

## Data Mining Untuk Memprediksi Status Kelulusan Mahasiswa

Sumiyatun\*<sup>1</sup>, Yagus Cahyadi<sup>2</sup>, Edi Iskandar<sup>3</sup>

<sup>1,2</sup>Prodi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia

<sup>3</sup>Prodi Informatika, Fakultas Teknologi Informasi, Universitas Teknologi Digital Indonesia

e-mail: \*[sumiyatun@utdi.ac.id](mailto:sumiyatun@utdi.ac.id), [yagus.cahyadi@utdi.ac.id](mailto:yagus.cahyadi@utdi.ac.id), [edi\\_iskadar@utdi.ac.id](mailto:edi_iskadar@utdi.ac.id)

Correspondence author email: \*[sumiyatun@utdi.ac.id](mailto:sumiyatun@utdi.ac.id)

### Abstrak

Semakin lama operasional sebuah perguruan tinggi, semakin banyak lulusan (alumni) yang dihasilkan. Data alumni serta mahasiswa aktif, merupakan aset yang dapat digunakan untuk keperluan tertentu, salah satunya adalah digunakan untuk memprediksi status kelulusan. Pendekatan yang umum digunakan adalah teknik penambangan data (*data mining*). Penelitian ini akan menerapkan teknik data mining untuk memprediksi status kelulusan mahasiswa. Pengolahan data akan memanfaatkan aturan asosiatif (*association rule mining*) dan menggunakan algoritma apriori. Algoritma apriori adalah algoritma dasar untuk penentuan frequent itemsets aturan asosiasi boolean. Algoritma ini mengontrol berkembangnya kandidat itemset dari hasil frequent itemsets dengan support-based pruning untuk menghilangkan itemset yang tidak menarik dengan menetapkan minsup. Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa aplikasi data mining dengan menggunakan pendekatan aturan asosiasi dan algoritma apriori dapat digunakan untuk menampilkan informasi aturan status kelulusan. Informasi yang ditampilkan berupa nilai support dan confidence. Semakin tinggi nilai confidence dan support maka semakin kuat nilai hubungan antar atribut/item. Data alumni dan data mahasiswa yang diproses meliputi IPS2, IPS4, ORG, PRG dan SL.

**Kata kunci**—prediksi, kelulusan, data mining, association, apriori.

## 1. PENDAHULUAN

Universitas Teknologi Digital Indonesia (UTDI) adalah perguruan tinggi swasta di D.I. Yogyakarta yang merupakan perubahan bentuk STMIK Akakom. Sejarah berdirinya UTDI dimulai dari Akademi Aplikasi Komputer (AKAKOM) tahun 1979, lalu berkembang menjadi AMIK Akakom pada tahun 1985, selanjutnya menjadi STMIK Akakom pada tahun 1992, dan sekarang menjadi UTDI mulai tahun 2021. UTDI dari awal berdirinya telah menghasilkan banyak lulusan yang berkontribusi di berbagai bidang, khususnya industri teknologi informasi. Saat ini, UTDI memiliki 9 program studi, yaitu : S2-Teknologi Informasi, S1-Informatika, S1-Sistem Informasi, S1-Teknik Komputer, S1-Manajemen Ritel, S1-Bisnis Digital, D3-Rekayasa Perangkat Lunak Aplikasi, D3-Teknologi Komputer, dan D3-Sistem Informasi Akuntansi [1].

Semakin lama operasional sebuah perguruan tinggi, semakin banyak lulusan (alumni) yang dihasilkan. Alumni UTDI tersebar diseluruh nusantara, bahkan beberapa alumni berkiprah di luar negeri. Data penting alumni serta mahasiswa aktif, baik mahasiswa lama maupun mahasiswa baru tersimpan di server perguruan tinggi. Semakin lama, data akan semakin meningkat baik jumlah maupun jenisnya. Data tersebut merupakan aset yang dapat digunakan untuk keperluan tertentu, salah satunya adalah digunakan untuk memprediksi tingkat kelulusan. Pendekatan yang umum digunakan adalah teknik penambangan data (*data mining*).

