan adanya manfaat dari aturan asosiasi tersebut. Semakin tinggi nilai *lift ratio* maka semakin besar pula kekuatan asosiasinya (Santosa, 2007).

Untuk menghitung nilai *lift ratio* dapat dihitung dengan rumus dibawah ini:

$$\text{Lift Rasio} = \frac{\text{Confidence}(A,B)}{\text{Benchmark Confidence}(A,B)} \quad (2.5)$$

Untuk mendapatkan nilai *benchmark confidence* dapat dihitung dengan rumus berikut:

$$\text{Benchmark Confidence} = \frac{Nc}{N} \quad (2.6)$$

Keterangan:

- $Nc$  = jumlah transaksi dengan item yang menjadi *consequent*
- $N$  = jumlah transaksi basis data

## 2.8 Penelitian Terkait

Pada bab ini membahas penelitian-penelitian yang sudah dilakukan sebagai perbandingan dan acuan untuk penelitian ini.

Dalam penelitian “Implementasi Algoritma FP-Growth dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas” (Fitria, Nengsih, & Qudsi, 2017) ini menggunakan teknik *data mining* untuk menemukan informasi berupa pola yang dapat menjadi acuan dalam pengambilan kebijakan di Satuan Lalu Linta Polresta Pekanbaru. Berdasarkan pengujian lift ratio didapatkan pola kecelakaan yang paling sering terjadi, yaitu faktor kecelakaan dengan jenis\_luka adalah luka ringan, jenis\_jalan adalah jalan arteri, waktu adalah padat kendaraan, dan jenis\_kelamin adalah perempuan dengan nilai *lift ratio* 1.20%. Sedangkan pada hasil pengujian skala *Likert* didapatkan bahwa 88.09% pengguna merasa dimudahkan dengan sistem ini dan dapat membantu mereka dalam mengetahui pola kecelakaan yang sering terjadi.

Menurut (Soelaiman & Arini WP, 2006) pada penelitian “Analisis Kinerja Algoritma *Fold-Growth* dan *FP-Growth* pada Penggalian Pola Asosiasi” dijelaskan bahwa uji coba yang melibatkan dataset sintetik dapat disimpulkan durasi eksekusi, skalabilitas, reliabilitas, dan utilisasi memori *Fold-Growth* lebih baik daripada *FP-Growth*. Hal ini disebabkan pada algoritma *Fold-Growth* telah melalui proses pemangkasan terlebih dahulu sehingga lebih efisien dalam penyimpanan data.

Pada penelitian (Zahrotun et.al, 2018) “Aplikasi *Data Mining* untuk Mencari Pola Asosiasi *Tracer Study* Menggunakan Algoritma *FOLDARM*” ini

menggunakan pendekatan *SOTrieIT* dan dengan data sampel dari alumni Universitas Ahmad Dahlan yang diambil secara acak. Nilai terbesar dari daftar calon asosiasi yaitu *threshold* = 3, nilai *support* 24% dan *confidence* = 45% dan menghasilkan aturan asosiasi yang memenuhi nilai *support* dan *confidence* yang telah diberikan tetapi nilai *lift rationya* < 1 sehingga hubungannya tidak kuat yaitu jika masa tunggu mencari kerjanya 3-5,9 bulan dan gaji pertama kali bekerja gaji < 1.572.200 dan nilai IPK  $\geq 2.75$ -3.50 dan umur ketika lulus umur 15-24 tahun maka lulus tidak tepat waktu.

Dalam penelitian “Penerapan *data mining* untuk analisis pengaruh lama studi mahasiswa Teknik Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta menggunakan metode *apriori*” (Sutradana & Wahyudi, 2017) peneliti menerapkan metode *data mining* menggunakan metode *apriori* untuk menganalisa dan mencari pola-pola yang mempengaruhi lama studi mahasiswa Teknik Informatika berdasarkan beberapa kategori kelulusan yang dapat diukur dari semester kelulusannya. Penelitian ini dapat menerapkan data mining dengan algoritma *apriori* pada data alumni mahasiswa Teknik Informatika untuk mencari pola-pola yang mempengaruhi lama studi mahasiswa berdasarkan 3 kategori kelulusan yang dapat diukur dari semester lulus.

Analisa pola asosiasi jalur masuk terhadap kelulusan mahasiswa dengan menggunakan metode *FOLD-Growth* (Irfan, 2015) peneliti menggunakan metode *FOLD-Growth* ini untuk mendapatkan beberapa informasi yang bermanfaat tentang hubungan antara tingkat kelulusan mahasiswa dengan jalur masuk perguruan tinggi mahasiswa Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Gunung

Djati Bandung menggunakan teknik data *mining*. Hasil dari penelitian ini terbentuklah *rule* jika jalur masuknya SPMB maka kategori kelulusannya B2 (lama studi lebih dari 4 tahun dan IPK 3,00-3,49) dengan nilai *confident* 0,5 dan nilai *support* 0,2.

Menurut penelitian (Hardianti et al, 2014) “Implementasi Algoritma *FOLD-Growth* pada Pemodelan Pola Pembelian Barang sebagai Pendukung Penentuan Letak Barang”, algoritma *FOLD-Growth* dapat diimplementasikan dalam kasus *Market Basket Analysis* untuk menentukan *rule* (aturan terkait) peletakan barang pada toko atau supermarket. Dari data transaksi yang diuji cobakan berdasarkan jumlah data serta kombinasi parameter *minimum support* 0.1% - 0.5% dan *minimum confidence* 10% - 50%. *Rule* yang terbaik didapatkan dari kombinasi nilai *minimum support* 0,1% dan *minimum confidence* 10%. Hasil penggalian data dari sejumlah percobaan menunjukkan bahwa banyaknya kombinasi dari item-item yang dipasangkan pada satu rak atau kelompok menghasilkan jumlah *rule* yang sesuai dan dapat dijadikan acuan dalam peletakan barang.

**Tabel 4.2** Penelitian Terkait Pola Asosiasi

No	Parameter / Pengujian	Rule Asosiasi	Algoritma	Pengujian
1	Fitria, Nengsih, & Qudsi, 2017	✓	<i>FP-Growth</i>	<i>Lift Ratio</i>
2	Soelaiman & Arini WP, 2006	-	<i>Fold-Growth</i> dan <i>FP-Growth</i>	Uji Coba Kebenaran

				dan Uji Coba Kinerja
3	Zahrotun et.al, 2018	✓	<i>Foldarm</i>	<i>Lift Ratio</i>
4	Sutradana & Wahyudi, 2017	✓	<i>Apriori</i>	-
5	Irfan, 2015	✓	<i>Fold-Growth</i>	-
6	Hardianti et al, 2014	✓	<i>Fold-Growth</i>	Uji Coba Pengaruh Nilai <i>Min. Sup.</i> dan <i>Min. Conf.</i>
7	Penelitian ini	✓	<i>Fold-Growth</i>	<i>Lift Ratio</i> dan Uji Coba Pengaruh Nilai <i>Min. Sup.</i> dan <i>Min. Conf.</i>

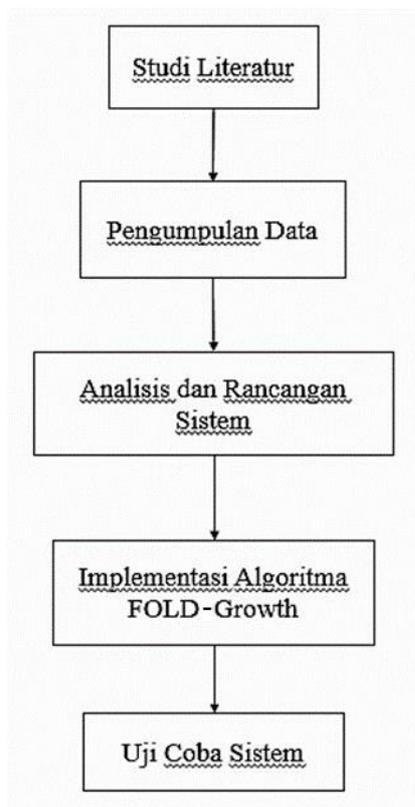
Pada tabel 2.2 dapat dilihat perbandingan parameter peneliti sebelumnya dan penulis gunakan pada penelitian pola asosiasi. Rata-rata penulis sebelumnya dan penulis menggunakan *rule* asosiasi sebagai latar belakang pengembangan dan pembaharuan pada penelitian pola asosiasi. Penulis menggunakan algoritma *Fold-Growth* dengan pengujian *Lift Ratio* dan Uji Coba Pengaruh Nilai *Minimum Support* dan *Minimum Confidence*.

### **BAB III**

#### **METODE PENELITIAN**

Adapun langkah-langkah dalam menyelesaikan tujuan penelitian ini sebagai berikut:

1. Melakukan studi literatur tentang association rule dan algoritma *FOLD-Growth*.
2. Mengumpulkan data-data yang diperlukan untuk penelitian ini yaitu data mahasiswa ketika pertama kali masuk universitas dan data kelulusan yang menempuh studi mulai tahun 2013 sampau tahun 2016. Data tersebut nantinya akan dilakukan *preprocessing* agar dapat diolah dalam sistem.
3. Melakukan analisis dan merancang sistem untuk menggali *association rule* dari data mahasiswa tersebut.
4. Mengimplementasikan rancangan yang sudah dibuat menjadi sebuah sistem untuk menganalisis *association rule* dari data mahasiswa menggunakan algoritma *FOLD-Growth*.
5. Melakukan uji coba terhadap sistem yang sudah dibangun.



**Gambar 5.1** Prosedur Perancangan Sistem

### 3.1 Pengumpulan Data

Penulis menggunakan metode dokumentasi dalam mengumpulkan data untuk penelitian ini. Metode dokumentasi merupakan salah satu metode pengumpulan data kualitatif dengan menganalisis dokumen-dokumen yang dibuat oleh orang itu sendiri ataupun oleh orang lain. Metode ini adalah salah satu cara bagi peneliti untuk mendapatkan data dari suatu media tertulis maupun dokumen tertulis yang lainnya.

Sumber data pada penelitian ini adalah dari Bagian Akademik dan Kemahasiswaan Universitas Islam Negeri Maulana Malik Ibrahim Malang. Data yang digunakan adalah data mahasiswa pertama masuk ke universitas dan data

kelulusan mahasiswa jurusan Teknik Informatika yang menempuh Pendidikan dari tahun 2013 sampai tahun 2016 sebanyak 95 data mahasiswa. Dalam penelitian ini data yang akan diproses terlebih dahulu dilakukan *preprocessing* dengan menghapus atribut yang tidak diperlukan dalam sistem.

### 3.1.1 Data Induk Mahasiswa

Data induk mahasiswa merupakan data mahasiswa yang di data ketika mahasiswa pertama masuk ke Universitas Islam Negeri Maulana Malik Ibrahim Malang setelah selesai melakukan registrasi ulang. Data yang digunakan berupa identitas pribadi mahasiswa, yaitu Nomor Induk Mahasiswa (NIM), Nama Lulus, Jenis Kelamin, dan Jalur Masuk.

**Tabel 5.1** Atribut Data Mahasiswa

Atribut	Keterangan
Nomor Induk Mahasiswa (NIM)	Nomor identitas mahasiswa yang diberikan pihak universitas
Nama Lulus	Nama mahasiswa yang lulus seleksi masuk universitas
Jenis Kelamin	Jenis kelamin mahasiswa yang bersangkutan
Jalur Masuk	Jalur masuk tes mahasiswa yang bersangkutan pada saat pertama masuk ke universitas

### 3.1.2 Data Kelulusan

Data kelulusan adalah data mahasiswa yang telah dinyatakan lulus dan telah menyelesaikan pendidikannya di Jurusan Teknik Informatika Universitas Islam

Negeri Maulana Malik Ibrahim Malang. Pada data kelulusan mahasiswa terdapat atribut data kelengkapan kelulusan mahasiswa meliputi Nomor Pokok Ijazah, Nomor Induk Mahasiswa (NIM), Nama Mahasiswa, dan Indeks Prestasi Kumulatif (IPK).

**Tabel 5.2** Atribut Data Kelulusan

<b>Atribut</b>	<b>Keterangan</b>
No. Pokok Ijazah	Nomor pokok yang terdapat di ijazah pada saat mahasiswa yang bersangkutan telah menyelesaikan pendidikan di universitas
NIM	Nomor Induk Mahasiswa (NIM) adalah nomor identitas mahasiswa yang diberikan pihak universitas
Nama Mahasiswa	Nama mahasiswa yang bersangkutan
Satuan Kredit Semester (SKS) Total	Satuan Kredit Semester (SKS) selama menempuh pendidikan di universitas
Indeks Prestasi Kumulatif (IPK)	Indeks prestasi yang dicapai mahasiswa yang bersangkutan selama menempuh pendidikan di universitas
Status studi	Status studi mahasiswa yang bersangkutan selama menempuh pendidikan di universitas

Kondisi Status	Kondisi status mahasiswa yang bersangkutan telah menyelesaikan studi atau dalam masa kritis di universitas
----------------	--

Berikut kategori Indeks Prestasi Kumulatif (IPK) berdasarkan kelulusan Universitas Islam Negeri Maulana Malik Ibrahim Malang :

**Tabel 5.3** Indeks Prestasi Kumulatif (IPK)

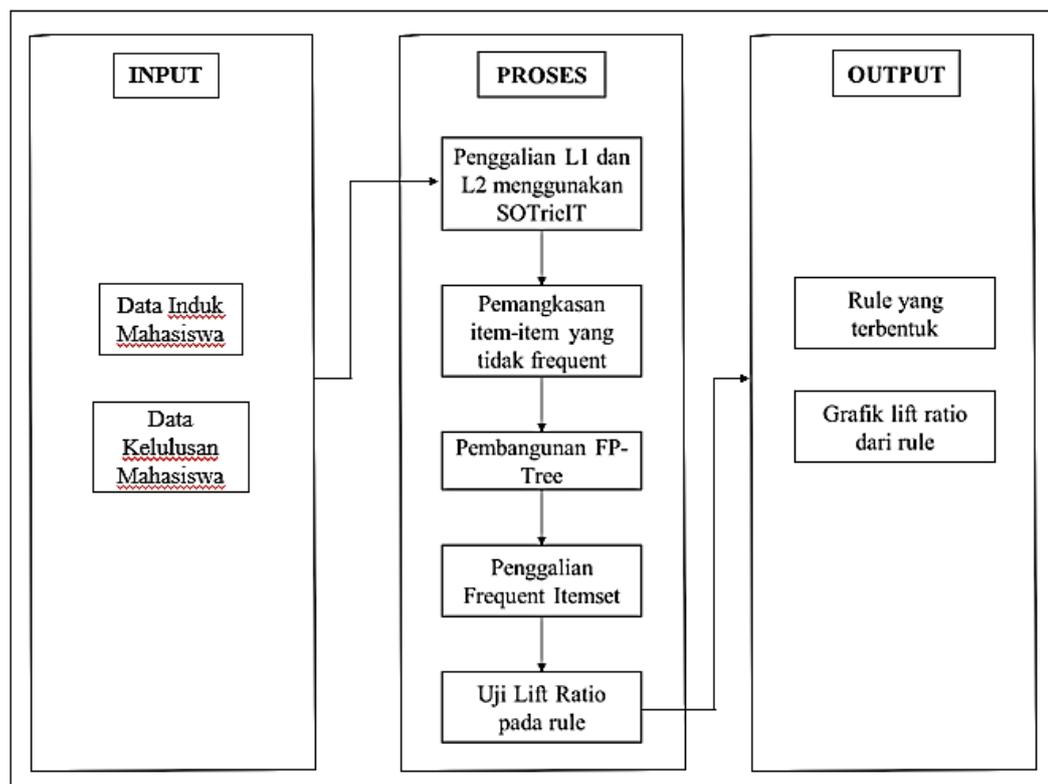
No.	Indeks Prestasi	Predikat Kelulusan
1	3.51-4.00	Dengan Pujian
2	3.00-3.50	Sangat Baik
3	2.51-2.99	Baik
4	2.00-2.50	Cukup

### 3.2 Desain Sistem

Menurut George M. Scott dalam buku Principles of Management Information Systems bahwa "Desain sistem menentukan bagaimana suatu sistem akan menyelesaikan apa yang harus diselesaikan, tahap ini menyangkut mengkonfigurasi dari komponen-komponen perangkat lunak dan perangkat keras dari suatu sistem sehingga setelah instansi dari sistem akan benar-benar memuaskan rancang bangun yang telah di tetapkan pada akhir tahap analisis sistem."

