e-ISSN: 2809-3399, p-ISSN: 2338-5197
http://journal.uad.ac.id/index.php/JSTIF/index

Analisis Data Masuk Kuliah Dengan Data Kelulusan Mahasiswa Menggunakan Metode Apriori

Rizky Aulia Dwiyanti [a,1], Lisna Zahrotun [b,2,*], Utaminingsih Linarti [c,3]

[a,b,c] Universitas Ahmad Dahlan, Indonesia
[1] rizkyauliadwiyanti@gmail.com; [2] lisna.zahrotun@tif.uad.ac.id; [3] utaminingsih.linarti@ie.uad.ac.id
* Penulis Korespondensi

ABSTRAK

Evaluasi keberhasilan penyelenggara program studi di perguruan tinggi merupakan aktivitas yang sering dilakukan oleh Fakultas Y Universitas X. Fakultas melakukan evaluasi dengan bentuk evaluasi studi terhadap mahasiswa. Namun selama adanya proses evaluasi studi tersebut, jumlah kelulusan mahasiswa Fakultas Y setiap tahunnya masih di bawah 50%, sehingga perlu dilakukan identifikasi penyebab mahasiswa tidak lulus tepat waktu. Identifikasi tersebut dilakukan dengan pencarian hubungan antara data mahasiswa sebelum masuk kuliah dan data kelulusan mahasiswa. Penelitian ini melakukan analisis pola asosiasi antara data masuk kuliah dengan data kelulusan mahasiswa Fakultas Y tahun 2014-2015 dengan data kelulusan mahasiswa tahun 2018-2019. Penelitian ini memperoleh pola asosiasi yaitu pada dataset prodi U. Tahapan penelitian meliputi seleksi data, pembersihan data, transformasi data, penerapan metode apriori dan pengujian menggunakan lift rasio. Penelitian ini menghasilkan pola asosiasi data masuk kuliah dengan data kelulusan yang dapat dijadikan rekomendasi prodi U. Data masuk yang saling berelasi dengan data kelulusan lama studi diantaranya adalah nilai rata-rata mata Pelajaran Bahasa Indonesia, nilai rata-rata kimia, nilai computer range, nilai rata-rata fisika dibawah 50 dan nilai rata-rata matematika range 50-<60. Sedangkan datamasuk yang saling berelasi data kelulusan nilai IPK diantaranya nilai rata-rata elektronika, nilai rata-rata computer dibawah 50 dan berasal dari SMA. Dengan nilai *lift* yang dihasilkan yaitu rata-rata 1,01-2,20 yang artinya berkorelasi positif.

Riwayat Artikel
Diterima 10 Januari 2024
Diperbaiki 15 Februari 2024
Diterbitkan 20 Oktober 2024

Kata Kunci
Data Mining
Pola Asosiasi
Algoritma Apriori
Lift Rasio

This is an open-access article under the CC–BY-SA license

1. Pendahuluan

Universitas X merupakan salah satu perguruan tinggi yang ada di Yogyakarta. Sebagai perguruan tinggi, Universitas X memiliki proses penerimaan mahasiswa baru pada semua program studi melalui Biro Akademik dan Admisi Universitas X berkoordinasi dengan program studi dan fakultas. Pada proses penerimaan mahasiswa tersebut, diperoleh data masuk kuliah mahasiswa diantaranya jalur masuk, nilai rapor, dan asal sekolah. Selain proses penerimaan mahasiswa baru, juga terdapat proses kelulusan mahasiswa yang juga menghasilkan data kelulusan mahasiswa diantaranya IPK dan lama studi. Pada proses kelulusan mahasiswa tersebut, berdasarkan peraturan Rektor Universitas X Nomor 3 Tahun 2016 tentang peraturan akademik dalam pasal 25 Bab VII Kelulusan dan Wisuda, universitas X memiliki beberapa syarat kelulusan diantaranya telah lulus semua mata kuliah, memiliki IPK minimal 2,0 untuk program diploma dan program sarjana, tidak ada nilai E, dan jumlah SKS mata kuliah dengan nilai D maksimal 10%. Universitas X memiliki beberapa fakultas salah satunya yaitu Fakultas Y. Setiap tahun jumlah pendaftar di Universitas X khususnya di Fakultas S mengalami kenaikan. Hal tersebut dapat diliat dari data mahasiswa pada Tabel 1.

Jurnal Sarjana Teknik Informatika
Vol. 12., No. 3, Oktober 2024, pp. 99-109

Tabel 1. Jumlah Data Lulusan Mahasiswa Fakultas Y

Tahun Angkatan	Jumlah Mahasiswa	Jumlah Lulusan Tepat Waktu
2012	360	14 (4%)
2013	340	22 (6%)
2014	591	55 (9%)
2015	837	181 (22%)

Berdasarkan data pada Tabel 1.1 dapat dilihat bahwa jumlah mahasiswa pertahun mengalami kenaikan yang signifikan sejak tahun 2013. Namun, berdasarkan data tersebut dapat dilihat juga bahwa lulusan mahasiswa Fakultas Y belum mencapai target kuliah tepat waktu maksimal 50% per angkatan sesuai standar Badan Akreditasi Nasional (BAN-PT) [1]. Mahasiswa merupakan salah satu aspek penting dalam evaluasi keberhasilan penyelenggara program studi di perguruan tinggi [2]. Untuk memacu mahasiswa Universitas X agar lulus tepat waktu makan diselenggarakan Prosedur Evaluasi Mahasiswa, yang disebut Evaluasi Studi. Hal ini diatur dalam SK Rektor Universitas X No. 32/2008 tentang Prosedur Evaluasi Studi Mahasiswa. Namun selama adanya proses evaluasi studi tersebut, jumlah kelulusan mahasiswa Fakultas Y setiap tahunnya masih dibawah 50%. [1].