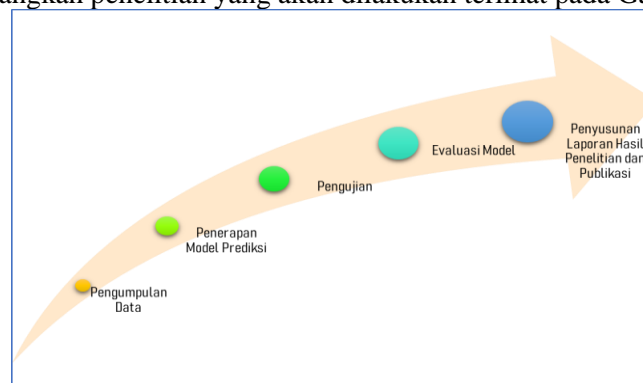
*Data mining* adalah suatu teknik menggali informasi berharga yang terpendam atau tersembunyi pada suatu koleksi data (*database*) yang sangat besar sehingga ditemukan suatu pola yang menarik yang sebelumnya tidak diketahui. Sedangkan definisi yang lain menerangkan bahwa, *data mining* adalah suatu pencarian dan analisa dari jumlah data yang sangat besar dan bertujuan untuk mencari arti dari pola dan aturan [2]. *Data mining* sudah ada sejak lama dan teori-teorinya pun sudah banyak dibahas dalam literatur. Teori-teori tersebut antara lain *naive-bayes* dan *nearest neighbour*, *decission tree*, *association rule*, *k-means* dan *text mining* [3].

Menurut [4], *association rule mining* merupakan teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*, atau untuk menemukan hubungan hal tertentu dalam suatu transaksi data dengan hal lain di dalam transaksi, yang digunakan untuk memprediksi pola, seperti halnya prediksi kelulusan. Kumpulan data alumni dan data mahasiswa pada program studi sistem informasi S1 UTDI, selama ini hanya di gunakan sebagai aset data saja. Belum dilakukan analisis dan pengolahan dengan pendekatan tertentu untuk mendapatkan hal yang lebih bermakna dari kumpulan data yang besar dan beragam tersebut.

Berdasarkan uraian diatas, penelitian ini akan menerapkan teknik *data mining* untuk memprediksi tingkat kelulusan mahasiswa pada program studi sistem informasi S1 UTDI. Pengolahan data akan memanfaatkan aturan asosiatif (*association rule mining*) dan menggunakan algoritma apriori. Algoritma apriori adalah algoritma dasar untuk penentuan *frequent itemsets* aturan asosiasi *boolean*. Algoritma ini mengontrol berkembangnya kandidat itemset dari hasil *frequent itemsets* dengan *support-based pruning* untuk menghilangkan *itemset* yang tidak menarik dengan menetapkan *minsup* [5].

## 2. METODE PENELITIAN

Subjek penelitian adalah penerapan *data mining* untuk memprediksi tingkat kelulusan mahasiswa. Metode yang digunakan adalah *assosiation rule ming* dengan menggunakan algoritma apriori. Langkah penelitian yang akan dilakukan terlihat pada Gambar 1.



**Gambar 1.** Langkah penelitian

### *Pengumpulan Data*

Data yang diperlukan untuk membuat sistem ini memerlukan dua jenis data, yaitu data primer (data yang diperoleh dari pengumpulan data mahasiswa aktif dan lulusan/alumni) dan data sekunder (data yang diperoleh dari tinjauan pustaka maupun data-data yang disediakan dengan melakukan studi literatur, mencangkup buku-buku teks, diktat, makalah, artikel dan buku petunjuk teknis).