Desain sistem merupakan gambaran kebutuhan-kebutuhan fungsional dan persiapan perencanaan implementasi suatu sistem yang akan kita buat. Data-data

dari Bagian Akademik dan Kemahasiswaan Universitas Islam Negeri Maulana Malik Ibrahim Malang yaitu data mahasiswa pertama masuk dan data kelulusan mahasiswa akan diolah menggunakan Algoritma *FOLD-Growth* hingga menghasilkan beberapa pola asosiasi hubungan antara kedua data tersebut.



**Gambar 5.2** Desain Sistem

### 3.2.1 Desain Sistem Bagian Input

Pada penelitian ini peneliti membutuhkan data input berupa data induk mahasiswa dimana data tersebut merupakan data pertama masuk ke universitas dan data kelulusan mahasiswa meliputi data mahasiswa yang telah dinyatakan lulus dan sudah menyelesaikan pendidikan di Jurusan Teknik Informatika.

Data yang diambil adalah data mahasiswa angkatan 2013 yang telah menempuh pendidikan mulai tahun 2013 sampai dengan tahun 2016 dari Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.

### 3.2.2 Desain Sistem Bagian Proses

Setelah data terkumpul, langkah selanjutnya adalah implementasi algoritma *FOLD-Growth*. Tahap pertama yaitu membaca data-data mahasiswa dan data kelulusan. Setiap datanya akan dibangkitkan dengan kemungkinan *1-itemset* dan *2-itemset* lalu dicatat kembali dalam struktur data *SOTrieIT*.

Tahap kedua merupakan pemangkasan item yang tidak *frequent* yaitu jika nilai *support* dan nilai *count* nya kurang dari Batasan *minimum support* yang sudah ditentukan. Apabila *minimum support* dan *minimum confidence* tidak sesuai maka akan dipangkas.

Setelah melakukan pemangkasan item yang tidak *frequent*, selanjutnya *item-item* tersebut diurutkan berdasarkan nilai *support* dan nilai *confidence* yang paling besar dan yang sudah di pangkas.

Tahap keempat yaitu pembangunan *FP-Tree* menggunakan *item* yang telah dipangkas berdasarkan urutan nilai *support* dan nilai *confidence*. Selanjutnya dilakukan proses penggalian *itemset* yang *frequent*.

Tahap berikutnya adalah *conditional FP-Tree* yaitu *support count* dari setiap *item* dijumlahkan. Item yang dibangkitkan adalah nilai *support* lebih besar atau sama dengan nilai minimum *support*.

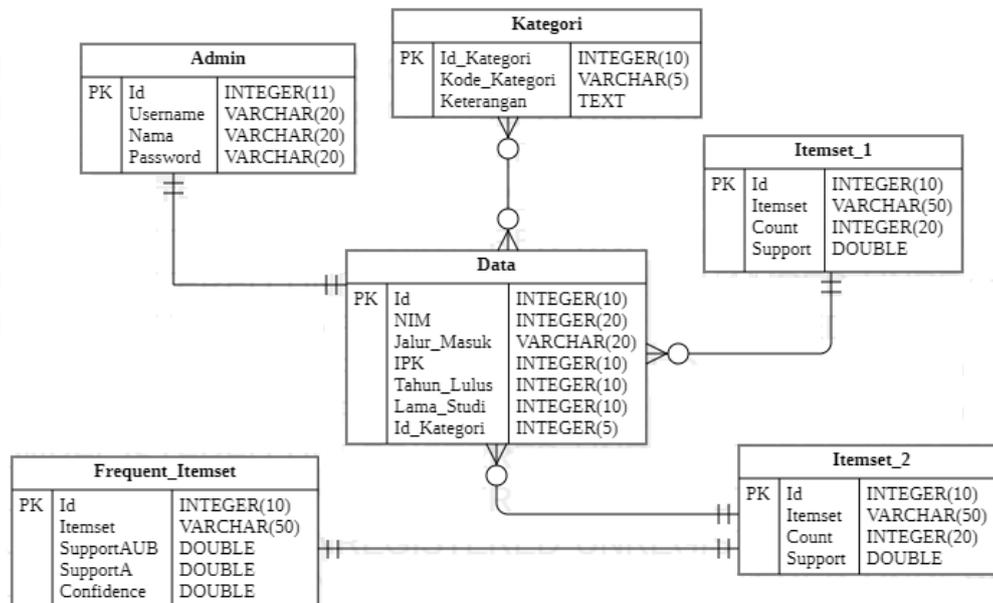
Tahap yang terakhir adalah pengujian lift ratio pada hasil rule yang telah terbentuk memiliki nilai lebih besar dari 1. Hal ini menunjukkan bahwa rule tersebut memiliki sifat yang kuat dan valid untuk digunakan sebagai pola asosiasi hubungan antara data mahasiswa dan data kelulusan.

### **3.2.3 Desain Sistem Bagian Output**

Output dari proses perhitungan menggunakan algoritma *FOLD-Growth* yaitu berupa hasil rule yang terbentuk dan grafik lift ratio dari rule yang sudah terbentuk.

### **3.2.4 Perancangan Database**

Perancangan database sistem pada penelitian ini menggunakan ERD (Entity Relation Diagram) merupakan diagram yang menjelaskan tentang objek-objek yang akan dibangun pada sistem yang terdapat pada gambar 3.3.



**Gambar 5.3** (ERD) Entity Relation Diagram

### 3.3 Perhitungan Manual

Pada perhitungan manual ini dilakukan dengan data yang akan diambil dan diolah sebanyak 95 data mahasiswa dengan *minimum support* sebesar 4% dan *minimum confidence* 30%. Perhitungan data pada penelitian ini, data disesuaikan dengan kebutuhan dalam mengolah data, dengan melakukan *preprocessing*.

Langkah pertama dalam *preprocessing* ini adalah melakukan seleksi data sehingga dari 8 atribut yang ada pada data mahasiswa pertama masuk universitas dan data kelulusan hanya diambil Nomor Induk Mahasiswa (NIM), jalur masuk universitas, Indeks Prestasi Kumulatif (IPK), tahun kelulusan, dan lama studi.

Langkah kedua, data hasil seleksi sebelumnya ditransformasikan ke dalam bentuk yang sesuai dengan kebutuhan pengolahan data. Transformasi data dilakukan pada data kelulusan mahasiswa, dengan melihat dari lama studi dan

Indeks Prestasi Kumulatif (IPK). Dua kategori tersebut dapat dibentuk kategori berdasarkan kombinasi keduanya, seperti pada tabel 3.5. Data mahasiswa pertama masuk dan data kelulusan yang akan diolah dirubah menjadi 1 tabel yaitu tabel mahasiswa. Tabel mahasiswa akan menyimpan informasi Nomor induk Mahasiswa, Jalur Masuk Mahasiswa, dan Kategori Tingkat Kelulusan. Hasil dari *preprocessing* terdapat pada tabel 3.4 dan hasil dari transformasi data terdapat pada tabel 3.5 dan 3.6.

**Tabel 5.4** Sample Data Mahasiswa Pertama Masuk dan Data Kelulusan Hasil *Preprocessing*

<b>NIM</b>	<b>Jalur Masuk</b>	<b>IPK</b>	<b>Tahun Lulus</b>	<b>Lama Studi</b>
13650002	SPAN-PTKIN	3.41	2018	5
13650005	SNMPTN	3.53	2018	5
13650006	SNMPTN	3.18	2019	6
13650007	SNMPTN	2.84	2020	7
13650008	SNMPTN	3.44	2018	5
13650009	SNMPTN	3.41	2019	6
13650010	SNMPTN	3.53	2018	5
13650011	SNMPTN	3.1	2019	6
13650012	SPAN-PTKIN	3.32	2019	6
13650013	SPAN-PTKIN	3.38	2018	5
13650014	SPAN-PTKIN	3.22	2020	7

13650015	SNMPTN	3.65	2018	5
13650017	SNMPTN	3.27	2019	6
13650019	SPAN-PTKIN	3.52	2018	5
13650022	SPAN-PTKIN	3.57	2018	5
13650023	SNMPTN	3.3	2020	7
13650024	SNMPTN	3.36	2018	5
13650025	SNMPTN	3.45	2019	6
13650026	SNMPTN	3.36	2019	6
13650027	SPAN-PTKIN	3.64	2017	4

**Tabel 5.5** Kategori Tingkat Kelulusan

<b>Kategori</b>	<b>Keterangan</b>
A1	Lama studi kurang dari 4 tahun atau tepat 4 tahun dan IPK 3,51 - 4,00
A2	Lama studi kurang dari 4 tahun atau tepat 4 tahun dan IPK 2,76 - 3,50
A3	Lama studi kurang dari 4 tahun atau tepat 4 tahun dan IPK 2,00 - 2,75
B1	Lama studi lebih dari 4 tahun dan IPK 3,51 - 4,00
B2	Lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50

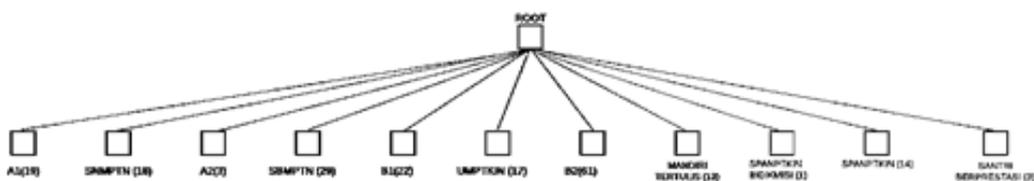
<b>Kategori</b>	<b>Keterangan</b>
B3	Lama studi lebih dari 4 tahun dan IPK 2,00 - 2,75

**Tabel 5.6** Sample Data Mahasiswa dan Kategori Kelulusan

<b>NIM</b>	<b>Jalur Masuk</b>	<b>Kategori Tingkat Kelulusan</b>
13650002	SPAN-PTKIN	B2
13650005	SNMPTN	B1
13650006	SNMPTN	B2
13650007	SNMPTN	B2
13650008	SNMPTN	B2
13650009	SNMPTN	B2
13650010	SNMPTN	B1
13650011	SNMPTN	B2
13650012	SPAN-PTKIN	B2
13650013	SPAN-PTKIN	B2
13650014	SPAN-PTKIN	B2
13650015	SNMPTN	B1
13650017	SNMPTN	B2
13650019	SPAN-PTKIN	B1

13650022	SPAN-PTKIN	B1
13650023	SNMPTN	B2
13650024	SNMPTN	B2
13650025	SNMPTN	B2
13650026	SNMPTN	B2
13650027	SPAN-PTKIN	A1

Langkah ketiga adalah penggalan 1-2 *itemset* dengan menggunakan struktur data *SOTrieIT* dan melakukan pembacaan dari data mahasiswa yang telah ditransformasi sebelumnya. Dalam penggalan ini akan didapatkan 1-*itemset* dan 2-*itemset* seperti pada tabel 3.4 dan tabel 3.5. Maksud dari *itemset* ini adalah jalur masuk dan kategori tingkat kelulusan dari setiap mahasiswa. Pada 1-*itemset* terdapat masing-masing jalur masuk dan kategori tingkat kelulusan, sedangkan dalam 2-*itemset* terdapat kombinasi antara jalur masuk dengan kategori tingkat kelulusan. Hasil dari ekstrak 1-*itemset* terdapat pada gambar 3.4.

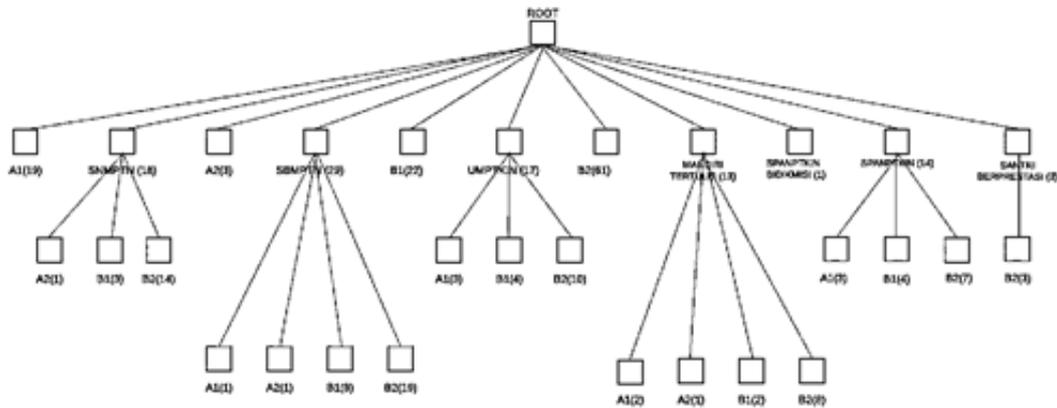


**Gambar 5.4** Pembentukan 1-*itemset* pada *SOTrieIT*

Setelah di ekstrak 1-*itemset* maka selanjutnya adalah mengekstrak 2-*itemset* berdasarkan item pada data mahasiswa. Setiap kombinasi item pada 2-*itemset*

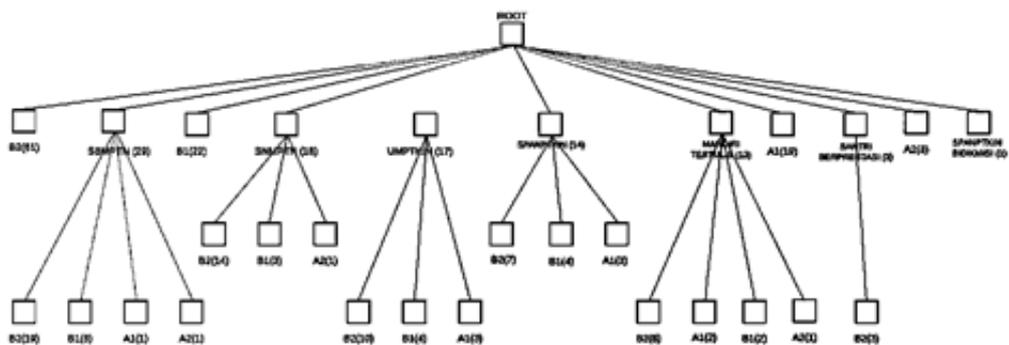
dimasukkan ke dalam *tree* satu per satu sesuai dengan induknya yaitu *1-itemset*.

Hasil *tree* setelah *2-itemset* dimasukkan ditunjukkan pada gambar 3.5.



**Gambar 5.5** Pembentukan *2-itemset* pada *SOTrieIT*

Gambar 3.6 menunjukkan bangunan *SOTrieIT* yang telah diurutkan berdasarkan *support count* terbesar baik pada *1-itemset* maupun pada *2-itemset*.



**Gambar 5.6** Pembentukan *SOTrieIT* dan diurutkan dari *support count* terbesar

Langkah selanjutnya yaitu membaca setiap *itemset* yang terdapat pada gambar 3.6 dan menghitung *support count* pada *1-itemset* dan pada *2-itemset* menggunakan rumus yang telah dijelaskan sebelumnya. Pada *1-itemset* diambil dari seluruh *itemset* dalam level 1 terdapat pada gambar 3.3, sedangkan *2-itemset* diambil dari kombinasi *itemset* pada level 2 dan jumlah *count* dari setiap *itemset*

untuk dihitung *support count*nya. Hasil *support count 1-itemset* terdapat pada tabel 3.7 dan *support count 2-itemset* terdapat pada tabel 3.8.

**Tabel 5.7** Hasil Perhitungan *Support Count 1-itemset*

<i>Itemset</i>	<i>Count</i>	<i>Support Count</i>
SNMPTN	18	0.19
SBMPTN	29	0.31
UM-PTKIN	17	0.18
SPAN-PTKIN	14	0.15
SPAN-PTKIN BIDIKMISI	1	0.01
MANDIRI TERTULIS	13	0.14
SANTRI BERPRESTASI	3	0.03
A1	9	0.09
A2	3	0.03
B1	22	0.23
B2	61	0.64

**Tabel 5.8** Hasil Perhitungan *Support 2-itemset*

<i>Itemset</i>	<i>Count</i>	<i>Support Count</i>
SNMPTN,A2	1	0.01
SNMPTN,B1	3	0.03

SNMPTN,B2	14	0.15
SBMPTN,A1	1	0.01
SBMPTN,A2	1	0.01
SBMPTN,B1	8	0.08
SBMPTN,B2	19	0.20
UM-PTKIN,A1	3	0.03
UM-PTKIN,B1	4	0.04
UM-PTKIN,B2	10	0.11
SPAN-PTKIN,A1	3	0.03
SPAN-PTKIN,B1	4	0.04
SPAN-PTKIN,B2	7	0.07
SPAN-PTKIN BIDIKMISI,B1	1	0.01
MANDIRI TERTULIS,A1	2	0.02
MANDIRI TERTULIS,A2	1	0.01
MANDIRI TERTULIS,B1	2	0.02
MANDIRI TERTULIS,B2	8	0.08
SANTRI BERPRESTASI,B2	3	0.03

Langkah selanjutnya adalah melakukan pemangkasan terhadap *itemset* yang tidak *frequent* atau tidak memenuhi syarat *minimum support*. *Minimum support* yang digunakan adalah 4% atau 0,04 sehingga ketika hasil *support* dari *itemset* tidak

memenuhi *minimum support* yang telah ditentukan maka akan dilakukan pemangkasan. Pada hasil dua tabel diatas maka bisa ditentukan *itemset* mana yang harus dipangkas. Untuk 1-*itemset*, yang tidak *frequent* adalah *itemset* SPAN-PTKIN BIDIKMISI, SANTRI BERPRESTASI, dan A2. Sementara 2-*itemset* yang tidak *frequent* adalah *itemset* SNMPTN,A2 ; SNMPTN,B1 ; SBMPTN,A1 ; SBMPTN,A2 ; UM-PTKIN,A1 ; SPAN-PTKIN,A1 ; SPAN-PTKIN BIDIKMISI,B1 ; MANDIRI TERTULIS,A1 ; MANDIRI TERTULIS,A2 ; MANDIRI TERTULIS,B1 dan SANTRI BERPRESTASI,B2. *Itemset* yang dipangkas itu karena memiliki nilai *support* dibawah 0,04 atau 4%.