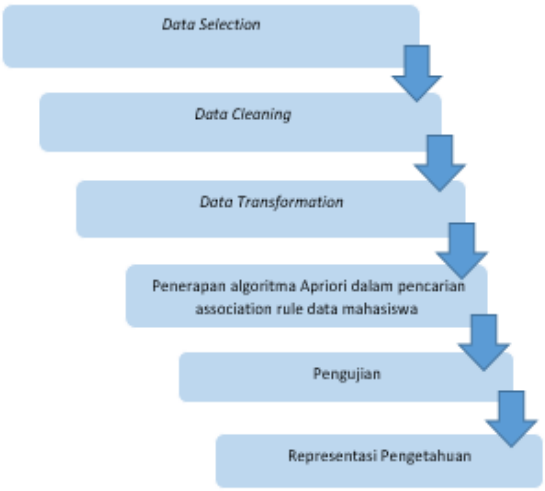
Terdapat beberapa penelitian yang dilakukan oleh para peneliti dalam pencarian pola asosiasi data masuk kuliah dengan data kelulusan. Penelitian yang dilakukan [3] dan [4] yang menggunakan metode *Apriori* menghasilkan kesimpulan bahwa data masuk kuliah dengan data kelulusan mahasiswa menjadi faktor-faktor yang mempengaruhi kelulusan tepat waktu. Penelitian selanjutnya dilakukan oleh [5] menggunakan data mahasiswa untuk menemukan aspek yang mempengaruhi lama studi mahasiswa menggunakan metode *Apriori*. Penelitian dibidang lain diantaranya terkait rekomendasi produk bagi pelanggan [6] [7] [8][9], pemilihan menu [10], mengthuai tingkata kelulusan[11], [12]. Sejauh ini, fakultas belum memanfaatkan faktor-faktor tersebut dalam mencari hubungan data masuk kuliah dengan data kelulusan mahasiswa dalam menemukan pola mahasiswa yang dapat lulus tepat waktu dan yang tidak dapat lulus tepat waktu. Salah satu cara untuk mengetahui hubungan antar data adalah dengan cara menerapkan *Data Mining* antara lain teknik *Association Rules.*

Association Rules sering digunakan untuk menemukan aturan asosiatif antara kombinasi item. *Association rule* adalah salah satu tahap analisis asosiasi yang menarik perhatian banyak peneliti untuk menghasilkan algoritma yang efisien, yaitu analisis pola frekuensi tinggi [10]. Berdasarkan hasil penelitian yang telah dipaparkan, maka penelitian ini akan menerapkan metode Apriori.

Berdasarkan permasalahan tersebut, maka dalam penelitian ini dilakukan pencarian pola asosiasi data masuk kuliah dengan data kelulusan mahasiswa menggunakan metode Apriori dengan harapan untuk membantu fakultas dalam evaluasi untuk mengetahui probabilitas sebagai dasar proses klasifikasi kelulusan mahasiswa. *Output* dari penelitian ini berupa informasi tentang pola hubungan antara data masuk kuliah dengan data kelulusan mahasiswa.

2. Metode

Berdasarkan data yang dihimpun melalui studi literatur dan wawancara dengan narasumber, maka akan dilakukan pencarian pola asosiasi Fakultas Y tahun 2014-2015 dengan data kelulusan mahasiswa tahun 2018-2019 menggunakan metode *Apriori*. Adapun tahapan penelitian utama yang digunakan dalam penelitian ini adalah sebagaimana pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 *Data Selection*

Tahapan awal penelitian ini adalah *data selection*. Data masuk kuliah dan data kelulusan mahasiswa Fakultas Y Program Studi R angkatan 2014-2015 yang telah terkumpul diseleksi dengan mengambil beberapa kolom agar data mahasiswa dapat diolah pada tahapan selanjutnya.

2.2 *Data Cleaning*

Data cleaning merupakan proses yang digunakan untuk membuang data yang tidak konsisten dan bersifat noise. Data yang dibuang yaitu data yang salah, rusak, tidak akurat, tidak lengkap dan salah format. Pada proses ini digunakan fitur *dropna* yang ada pada bahasa pemrograman *Python.*

2.3 Data Transformation

Transformasi yaitu proses pengkategorian atau perubahan data ke format yang sesuai untuk proses *data mining*. Transformasi dilakukan dengan cara mengubah format dan pengkategorian data sehingga lebih mudah untuk diolah. Tahap transformasi yang dilakukan data masuk kuliah dengan data kelulusan diubah dari format *string* ke *float*. Hal tersebut dilakukan agar data mudah diolah. Setelah mengubah format dataset, kemudian dilakukan pengelompokkan rata-rata nilai matematika, rata-rata nilai fisika, rata-rata kimia, rata-rata biologi, rata-rata nilai bahasa inggris, rata-rata nilai bahasa indonesia, rata-rata nilai elektronika, rata-rata nilai komputer, asal sekolah, dan lama studi atau IPK.

2.4 Algoritma Apriori

Algoritma Apriori merupakan algoritma untuk mengekstrak frekuensi itemset dari *association rule*[13]. Proses dalam algoritma apriori dimulai dari pencarian frekuensi 1-*itemsets* dengan menghitung frekuensi kemunculan masing-masing item dan mengambil item yang memenuhi *minimum support*, hasil tersebut dinyatakan sebagai L_1. Lproses kedua dinyatakan sebagai L_2 dan seterusnya. Langkah tersebut dilakukan berulang hingga dan tidak ada lagi item yang memenuhi *minimum support* atau hingga tidak ada itemset yang ditemukan.

$$Support\,(\mathrm{A,B}) = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Total\ Transaksi\ Yang\ Terjadi}$$

$$Confidence = \mathrm{P(B \mid A)} = \frac{\sum Transaksi\ mengandung\ A\ dan\ B}{\sum Transaksi\ mengandung\ A}$$

2.5 Pengujian

Pengujian akan dilakukan untuk mengetahui sebaik apakah algoritma Apriori dalam hal akurasi. Pengujian dilakukan dengan menggunakan perhitungan *lift ratio*. Jika nilai *lift ratio* jika lift≥1, maka menunjukkan proses transaksi dapat dikatakan valid karena korelasi antar item bersifat positif. Jika nilai *lift ratio* jika lift=1, maka rule bersifat independen. Jika nilai *lift ratio* jika lift<1, maka rule saling berlawanan atau negatif[8].