Pada teknik pengumpulan data, peneliti melakukan pengumpulan data-data yang terkait langsung sesuai dengan kebutuhan dan perumusan masalah. Metode yang digunakan adalah Observasi, Interview dan Studi Literatur

### *Pengembangan Model*

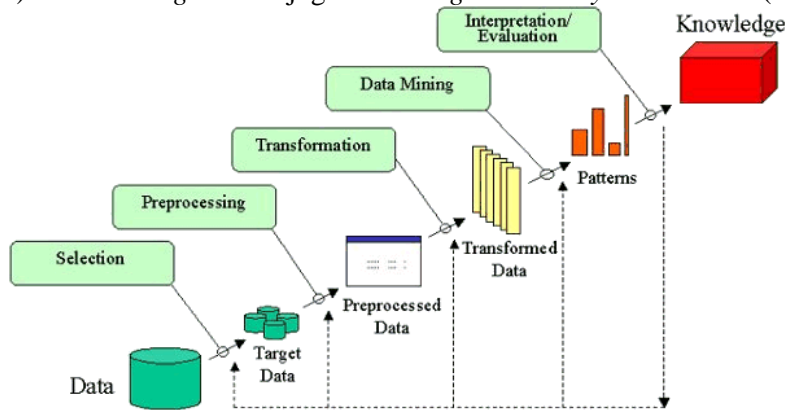
Tahap ini didahului tahap penganalisaan model yang digunakan untuk mengetahui kelebihan dan kekurangan model. Pada saat perancangan model, terdapat beberapa tahapan pemodelan dalam melakukan prediksi menggunakan algoritma apriori. Tahapan tersebut meliputi data scanning, menentukan nilai item, menghitung *support minimum*, Penggabungan *itemset*, menghitung nilai *confidence*, menghitung nilai *minimum frekuensi* dan langkah terakhir adalah penentuan *final assosiation rule*.

### *Data mining*

Menurut [6], *data mining* adalah proses pengumpulan dan pengolahan data yang bertujuan untuk mengekstrak informasi penting pada data. *Data mining* memiliki tiga tujuan

yaitu sebagai sarana untuk menjelaskan atau *explanatory*, untuk konfirmasi atau *confirmatory*, dan untuk eksplorasi atau *exploratory*. *Data mining* juga memiliki beberapa metode seperti *Association*, *Classification*, *Regression*, dan *Clustering*.

Proses pengumpulan dan ekstraksi informasi tersebut dapat dilakukan menggunakan perangkat lunak dengan bantuan perhitungan statistika, matematika, ataupun teknologi *Artificial Intelligence* (AI). *Data mining* disebut juga *Knowledge Discovery in Database* (Gambar 2).



**Gambar 2.** Knowledge Discovery in Database (KDD)

Secara umum, terdapat beberapa metode yang digunakan untuk melakukan *data mining*. Beberapa metode tersebut adalah *association*, *classification*, *regression* dan *clustering*. *Association* adalah metode berbasis aturan yang digunakan untuk menemukan asosiasi dan hubungan variabel dalam satu set data. Biasanya analisis ini terdiri dari pernyataan “*if* atau *then*” sederhana. *Association* banyak digunakan dalam mengidentifikasi korelasi produk dalam keranjang belanja untuk memahami kebiasaan konsumsi pelanggan. Sehingga, perusahaan dapat mengembangkan strategi penjualan dan membuat sistem rekomendasi yang lebih baik.

*Classification* adalah metode yang paling umum digunakan dalam *data mining*. *Classification* adalah tindakan untuk memprediksi kelas suatu objek. Sedangkan *regression* adalah teknik yang menjelaskan variabel dependen melalui proses analisis variabel independen. Sebagai contoh, prediksi penjualan suatu produk berdasarkan korelasi antara harga produk dengan tingkat pendapatan rata-rata pelanggan. metode *data mining* lainnya adalah *clustering*. *Clustering* digunakan dalam membagi kumpulan data menjadi beberapa kelompok berdasarkan kemiripan atribut yang dimiliki. Contoh kasusnya adalah *Customer Segmentation*. Ia membagi pelanggan ke dalam beberapa grup berdasarkan tingkat kemiripannya.