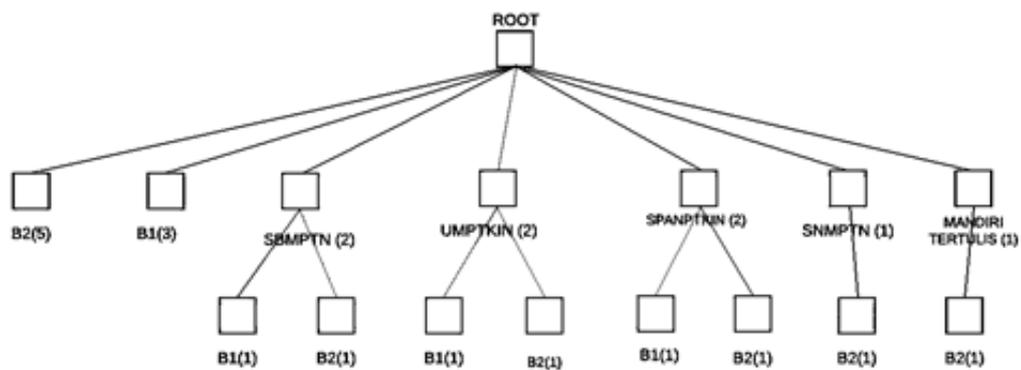
**Tabel 5.9** Pemangkasan *Support* 1-*itemset*

<b>Itemset</b>	<b>Count</b>	<b>Support Count</b>
SNMPTN	18	0.19
SBMPTN	29	0.31
UM-PTKIN	17	0.18
SPAN-PTKIN	14	0.15
MANDIRI TERTULIS	13	0.14
A1	9	0.09
B1	22	0.23
B2	61	0.64

**Tabel 5.10** Pemangkasan *Support 2-itemset*

<i>Itemset</i>	<i>Count</i>	<i>Support Count</i>
SNMPTN,B2	14	0.15
SBMPTN,B1	8	0.08
SBMPTN,B2	19	0.20
UM-PTKIN,B1	4	0.04
UM-PTKIN,B2	10	0.11
SPAN-PTKIN,B1	4	0.04
SPAN-PTKIN,B2	7	0.07
MANDIRI TERTULIS,B2	8	0.08

Hasil dari pemangkasan sebelumnya, *itemset* yang masih tersisa disusun kembali menjadi struktur *SOTrieIT* dan diurutkan berdasarkan besar nilai *support* yang paling besar, sehingga terbentuk *Ordered Frequent Items*. Hasil dari *Ordered Frequent Items* terdapat pada gambar 3.7.

**Gambar 5.7** Struktur *SOTrieIT* yang telah dipangkas dan diurutkan

Struktur *SOTrieIT* yang telah dipangkas sebelumnya bisa didapatkan kandidat 1-*itemset* dan 2-*itemset* untuk proses pembuatan *association rule* pada data mahasiswa. Kandidat 1-*itemset* terdapat pada tabel 3.11 dan kandidat 2-*itemset* terdapat pada tabel 3.12.

**Tabel 5.11** Kandidat 1-*itemset*

<i>Itemset</i>	<i>Count</i>	<i>Support Count</i>
SNMPTN	18	0.19
SBMPTN	29	0.31
UM-PTKIN	17	0.18
SPAN-PTKIN	14	0.15
MANDIRI TERTULIS	13	0.14
B1	22	0.23
B2	61	0.64

**Tabel 5.12** Kandidat 2-*itemset*

<i>Itemset</i>	<i>Count</i>	<i>Support Count</i>
SNMPTN,B2	14	0.15
SBMPTN,B1	8	0.08
SBMPTN,B2	19	0.20
UM-PTKIN,B1	4	0.04
UM-PTKIN,B2	10	0.11

SPAN-PTKIN,B1	4	0.04
SPAN-PTKIN,B2	7	0.07
MANDIRI TERTULIS,B2	8	0.08

Apabila kandidat 1-*itemset* dan 2-*itemset* sudah diketahui, langkah selanjutnya yaitu membuat *FP-Tree*. Ketika membuat *FP-Tree* diawali dengan nilai *null* pada *root*, *header table* dan kolom *head of table link* untuk semua *itemset*. Pembentukan *FP-Tree* diawal dapat dilihat pada gambar 3.8.

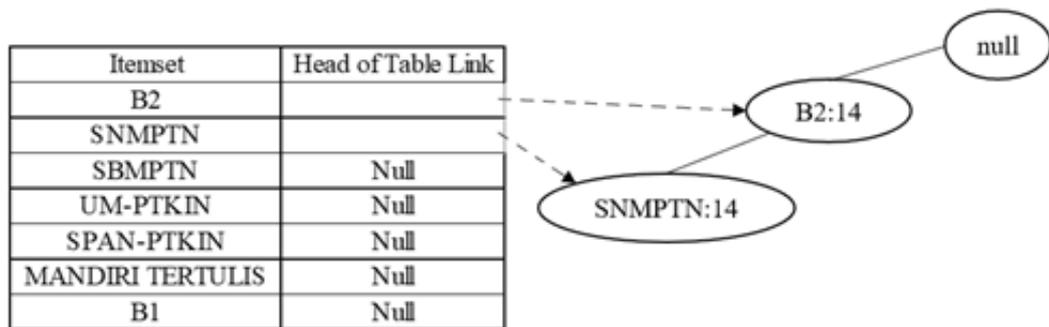
*Header Table*

Itemset	Head of Table Link
B2	Null
SNMPTN	Null
SBMPTN	Null
UM-PTKIN	Null
SPAN-PTKIN	Null
MANDIRI TERTULIS	Null
B1	Null



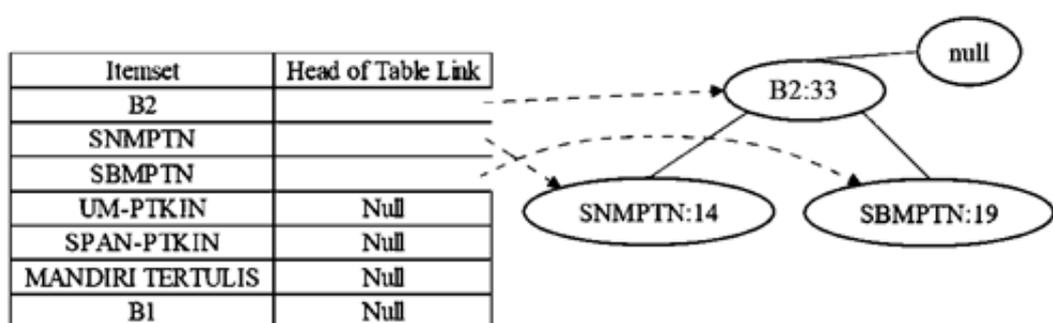
**Gambar 5.8** *FP-Tree* Awal

Kandidat pertama yaitu kategori kelulusan B2 dan SNMPTN. Kategori kelulusan B2 merupakan *prefix* sehingga simpul anak root berisi kategori kelulusan B2 dengan nilai *support count* 1, ketika SNMPTN melewati lintasan tersebut sebanyak 1 kali. SNMPTN menjadi simpul anak dari simpul B2 dan memiliki nilai *support count* 14 seperti yang terdapat pada tabel 2-*itemset*. Simbol garis panah putus-putus menunjukkan letak dimana kategori tersebut berada. *FP-Tree* dari kandidat pertama dapat dilihat pada gambar 3.9.



**Gambar 5.9** *FP-Tree* pada kandidat pertama

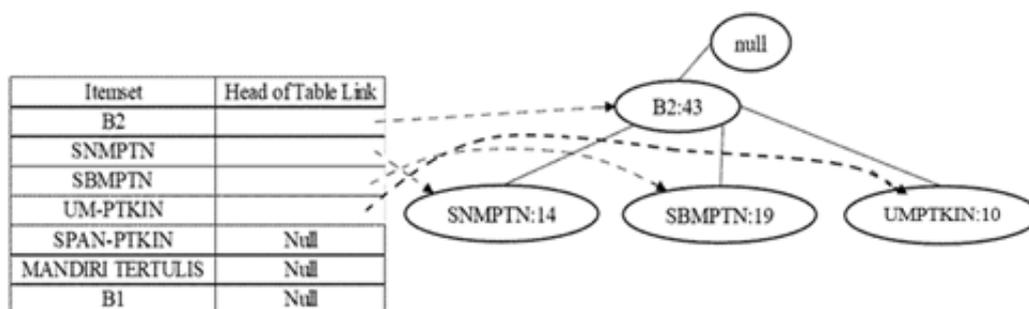
Kandidat kedua adalah transaksi dengan kategori kelulusan B2 dan SBMPTN. Kategori kelulusan B2 menjadi *prefix* dan nilai *support count* bertambah 1, sehingga nilai *support count*-nya bertambah menjadi 33. SBMPTN menjadi simpul anak dari simpul berlabel B2 dan nilai *support*-nya sebanyak 19 seperti yang terdapat pada tabel 2-*itemset*. Setelah SBMPTN masuk *tree*, jumlah *support* dari kategori kelulusan B2 menjadi 33 karena *support* dari simpul SNMPTN dan SBMPTN masing-masing memiliki *support* 14 dan 19. *FP-Tree* pada kandidat kedua dapat dilihat pada gambar 3.10.



**Gambar 5.10** *FP-Tree* pada kandidat kedua

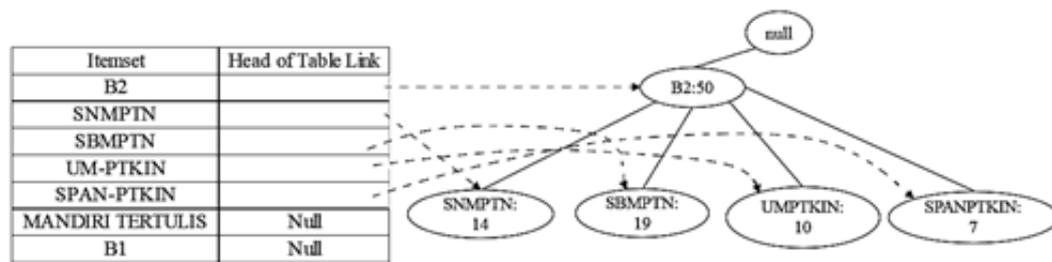
Kandidat ketiga adalah transaksi dengan kategori kelulusan B2 dan UM-PTKIN. Kategori kelulusan B2 menjadi *prefix* dan nilai *support count* bertambah 1, sehingga nilai *support count*-nya bertambah menjadi 43. UM-PTKIN menjadi

simpul anak dari simpul berlabel B2 dan nilai *support*-nya sebanyak 10 seperti yang terdapat pada tabel 2-*itemset*. Setelah UM-PTKIN masuk *tree*, jumlah *support* dari kategori kelulusan B2 menjadi 43 karena *support* dari simpul SNMPTN, SBMPTN, dan UM-PTKIN masing-masing memiliki *support* 14, 19, dan 10. Seperti pada proses sebelumnya *itemset* yang menjadi *prefix* memiliki *support* yang merupakan jumlah dari *support* masing-masing anak simpul. *FP-Tree* pada kandidat ketiga dapat dilihat pada gambar 3.11.



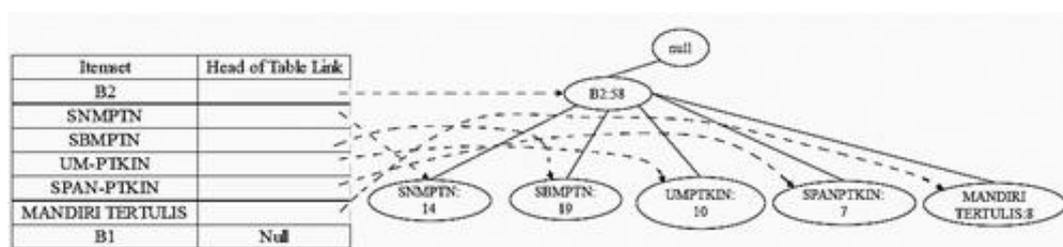
**Gambar 5.11** *FP-Tree* pada kandidat ketiga

Kandidat keempat adalah transaksi dengan kategori kelulusan B2 dan SPAN-PTKIN. Kategori kelulusan B2 menjadi *prefix* dan nilai *support count* bertambah 1, sehingga nilai *support count*-nya bertambah menjadi 50. SPAN-PTKIN menjadi simpul anak dari simpul berlabel B2 dan nilai *support*-nya sebanyak 7 seperti yang terdapat pada tabel 2-*itemset*. Setelah SPAN-PTKIN masuk *tree*, jumlah *support* dari kategori kelulusan B2 menjadi 50 karena *support* dari simpul SNMPTN, SBMPTN, UM-PTKIN, dan SPAN-PTKIN masing-masing memiliki *support* 14, 19, 10, dan 7. *FP-Tree* pada kandidat keempat dapat dilihat pada gambar 3.12.



**Gambar 5.12** *FP-Tree* pada kandidat keempat

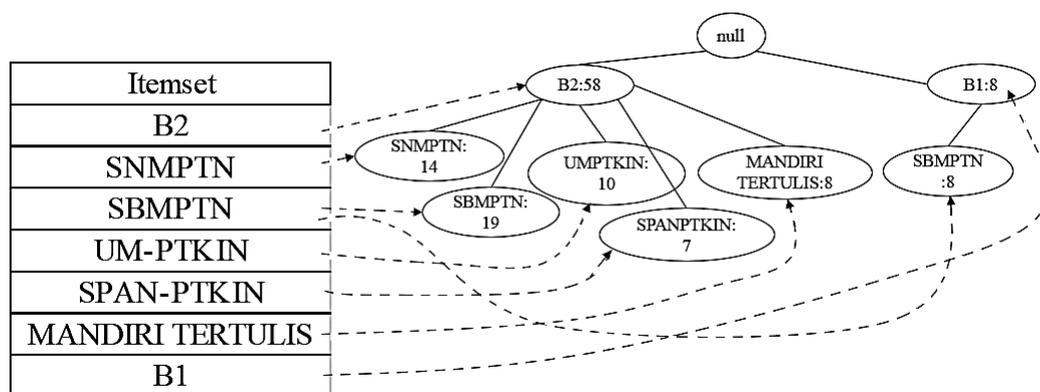
Kandidat kelima adalah transaksi dengan kategori kelulusan B2 dan MANDIRI TERTULIS. Kategori kelulusan B2 menjadi *prefix* dan nilai *support count* bertambah 1, sehingga nilai *support count*-nya bertambah menjadi 58. MANDIRI TERTULIS menjadi simpul anak dari simpul berlabel B2 dan nilai *support*-nya sebanyak 8 seperti yang terdapat pada tabel 2-*itemset*. Setelah MANDIRI TERTULIS masuk *tree*, jumlah *support* dari kategori kelulusan B2 menjadi 58 karena *support* dari simpul SNMPTN, SBMPTN, UM-PTKIN, SPAN-PTKIN dan MANDIRI TERTULIS masing-masing memiliki *support* 14, 19, 10, 7, dan 8. *FP-Tree* pada kandidat kelima dapat dilihat pada gambar 3.13.