Jurnal Sarjana Teknik Informatika
Vol. 12., No. 3, Oktober 2024, pp. 99-109

$$Nilai\ lift = \frac{confidence\ (A \rightarrow B)}{Support(B)}$$

2.6 Representasi Pengetahuan

Representasi pengetahuan adalah visualisasi hasil mining. Representasi pengetahuan merupakan tahap visualisasi dan penyajian hasil yang diperoleh kepada user. Tahapan ini mencakup semua proses yang dilakukan dimulai dari membaca data hingga data ditampilkan kedalam bentuk program *Python.*

3. Hasil dan Pembahasan

Pencarian pola asosiasi dengan menggunakan metode *Apriori* yang dapat dijadikan rekomendasi untuk Fakultas Y dan rpograsm studi R dalam acuan penerimaan mahasiswa baru. Data yang digunakan dalam penelitian ini. Proses pencarian pola yaitu sebagai berikut.

3.1 Data Selection

Pada proses *data selection* adalah memilih variabel yang akan digunakan dalam penelitian. Dalam penelitin ini hasil seleksi ditampilkan pada Tabel 2.

Tabel 2. Hasil proses Seleksi Data

No	Nama SLTA	Provinsi	Rata Mtk	Rata fisika	Rata biologi	Rata kimia	Rata Bhs Ing	Rata Bhs Indo	Rata elektronika	Rata kompuuter	Lama Studi	IPK
1	SMK Negeri 2, Pekanbaru	Riau	81	77	0	82	84	82	90	81	3 Th, 9 Bln, 27 Hr	3,41
2	smk muhammadiyah 3	lain-lain	52	0	0	75	0	0	0	0	4 Th, 9 Bln, 18 Hr	3,32
3	SMA Negeri 2, Kalianda	Lampung	80	80	0	0	0	0	0	0	4 Th, 7 Bln, 9 Hr	2,83
4	SMK Hasanah Pekanbaru	lain-lain	79	83	0	0	0	0	0	0	4 Th, 8 Bln, 26 Hr	3,19
5	sma negeri 1 bandar sribhawono	Lampung	79	77	0	0	0	0	0	0	4 Th, 1 Bln, 23 Hr	3,26
6	SMA NEGERI 1 CIAWIGEBANG	Jawa Barat	79	82	78	81	79	84	0	79	3 Th, 11 Bln, 26 Hr	3,34
7	SMA Negeri 1, Marabahan	Kalimantan Selatan	85	92	88	87	84	87	0	84	3 Th, 11 Bln, 30 Hr	3,44
8	SMA N 2 FAKFAK	Papua Barat	88	86	90	88	87	86	0	88	4 Th, 9 Bln, 18 Hr	3,79

Data masuk kuliah dan data kelulusan mahasiswa Fakultas Y angkatan 2014-2015 yang telah terkumpul diseleksi dengan mengambil beberapa kolom agar data mahasiswa dapat diolah pada tahapan selanjutnya. Data yang diseleksi yaitu rata-rata nilai matematika, rata-rata nilai fisika, rata-rata kimia, rata-rata biologi, rata-rata nilai bahasa inggris, rata-rata nilai bahasa indonesia, rata-rata nilai elektronika, rata-rata nilai komputer, asal sekolah, provinsi, dan lama studi atau nilai IPK.

3.2 *Data Cleaning*

Data cleaning merupakan proses yang digunakan untuk membuang data yang tidak konsisten dan bersifat noise. Data yang dibuang yaitu data yang salah, rusak, tidak akurat, tidak lengkap dan salah format. Pada proses ini digunakan fitur *dropna* yang ada pada bahasa pemrograman *Python*. Pada penelitian ini, dataset tidak melalui proses cleaning karena dataset prodi R bersih dari data yang kosong maupun yang hanya berisi (-).

3.3 *Data Transformation*

Data yang telah melalui proses cleaning selanjutnya akan dilakukan transformasi. Pada tahap transformasi, format dataset yaitu data masuk kuliah dengan data kelulusan diubah, dari format *string* ke *float*. Hal tersebut dilakukan agar data mudah diolah. Setelah mengubah format dataset, kemudian dilakukan pengelompokkan rata-rata nilai matematika, rata-rata nilai fisika, rata-rata kimia, rata-rata biologi, rata-rata nilai bahasa inggris, rata-rata nilai bahasa indonesia, rata-rata nilai elektronika, rata-rata nilai komputer, asal sekolah, dan lama studi. Pada tahap ini dilakukan transformasi pada *sample dataset* R.

Tabel 3. Data Transformasi

Rata-rata Nilai Rapot	Asala Sekolah	Lama Studi	Nilai IPK
Range 90-100	SMA	3-<4 tahun	Cumlaude
Range 70-<80	SMK	4-<5 tahun	Sangat Memuaskan
Range 60-<70	MA	5-<6 tahun	Memuaskan
Range 50-<60		>6 tahun	
Range dibawah 50			

3.4 Algoritma Apriori

Proses awal Apriori adalah menentukan pencarian frekuensi 1-n *itemsets* dengan menghitung frekuensi kemunculan masing-masing item dan mengambil item yang memenuhi *minimum support* dan *minimum confidence* yang telah ditentukan. Penentuan minimum support yaitu 0,3 dan *minimum confidence* yaitu 0,6. Minimum support 0,3 dan minimum confidence 0,6 tersebut ditentukan karena minimum tersebut paling optimal dalam menghasilkan rule yang saling berelasi dan memiliki akurasi yang baik. Hasil proses penghitungan nilai support untuk 2-item set ditunjukkan pada Tabel 4, 3 itemset ditunjukkan pada Tabel 5, 4-Itemset ditunjukkan pada Tabel 6, 5 Itemset ditunjukkan pada Tabel 7, 6 Itemset ditunjukkan pada Tabel 8, 7 Itemset ditunjukkan pada Tabel 9 dan hasil penghitungan nilai confident ditunjukkan pada Tabel 10.