#### *Association rule mining*

*Association rules* adalah salah satu *task data mining* deskriptif yang bertujuan untuk menemukan aturan asosiatif antara item-item data. Langkah utama yang perlu dalam *association rules* adalah mengetahui seberapa sering kombinasi item muncul dalam database, yang disebut sebagai *frequent patterns*. Penting tidaknya suatu aturan asosiatif dapat diketahui dengan dua parameter, *support* yaitu persentase kombinasi item dalam *database* dan *confidence* yaitu kuatnya hubungan antar item dalam aturan asosiatif. *Support* dan *confidence* dituliskan sebagai Persamaan 1 dan Persamaan 2.

Jika *support itemset* dari *itemset I* memenuhi minimum *support threshold* yang sudah ditentukan, maka *I* adalah *frequent G-itemset*. Secara umum *frequent G-itemset* dilambangkan dengan *Lk*. Berdasarkan Persamaan (2) diperoleh dengan persamaan (3).

$$\text{support}(A \Rightarrow B) = P(A \cup B) \quad (1)$$

$$\text{confidence}(A \Rightarrow B) = P(B | A) \quad (2)$$

$$\text{confidence}(A \Rightarrow B) = P(B|A) = \frac{\text{support}(A \cup B)}{\text{support}(A)} \quad (3)$$

### Algoritma apriori

Algoritma Apriori adalah salah satu algoritma untuk melakukan pencarian *frequent itemset* dengan *association rules*. Algoritma Apriori menggunakan pendekatan *level-wise search*, dimana k-itemset digunakan untuk memperoleh (k+1)-itemset. Proses ini dilakukan hingga tidak ada lagi kombinasi yang dapat dibentuk.

Berdasarkan persamaan (1), (2), dan (3), diperoleh persamaan (4), (5), dan (6). Persamaan (4) digunakan untuk menghitung nilai *support* untuk sebuah item. Persamaan (5) digunakan untuk menghitung nilai *support* dari 2 item. Persamaan (6) digunakan untuk menghitung nilai *confidence* dari rule A:B.

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi yang mengandung A}}{\text{Total Transaksi}} \quad (4)$$

$$\text{Support (A, B)} = \frac{\text{Jumlah Transaksi yang mengandung A dan B}}{\text{Total Transaksi}} \quad (5)$$

$$\text{Confidence (A} \rightarrow \text{B)} = \frac{\text{Jumlah Transaksi yang mengandung A dan B}}{\text{Jumlah Transaksi yang mengandung A}} \quad (6)$$

## 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini berupa data alumni dan data mahasiswa aktif yang diperoleh dari kumpulan data akademik program studi sistem informasi UTDI. Data alumni terdiri dari beberapa item data yang merupakan rekapitulasi alumni (lulusan) dalam rentang waktu 1 tahun pada dua periode wisuda di tahun 2022. Beberapa item data alumni yang dikumpulkan antara lain nim, nama, IPS2, IPS4, organisasi (ORG), pemrograman (PRG), tahun masuk (TM), masa studi (MS) dan status lulus (SL).

Nilai atribut nim dan nama tidak dituliskan rill tetapi nilai yang mewakili (contoh: nim001 dan nama001). Atribut IPS2 dan IPS4 adalah index prestasi yang diperoleh pada semester 2 dan semester 4. Atribut organisasi adalah tingkat keaktifan kegiatan organisasi kampus (Aktif/Tidak Aktif), Sedangkan atribut pemrograman menunjukkan level penguasaan terhadap salah satu atau beberapa bahasa pemrograman (Tidak Bisa/Bisa). Tahun masuk adalah tahun saat mendaftar menjadi mahasiswa, sedangkan masa studi adalah waktu penyelesaian studi (tahun, bulan). Status lulus adalah ketepatan waktu studi berdasarkan lama studi ideal (Tepat waktu  $\leq 4$  tahun 0 bulan; Terlambat  $> 4$  tahun 0 bulan). Data alumni secara lengkap sebagaimana disajikan pada Tabel 1. Selain data alumni, data lain yang diperlukan untuk pengujian (testing) model prediksi kelulusan adalah data mahasiswa aktif yang disajikan pada Tabel 2.

