

APLIKASI DATA MINING MENGGUNAKAN ALGORITME APRIORI MENAMPILKAN TINGKAT KELULUSAN MAHASISWA TEKNOLOGI INFORMASI UNIVERSITAS INDRAPRASTA PGRI

Lukas Arief Prasetyo¹, Redo Abeputra Sihombing², Imam Himawan³.

¹ Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Indraprasta PGRI

² Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Indraprasta PGRI

³ Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Indraprasta PGRI

¹²³ JL. Raya Tengah No.80, RT.6/RW.1 Gedong Kec. Ps. Rebo Kota Jakarta Timur

¹ lukasarieprasetyo@gmail.com, ² redoabe@gmail.com, ³ imamhimawann@gmail.com

Abstrak :Latar belakang masalah Kelulusan merupakan item penilaian dalam proses akreditasi perguruan tinggi. Data mahasiswa Fakultas Teknik Informatika, Universitas Indraprasta angkatan 2013-2015 yang lulus dari waktu tahun 2016-2019 dimana 606 dari 1755 mahasiswa menempuh masa studi lebih dari 8 semester. Menunjukkan bahwa 34,53% mahasiswa (S1) menempuh lama studi lebih dari yang dijadwalkan yaitu 8 semester. Memanfaatkan data mahasiswa diharapkan dapat menghasilkan informasi kelulusan melalui teknik data mining. Metode Association Rules merupakan metode yang tepat untuk memecahkan masalah keterkaitan antar suatu item, dengan menggunakan algoritme apriori dapat ditemukan pola-pola data mahasiswa dengan tingkat kelulusan. Penulis melakukan penelitian ini dengan tujuan agar dapat menerapkan algoritme apriori pada data mahasiswa untuk mencari kombinasi item terbanyak berdasarkan data mahasiswa dan kemudian membentuk pola asosiasi dari kombinasi item tersebut. Data hasil pengujian dengan menggunakan sistem aplikasi data mining algoritme Apriori terbukti menghasilkan 2 aturan asosiasi dari 341 data yang diuji dengan nilai minimum support 50 dan confidence 90% yaitu jika Mahasiswa kota asal sekolah nya di Kodya Tangerang dan kategori kelulusannya A2 (Lama studi 4 tahun atau kurang dari 4 tahun dan IPK 2,76-3,50) maka Mahasiswa proses masuk nya dengan Nilai Rapor ≥ 7.00 dengan nilai support 19,35% , confidence 91,67%, dan lift ratio 1,06 dan Jika siswa asal sekolahnya dari SMK dan kategori kelulusannya A2 (Lama 4 tahun atau kurang dari 4 tahun dan IPK 2,76-3,50) kesimpulan Mahasiswa masuk nya dengan Nilai Rapor ≥ 7.00 dengan nilai support 16,13%, confidence 92,96% dan lift ratio 1,07.

Kata kunci : data mining, algoritme apriori, kelulusan, data.

Abstract: Background to the problem Graduation is an assessment item in the higher education accreditation process. Data on students from the Faculty of Information Technology, Indraprasta University class of 2013-2015 who graduated from 2016-2019, where 606 out of 1755 students took a study period of more than 8 semesters. Shows that 34.53% of students (S1) took more study time than scheduled, namely 8 semesters. Utilizing student data is expected to produce graduation information through data mining techniques. The Association Rules method is an appropriate method for solving the problem of relationships between items. By using an a priori algorithm, student data patterns with graduation rates can be found. The author conducted this research with the aim of being able to apply an a priori algorithm to student data to find the most combinations of items based on student data and then form

association patterns from these item combinations. Data from testing using the Apriori algorithm data mining application system was proven to produce 2 association rules from 341 data tested with a minimum support value of 50 and confidence of 90%, namely if the student's hometown school is in Kodya Tangerang and the graduation category is A2 (length of study 4 years or less than 4 years old and a GPA of 2.76-3.50) then students are admitted with a report card score of ≥ 7.00 with a support value of 19.35%, confidence of 91.67%, and a lift ratio of 1.06 and if the student is from the school from vocational school and the graduation category is A2 (length of 4 years or less than 4 years and GPA 2.76-3.50) conclusion: Students enter with a report card score ≥ 7.00 with a support value of 16.13%, confidence 92.96% and lift ratio 1.07.

Keywords: data mining, a priori algorithm, graduation, data

PENDAHULUAN

Perguruan Tinggi adalah tempat Pendidikan bagi seseorang yang telah selesai di sekolah menengah atas. pertumbuhan Mahasiswa merupakan asset yang sangat penting disebuah perguruan tinggi. Maka, faktor yang menentukan kualitas perguruan tinggi adalah persentasi kemampuan mahasiswa untuk menyelesaikan studi tepat waktu, Algoritma yang digunakan adalah algoritma apriori informasi yang ditampilkan berupa nilai support dan confidence (Saputro, 2015) . Data tentang mahasiswa yang lulus dapat memberikan informasi yang berguna bagi universitas jika dimanfaatkan dengan maksimal. Salah satu cara untuk memanfaatkan data tentang mahasiswa yang lulus ini dengan mengolahnya menggunakan data mining. Data – data yang tercipta dapat dimanfaatkan untuk hal – hal dimasa yang akan datang baik yang bersifat prediksi. mahasiswa yang lulus melebihi masa studi yang ditargetkan. Dalam hal ini universitas atau program studi dapat memanfaatkan data diri mahasiswa untuk memprediksi tingkat kelulusan mahasiswanya (Indah Puji Astuti, 2019). klasifikasi maupun dalam mengasosiasikan sesuatu, hal ini menjadi salah satu faktor pendorong terciptanya suatu bidang ilmu Data Mining. Beban Studi Program Strata 1 ditetapkan