**Gambar 5.13** *FP-Tree* pada kandidat kelima

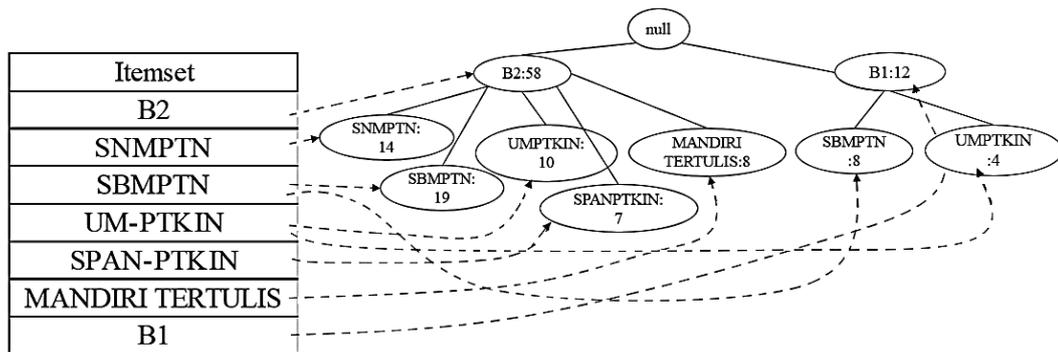
Kandidat kelima adalah transaksi dengan kategori kelulusan B1 dan SBMPTN. Kategori kelulusan B1 menjadi *prefix* dan nilai *support count* 1. SBMPTN menjadi simpul anak dari simpul berlabel B1 dan nilai *support*-nya sebanyak 8 seperti yang terdapat pada tabel 2-*itemset*. Setelah SBMPTN masuk

*tree*, jumlah *support* dari kategori kelulusan B1 menjadi 8. Pada pembacaan sebelumnya kategori kelulusan B1 belum terbentuk simpul, maka dibuat simpul baru dan menjadi anak simpul *root* dan SBMPTN menjadi anak simpul dari simpul B1. *FP-Tree* pada kandidat keenam dapat dilihat pada gambar 3.14.



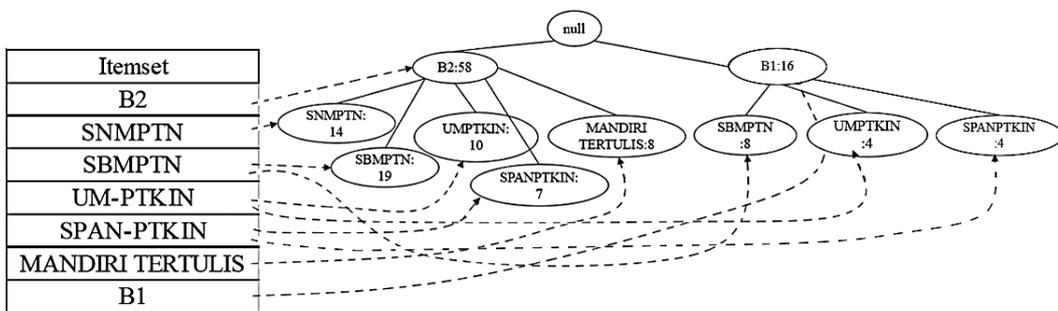
**Gambar 5.14** *FP-Tree* pada kandidat keenam

Kandidat ketujuh adalah transaksi dengan kategori kelulusan B1 dan UM-PTKIN. Kategori kelulusan B1 menjadi *prefix* dan nilai *support count* bertambah 1. UM-PTKIN menjadi simpul anak dari simpul berlabel B1 dan nilai *support*-nya sebanyak 4 seperti yang terdapat pada tabel 2-*itemset*. Setelah UM-PTKIN masuk *tree*, jumlah *support* dari kategori kelulusan B1 menjadi 12 karena *support* dari simpul SBMPTN dan UM-PTKIN masing-masing memiliki *support* 8 dan 4. *FP-Tree* pada kandidat ketujuh dapat dilihat pada gambar 3.15.



**Gambar 5.15** FP-Tree pada kandidat ketujuh

Kandidat kedelapan adalah transaksi dengan kategori kelulusan B1 dan SPAN-PTKIN. Kategori kelulusan B1 menjadi *prefix* dan nilai *support count* bertambah 1. SPAN -PTKIN menjadi simpul anak dari simpul berlabel B1 dan nilai *support*-nya sebanyak 4 seperti yang terdapat pada tabel 2-*itemset*. Setelah SPAN -PTKIN masuk *tree*, jumlah *support* dari kategori kelulusan B1 menjadi 16 karena *support* dari simpul SBMPTN, UM-PTKIN SPAN-PTKIN masing-masing memiliki *support* 8, 4, dan 4. FP-Tree pada kandidat kedelapan dapat dilihat pada gambar 3.16.



**Gambar 5.16** FP-Tree pada kandidat kedelapan

Tahap selanjutnya setelah terbentuk *FP-Tree* dari 1-*itemset* dan 2-*itemset* yaitu proses penggalian *itemset frequent* dengan menggunakan algoritma *FP-*

*Growth*. Langkah pertama adalah *conditional pattern base* yaitu mencari *suffix pattern* dan *prefix path* pada *FP-Tree* yang sebelumnya sudah terbentuk.

**Tabel 5.13** *Conditional Pattern Base* yang terbentuk

<i>Suffix</i>	<i>Conditional Pattern Base</i>
SPAN-PTKIN	{B1:4},{B2:7}
UM-PTKIN	{B1:4},{B2:10}
SBMPTN	{B1:8},{B2:19}
MANDIRI TERTULIS	{B2:8}
SNMPTN	{B2:14}

Setelah *conditional pattern base*-nya terbentuk, tahap selanjutnya adalah *conditional fp-tree* yaitu tahap dimana *support itemset* dalam *prefix path* setiap *suffix* akan dijumlahkan.

$$\text{Suffix } \{\text{SNMPTN}\} : \quad \text{Support } \{\text{B2}\} \quad = \frac{14}{95} \quad = 0.15$$

$$\text{Suffix } \{\text{MANDIRI TERTULIS}\} : \quad \text{Support } \{\text{B2}\} \quad = \frac{8}{95} \quad = 0.08$$

$$\text{Suffix } \{\text{SBMPTN}\} : \quad \text{Support } \{\text{B2}\} \quad = \frac{19}{95} \quad = 0.2$$

$$\text{Suffix } \{\text{UM -PTKIN}\} : \quad \text{Support } \{\text{B2}\} \quad = \frac{10}{95} \quad = 0.1$$

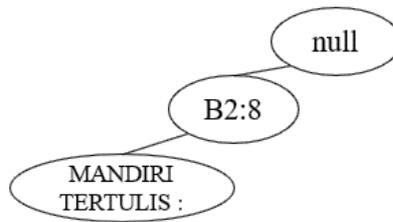
$$\text{Suffix } \{\text{SPAN-PTKIN}\} : \quad \text{Support } \{\text{B2}\} \quad = \frac{7}{95} \quad = 0.07$$

$$\text{Suffix } \{\text{SBMPTN}\} : \quad \text{Support } \{\text{B1}\} \quad = \frac{8}{95} \quad = 0.08$$

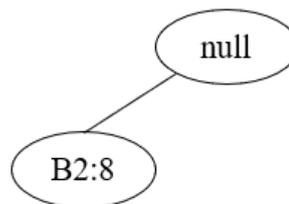
$$\text{Suffix } \{\text{UM -PTKIN}\} : \quad \text{Support } \{\text{B1}\} \quad = \frac{4}{95} \quad = 0.04$$

$$\text{Suffix } \{\text{SPAN-PTKIN}\} : \quad \text{Support } \{\text{B1}\} \quad = \frac{4}{95} \quad = 0.04$$

Gambar 3.17 merupakan gambaran lintasan yang memiliki *suffix* MANDIRI TERTULIS dan gambar 3.18 menunjukkan *conditional FP-Tree* yang dibangkitkan untuk *suffix* MANDIRI TERTULIS.



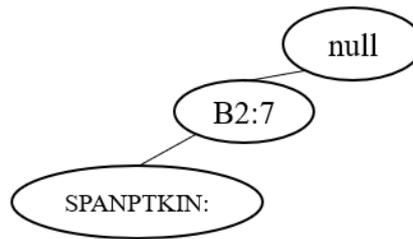
**Gambar 5.17** Lintasan yang memiliki *suffix* {MANDIRI TERTULIS}



**Gambar 5.18** Conditional *FP-Tree* yang terbentuk untuk *suffix* {MANDIRI TERTULIS}

Pada gambar 3.18, path yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B2, MANDIRI TERTULIS} dengan frekuensi sebanyak 8 kali.

Gambar 3.19 merupakan gambaran lintasan yang memiliki *suffix* SPAN-PTKIN dan gambar 3.20 menunjukkan *conditional FP-Tree* yang dibangkitkan untuk *suffix* SPAN-PTKIN.



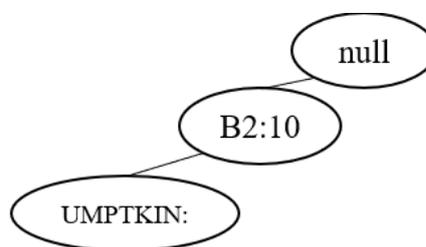
**Gambar 5.19** Lintasan yang memiliki *suffix* {SPAN-PTKIN}



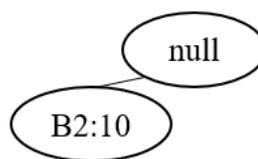
**Gambar 5.20** *Conditional FP-Tree* yang terbentuk untuk *suffix* {SPAN-PTKIN}

Pada gambar 3.20, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B2, SPAN-PTKIN} dengan *frekuensi* sebanyak 7 kali.

Gambar 3.21 merupakan gambaran lintasan yang memiliki *suffix* UM-PTKIN dan gambar 3.22 menunjukkan *conditional FP-Tree* yang dibangun untuk *suffix* UM-PTKIN.



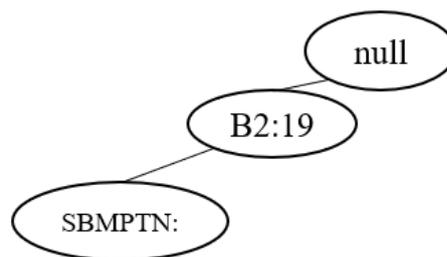
**Gambar 5.21** Lintasan yang memiliki *suffix* {UM-PTKIN}



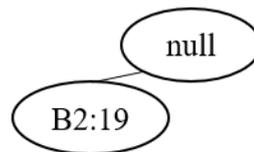
**Gambar 5.22** *Conditional FP-Tree* yang terbentuk untuk *suffix* {UM-PTKIN}

Pada gambar 3.22, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B2, UM-PTKIN} dengan *frekuensi* sebanyak 10 kali.

Gambar 3.23 merupakan gambaran lintasan yang memiliki *suffix* SBMPTN dan gambar 3.24 menunjukkan *conditional FP-Tree* yang dibangkitkan untuk *suffix* SBMPTN.



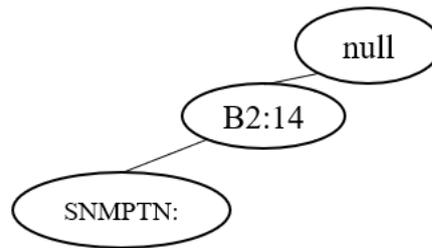
**Gambar 5.23** Lintasan yang memiliki *suffix* {SBMPTN}



**Gambar 5.24** *Conditional FP-Tree* yang terbentuk untuk *suffix* {SBMPTN}

Pada gambar 3.24, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B2, UM-PTKIN} dengan *frekuensi* sebanyak 10 kali.

Gambar 3.25 merupakan gambaran lintasan yang memiliki *suffix* SNMPTN dan gambar 3.26 menunjukkan *conditional FP-Tree* yang dibangkitkan untuk *suffix* SNMPTN.



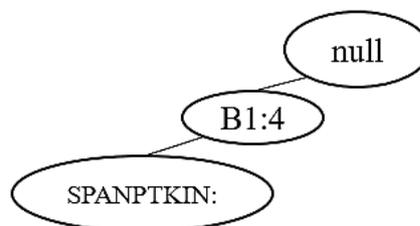
**Gambar 5.25** Lintasan yang memiliki *suffix* {SNMPTN}



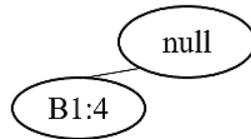
**Gambar 5.26** *Conditional FP-Tree* yang terbentuk untuk *suffix* {SNMPTN}

Pada gambar 3.26, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B2, SNMPTN} dengan *frekuensi* sebanyak 14 kali.

Gambar 3.27 merupakan gambaran lintasan yang memiliki *suffix* SPAN-PTKIN dan gambar 3.28 menunjukkan *conditional FP-Tree* yang dibangun untuk *suffix* SPAN-PTKIN.



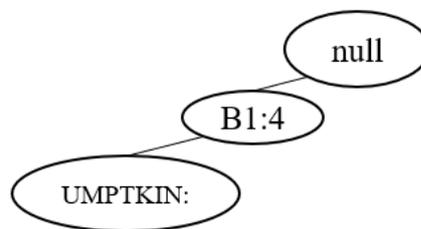
**Gambar 5.27** Lintasan yang memiliki *suffix* {SPAN-PTKIN}



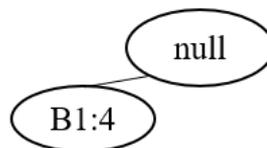
**Gambar 5.28** *Conditional FP-Tree* yang terbentuk untuk *suffix* {SPAN-PTKIN}

Pada gambar 3.28, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B1, SPAN-PTKIN} dengan *frekuensi* sebanyak 4 kali.

Gambar 3.29 merupakan gambaran lintasan yang memiliki *suffix* UM-PTKIN dan gambar 3.30 menunjukkan *conditional FP-Tree* yang dibangun untuk *suffix* UM-PTKIN.



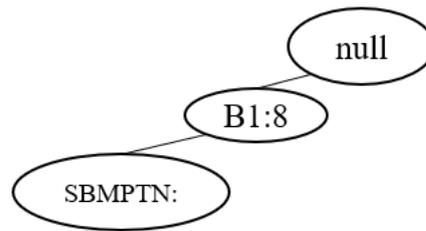
**Gambar 5.29** Lintasan yang memiliki *suffix* {UM-PTKIN}



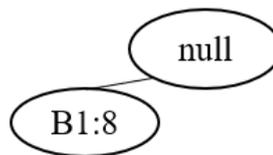
**Gambar 5.30** *Conditional FP-Tree* yang terbentuk untuk *suffix* {UM-PTKIN}

Pada gambar 3.30, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B1, UM-PTKIN} dengan *frekuensi* sebanyak 4 kali.

Gambar 3.31 merupakan gambaran lintasan yang memiliki *suffix* SBMPTN dan gambar 3.32 menunjukkan *conditional FP-Tree* yang dibangun untuk *suffix* SBMPTN.



**Gambar 5.31** Lintasan yang memiliki *suffix* {SBMPTN}



**Gambar 5.32** *Conditional FP-Tree* yang terbentuk untuk *suffix* {SBMPTN}

Pada gambar 3.32, *path* yang terbentuk adalah *single path* maka didapat *frequent itemset*-nya yaitu {B1, SBMPTN } dengan *frekuensi* sebanyak 8 kali.

Berikut adalah hasil penyusunan *frequent itemset* dari seluruh proses *FP-Tree* dapat dilihat pada tabel 3.14.

**Tabel 5.14** *Frequent Itemset* dari data

<i>Frequent Itemset</i>	<b>Frekuensi</b>	<b>Support</b>
MANDIRI TERTULIS → B2	8	0.08
SPAN-PTKIN → B2	7	0.07
UM-PTKIN → B2	10	0.11
SBMPTN → B2	19	0.20
SNMPTN → B2	14	0.15
SPAN-PTKIN → B1	4	0.04

UM-PTKIN → B1	4	0.04
SBMPTN → B1	8	0.08

Langkah selanjutnya yaitu pengecekan *frequent itemset* yang memenuhi syarat untuk dijadikan kandidat aturan/*rule* dengan cara menghitung nilai *confidence*. Jika nilai *confidence* sama atau lebih besar dari *minimum confidence* maka *frequent itemset* tersebut bisa dijadikan sebuah *rule*. Begitu juga sebaliknya, jika nilai *confidence* kurang dari *minimum confidence* yang telah ditentukan, maka *frequent itemset* tersebut tidak bisa dijadikan sebuah *rule*.

$$\text{Confidence } \{\text{MANDIRI TERTULIS}\} \rightarrow \{\text{B2}\} = 0.08/0.14 = 0.6$$

$$\text{Confidence } \{\text{SPAN-PTKIN}\} \rightarrow \{\text{B2}\} = 0.07/0.15 = 0.5$$

$$\text{Confidence } \{\text{UM-PTKIN}\} \rightarrow \{\text{B2}\} = 0.11/0.18 = 0.6$$

$$\text{Confidence } \{\text{SBMPTN}\} \rightarrow \{\text{B2}\} = 0.2/0.31 = 0.7$$

$$\text{Confidence } \{\text{SNMPTN}\} \rightarrow \{\text{B2}\} = 0.15/0.19 = 0.8$$

$$\text{Confidence } \{\text{SPAN-PTKIN}\} \rightarrow \{\text{B1}\} = 0.04/0.15 = 0.3$$

$$\text{Confidence } \{\text{UM-PTKIN}\} \rightarrow \{\text{B1}\} = 0.04/0.18 = 0.2$$

$$\text{Confidence } \{\text{SBMPTN}\} \rightarrow \{\text{B1}\} = 0.08/0.31 = 0.3$$

Setelah perhitungan selesai dilakukan, hasil dari perhitungan nilai *confidence* terdapat pada tabel 3.15.