**Tabel 1. Data alumni**

<i>NIM</i>	<i>NAMA</i>	<i>IPS2</i>	<i>IPS4</i>	<i>ORG</i>	<i>PRG</i>	<i>TM</i>	<i>MS</i>	<i>SL</i>
nim001	nama002	3,43	2,54	Tidak Aktif	Bisa	2017	4,20	Terlambat
nim002	nama003	3,21	3,47	Aktif	Bisa	2018	4,00	Tepat Waktu
nim003	nama005	2,85	2,50	Tidak Aktif	Bisa	2017	4,10	Terlambat
nim004	nama009	3,93	2,45	Tidak Aktif	Bisa	2017	4,50	Terlambat
nim005	nama011	3,58	2,89	Tidak Aktif	Bisa	2017	4,60	Terlambat
nim006	nama012	3,06	3,69	Aktif	Tidak Bisa	2018	4,00	Tepat Waktu
...	...	...	...	...	...	...	...	...
nim100	nama097	3,17	2,45	Aktif	Bisa	2018	4,00	Tepat Waktu

**Tabel 2. Data mahasiswa aktif**

<i>NIM</i>	<i>NAMA</i>	<i>IPS2</i>	<i>IPS4</i>	<i>ORG</i>	<i>PRG</i>	<i>TM</i>
nimA001	mhs001	2,71	3,1	Tidak Aktif	Tidak Bisa	2018
nimA002	mhs002	3,08	4	Aktif	Tidak Bisa	2018

nimA003	mhs003	2,67	2,97	Tidak Aktif	Tidak Bisa	2018
nimA004	mhs004	2,52	3,68	Aktif	Bisa	2019
nimA005	mhs005	3,44	2,99	Tidak Aktif	Tidak Bisa	2019
nimA006	mhs006	2,7	2,3	Aktif	Bisa	2019
nimA007	mhs007	3,5	2,9	Aktif	Tidak Bisa	2019
nimA008	mhs008	3,32	3,63	Tidak Aktif	Tidak Bisa	2019
nimA009	mhs009	3,75	3,43	Aktif	Bisa	2019
nimA010	mhs010	3,71	3,62	Tidak Aktif	Tidak Bisa	2019
nimA011	mhs011	2,24	2,48	Tidak Aktif	Bisa	2019
nimA012	mhs012	2,78	2,37	Aktif	Tidak Bisa	2019
nimA013	mhs013	3,54	2,6	Aktif	Tidak Bisa	2019
nimA014	mhs014	2,42	3,71	Aktif	Tidak Bisa	2019
...	...	...	...	...	...	...
nimA100	mhs100	4	3,61	Aktif	Bisa	2021

### Pengembangan Model

Pengembangan model dilakukan melalui beberapa tahapan. Tahap pertama adalah melakukan kategorisasi atau pengelompokan data. Selanjutnya menghitung frekuensi kemunculan setiap kategori variabel. Tahapan terakhir adalah pembentukan frequent itemset dan pembentukan aturan asosiasi menggunakan algoritma apriori.

#### a. Kategorisasi data

Kategorisasi atau pengelompokan data dilakukan untuk mengurangi jumlah sebaran nilai. Pengelompokan data berdasarkan rentang tertentu akan mempermudah dalam penghitungan frekuensi kemunculan data. Proses pengelompokan data dilakukan terhadap faktor IPS (2 dan 4), ORG, PRG dan SL. Kategorisasi data sebagaimana disajikan pada Tabel 3, Tabel 4, Tabel 5 dan Tabel 6.