Tabel 4 Hasil 2-Item Set

No	Kategori	Jumlah	Nilai Support
1.	Kimia Range dibawah 50, Biologi Range dibawah 50	3	0,375
2.	Kimia Range dibawah 50, Bahasa Inggris Range dibawah 50	3	0,375
3.	Kimia Range dibawah 50, Bahasa Indonesia Range dibawah 50	3	0,375
4.	Kimia Range dibawah 50, Elektronika Range dibawah 50	3	0,375
5.	Kimia Range dibawah 50, Komputer Range dibawah 50	3	0,375
6.	Kimia Range dibawah 50, Lama Studi 4 – < 5 Tahun	3	0,375
…			
47.	SMA, Lama Studi 4 – < 5 Tahun	4	0,5

Tabel 5. Hasil 3-Item Set

No	Rule	Jumlah	Nilai *Support*
1.	nilai kimia range 80-<90, nilai bahasa indonesia range 80-<90, Lama Studi 3 sampai < 4 tahun	3	0,375
2.	bahasa inggris range dibawah 50, bahasa indonesia range dibawah 50, Lama Studi 4 sampai < 5 tahun	4	0,5
3.	bahasa indonesia range dibawah 50, biologi range dibawah 50, Lama Studi 4 sampai < 5 tahun	4	0,5

4.	bahasa indonesia range dibawah 50, komputer range dibawah 50, Lama Studi 4 sampai < 5 tahun	4	0,5
5.	bahasa indonesia range dibawah 50, kimia range dibawah 50, Lama Studi 4 sampai < 5 tahun	3	0,375
6.	bahasa inggris range dibawah 50, kimia range dibawah 50, Lama Studi 4 sampai < 5 tahun	3	0,375
…	..	..	…
17.	fisika range 80-<90, elektronika range dibawah 50, Lama Studi 4 sampai < 5 tahun	3	0,375

Tabel 6. Hasil 4-Itemset

Kategori	Jumlah	Nilai *Support*
Biologi range dibawah 50, bahasa inggris range dibawah 50, bahasa indonesia range dibawah 50, Lama Studi 4 sampai < 5 tahun	**4**	**0,5**
kimia range dibawah 50, nilai bahasa inggris range dibawah 50, komputer range dibawah 50, Lama Studi 4 sampai < 5 tahun	**3**	**0,375**
nilai bahasa inggris range dibawah 50, nilai bahasa indonesia range dibawah 50, nilai elektronika range dibawah 50, Lama Studi 4 sampai < 5 tahun	**4**	**0,5**
Bahasa inggris range dibawah 50, elektronika range dibawah 50, komputer range dibawah 50, Lama Studi 4 sampai < 5 tahun	**4**	**0,5**

Tabel 7. Hasil 5 Itemset

Kategori	Jumlah	Nilai *Support*
Biologi range dibawah 50, bahasa inggris range dibawah 50, bahasa indonesia range dibawah 50, elektronika rang dibawah 50, Lama Studi 4 sampai < 5 tahun	**4**	**0,5**
nilai bahasa inggris range dibawah 50, nilai bahasa indonesia range dibawah 50, nilai elektronika range dibawah 50, komputer dibawah range 50, Lama Studi 4 sampai < 5 tahun	**4**	**0,5**
Kimia range dibawah 50, Bahasa inggris range dibawah 50, elektronika range dibawah 50, komputer range dibawah 50, Lama Studi 4 sampai < 5 tahun	**3**	**0,375**

Tabel 8. Hasil 6-Itemset

Kategori	Jumlah	Nilai *Support*
Kimia range dibawah 50, Biologi range dibawah 50, bahasa inggris range dibawah 50, bahasa indonesia range dibawah 50, elektronika rang dibawah 50, Lama Studi 4 sampai < 5 tahun	3	0,375
Kimia range dibawah 50, nilai bahasa inggris range dibawah 50, nilai bahasa indonesia range dibawah 50, nilai elektronika range dibawah 50, komputer dibawah range 50, Lama Studi 4 sampai < 5 tahun	3	0,375
Kimia range dibawah 50, Biologi range dibawah 50, Bahasa inggris range dibawah 50, elektronika range dibawah 50, komputer range dibawah 50, Lama Studi 4 sampai < 5 tahun	3	0,375

Tabel 9. Hasil 7-Itemset

Kategori	Jumlah	Nilai *Support*
Kimia range dibawah 50, Biologi range dibawah 50, bahasa inggris range dibawah 50, bahasa indonesia range dibawah 50, elektronika rang dibawah 50, komputer dibawah range 50, Lama Studi 4 sampai < 5 tahun	3	0,375