**Tabel 5.15** Nilai *Confidence* dari *frequent itemset*

<i>Frequent Itemset</i>	<i>Support (A,B)</i>	<i>Support (A)</i>	<i>Confidence</i>
MANDIRI TERTULIS → B2	0.08	0.14	0.6
SPAN-PTKIN → B2	0.07	0.15	0.5
UM-PTKIN → B2	0.11	0.18	0.6
SBMPTN → B2	0.20	0.31	0.7
SNMPTN → B2	0.15	0.19	0.8
SPAN-PTKIN → B1	0.04	0.15	0.3
UM-PTKIN → B1	0.04	0.18	0.2
SBMPTN → B1	0.08	0.31	0.3

Hasil dari perhitungan nilai *confidence* menunjukkan bahwa *frequent itemset* yang lolos adalah MANDIRI TERTULIS → B2, SPAN-PTKIN → B2, UM-PTKIN → B2, SBMPTN → B2, SNMPTN → B2, SPAN-PTKIN → B1, dan SBMPTN → B1 karena masing-masing *frequent itemset* memenuhi syarat untuk dibangkitkan menjadi sebuah *rule*, yaitu nilai *confidence*-nya memenuhi *minimum support* yang telah ditentukan. *Rule* yang terbentuk terdapat pada tabel 3.16.

**Tabel 5.16** *Rule* yang bisa dibangkitkan

<i>Rules</i>	<i>Keterangan</i>

MANDIRI TERTULIS → B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2
SPAN-PTKIN → B2	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B2
UM-PTKIN → B2	Jika jalur masuk UM-PTKIN maka kategori kelulusannya B2
SBMPTN → B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2
SNMPTN → B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2
SPAN-PTKIN → B1	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B1
SBMPTN → B1	Jika jalur masuk SBMPTN maka kategori kelulusannya B1

Langkah berikutnya yaitu *rule* yang sudah terbentuk akan dilakukan pengujian kekuatan *rule* menggunakan *lift ratio*. *Lift ratio* didapatkan dari membandingkan antara nilai *confidence* dengan *benchmark confidence*. Hasil dari pengujian *rule* terdapat pada tabel 3.17.

$$Lift\ Ratio\ \{MANDIRI\ TERTULIS\} \rightarrow \{B2\} = \frac{0.6}{0.08} = 7.31$$

$$Lift\ Ratio\ \{SPAN-PTKIN\} \rightarrow \{B2\} = \frac{0.5}{0.07} = 6.79$$

$$Lift\ Ratio\ \{UM-PTKIN\} \rightarrow \{B2\} = \frac{0.6}{0.11} = 5.59$$

$$\text{Lift Ratio } \{SBMPTN\} \rightarrow \{B2\} = 0.7/0.2 = 3.28$$

$$\text{Lift Ratio } \{SNMPTN\} \rightarrow \{B2\} = 0.8/0.15 = 5.28$$

$$\text{Lift Ratio } \{SPAN-PTKIN\} \rightarrow \{B1\} = 0.3/0.04 = 6.79$$

$$\text{Lift Ratio } \{SBMPTN\} \rightarrow \{B1\} = 0.3/0.08 = 3.28$$

**Tabel 5.17** *Lift Ratio Rules*

<i>Rules</i>	<i>Confidence</i>	<i>Frekuensi Item Consequent</i>	<i>Benchmark Confidence</i>	<i>Lift Ratio</i>
MANDIRI TERTULIS → B2	0.6	8	0.08	7.31
SPAN-PTKIN → B2	0.5	7	0.07	6.79
UM-PTKIN → B2	0.6	10	0.11	5.59
SBMPTN → B2	0.7	19	0.20	3.28
SNMPTN → B2	0.8	14	0.15	5.28
SPAN-PTKIN → B1	0.3	4	0.04	6.79
SBMPTN → B1	0.3	8	0.08	3.28

Pada tabel 3.17 maka dapat dilihat bahwa semua *rule* yang dihasilkan mempunyai tingkat kekuatan yang bagus karena *lift ratio*-nya lebih besar dari 1, jadi dapat disimpulkan bahwa *rule* memberikan manfaat.

## **BAB IV**

### **HASIL DAN PEMBAHASAN**

Pada bab ini menjelaskan tentang implementasi algoritma *FOLD-Growth* dalam membangun sistem Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa pada Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang.

#### **6.1 Implementasi**

Pada tahap ini proses penerapan perancangan ke dalam sistem secara utuh yang dibuat berdasarkan algoritma yang digunakan dengan bahasa pemrograman komputer, berbasis *website*, dan sesuai dengan kebutuhan perangkat lunak (*Software*) dan perangkat keras (*Hardware*).

##### **6.1.1 Implementasi Sistem**

Dibawah ini adalah kebutuhan yang diperlukan dalam tahap implementasi sistem.

###### **6.1.1.1 Kebutuhan Perangkat Keras (*Hardware*)**

- Prosesor : Intel(R) Core(TM) i3-7020U
- RAM : 4.00 GB
- HDD : 1 TB

###### **6.1.1.2 Kebutuhan Perangkat Lunak (*Software*)**

- *MySQL*
- *PHP*

- *Sublime Text*
- *Web Browser*

### 6.1.2 Implementasi Algoritma *FOLD-Growth*

Pada implementasi algoritma *FOLD-Growth* ini akan dijelaskan sesuai dengan alur program sebagai berikut.

1. *Source code* menampilkan data mahasiswa dengan import menggunakan data *Excel* ke dalam *MySQL*.

```
$target = basename($_FILES['file']['name']);
move_uploaded_file($_FILES['file']['tmp_name'], $target);

chmod($_FILES['file']['name'], 0777);

$data = new
    Spreadsheet_Excel_Reader($_FILES['file']['name'], false);
$jumlah_baris = $data->rowcount($sheet_index = 0);
```

2. *Source code* penggalan L1 dengan *SOTrieIT*.

```
$data = mysqli_query($koneksi, "SELECT * FROM itemset_1
    ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
    $id = $row[0];
    $itemset = $row[1];
    $count = $row[2];
    $support = $row[3];
    $persen = $support * 100;
```

3. *Source code* penggalan L2 dengan *SOTrieIT*.

```
$data = mysqli_query($koneksi, "SELECT * FROM itemset_2
    ORDER BY id ASC ;");
```

```

no = 1;
while ($row = mysqli_fetch_array($data)) {
    $id = $row[0];
    $itemset = $row[1];
    $count = $row[2];
    $support = $row[3];
    $persen = $support * 100;
}

```

4. *Source code pemangkasan 1-itemset yang tidak frequent dengan min. sup 0,04.*

```

$data = mysqli_query($koneksi, "SELECT * FROM itemset_1
where support>=0.04 ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
    $id = $row[0];
    $itemset = $row[1];
    $count = $row[2];
    $support = $row[3];
    $persen = $support * 100;
}

```

5. *Source code pemangkasan 2-itemset yang tidak frequent dengan min. sup 0,04.*

```

$data = mysqli_query($koneksi, "SELECT * FROM itemset_2
where support_a_b>=0.04 ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
    $id = $row[0];
    $itemset = $row[1];
    $count = $row[2];
    $support = $row[3];
    $persen = $support * 100;
}

```

6. *Source code* kandidat 1-*itemset* dari pemangkasan struktur *SOTrieIT*.

```

$data = mysqli_query($koneksi, "SELECT * FROM itemset_1
  where support>=0.04 AND itemset NOT LIKE 'A1' ORDER BY
  id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
  $id = $row[0];
  $itemset = $row[1];
  $count = $row[2];
  $support = $row[3];
  $persen = $support * 100;

```

7. *Source code* kandidat 2-*itemset* dari pemangkasan struktur *SOTrieIT*.

```

$data = mysqli_query($koneksi, "SELECT * FROM itemset_2
  where support_a_b>=0.04 ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
  $id = $row[0];
  $itemset = $row[1];
  $count = $row[2];
  $support = $row[3];
  $persen = $support * 100;

```

8. *Source code* pengecekan *frequent itemset* yang memenuhi syarat (kandidat aturan/*rule*) menghitung *nilai confidence*.

```

$data = mysqli_query($koneksi, "SELECT * FROM itemset_2
  where support_a_b>=0.04 ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
  $id = $row[0];
  $itemset = $row[1];
  $count = $row[2];
  $support_a_b = $row[3];

```

```

$support_a = $row[4];
$confidence = number_format($support_a_b / $support_a,
2);
$tambah_conf = "INSERT INTO conf (id, itemset,
frekuensi, support_a_b, support_a, conf) VALUE
('$no', '$itemset', '$count', '$support_a_b', '$support_a', '
$confidence')";
$conf = mysqli_query($koneksi, $tambah_conf);

```

### 9. Source code rule yang terbentuk dengan *min. confidence* 0.3.

```

$data = mysqli_query($koneksi, "SELECT * FROM conf where
conf>=0.3 ORDER BY id ASC ;");
$no = 1;
while ($row = mysqli_fetch_array($data)) {
    $id = $row[0];
    $itemset = $row[1];
    $count = $row[2];
    $support_a_b = $row[3];
    $support_a = $row[4];
    $conf = $row[5];
}

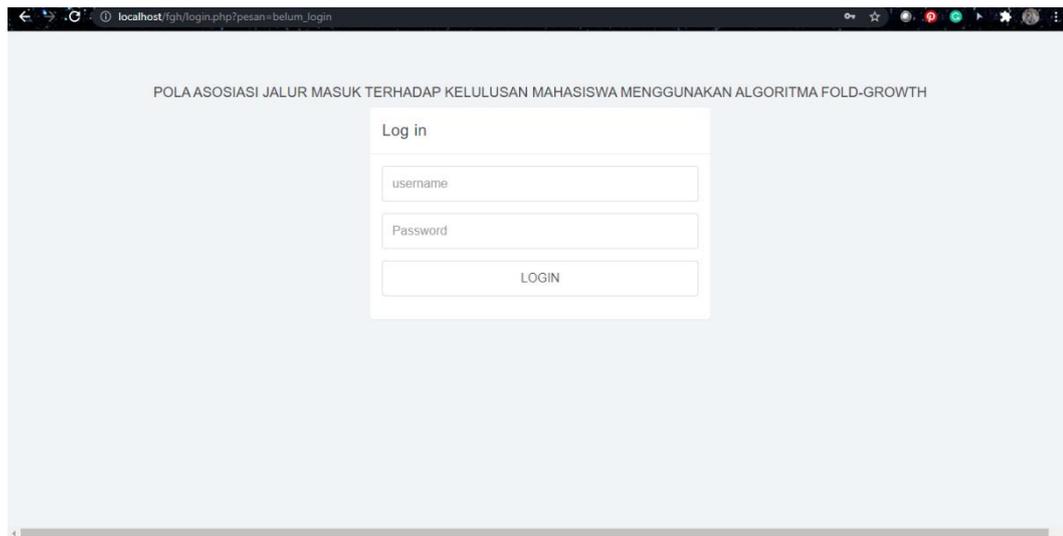
```

## 6.1.3 Implementasi Interface

Pada implementasi ini akan ditampilkan interface sistem Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa pada Jurusan Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang

### 6.1.3.1 Bagian Login

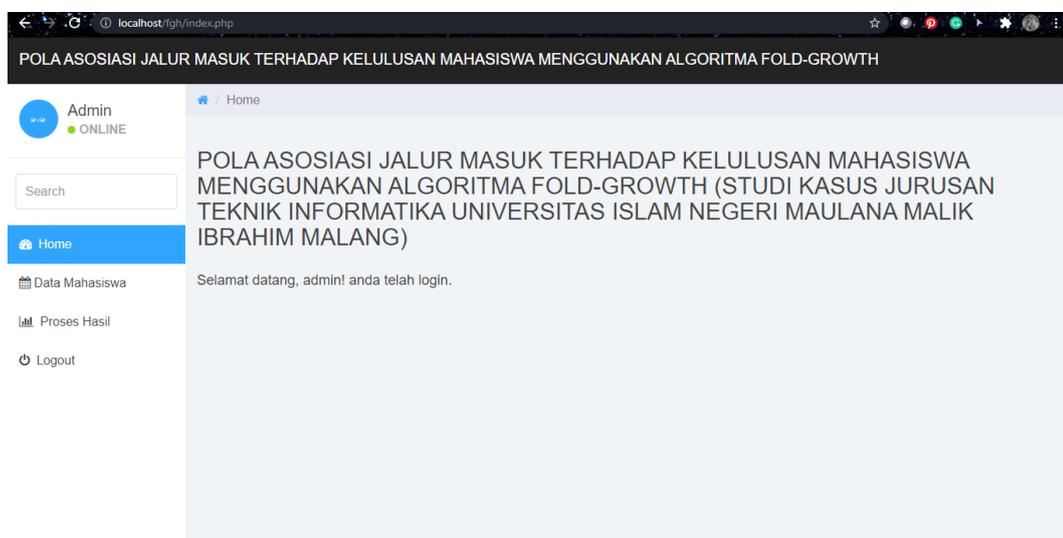
Pada sistem ini memiliki form login yang harus diisi terlebih dahulu sebelum masuk ke dalam sistem oleh admin. Admin harus menginputkan *username* dan *password* yang telah tersedia. Adanya form *login* ini agar sistem terjaga keamanan datanya.



**Gambar 6.1** Tampilan *Login*

### 6.1.3.2 Tampilan Home

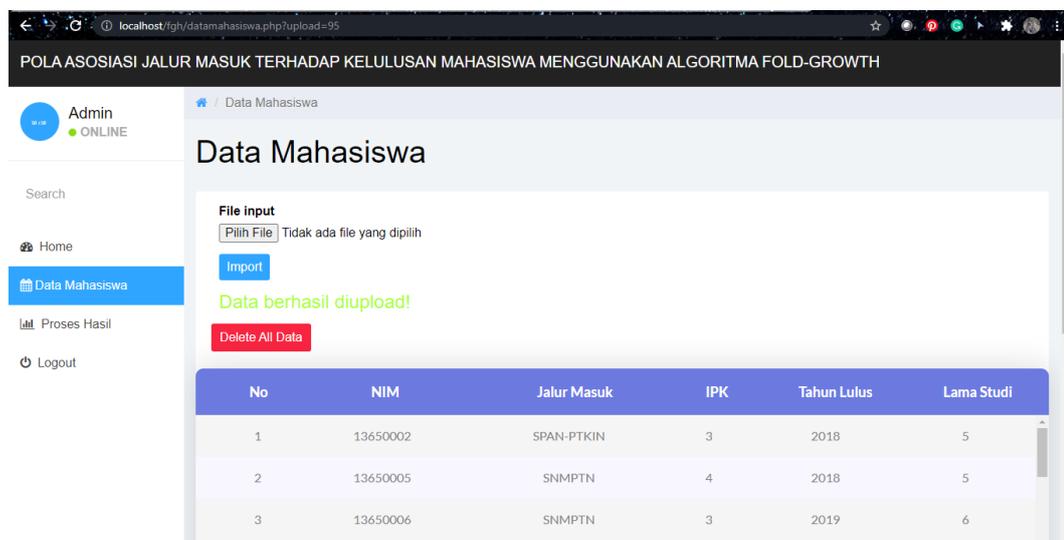
Setelah login, maka sistem akan otomatis masuk ke halaman awal atau *home*. Pada halaman awal ini menampilkan informasi sistem ini, dan *menu bar* yang terdiri dari *button* Data Mahasiswa, Proses Hasil, dan *Logout* untuk keluar dari sistem.



**Gambar 6.2** Tampilan *Home*

### 6.1.3.3 Tampilan Data Mahasiswa

Tampilan Data Mahasiswa ini terdapat button Pilih File untuk menginputkan file data excel, lalu klik button Import untuk mengimportkan data ke dalam database dan ditampilkan pada tabel dibawah. Apabila data berhasil di import ke dalam database maka akan muncul pesan “Data berhasil diupload!”. Button Delete All Data berfungsi untuk menghapus data yang sudah diimport di database.



**Gambar 6.3** Tampilan Data Mahasiswa

### 6.1.3.4 Tampilan Proses Hasil

Data yang sudah diimport sebelumnya akan diolah untuk menghasilkan rule, dengan klik button Proses pada tab Proses Hasil. Hasil dari proses *FOLD-Growth* dapat dilihat pada gambar 4.5.

POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA FOLD-GROWTH

2	13650005	SNMPTN	4	2018	5
3	13650006	SNMPTN	3	2019	6
4	13650007	SNMPTN	3	2020	7
5	13650008	SNMPTN	3	2018	5
6	13650009	SNMPTN	3	2019	6
7	13650010	SNMPTN	4	2018	5
8	13650011	SNMPTN	3	2019	6
9	13650012	SPAN-PTKIN	3	2019	6
10	13650013	SPAN-PTKIN	3	2018	5
11	13650014	SPAN-PTKIN	3	2020	7

Proses

**Gambar 6.4** Tampilan Proses Hasil

POLA ASOSIASI JALUR MASUK TERHADAP KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA FOLD-GROWTH

Proses Fold-Growth

Step, Penggalan L1 dengan SOTriEIT

No	Item Set	Count	Support
1	SNMPTN	18	19%
2	SBMPTN	29	31%
3	UM-PTKIN	17	18%
4	SPAN-PTKIN	14	15%
5	SPAN-PTKIN BIDIKMISI	1	1%
6	MANDIRI TERTULIS	13	14%
7	SANTRI BERPRESTASI	3	3%

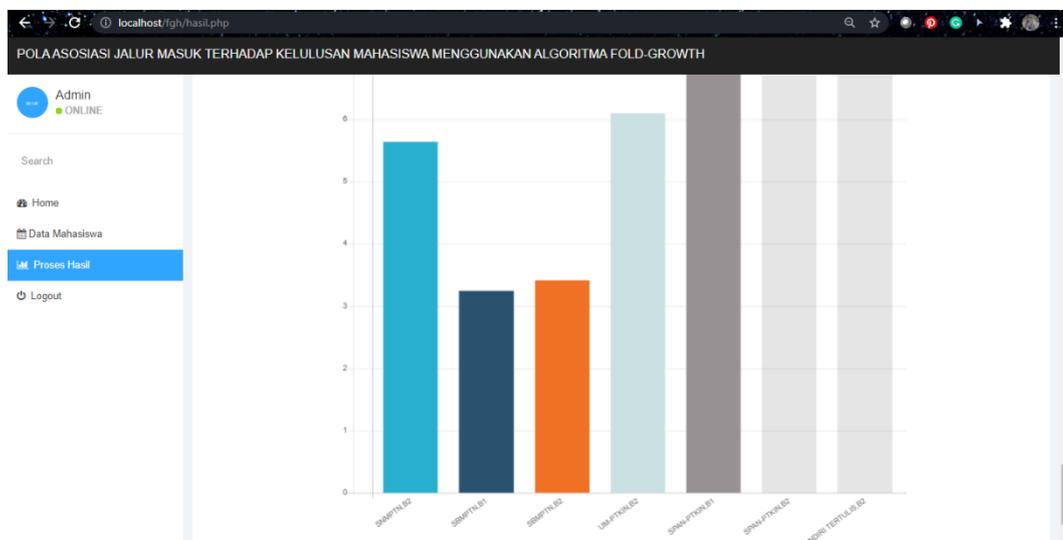
**Gambar 6.5** Tampilan Hasil Proses *FOLD-Growth*

Hasil dari proses *FOLD-Growth* berupa tabel *rule* dan keterangan dari *rule* tersebut dan grafik *rule* dengan nilai *lift ratio* dapat dilihat pada gambar 4.6 dan 4.7.



No	Rules	Keterangan
1	SNMPTN => B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2
2	SBMPTN => B1	Jika jalur masuk SBMPTN maka kategori kelulusannya B1
3	SBMPTN => B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2
4	UM-PTKIN => B2	Jika jalur masuk UM-PTKIN maka kategori kelulusannya B2
5	SPAN-PTKIN => B1	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B1
6	SPAN-PTKIN => B2	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B2
7	MANDIRI TERTULIS => B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2

**Gambar 6.6** Tampilan Hasil *Rule*



**Gambar 6.7** Tampilan Hasil Grafik *Rule* dan Nilai *Lift Ratio*

## 6.2 Pengujian

Pada penelitian ini pengujian dilakukan terhadap 95 data mahasiswa yang terdapat data pertama masuk universitas dan data kelulusan mahasiswa jurusan Teknik Informatika angkatan 2013. Pengujian pada penelitian ini ada 2 tahap. Pengujian pertama yaitu pengujian untuk mengetahui pengaruh *minimum support*

dan *minimum confidence* terhadap jumlah *rule* yang dihasilkan. Pengujian kedua yaitu menguji kekuatan *rule* yang dihasilkan dengan *lift ratio* yang didapat.

Pengujian pertama, *minimum support* yang digunakan adalah sejumlah 4% dan 8%, sedangkan untuk *minimum confidence* yang digunakan adalah 30% dan 60%. Penetapan *minimum support* dan *minimum confidence* dalam pengujian ini diharapkan dapat memberikan suatu kesimpulan pada *association rule* yang terbentuk dan supaya hasil dari *itemset* yang didapatkan cukup banyak untuk dianalisis.

Pada pengujian kedua ini dilakukan dengan membandingkan *confidence* dengan *benchmark confidence*. Jika nilai dari *lift ratio* tersebut lebih dari 1, maka menunjukkan *rule* yang dihasilkan memiliki manfaat. Semakin tinggi nilai *lift ratio* yang dihasilkan maka semakin besar kekuatan *association rule*.

### **6.2.1 Pengujian Pengaruh *Minimum Support* dan *Minimum Confidence***

*Association rule* yang dihasilkan dapat diketahui seberapa pentingnya dapat dilihat dari 2 parameter, yaitu *support* dan *confidence*. *Support* atau nilai penunjang merupakan persentase hasil kombinasi *item* yang memiliki fungsi untuk mengukur seberapa tingkat dukungan data terhadap validitas *rule* yang akan dikembangkan. Sedangkan *confidence* atau nilai kepastian merupakan persentase kuatnya suatu hubungan antar item dalam *association rule*. Hasil uji pengaruh nilai *minimum support* dan nilai *minimum confidence* terhadap jumlah *rule* yang dihasilkan terdapat pada tabel 4.1.

**Tabel 6.1** Hasil Pengujian Pengaruh *Minimum Support* dan *Minimum Confidence*

<b>Jumlah Mahasiswa</b>	<b><i>Minimum Support</i> (%)</b>	<b><i>Minimum Confidence</i> (%)</b>	<b>Jumlah Rule yang dihasilkan</b>
95	4	30	7
		60	4
	8	30	4
		60	3

Hasil pengujian diatas *rule* yang terbentuk pada setiap masing-masing *minimum support* dan *minimum confidence* terdapat pada tabel 4.2.

**Tabel 6.2** Hasil *Rule* dari Masing-Masing *Minimum Support* dan *Minimum Confidence*

<b><i>Minimum Support</i></b>	<b><i>Minimum Confidence</i></b>	<b><i>Rule</i> yang terbentuk</b>	<b>Keterangan</b>
4	30	MANDIRI TERTULIS → B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2
		SPAN-PTKIN → B2	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B2
		UM-PTKIN → B2	Jika jalur masuk UM-PTKIN maka kategori kelulusannya B2

		SBMPTN → B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2
		SNMPTN → B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2
		SPAN-PTKIN → B1	Jika jalur masuk SPAN-PTKIN maka kategori kelulusannya B1
		SBMPTN → B1	Jika jalur masuk SBMPTN maka kategori kelulusannya B1
	60	MANDIRI TERTULIS → B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2
		UM-PTKIN → B2	Jika jalur masuk UM-PTKIN maka kategori kelulusannya B2
		SBMPTN → B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2

		SNMPTN → B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2
8	30	SBMPTN → B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2
		SNMPTN → B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2
		UM-PTKIN → B2	Jika jalur masuk UM-PTKIN maka kategori kelulusannya B2
		MANDIRI TERTULIS → B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2
	60	SBMPTN → B2	Jika jalur masuk SBMPTN maka kategori kelulusannya B2
		SNMPTN → B2	Jika jalur masuk SNMPTN maka kategori kelulusannya B2

		MANDIRI TERTULIS → B2	Jika jalur masuk MANDIRI TERTULIS maka kategori kelulusannya B2
--	--	--------------------------	--

### 6.2.2 Pengujian *Lift Ratio*

Dalam pengujian ini *rule* yang akan diuji yaitu *rule* yang telah memenuhi *minimum confidence* yang telah ditentukan pada tabel diatas. Hasil pengujian *lift ratio* terdapat pada tabel 4.3.

**Tabel 6.3** Hasil Pengujian *Lift Ratio*

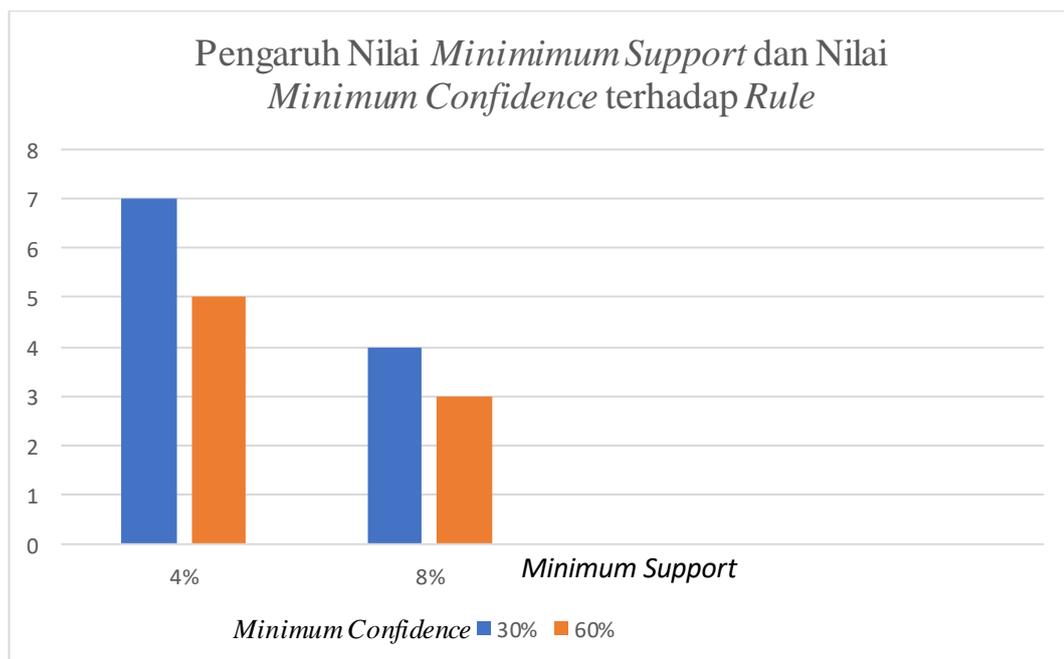
<i>Rules</i>	<i>Confidence</i>	<i>Lift Ratio</i>
MANDIRI TERTULIS → B2	0.6	7.31
SPAN-PTKIN → B2	0.5	6.79
UM-PTKIN → B2	0.6	5.59
SBMPTN → B2	0.7	3.28
SNMPTN → B2	0.8	5.28
SPAN-PTKIN → B1	0.3	6.79
SBMPTN → B1	0.3	3.28

### 6.3 Pembahasan

Hasil dari pengujian di atas tentang pengaruh *minimum support* dan *minimum confidence* yang telah dilakukan pada 95 data mahasiswa didapatkan bahwa jumlah yang dihasilkan berbanding terbalik. Semakin tinggi nilai *minimum*

*support* dan *minimum confidence* maka semakin rendah pula jumlah *association rule* yang dihasilkan terdapat pada gambar 4.8.

*Minimum support* digunakan untuk menetapkan batas persentase kemunculan kombinasi *item* yang harus dicapai, sedangkan *minimum confidence* untuk menetapkan batas *confidence* terhadap *frequent itemset* agar dapat menghasilkan *association rule*. Semakin tinggi nilai *minimum support* dan *minimum confidence* maka semakin tinggi pula nilai dari batas minimal dari kombinasi *item* yang dilakukan pada proses penggalian data sehingga semakin rendah mencapai *frequent itemset* dan *rule* yang dihasilkan akan semakin sedikit. Hasil *rule* yang dihasilkan dapat dilihat pada tabel 4.4.



**Gambar 6.8** Grafik Pengaruh Nilai *Minimum Support* dan Nilai *Minimum Confidence* Terhadap Jumlah *Rule*

Tabel 6.4 Rule yang dihasilkan

<i>Minimum Support (%)</i>	<i>Minimum Confidence (%)</i>	<i>Jumlah Rule</i>	<i>Rule</i>	<i>Support</i>	<i>Confidence</i>
4	30	7 rule	MANDIRI TERTULIS → B2	8%	62%
			SPAN- PTKIN → B2	7%	50%
			UM-PTKIN → B2	11%	59%
			SBMPTN → B2	20%	66%
			SNMPTN → B2	15%	78%
			SPAN- PTKIN → B1	4%	29%
			SBMPTN → B1	8%	28%
	60	4 rule	MANDIRI TERTULIS → B2	8%	62%

			UM-PTKIN → B2	11%	59%
			SBMPTN → B2	20%	66%
			SNMPTN → B2	15%	78%
8	30	4 rule	SBMPTN → B2	20%	66%
			SNMPTN → B2	15%	78%
			UM-PTKIN → B2	11%	59%
			MANDIRI TERTULIS → B2	8%	62%
	60	3 rule	SBMPTN → B2	20%	66%
			SNMPTN → B2	15%	78%
			MANDIRI TERTULIS → B2	8%	62%

Pada tabel 4.4 dapat diketahui bahwa *rule* yang dihasilkan dari nilai *minimum support* 4% dan nilai *minimum confidence* 30% berjumlah 7 *rule*. Dapat diartikan bahwa dalam *database* memuat minimal 4% dari masing-masing *item*

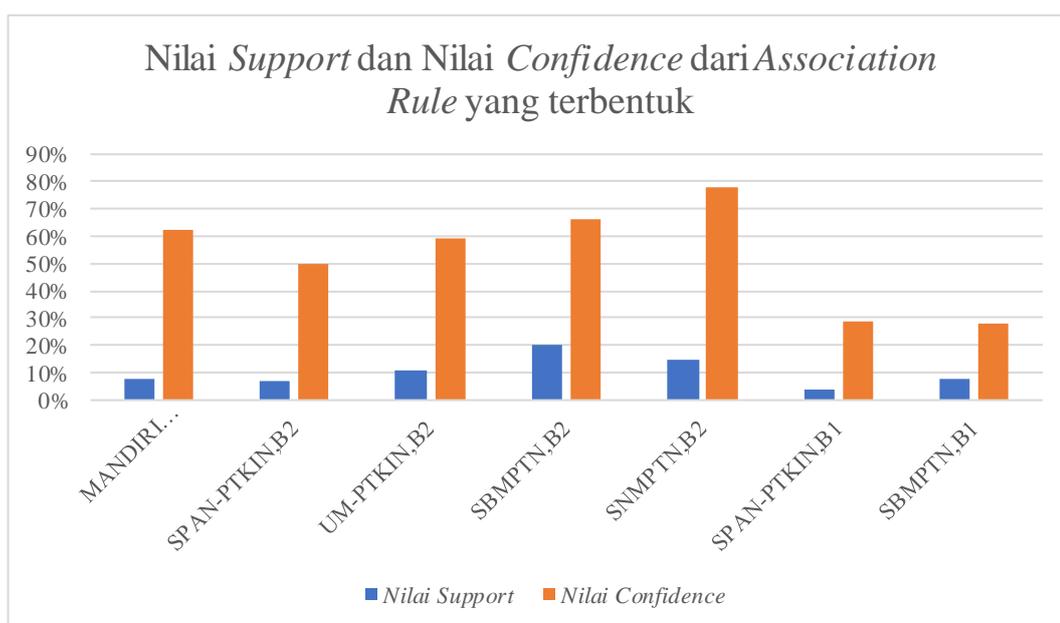
dalam *rule* tersebut dan 30% dari keseluruhan kombinasi dalam *database* yang memuat kombinasi 2 *item* tersebut. Contohnya pada *rule* MANDIRI TERTULIS → B2 yang memiliki nilai *support* 8% dan nilai *confidence* 62%, dapat diartikan ada 8% dari kombinasi data mahasiswa yang terdapat dalam *database* yang memuat *item* dan MANDIRI TERTULIS dan B2. Nilai *confidence* 62% menunjukkan kemunculan kombinasi 2 *item* tersebut dari keseluruhan data dalam *database*. Dapat juga diartikan bahwa jika mahasiswa masuk jalur MANDIRI TERTULIS memiliki kemungkinan masuk dalam kategori kelulusan B2 sebesar 8% dan *rule* ini cukup relevan karena mewakili 62% dari *database* mahasiswa.