**Tabel 2. Kategorisasi IPS**

Kategori	Rentang Nilai	Keterangan Nilai
0	$IPS \leq 2,50$	Kurang
1	$IPS \geq 2,5$	Baik

**Tabel 4. Kategorisasi ORG**

Kategori	Keterangan Nilai
0	Tidak Aktif
1	Aktif

**Tabel 5. Kategorisasi PRG**

Kategori	Keterangan Nilai
0	Tidak Bisa
1	Bisa

**Tabel 6. Kategorisasi SL**

Kategori	Keterangan Nilai
0	Terlambat
1	Tepat Waktu

#### b. Pembentukan Kandidat Itemset

Berdasarkan data alumni (Tabel 1) dan data mahasiswa aktif (Tabel 2), ditentukan variabel (item) yang berpengaruh terhadap aturan status kelulusan. Item tersebut sebagaimana disajikan pada Tabel 7.

**Tabel 7. Daftar Item**

No	Item	Keterangan
1	IPS2	Index prestasi semester 2 dengan nilai baik dan kurang
2	IPS4	Index prestasi semester 4 dengan nilai baik dan kurang
3	ORG	Keaktifan organisasi

4	PRG	Kemampuan pemrograman
5	SL	Status kelulusan

c. Pembentukan frequent itemset (K-Itemset)

Setiap item yang berpengaruh terhadap status kelulusan akan di hitung nilai *support* berdasarkan kemunculan item (*frequent itemset*) tersebut. Pada tahap awal ditentukan tingkat kemunculan terhadap 1 itemset. Selanjutnya perhitungan 2-itemset, 3-itemset sampai dengan K-itemset ditentukan berdasarkan kombinasi itemset dengan minimum *support* 0,25 atau 25%. Penentuan 1-itemset hingga 4-itemset dengan mengacu pada kategorisasi data diperoleh kombinasi sebagaimana disajikan pada Tabel 8, 9, 10 dan Tabel 11.

**Tabel 8.** 1-itemset

Item	Jumlah	Support	% Support
IPS2=1	72	0,72	72%
IPS2=0	28	0,28	28%
IPS4=1	79	0,79	79%
IPS4=0	21	0,21	21%
ORG=0	63	0,63	63%
ORG=1	37	0,37	37%
PRG=0	33	0,33	33%
PRG=1	67	0,67	67%
SL=0	24	0,24	24%
SL=1	76	0,76	76%

**Tabel 9.** 2-itemset

Item	Jumlah	Support	% Support
IPS2=1 IPS4=1	59	0,59	59%
IPS2=1 ORG=0	43	0,43	43%
IPS2=1 ORG=1	29	0,29	29%
IPS2=1 PRG=0	25	0,25	25%
IPS2=1 PRG=1	47	0,47	47%
IPS2=1 SL=0	16	0,16	16%
IPS2=1 SL=1	56	0,56	56%
IPS2=0 IPS4=1	20	0,20	20%
IPS2=0 ORG=0	20	0,20	20%
IPS2=0 PRG=1	20	0,20	20%
IPS2=0 SL=1	20	0,20	20%
IPS4=1 ORG=0	50	0,50	50%
IPS4=1 ORG=1	29	0,29	29%
IPS4=1 PRG=0	29	0,29	29%
IPS4=1 PRG=1	50	0,50	50%
IPS4=1 SL=0	20	0,20	20%
IPS4=1 SL=1	59	0,59	59%
IPS4=0 PRG=1	17	0,17	17%
IPS4=0 SL=1	17	0,17	17%
ORG=0 PRG=0	22	0,22	22%
ORG=0 PRG=1	41	0,41	41%
ORG=0 SL=0	22	0,22	22%
ORG=0 SL=1	41	0,41	41%