Tabel 10. Hasil Nilai Confidence

No.	Kategori	*Confidence*
1	Jika Lama Studi 3 sampai < 4 tahun, nilai kimia range 80-<90, maka nilai bahasa indonesia range 80-<90	1
2	Jika Lama Studi 4 sampai < 5 tahun bahasa inggris range dibawah 50, maka bahasa indonesia range dibawah 50	1
3	Jika Lama Studi 4 sampai < 5 tahun, biologi range dibawah 50, maka bahasa indonesia range dibawah 50	1
4	Jika Lama Studi 4 sampai < 5 tahun, komputer range dibawah 50, maka bahasa indonesia range dibawah 50	1
5	Jika bahasa indonesia range dibawah 50, kimia range dibawah 50, maka Lama Studi 4 sampai < 5 tahun	1
6	Jika bahasa inggris range dibawah 50, kimia range dibawah 50, maka Lama Studi 4 sampai < 5 tahun	1
7	Jika biologi range dibawah 50, kimia range dibawah 50, maka Lama Studi 4 sampai < 5 tahun	1
8	Jika komputer range dibawah 50, kimia range dibawah 50, maka Lama Studi 4 sampai < 5 tahun	1
9	Jika Lama Studi 4 sampai < 5 tahun, nilai elektronika range dibawah 50, maka nilai fisika rnge 80-<90	0,6
10	Jika nilai elektronika range dibawah 50, nilai fisika rnge 80-<90, maka Lama Studi 4 sampai < 5 tahun	0,75
...	...	...
162	fisika range 80-<90, elektronika range dibawah 50, Lama Studi 4 sampai < 5 tahun	0,75

Jika ditentukan minimum support dan *minimum* tersebut pada perhitungan data masuk dengan data kelulusan yaitu lama studi R, maka akan diperoleh 72 rule. Sedangkan perhitungan data masuk dengan data kelulusan yaitu nilai IPK R, maka akan diperoleh 111 rule. Jika ditentukan minimum support dan *minimum* tersebut pada perhitungan data masuk dengan data kelulusan yaitu lama studi T, maka akan diperoleh 230 rule. Sedangkan perhitungan data masuk dengan data kelulusan yaitu nilai IPK T, maka akan diperoleh 228 rule. Jika ditentukan minimum support dan *minimum* tersebut pada perhitungan data masuk dengan data kelulusan yaitu lama studi S, maka akan diperoleh 214 rule. Sedangkan perhitungan data masuk dengan data kelulusan nilai IPK S, maka akan diperoleh 222 rule. Jika ditentukan minimum support dan *minimum* tersebut pada perhitungan data masuk dengan data kelulusan yaitu lama studi U, maka akan diperoleh 237 rule. Sedangkan perhitungan data masuk dengan data kelulusan yaitu nilai IPK U, maka akan diperoleh 165 rule. Berdasarkan jumlah rule setiap dataset yang telah diperoleh, maka akan dihasilkan rule yang data masuk yang berelasi dengan data kelulusan.

3.5 Pengujian

Pada proses pengujian lift ratio dihasilkan nilai-nilai rule yang memiliki korelasi negatif, independen dan korelasi positif. Evaluasi ini berupa hasil *rule* dari dataset yang memiliki atribut rata-rata nilai matematika, rata-rata nilai fisika, rata-rata kimia, rata-rata biologi, rata-rata nilai bahasa inggris, rata-rata nilai bahasa indonesia, rata-rata nilai elektronika, rata-rata nilai komputer, asal sekolah, provinsi yang berelasi dengan lama studi atau IPK. Evaluasi masing-masing data prodi adalah sebagai berikut:

Hasil kombinasi itemset yang menggunakan kombinasi atribut lama studi, dihasilkan 72 rule yang memiliki 70 rule berkorelasi positif, tidak ada rule bersifat independen dan 2 rule berkorelasi negatif. Pada proses pengujian data yang menggunakan kombinasi atribut IPK, dihasilkan 111 rule yang memiliki 110 rule

Jurnal Sarjana Teknik Informatika
Vol. 12., No. 3, Oktober 2024, pp. 99-109

berkorelasi positif, tidak ada rule bersifat independen dan 1 rule berkorelasi negatif. Pada pengujian tersebut diperoleh pola asosiasi data masuk kuliah dengan data kelulusan dengan kualitas terbaik. Berikut ditampilkan beberapa rule yang berkorelasi positif yang telah diuji pada sistem ditunjukkan pada Tabel 11 dan Tabel 12.

Tabel 11. Rule Kombinasi Data Masuk dengan Lama Studi

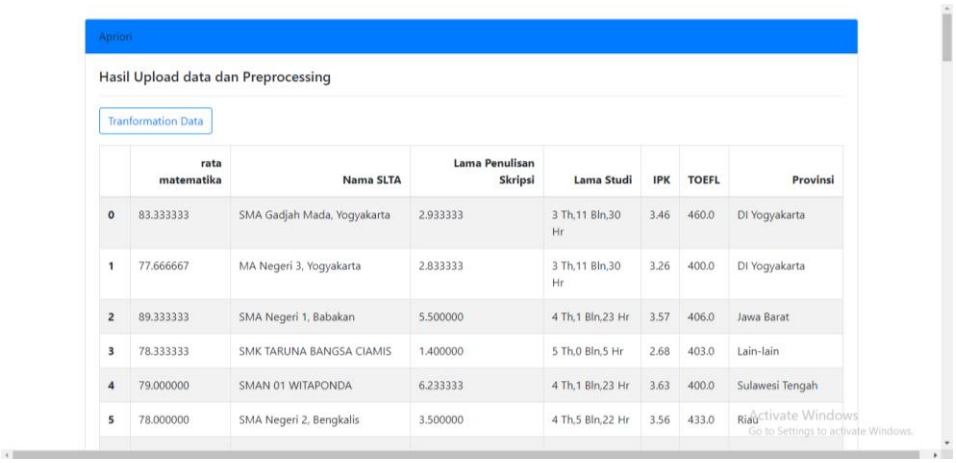
Rules	*Lift ratio*
Jika Lama Studi 3 sampai < 4 tahun, nilai elektronika range dibawah 50, maka nilai fisika range 80-<90	1,42
Jika Lama Studi 3 sampai < 4 tahun, nilai elektronika range dibawah 50, maka SMA	1,26
Jika nilai elektronika range dibawah 50, nilai fisika range 80-<90, maka Lama Studi 3 sampai < 4 tahun	1,15
Jika Lama Studi 3 sampai < 4 tahun, SMA, maka nilai elektronika range dibawah 50	1,14
Jika Lama Studi 3 sampai < 4 tahun, nilai fisika range 80-<90, maka nilai elektronika range dibawah 50	1,14