Apabila nilai *support* dinaikkan ke *minimum support* 8%, maka *rule* yang dihasilkan makin sedikit menjadi 4 *rule*. Bahkan jika nilai *minimum confidence* dinaikkan ke 60% maka *rule* yang dihasilkan berjumlah 3 *rule*. Hal ini dikarenakan kombinasi dari 2 *item* ini tidak ada yang memenuhi *minimum support* dan *minimum confidence*.

Pada tabel 4.4 dapat dilihat dari keseluruhan *association rule* yang dihasilkan terdapat 18 *rule* dengan penentuan *minimum support* 4% dan 8%. Sedangkan untuk *minimum confidence* 30% dan 60%. Dari keseluruhan *rule* yang dihasilkan, *rule* yang memiliki *support* tertinggi adalah kombinasi SBMPTN → B2 yang memiliki nilai *support* 20% dan nilai *confidence* 66%. Dapat disimpulkan bahwa *rule* kombinasi SBMPTN → B2 adalah yang paling banyak dalam *database* daripada *rule* yang lain. *Rule* tersebut yaitu “Jika jalur masuk mahasiswa SBMPTN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50 (B2)”. *Rule* SNMPTN → B2 memiliki nilai *support* 15% dan nilai *confidence* tertinggi yaitu 78%, ini

menjelaskan bahwa pada *rule* ini memiliki tingkat kepercayaan paling tinggi daripada *rule* yang lain. *Rule* tersebut yaitu “Jika jalur masuk mahasiswa SNMPTN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50 (B2)”.

Grafik nilai *support* dan nilai *confidence* dari *association rule* lebih rincinya terdapat pada gambar 4.9.

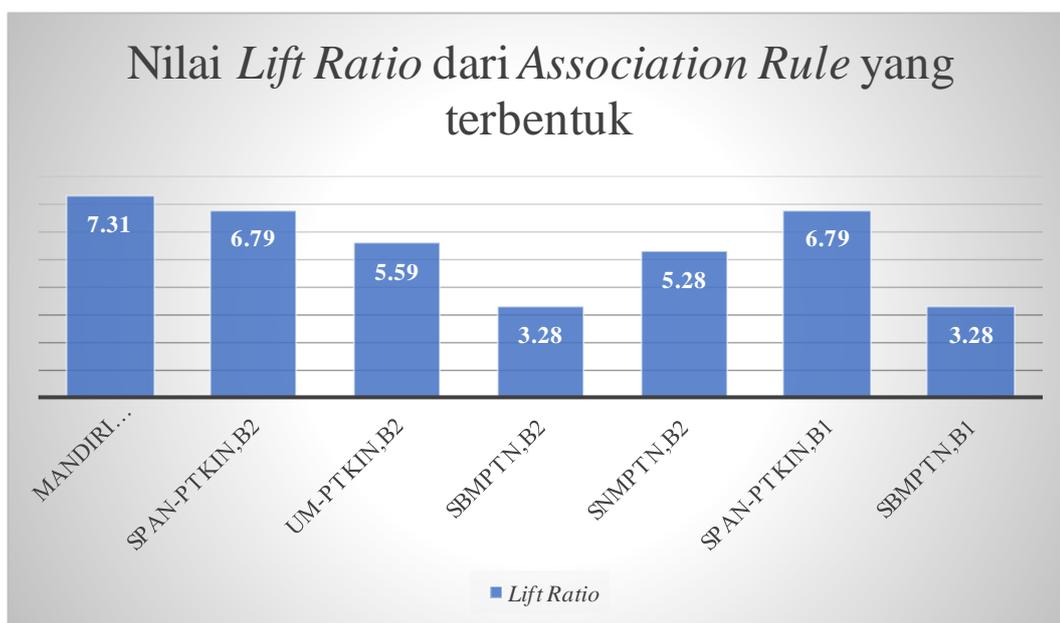


**Gambar 6.9** Grafik Nilai *Support* dan Nilai *Confidence* dari *Association Rule* yang terbentuk

Pada pengujian kekuatan *rule* menggunakan *lift ratio* yang terdapat dalam tabel 4.3 dapat dilihat hasil nilai *lift ratio* dari masing-masing *rule* yang dihasilkan. Nilai *lift ratio* dapat diambil dari membandingkan *confidence* suatu *rule* dengan *benchmark confidence* atau *support* dari *consequent*. Nilai *lift ratio* lebih dari 1 menunjukkan bahwa keterikatan antara dua *item*, yaitu *consequent* dan *antecedent* yang kuat. Dikarenakan nilai *confidence rule* yang melebihi *minimum confidence*

dilakukan pengujian lagi pada persentase dengan kombinasi *antecedent* (Sari et al, 2012).

Dalam tabel 4.3 dapat dilihat bahwa nilai *lift ratio* tertinggi yaitu pada *rule* MANDIRI TERTULIS, B2 dengan nilai *lift ratio* sebesar 7,31 dan yang terendah adalah pada *rule* SBMPTN, B2 dan SBMPTN, B1 dengan nilai *lift ratio* sebesar 3,28. Pada tabel 4.3 juga menunjukkan bahwa semua *rule* memiliki nilai *lift ratio* lebih dari 1, sehingga semua *rule* yang dihasilkan menunjukkan bahwa *item consequent* dan *antecedent* memiliki keterikatan yang kuat. Grafik nilai *lift ratio* dari setiap *rule* terdapat pada gambar 4.10.



**Gambar 6.10** Grafik Lift Ratio dari Association Rule yang terbentuk

Hasil *rule* yang dihasilkan sejumlah 7 *rule* yang terdapat pada tabel 4.5. *Rule* tersebut didapatkan dari *rule* yang memiliki nilai *lift ratio* lebih dari 1 maka dapat dipastikan bahwa ada keterikatan antara *item consequent* dan *antecedent* dan dapat juga disimpulkan *rule* tersebut bermanfaat.

**Tabel 6.5** Hasil *Rule* yang dihasilkan

<i>Rule</i>	<b>Keterangan</b>
MANDIRI TERTULIS → B2	Jika jalur masuk mahasiswa MANDIRI TERTULIS maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50
SPAN-PTKIN → B2	Jika jalur masuk mahasiswa SPAN-PTKIN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50
UM-PTKIN → B2	Jika jalur masuk mahasiswa UM-PTKIN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50
SBMPTN → B2	Jika jalur masuk mahasiswa SBMPTN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50
SNMPTN → B2	Jika jalur masuk mahasiswa SNMPTN maka lama studi lebih dari 4 tahun dan IPK 2,76 - 3,50
SPAN-PTKIN → B1	Jika jalur masuk mahasiswa SPAN-PTKIN maka Lama studi lebih dari 4 tahun dan IPK 3,51 - 4,00
SBMPTN → B1	Jika jalur masuk mahasiswa SBMPTN maka Lama studi lebih dari 4 tahun dan IPK 3,51 - 4,00

#### 6.4 Integrasi Islam

Integrasi penelitian ini terhadap Islam yaitu adanya sistem pola asosiasi jalur masuk mahasiswa terhadap kelulusan ini dapat membantu dan memberikan manfaat kepada sesama, sesuai dalam potongan ayat Al-Qur'an Surat Al-Isra' ayat 7

...إِنْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنْفُسِكُمْ وَإِنْ أَسَأْتُمْ فَلَهَا

*“Jika kalian berbuat baik, sesungguhnya kalian berbuat baik bagi diri kalian sendiri. Dan jika kamu berbuat jahat, maka (kerugian kejahatan) itu untuk dirimu sendiri.” (QS. Al-Isra’ ayat 7)*

Maksud dari potongan ayat diatas menurut Tafsir Jalalayn adalah Kemudian Kami katakan (Jika kalian berbuat baik) dengan mengerjakan ketaatan (berarti kalian berbuat baik bagi diri kalian sendiri) karena sesungguhnya pahala kebaikan itu untuk diri kalian sendiri (dan jika kalian berbuat jahat) dengan menimbulkan kerusakan (maka kejahatan itu bagi diri kalian sendiri) sebagai pembalasan atas kejahatan kalian. Sedangkan menurut Tafsir Ibnu Katsir Artinya, jika kalian berbuat kejahatan, maka akibatnya akan menimpa diri kalian sendiri. Makna ayat ini sama dengan apa yang disebutkan dalam ayat lain melalui firman-Nya : Barang siapa yang mengerjakan amal saleh, maka (pahalanya) untuk dirinya sendiri, dan barang siapa yang berbuat jahat, maka (dosanya) atas dirinya sendiri. (QS. Fushshilat [41]: 46)

Terdapat juga hadits tentang membantu sesama saudara Rasulullah ﷺ bersabda :

وَمَنْ كَانَ فِي حَاجَةِ أَخِيهِ كَانَ اللَّهُ فِي حَاجَتِهِ ...

*“ ... dan barangsiapa (yang bersedia) membantu keperluan saudaranya, maka Allah (akan senantiasa) membantu keperluannya.” (Hadits Riwayat Bukhari, Shahîh al-Bukhâriy, juz III, hal. 168, hadits no. 2442 dan Muslim, Shahîh Muslim, juz VIII, hal. 18, hadits no. 6743 dari Abdullah bin Umar رَضِيَ اللهُ عَنْهُ).*

Rasulullah ﷺ bersabda dalam hadits yang lain:

مَنْ نَفَسَ عَنْ مُؤْمِنٍ كُرْبَةً مِنْ كُرْبِ الدُّنْيَا نَفَسَ اللَّهُ عَنْهُ كُرْبَةً مِنْ كُرْبِ يَوْمِ الْقِيَامَةِ وَمَنْ يَسَّرَ عَلَى مُعْسِرٍ يَسَّرَ اللَّهُ عَلَيْهِ فِي الدُّنْيَا وَالْآخِرَةِ وَمَنْ سَتَرَ مُسْلِمًا سَتَرَهُ اللَّهُ فِي الدُّنْيَا وَالْآخِرَةِ وَاللَّهُ فِي عَوْنِ الْعَبْدِ مَا كَانَ الْعَبْدُ فِي عَوْنِ أَخِيهِ وَمَنْ سَلَكَ طَرِيقًا يَلْتَمِسُ فِيهِ عِلْمًا سَهَّلَ اللَّهُ لَهُ بِهِ طَرِيقًا إِلَى الْجَنَّةِ وَمَا اجْتَمَعَ قَوْمٌ فِي بَيْتٍ مِنْ بُيُوتِ اللَّهِ يَتْلُونَ كِتَابَ اللَّهِ وَيَتَدَارَسُونَ بَيْنَهُمْ إِلَّا نَزَلَتْ عَلَيْهِمُ السَّكِينَةُ وَعَشِيَتْهُمُ الرَّحْمَةُ وَحَفَّتْهُمُ الْمَلَائِكَةُ وَذَكَرَهُمُ اللَّهُ فِيمَنْ عِنْدَهُ وَمَنْ بَطَأَ بِهِ عَمَلُهُ لَمْ يُسْرِعْ بِهِ نَسَبُهُ

*“Barangsiapa membebaskan seorang mukmin dari suatu kesulitan dunia, maka Allah akan membebaskannya dari suatu kesulitan pada hari kiamat. Barangsiapa memberi kemudahan kepada orang yang berada dalam kesulitan, maka Allah akan memberikan kemudahan di dunia dan akhirat. Barangsiapa menutupi aib seorang muslim, maka Allah akan menutup aibnya di dunia dan akhirat. Allah akan selalu menolong hamba-Nya selama hamba tersebut menolong saudaranya sesama muslim. Barangsiapa menempuh jalan untuk mencari ilmu, maka Allah akan memudahkan jalan ke surga baginya. Tidaklah sekelompok orang berkumpul di suatu masjid (rumah Allah) untuk membaca al-Qur’an, melainkan mereka akan diliputi ketenangan, rahmat, dan dikelilingi para malaikat, serta Allah akan menyebut-nyebut mereka pada malaikat-malaikat yang berada di sisi-Nya. Barangsiapa yang ketinggalan amalnya, maka nasabnya tidak juga meninggikannya.” (Hadits Riwayat Muslim, Shahîh Muslim, juz VIII, hal. 71, hadits no. 7028, dari Abu Hurairah ﷺ).*

Integrasi penelitian ini terhadap Islam yaitu adanya sistem pola asosiasi jalur masuk mahasiswa terhadap kelulusan ini dapat membantu dan memberikan manfaat kepada sesama, sesuai dalam potongan ayat Al-Qur’an Surat Al-Isra’ ayat 7:

...إِنْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنْفُسِكُمْ وَإِنْ أَسَأْتُمْ فَلَهَا

*“Jika kalian berbuat baik, sesungguhnya kalian berbuat baik bagi diri kalian sendiri. Dan jika kamu berbuat jahat, maka (kerugian kejahatan) itu untuk dirimu sendiri.” (QS. Al-Isra’ ayat 7)*

Terdapat juga hadits tentang membantu sesama saudara Rasulullah ﷺ

bersabda :

وَمَنْ كَانَ فِي حَاجَةِ أَخِيهِ كَانَ اللَّهُ فِي حَاجَتِهِ ...

“.... dan barangsiapa (yang bersedia) membantu keperluan saudaranya, maka Allah (akan senantiasa) membantu keperluannya.” (Hadits Riwayat Bukhari, *Shahîh al-Bukhâriy*, juz III, hal. 168, hadits no. 2442 dan Muslim, *Shahîh Muslim*, juz VIII, hal. 18, hadits no. 6743 dari Abdullah bin Umar رضي الله عنه).

## BAB V

### KESIMPULAN DAN SARAN

#### 7.1 Kesimpulan

Proses penggalian data pada penelitian ini dengan melakukan pengujian dapat diambil manfaat dari penerapan algoritma *FOLD-Growth. Rule* yang dihasilkan dapat dijadikan acuan sebagai pedoman untuk melakukan peningkatan keunggulan dari suatu perguruan tinggi. Sehingga dapat disimpulkan dari hasil pengujian dan analisis pada bab 5 bahwa :

1. Penerapan algoritma *FOLD-Growth* pada data mahasiswa untuk mengetahui pola asosiasi dari data jalur masuk dan data kelulusan mahasiswa berdasarkan kategori kelulusan mahasiswa terdapat 4 langkah yaitu Penggalian L1 dan L2 dengan menggunakan struktur data *SOTrieIT*, pemangkasan item-item yang tidak *frequent*, Membangun *FP-Tree* dan yang terakhir penggalian *itemset frequent*. Parameter *input* yang digunakan dalam *minimum support* dan *minimum confidence* sangat berpengaruh untuk mendapatkan *rule* dengan nilai *confidence* dan *lift ratio* yang tinggi.
2. Penerapan algoritma *FOLD-Growth* pada data mahasiswa ini menghasilkan *rule* yang memiliki *lift ratio* (tingkat kekuatan) *association rule* terbesar yang dihasilkan yaitu 7,31 dan yang terkecil yaitu 3,28. *Lift ratio* yang dihasilkan dapat disimpulkan bahwa *rule* yang bermanfaat bisa dilihat dari 3 parameter yaitu nilai *minimum support*, nilai *minimum confidence*, dan nilai *lift ratio*. Pengaruh nilai *minimum support* dan nilai *minimum confidence* terhadap *rule* yang dihasilkan berbanding terbalik ini dapat dilihat dari jumlah pola asosiasi

yang didapatkan. Hal ini dikarenakan semakin tinggi batasan nilai *minimum support* dan nilai *minimum confidence* maka semakin sedikit pula *rule* yang dihasilkan. Jumlah pola asosiasi dari *rule* yang dihasilkan sebanyak 7 *rule*.

## 7.2 Saran

Berdasarkan dari penujian terhadap penelitian diatas, ada beberapa saran penulis untuk penelitian selanjutnya agar sistem ini dapat dikembangkan menjadi lebih baik, yaitu penggunaan data mahasiswa yang lebih besar lagi sehingga data bisa cukup untuk dianalisis dan *rule* yang dihasilkan akan memiliki nilai *confidence*, nilai *support*, dan nilai *lift ratio* yang tinggi.