ORG=1 PRG=1	26	0,26	26%
ORG=1 SL=1	35	0,35	35%
PRG=0 SL=0	24	0,24	24%
PRG=1 SL=1	67	0,67	67%

**Tabel 10.** 3-itemset

<i>Item</i>	<i>Jumlah</i>	<i>Support</i>	<i>% Support</i>
IPS2=1 IPS4=1 ORG=0	36	0,36	36%
IPS2=1 IPS4=1 PRG=1	36	0,36	36%
IPS2=1 IPS4=1 SL=1	45	0,45	45%
IPS2=1 ORG=0 PRG=1	27	0,27	27%
IPS2=1 ORG=0 SL=1	27	0,27	27%
IPS2=1 ORG=1 SL=1	29	0,29	29%
IPS2=1 PRG=1 SL=1	47	0,47	47%
IPS4=1 ORG=0 PRG=1	32	0,32	32%
IPS4=1 ORG=0 SL=1	32	0,32	32%
IPS4=1 ORG=1 SL=1	27	0,27	27%
IPS4=1 PRG=1 SL=1	50	0,50	50%
ORG=0 PRG=1 SL=1	41	0,41	41%
ORG=1 PRG=1 SL=1	26	0,26	26%

**Tabel 11.** 4-itemset

<i>Item</i>	<i>Jumlah</i>	<i>Support</i>	<i>% Support</i>
IPS2=1 IPS4=1 PRG=1 SL=1	36	0,36	36%
IPS2=1 ORG=0 PRG=1 SL=1	27	0,27	27%
IPS4=1 ORG=0 PRG=1 SL=1	32	0,32	32%

#### d. Pembentukan pola aturan asosiasi

Proses selanjutnya adalah menghitung nilai *confidence* berdasarkan n-itemset yang terbentuk. Pembentukan aturan asosiasi dilakukan untuk mendapatkan nilai *confidence* (persamaan 6) guna melihat relasi item-item dalam algoritma apriori. Ditentukan dahulu nilai *confidence* minimum yaitu 90% untuk menghitung nilai *confidence* terhadap kombinasi 2-itemset, 3-itemset dan 4-itemset hingga dapat ditentukan aturan (*rule*) untuk menentukan status kelulusan sebagaimana terlihat pada Tabel 12.

**Tabel 12.** Pehitungan nilai *confidence*

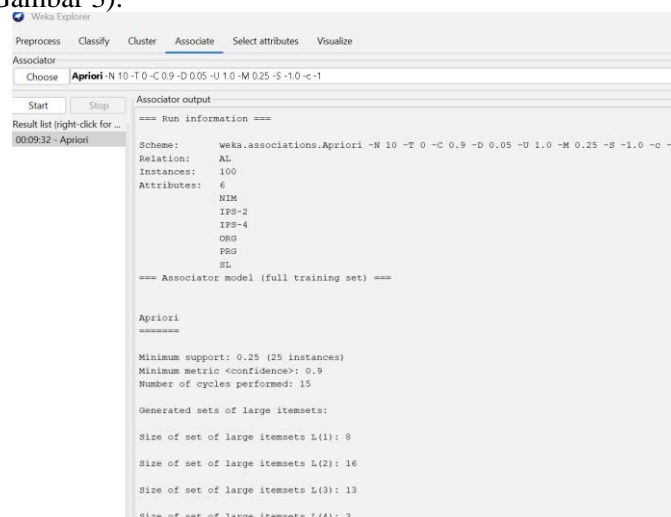
<i>Kombinasi itemset</i>	<i>Support</i>	<i>Confidence</i>	<i>%Confidence</i>
IPS-2=1 IPS-4=1	0,59	0,82	82%
IPS-2=1 ORG=0	0,43	0,60	60%
IPS-2=1 ORG=1	0,29	0,40	40%
IPS-2=1 PRG=0	0,25	0,35	35%
IPS-2=1 PRG=1	0,47	0,65	65%
IPS-2=1 SL=1	0,56	0,78	78%
IPS-4=1 ORG=0	0,5	0,63	63%
IPS-4=1 ORG=1	0,29	0,37	37%
IPS-4=1 PRG=0	0,29	0,37	37%
IPS-4=1 PRG=1	0,5	0,63	63%
IPS-4=1 SL=1	0,59	0,75	75%