Tabel 12. Rule Kombinasi Data Masuk Teknik Elektro dengan Nilai IPK

Rules	*Lift ratio*
Jika IPK Sangat Memuaskan, nilai bahasa indonesia range dibawah 50, maka ('nilai kimia range dibawah 50	2,20
Jika IPK Sangat Memuaskan, nilai bahasa inggris range dibawah 50, maka nilai bahasa indonesia range dibawah 50	1,06
Jika IPK Sangat Memuaskan, nilai kimia range dibawah 50, maka nilai bahasa indonesia range dibawah 50	2,06
Jika nilai biologi range dibawah 50, nilai komputer range dibawah 50, maka IPK Sangat Memuaskan	1,33
Jika nilai biologi range dibawah 50, nilai elektronika range dibawah 50, maka IPK Sangat Memuaskan	1,33

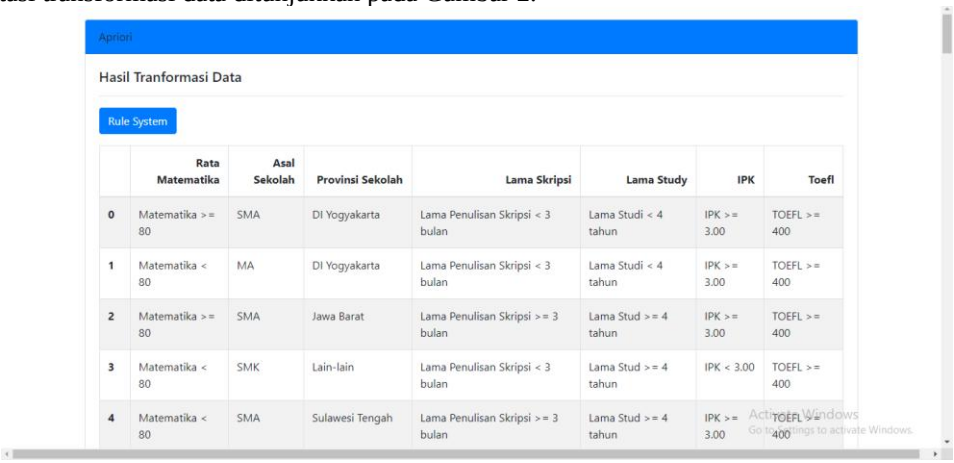
3.6 Implementasi

Aplikasi "Penerapan Data Mining Dalam Pencarian Pola Asosiasi Data Masuk Kuliah Dengan Data Kelulusan Mahasiswa Menggunakan Metode *Apriori"* untuk mencari pola asosiasi dari data masuk kuliah dan data kelulusan mahasiswa Fakultas Y. Aplikasi berbasis website dengan menggunakan *Framework Flask* dan bahasa pemrograman *Python.* Hasil implementasi upload data ditunjukkan pada Gambar 1.



Gambar 1. Implementasi Upload Data set

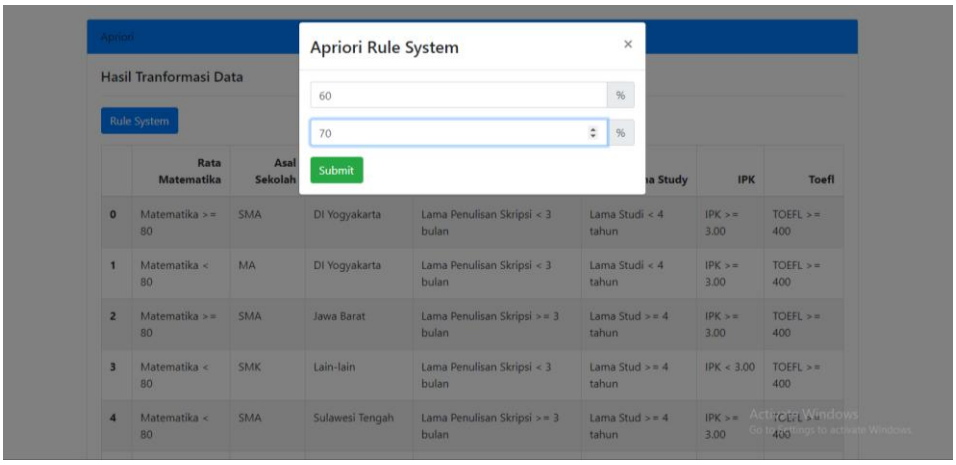
Setelah data berhasil diupload maka Langkah selanjutnya adalah preprocessing, yaitu transformasi data. Hasil implementasi transformasi data ditunjukkan pada Gambar 2.