## DAFTAR PUSTAKA

- Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rule. International Conference on Very Large Data Bases (hal. 487-499). San Jose: IBM Almaden Research Center.
- Attamimi, A. R., Nirmala, I., & Putri, D. V. (2018). Statistik Pendidikan Tinggi 2019. Jakarta: Pusat Data dan Informasi IPTEK DIKTI.
- Beynon, P., & Davies. (New York). Database Systems Third Edition. 2004: PALGRAVE MACMILLAN.
- Budhi, G. S., Santoso, L. W., & Susanto, E. (2006). Metode Market Basket Analysis menggunakan Algoritma Pincer Search untuk Sistem Pembantu Pengambilan Keputusan.
- Erwin. (2009). Analisis Market Basket Dengan Algoritma Apriori dan FP-Growth. JURNAL GENERIC, 26-30.
- Febrilia, B. A., Yani, H. M., & Anwar, S. (2020). Analisis Regresi Logistik Ordinal Faktor-Faktor yang Mempengaruhi Status Beasiswa Mahasiswa di Universitas Mataram. Jurnal Ilmu Matematika dan Terapan, 227-232.
- Fitria, R., Nengsih, W., & Qudsi, D. H. (2017). Implementasi Algoritma FP-Growth dalam Penentuan Pola Hubungan Kecelakaan Lalu Lintas. Jurnal Sistem Informasi (Journal of Information Systems), 118-124.
- Han, J., & Kamber, M. (2006). Data Mining Concepts and Techniques Second Edition. California: Morgan Kaufmann.
- Han, J., Kamber, M., & Pei, J. (2011). Data Mining : Concepts and Techniques (Third Edition). Waltham: Morgan Kaufmann.
- Hardianti, B. D., Sabariah, M. K., & Akbar, A. G. (2014). Implementasi Algoritma FOLD-Growth pada Pemodelan Pola Pembelian Barang sebagai Pendukung Penentuan Letak Barang. Seminar Nasional Ilmu Komputer dan Teknik Informatika.

- Irfan, M. (2015). Analisa Pola Asosiasi Jalur Masuk Terhadap Kelulusan Mahasiswa dengan Menggunakan Metode Fold-Growth (Studi Kasus Fakultas Sains dan Teknologi).
- Naskah Akademik Instrumen Akreditasi Perguruan Tinggi 3.0. (2019). Jakarta: Badan Akreditasi Nasional Perguruan Tinggi.
- Noranita, B., & Bahtiar, N. (2010). IMPLEMENTASI DATA MINING UNTUK MENEMUKAN POLA HUBUNGAN TINGKAT KELULUSAN MAHASISWA DENGAN DATA INDUK MAHASISWA. Seminar dan Call For Paper Munas Aptikom, 156-163.
- Pramudiono, I. (2003). Pengantar Data Mining:Menambang Permata Pengetahuan di Gunung Data.
- Rahmi, I., & Yozza, H. (2020). Analisis Kausal Masa Studi Mahasiswa Program Studi Matematika Universitas Andalas dengan Menggunakan Metode Cart. Jurnal Matematika Sains dan Teknologi, 22-34.
- Santosa, B. (2007). Teknik Pemanfaatan Data untuk Keperluan Bisnis. Yogyakarta: Graha Ilmu.
- Sari, B. N. (2012). Penerapan Data Mining untuk Mengetahui Pola Asosiasi Antara Data Mahasiswa dan Tingkat Kelulusan Menggunakan Algoritma FOLD-Growth.
- Sari, B. N., Rahman, M. A., & Mursityo, Y. T. (2012). Penerapan Data Mining Untuk Mengetahui Pola Asosiasi Antara Data Mahasiswa Dan Tingkat Kelulusan Menggunakan Algoritma FOLD-Growth.
- Scott, G. M. (1996). Principles Of Management Information Systems chapter 16. New York: McGraw-Hill.
- Soelaiman, R., & Arini WP, N. (2006). Analisis Kinerja Algoritma Fold Growth dan FP-Growth pada Penggalan Pola Asosiasi. Seminar Nasional Aplikasi Teknologi Informasi 2006 (SNATI 2006), 13-18.
- Soelaiman, R., & Arini WP, N. (2006). Analisis Kinerja Algoritma Fold-Growth dan FP-Growth pada PEnggalan Pola Asosiasi. Seminar Nasional Aplikasi dan Teknologi Informasi 2006 (SNATI 2006), 13-18.

- Soelaiman, R., & WP, N. A. (2006). Analisis Kinerja Algoritma FOLD-Growth dan FP-Growth pada Penggalian Pola Asosiasi. Seminar Nasional Aplikasi Teknologi Informasi .
- Sutradana, G. C., & Wahyudi, M. (2017). Penerapan Data Mining untuk Analisis Pengaruh Lama Studi Mahasiswa Teknik Informatika UIN Sunan Kalijaga Yogyakarta Menggunakan Metode Apriori. JISKA (Jurnal Informatika Sunan Kalijaga), 153-162.
- Woon, Y. K., Keong Ng, W., & Das, A. (2002). Fast Online Dynamic Association Rule Mining. Proc. Second Int'l Conference, 278-287.
- Woon, Y. K., Ng, W. K., & Lim, E. P. (2004). A Support-Ordered Trie for Fast Frequent Itemset Discovery. IEEE Transactions on Knowledge and Data Engineering, Vol. 16, No. 7, pg 875-879.
- Zahrotun, L., Setiadi, T., & Haryadi, T. M. (2018). Aplikasi Data Mining untuk Mencari Pola Asosiasi Tracer Study Menggunakan Algoritma FOLDARM. Jurnal Nasional Teknologi dan Sistem Informasi, 37-43.
- Zhao, Q., & Bhowmick, S. S. (2003). Association Rule Mining: A Survey.

## LAMPIRAN

### 1. Tabel Sample Data Mahasiswa

<b>NIM</b>	<b>Jalur Masuk</b>	<b>Status</b>	<b>No. Ijazah</b>	<b>SKS Total</b>	<b>IPK</b>	<b>Kondisi</b>
13650001	SNMPTN	Mutasi / Pernah Studi	-	151	3.20	Selesai
13650002	SPAN-PTKIN	Lulus	22646/UIN- S1/F.06/2018	160	3.41	Selesai
13650003	SNMPTN	Mutasi / Pernah Studi	-	111	1.88	Selesai
13650004	SNMPTN	Mutasi / Pernah Studi	-	19	3.08	Selesai
13650005	SNMPTN	Lulus	22647/UIN- S1/F.06/2018	160	3.53	Selesai
13650006	SNMPTN	Lulus	26219/UIN- S1/F.06/2019	159	3.18	Selesai
13650007	SNMPTN	Lulus	27715/UIN- S1/F.06/2020	154	2.84	Selesai
13650008	SNMPTN	Lulus	21870/UIN- S1/F.06/2018	158	3.44	Selesai

13650009	SNMPTN	Lulus	23963/UIN-S1/F.06/2019	156	3.41	Selesai
13650010	SNMPTN	Lulus	22648/UIN-S1/F.06/2018	160	3.53	Selesai
13650011	SNMPTN	Lulus	26220/UIN-S1/F.06/2019	157	3.10	Selesai
13650012	SPAN-PTKIN	Lulus	23964/UIN-S1/F.06/2019	158	3.32	Selesai
13650013	SPAN-PTKIN	Lulus	22649/UIN-S1/F.06/2018	159	3.38	Selesai
13650014	SPAN-PTKIN	Lulus	28057/UIN-S1/F.06/2020	158	3.22	Selesai
13650015	SNMPTN	Lulus	21863/UIN-S1/F.06/2018	159	3.65	Selesai
13650016	SNMPTN	Mutasi / Pernah Studi	-	57	1.93	Selesai
13650017	SNMPTN	Lulus	24776/UIN-S1/F.06/2019	159	3.27	Selesai
13650018	SNMPTN	Non Aktif	-	157	3.30	Kritis
13650019	SPAN-PTKIN	Lulus	21864/UIN-S1/F.06/2018	159	3.52	Selesai

13650020	SNMPTN	Non Aktif	-	149	3.17	Kritis
13650021	SNMPTN	Aktif	-	157	3.14	Kritis
13650022	SPAN-PTKIN	Lulus	21865/UIN-S1/F.06/2018	158	3.57	Selesai
13650023	SNMPTN	Lulus	26916/UIN-S1/F.06/2020	158	3.30	Selesai
13650024	SNMPTN	Lulus	22650/UIN-S1/F.06/2018	160	3.36	Selesai
13650025	SNMPTN	Lulus	23965/UIN-S1/F.06/2019	158	3.45	Selesai
13650026	SNMPTN	Lulus	23966/UIN-S1/F.06/2019	160	3.36	Selesai
13650027	SPAN-PTKIN	Lulus	20233/UIN-S1/F.06/2017	158	3.64	Selesai
13650028	SPAN-PTKIN	Mutasi / Pernah Studi	-	119	2.55	Selesai
13650029	SNMPTN	Lulus	26917/UIN-S1/F.06/2020	159	3.43	Selesai
13650030	SNMPTN	Lulus	20234/UIN-S1/F.06/2017	160	3.32	Selesai

## 2. Tabel Sample Data Mahasiswa Pertama Masuk dan Data Kelulusan hasil

*Preprocessing*

<b>NIM</b>	<b>Jalur Masuk</b>	<b>IPK</b>	<b>Tahun Lulus</b>	<b>Lama Studi</b>
13650002	SPAN-PTKIN	3.41	2018	5
13650005	SNMPTN	3.53	2018	5
13650006	SNMPTN	3.18	2019	6
13650007	SNMPTN	2.84	2020	7
13650008	SNMPTN	3.44	2018	5
13650009	SNMPTN	3.41	2019	6
13650010	SNMPTN	3.53	2018	5
13650011	SNMPTN	3.1	2019	6
13650012	SPAN-PTKIN	3.32	2019	6
13650013	SPAN-PTKIN	3.38	2018	5
13650014	SPAN-PTKIN	3.22	2020	7
13650015	SNMPTN	3.65	2018	5
13650017	SNMPTN	3.27	2019	6
13650019	SPAN-PTKIN	3.52	2018	5
13650022	SPAN-PTKIN	3.57	2018	5
13650023	SNMPTN	3.3	2020	7
13650024	SNMPTN	3.36	2018	5
13650025	SNMPTN	3.45	2019	6
13650026	SNMPTN	3.36	2019	6
13650027	SPAN-PTKIN	3.64	2017	4
13650029	SNMPTN	3.43	2020	7
13650030	SNMPTN	3.32	2017	4
13650031	SNMPTN	3.44	2019	6
13650032	SPAN-PTKIN	3.59	2017	4
13650033	SPAN-PTKIN	3.23	2018	5

13650036	SPAN-PTKIN BIDIKMISI	3.62	2018	5
13650037	SNMPTN	3.28	2020	7
13650039	SPAN-PTKIN	3.58	2018	5
13650041	SPAN-PTKIN	3.51	2018	5
13650043	SPAN-PTKIN	3.44	2018	5
13650046	SNMPTN	3.39	2020	7
13650047	SPAN-PTKIN	3.42	2018	5
13650049	SPAN-PTKIN	3.56	2017	4
13650051	SBMPTN	3.37	2019	6
13650052	SBMPTN	3.61	2018	5
13650053	SBMPTN	3.39	2018	5
13650054	SBMPTN	3.66	2018	5
13650055	SBMPTN	3.43	2018	5
13650056	SBMPTN	3.53	2020	7
13650058	SBMPTN	3.33	2020	7
13650060	SBMPTN	3.25	2020	7
13650061	SBMPTN	3.62	2018	5
13650063	SBMPTN	3.61	2018	5
13650064	SBMPTN	3.44	2018	5
13650065	UM-PTKIN	3.43	2018	5
13650066	SBMPTN	3.59	2017	4
13650067	SBMPTN	3.5	2018	5
13650069	UM-PTKIN	3.32	2019	6
13650070	UM-PTKIN	3.69	2017	4
13650071	SBMPTN	3.31	2019	6
13650072	SBMPTN	3.49	2018	5
13650073	SBMPTN	3.56	2018	5
13650074	SBMPTN	3.39	2018	5
13650075	UM-PTKIN	3.36	2019	6
13650076	SBMPTN	3.48	2019	6

13650078	UM-PTKIN	3.53	2019	6
13650079	UM-PTKIN	3.14	2019	6
13650081	UM-PTKIN	3.09	2019	6
13650082	SBMPTN	3.16	2020	7
13650083	UM-PTKIN	3.65	2018	5
13650084	SBMPTN	3.39	2019	6
13650085	UM-PTKIN	3.72	2018	5
13650087	UM-PTKIN	3.59	2018	5
13650090	SBMPTN	3.69	2018	5
13650092	UM-PTKIN	3.19	2019	6
13650093	SBMPTN	3.25	2019	6
13650094	UM-PTKIN	3.27	2019	6
13650095	SBMPTN	3.37	2018	5
13650096	UM-PTKIN	3.36	2019	6
13650098	UM-PTKIN	3.14	2019	6
13650099	SBMPTN	3.61	2018	5
13650100	UM-PTKIN	3.48	2019	6
13650104	UM-PTKIN	3.78	2017	4
13650105	UM-PTKIN	3.52	2017	4
13650108	SBMPTN	3.34	2019	6
13650109	SBMPTN	3.33	2017	4
13650110	SBMPTN	3.08	2019	6
13650113	MANDIRI TERTULIS	3.05	2019	6
13650114	MANDIRI TERTULIS	3.74	2018	5
13650115	MANDIRI TERTULIS	3.47	2018	5
13650116	MANDIRI TERTULIS	3.74	2017	4
13650118	MANDIRI TERTULIS	3.59	2017	4
13650119	MANDIRI TERTULIS	3.36	2018	5
13650120	SBMPTN	3.32	2018	5
13650121	MANDIRI TERTULIS	3.51	2018	5

## 3. Tabel Sample Data Mahasiswa dan Kategori Kelulusan

<b>NIM</b>	<b>Jalur Masuk</b>	<b>Kategori Tingkat Kelulusan</b>
13650002	SPAN-PTKIN	B2
13650005	SNMPTN	B1
13650006	SNMPTN	B2
13650007	SNMPTN	B2
13650008	SNMPTN	B2
13650009	SNMPTN	B2
13650010	SNMPTN	B1
13650011	SNMPTN	B2
13650012	SPAN-PTKIN	B2
13650013	SPAN-PTKIN	B2
13650014	SPAN-PTKIN	B2
13650015	SNMPTN	B1
13650017	SNMPTN	B2
13650019	SPAN-PTKIN	B1
13650022	SPAN-PTKIN	B1
13650023	SNMPTN	B2
13650024	SNMPTN	B2
13650025	SNMPTN	B2
13650026	SNMPTN	B2
13650027	SPAN-PTKIN	A1
13650029	SNMPTN	B2
13650030	SNMPTN	A2
13650031	SNMPTN	B2
13650032	SPAN-PTKIN	A1
13650033	SPAN-PTKIN	B2
13650036	SPAN-PTKIN BIDIKMISI	B1
13650037	SNMPTN	B2

13650039	SPAN-PTKIN	B1
13650041	SPAN-PTKIN	B1
13650043	SPAN-PTKIN	B2
13650046	SNMPTN	B2
13650047	SPAN-PTKIN	B2
13650049	SPAN-PTKIN	A1
13650051	SBMPTN	B2
13650052	SBMPTN	B1
13650053	SBMPTN	B2
13650054	SBMPTN	B1
13650055	SBMPTN	B2
13650056	SBMPTN	B1
13650058	SBMPTN	B2
13650060	SBMPTN	B2
13650061	SBMPTN	B1
13650063	SBMPTN	B1
13650064	SBMPTN	B2
13650065	UM-PTKIN	B2
13650066	SBMPTN	A1
13650067	SBMPTN	B2
13650069	UM-PTKIN	B2
13650070	UM-PTKIN	A1
13650071	SBMPTN	B2
13650072	SBMPTN	B2
13650073	SBMPTN	B1
13650074	SBMPTN	B2
13650075	UM-PTKIN	B2
13650076	SBMPTN	B2
13650078	UM-PTKIN	B1
13650079	UM-PTKIN	B2

13650081	UM-PTKIN	B2
13650082	SBMPTN	B2
13650083	UM-PTKIN	B1
13650084	SBMPTN	B2
13650085	UM-PTKIN	B1
13650087	UM-PTKIN	B1
13650090	SBMPTN	B1
13650092	UM-PTKIN	B2
13650093	SBMPTN	B2
13650094	UM-PTKIN	B2
13650095	SBMPTN	B2
13650096	UM-PTKIN	B2
13650098	UM-PTKIN	B2
13650099	SBMPTN	B1
13650100	UM-PTKIN	B2
13650104	UM-PTKIN	A1
13650105	UM-PTKIN	A1
13650108	SBMPTN	B2
13650109	SBMPTN	A2
13650110	SBMPTN	B2
13650113	MANDIRI TERTULIS	B2
13650114	MANDIRI TERTULIS	B1
13650115	MANDIRI TERTULIS	B2
13650116	MANDIRI TERTULIS	A1
13650118	MANDIRI TERTULIS	A1
13650119	MANDIRI TERTULIS	B2
13650120	SBMPTN	B2
13650121	MANDIRI TERTULIS	B1
13650122	MANDIRI TERTULIS	B2
13650124	MANDIRI TERTULIS	B2