ORG=0 PRG=1	0,41	0,65	65%
ORG=0 SL=1	0,41	0,65	65%
ORG=1 PRG=1	0,26	0,70	70%
ORG=1 SL=1	0,35	0,95	95%
PRG=1 SL=1	0,67	1,00	100%
IPS-2=1 IPS-4=1 ORG=0	0,36	0,50	50%
IPS-2=1 IPS-4=1 PRG=1	0,36	0,50	50%
IPS-2=1 IPS-4=1 SL=1	0,45	0,63	63%
IPS-2=1 ORG=0 PRG=1	0,27	0,38	38%
IPS-2=1 ORG=0 SL=1	0,27	0,38	38%
IPS-2=1 ORG=1 SL=1	0,29	0,40	40%
IPS-2=1 PRG=1 SL=1	0,47	0,65	65%
IPS-4=1 ORG=0 PRG=1	0,32	0,41	41%
IPS-4=1 ORG=0 SL=1	0,32	0,41	41%
IPS-4=1 ORG=1 SL=1	0,27	0,34	34%
IPS-4=1 PRG=1 SL=1	0,5	0,63	63%
ORG=0 PRG=1 SL=1	0,41	0,65	65%
ORG=1 PRG=1 SL=1	0,26	0,70	70%
IPS-2=1 IPS-4=1 PRG=1 SL=1	0,36	0,50	50%
IPS-2=1 ORG=0 PRG=1 SL=1	0,27	0,38	38%
IPS-4=1 ORG=0 PRG=1 SL=1	0,32	0,41	41%

Mengacu pada nilai minimum *confidence*, diperoleh 2 aturan yang memenuhi yaitu Jika ORG=1 (organisasi aktif) maka cenderung SL=1 (lulus tepat waktu) sebesar 95% dan Jika PRG=1 (bisa pemrograman) Maka cenderung SL=1 (lulus tepat waktu) sebesar 100%.

### Pengujian Model dan Evaluasi Model

Pengujian model dilakukan dengan aplikasi WeKa untuk mengetahui validitas model yang di hasilkan (Gambar 3).



**Gambar 3.** Pengujian menggunakan aplikasi WeKa

Proses pengujian ini dilakukan memberikan prediksi terhadap data mahasiswa aktif (Tabel 2) menggunakan aturan yang terbentuk.



#### 4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa aplikasi data mining dengan menggunakan pendekatan aturan asosiasi dan algoritma apriori dapat digunakan untuk menampilkan informasi aturan status kelulusan. Informasi yang ditampilkan berupa nilai support dan confidence. Semakin tinggi nilai confidence dan support maka semakin kuat nilai hubungan antar atribut/item. Data alumni dan data mahasiswa yang diproses meliputi IPS2, IPS4, ORG, PRG dan SL. Penelitian ini dapat dikembangkan lebih lanjut dengan menggunakan metode yang berbeda guna mengetahui bebarapa perbedaan yang mungkin muncul seperti kecepatan, akurasi dan kombinasi aturan yang dihasilkan.

#### DAFTAR PUSTAKA

- [1] Anoname. (2021). Profil Universitas Teknologi Digital Indonesia (UTDI). tersedia pada tautan: <https://www.utdi.ac.id/>, diakses pada 12 Agustus 2023
- [2] Linoff , G. S., & Berry, M. J. (2011). *Data mining Techniques for Marketing, Sales, Customer Relationship Management*. United States of America: Wiley Publishing, Inc.
- [3] Bramer, M. (2007). *Principles of Data mining*. London: Springer-Verlag.
- [4] David, O. & Yong, S. (2011). *Introduction to Business Data mining*. International Edision: Mc Graw Hill.
- [5] Sensuse, G.G. (2011). *Penerapan Metode Data mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (FP-Growth) : Studi Kasus Percetakan PT. Gramedia*. Jurnal TELEMATIKA , 4 (1), ISSN 118-132.
- [6] Setiawan, R. (2021). *Apa itu Data mining dan Bagaimana Metodenya?*, tersedia pada tautan: <https://www.dicoding.com/blog/apa-itu-data-mining/>, diakses pada 12 Agustus 2023