	Rata Matematika	Asal Sekolah	Provinsi Sekolah	Lama Skripsi	Lama Study	IPK	Toefl
0	Matematika >= 80	SMA	DI Yogyakarta	Lama Penulisan Skripsi < 3 bulan	Lama Studi < 4 tahun	IPK >= 3.00	TOEFL >= 400
1	Matematika < 80	MA	DI Yogyakarta	Lama Penulisan Skripsi < 3 bulan	Lama Studi < 4 tahun	IPK >= 3.00	TOEFL >= 400
2	Matematika >= 80	SMA	Jawa Barat	Lama Penulisan Skripsi >= 3 bulan	Lama Stud >= 4 tahun	IPK >= 3.00	TOEFL >= 400
3	Matematika < 80	SMK	Lain-lain	Lama Penulisan Skripsi < 3 bulan	Lama Stud >= 4 tahun	IPK < 3.00	TOEFL >= 400
4	Matematika < 80	SMA	Sulawesi Tengah	Lama Penulisan Skripsi >= 3 bulan	Lama Stud >= 4 tahun	IPK >= 3.00	TOEFL >= 400

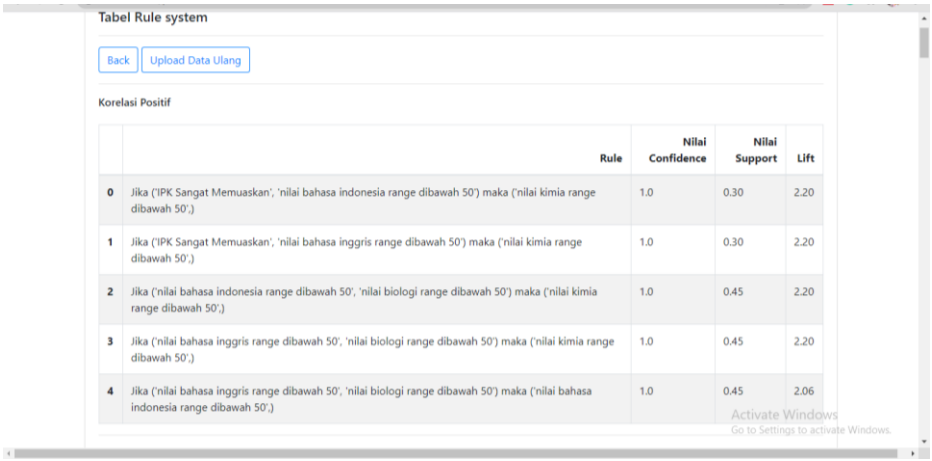
Gambar 2. Implementasi proses Transformasi Data

Tahap berikutnya adalah proses Apriori. Dalam proses ini user diminta memasukkan nilai minimum support dan minimum confident dengan nilai erntang 50- 100. Tampilan implementasi apriori ditunjukkan pada Gambar 3.



Gambar 3. Implementasi Hasil Algoritma Apriori

Tahap yang terakhir yaitu menampilkan hasil representasi pengetahuan atau rule yang dihasilkan dari proses Apriori. Hasil implementasi tampilan rule yang dihasilkan ditunjukkan pada Gambar 4.



	Rule	Nilai Confidence	Nilai Support	Lift
0	Jika ('IPK Sangat Memuaskan', 'nilai bahasa indonesia range dibawah 50') maka ('nilai kimia range dibawah 50',)	1.0	0.30	2.20
1	Jika ('IPK Sangat Memuaskan', 'nilai bahasa inggris range dibawah 50') maka ('nilai kimia range dibawah 50',)	1.0	0.30	2.20
2	Jika ('nilai bahasa indonesia range dibawah 50', 'nilai biologi range dibawah 50') maka ('nilai kimia range dibawah 50',)	1.0	0.45	2.20
3	Jika ('nilai bahasa inggris range dibawah 50', 'nilai biologi range dibawah 50') maka ('nilai kimia range dibawah 50',)	1.0	0.45	2.20
4	Jika ('nilai bahasa inggris range dibawah 50', 'nilai biologi range dibawah 50') maka ('nilai bahasa indonesia range dibawah 50',)	1.0	0.45	2.06

Gambar 4. Implementasi hasil rule Algoritma Apriori

Jurnal Sarjana Teknik Informatika
Vol. 12., No. 3, Oktober 2024, pp. 99-109

3.7 Representasi Pengetahuan

Berdasarkan hasil asosiasi setiap dataset, maka menghasilkan rule yang mengandung data masuk dengan data kelulusan mahasiswa yang pola asosiasinya diantaranya:

1. Jika Lama Studi 3 sampai < 4 tahun, nilai elektronika range dibawah 50, maka nilai fisika range 80-<90. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran elektronika dengan range dibawah 50 dan fisika range 80-<90 juga memiliki data kelulusan dengan lama studi 3 sampai kurang dari 4 tahun.
2. Jika Lama Studi 3 sampai < 4 tahun, nilai elektronika range dibawah 50, maka SMA. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran elektronika dengan range dibawah 50 dan berasal dari SMA juga memiliki data kelulusan dengan lama studi 3 sampai kurang dari 4 tahun.
3. Jika nilai elektronika range dibawah 50, nilai fisika range 80-<90, maka Lama Studi 3 sampai < 4 tahun. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran elektronika dengan range dibawah 50 dan fisika range 80-<90 juga memiliki data kelulusan dengan lama studi 3 sampai kurang dari 4 tahun.
4. Jika Lama Studi 3 sampai < 4 tahun, SMA, maka nilai elektronika range dibawah 50. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran elektronika dengan range dibawah 50 dan berasal dari SMA juga memiliki data kelulusan dengan lama studi 3 sampai kurang dari 4 tahun.
5. Jika Lama Studi 3 sampai < 4 tahun, nilai fisika range 80-<90, maka nilai elektronika range dibawah 50. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran fisika range 80-<90 dan elektronika dengan range dibawah 50 juga memiliki data kelulusan dengan lama studi 3 sampai kurang dari 4 tahun.
6. Jika IPK sangat memuaskan, nilai bahasa indonesia range dibawah 50, maka nilai kimia range dibawah 50. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran bahasa indonesia dan kimia dengan range dibawah 50 juga memiliki data kelulusan nilai IPK predikat sangat memuaskan.
7. Jika IPK sangat memuaskan, nilai bahasa inggris range dibawah 50, maka nilai bahasa indonesia range dibawah 50. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran bahasa inggris dan bahasa indonesia dengan range dibawah 50 juga memiliki data kelulusan nilai IPK predikat sangat memuaskan.
8. Jika IPK sangat memuaskan, nilai kimia range dibawah 50, maka nilai bahasa indonesia range dibawah 50. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran kimia dan bahasa indonesia dengan range dibawah 50 juga memiliki data kelulusan nilai IPK predikat sangat memuaskan.
9. Jika nilai biologi range dibawah 50, nilai komputer range dibawah 50, maka IPK sangat memuaskan. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran biologi dan komputer dengan range dibawah 50 juga memiliki data kelulusan nilai IPK predikat sangat memuaskan.
10. Jika nilai biologi range dibawah 50, nilai elektronika range dibawah 50, maka IPK sangat memuaskan. Berdasarkan pola tersebut, mahasiswa yang mayoritas memiliki nilai rata-rata mata pelajaran komputer dan biologi dengan range dibawah 50 juga memiliki data kelulusan nilai IPK predikat sangat memuaskan.

4. Kesimpulan

Berdasarkan penelitian ini, maka dapat disimpulkan bahwa :

1. Penelitian menghasilkan pola asosiasi data masuk kuliah dengan data kelulusan yang dapat dijadikan rekomendasi prodi U, Penelitian ini menghasilkan pola asosiasi data masuk kuliah dengan data kelulusan yang dapat dijadikan rekomendasi prodi U. Data masuk yang saling berelasi dengan data kelulusan lama studi diantaranya adalah nilai rata-rata mata Pelajaran Bahasa Indonesia, nilai rata-rata kimia, nilai computer range, nilai rata-rata fisika dibawah 50 dan nilai rata-rata matematika range 50-<60. Sedangkan datamasuk yang saling berelasi data kelulusan nilai IPK diantaranya nilai rata-rata elektronika, nilai rata-rata computer dibawah 50 dan berasal dari SMA.
2. Pengujian menggunakan lift ratio memperoleh rule dengan kualitas baik yaitu rata-rata 1,01-2,20.

5. Saran

Adapun saran dari penelitian adalah sebagai berikut:

1. Perlu penambahan dataset dan atribut data, sehingga terdapat rule yang lebih banyak juga dapat meningkatkan akurasi dari pengelompokkan dalam penelitian selanjutnya.
2. Penelitian ini dapat dikembangkan menggunakan konsep lain seperti *Desicion Tree* (pohon keputusan).

Daftar Pustaka

[1] N. Syahid, "Tahun 2019 'Lulusan Tepat Waktu Mahasiswa FTI UAD Angkatan 2015 Sebanyak 22%,'" *fti.uad.ac.id*, 2019.

[2] M. S. Mujib Ridwan, Hadi Suyono, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, vol. 4, no. 2, p. 151, 2018, doi: 10.24076/citec.2017v4i2.106.

[3] A. S. Saputro, "Penerapan Association Rule Dengan Algoritma Apriori Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa Teknik Informatika S1 Fakultas Ilmu," 2015.

[4] A. S. R. Sinaga, "DATA MINING INFORMASI TINGKAT KELULUSAN MAHASISWA STMIK PELITA NUSANTARA MEDAN." 2017.

[5] G. C. Sutradana and M. D. R. Wahyudi, "Penerapan Data Mining Untuk Analisis Pengaruh Lama Studi Mahasiswa Teknik Informatika Uin Sunan," *Penerapan Data Min. Untuk Anal. Pengaruh Lama Stud. Mhs. Tek. Inform. Uin Sunan Kalijaga Yogyakarta Menggunakan Metod. Apriori*, vol. 1, no. 3, pp. 153–162, 2017.

[6] A. R. Riszky and M. Sadikin, "Data Mining Menggunakan Algoritma Apriori untuk Rekomendasi Produk bagi Pelanggan," *J. Teknol. dan Sist. Komput.*, vol. 7, no. 3, pp. 103–108, Jul. 2019, doi: 10.14710/jtsiskom.7.3.2019.103-108.

[7] N. N. Merliani, N. I. Khoerida, N. T. Widiawati, L. A. Triana, and P. Subarkah, "Penerapan Algoritma Apriori Pada Transaksi Penjualan Untuk Rekomendasi Menu Makanan Dan Minuman," *J. Nas. Teknol. dan Sist. Inf.*, vol. 8, no. 1, 2022, doi: 10.25077/teknosi.v8i1.2022.9-16.

[8] A. Setiawan and F. P. Putri, "Implementasi Algoritma Apriori untuk Rekomendasi Kombinasi Produk Penjualan," *Ultim. J. Tek. Inform.*, vol. 12, no. 1, 2020, doi: 10.31937/ti.v12i1.1644.

[9] A. J. P. Sibarani, "Implementasi Data Mining Menggunakan Algoritma Apriori Untuk Meningkatkan Pola Penjualan Obat," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 7, no. 2, pp. 262–276, 2020, doi: 10.35957/jatisi.v7i2.195.

[10] F. Amsury, I. Kurniawati, and M. Rizki Fahdia, "Implementasi Association Rules Menentukan Pola Pemilihan Menu Di the Gade Coffee & Gold Menggunakan Algoritma Apriori," *INFOTECH J.*, vol. 9, no. 1, pp. 279–286, Jun. 2023, doi: 10.31949/infotech.v9i1.5357.

[11] Nurjoko and H. Kurniawan, "Aplikasi Data Mining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Apriori di IBI Darmajaya, Bandar Lampung," *J. TIM Darmajaya*, vol. 02, no. 01, pp. 79–93, 2016.

[12] H. Indriyawati and T. Winarty, "Analisa Data Mining Kemampuan Lulusan dengan Kebutuhan Stakeholder Menggunakan Algoritma Apriori (Studi Kasus Universitas Semarang)," *Proceeding SENDIU*, pp. 499–505, 2021.

[13] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. 2012. doi: 10.1016/B978-0-12-381479-1.00001-0.