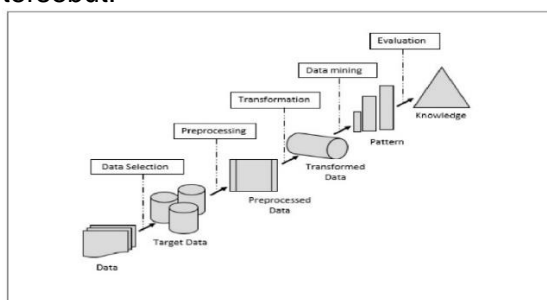
sekurang-kurangnya sebesar 144 SKS dalam masa studi kurangnya 7 semester dan selama-lamanya 14 semester, di jadwalkan dalam 8 semester dengan ketentuan tugas akhir harus sudah diambil selambat-lambatnya pada semester ke 13. Berdasarkan data mahasiswa Fakultas Teknik Informatika, Program Sarjana (S1) di Universitas Indraprasta angkatan 2013-2015 yang diasumsikan akan lulus dari rentang waktu tahun 2016-2019 dimana 606 dari 1755 mahasiswa menempuh masa studi lebih dari 8 semester. Hal ini menunjukkan bahwa 34,53% mahasiswa Program Sarjana (S1) menempuh lama studi lebih dari yang dijadwalkan yaitu 8 semester. Banyaknya kegiatan yang dilakukan secara komputerisasi sekarang ini mempermudah dalam ketersediaan data. Namun data yang tersedia seringkali hanya diperlakukan sebagai rekaman *record* tanpa pengolahan lebih lanjut sehingga tidak bernilai lebih untuk keperluan masa mendatang (Kusumo et al., 2019). Aturan asosiasi digunakan untuk penemuan pengetahuan dan pengambilan keputusan manajerial berdasarkan asosiasi antar data untuk membentuk system yang baik. Beberapa jenis analisis asosiasi pada data mining adalah algoritme Apriori merupakan aturan asosiasi dengan menggunakan frequent item. Salah satu aplikasi algoritme Apriori adalah dalam

bidang Pendidikan. Dalam bidang Pendidikan, data mahasiswa dapat digunakan untuk menentukan semua variabel yang memenuhi syarat minimum untuk support dan syarat minimum untuk *confidence*. aplikasi data mining menggunakan metode Algoritma apriori adalah untuk menentukaan pola (Saefudin & DN, 2019). Setelah proses mining akan disajikan hasil dari data mining berupa tabel hubungan kekuatan dengan nilai support dan *confidence* masing-masing atribut serta *threshold* yang digunakan.

KAJIAN PUSTAKA DAN LANDASAN TEORI

Data adalah segala fakta, angka, atau teks yang dapat diproses oleh komputer. Saat ini akumulasi pertumbuhan jumlah data berjalan dengan cepat dalam format dan basis yang berbeda (Syahra & Ginting, 2019). Walaupun sudah banyak perangkat lunak yang menawarkan kemampuan dalam proses data mining, keterlibatan manusia sangat dibutuhkan dalam setiap fase proses data mining itu sendiri. Seiring berjalannya waktu, perkembangan dan kemajuan infrastruktur di Indonesia bisa dikategorikan sangat pesat. Teknologi informasi dapat membantu pengguna untuk mempermudah dalam mendapatkan informasi secara cepat dan akurat. Dengan kelebihan tersebut, teknologi informasi telah memasuki dunia bisnis diberbagai bidang kehidupan (Mulyani, 2018).

Pendidikan merupakan salah satu faktor kemajuan dan kemandirian bangsa (Susanto & Sudyatno, 2014). Semakin maju pendidikan suatu bangsa, maka akan semakin maju dan mandiri bangsa tersebut.



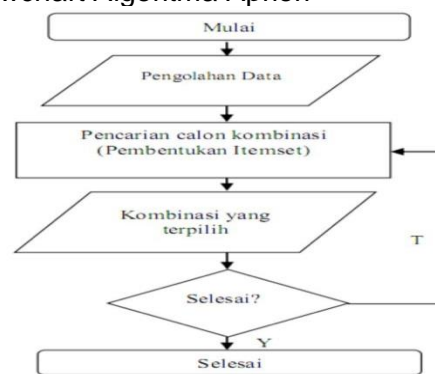
Gambar 1 Tahapan Data mining

Tahapan Data Mining

Tahap-tahap data mining yaitu :

1. Seleksi Data (Data Selection) Data yang ada pada database sering tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database.
2. Pembersihan Data/ Preprocessing (Data Cleaning) Pembersihan data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data tidak relevan.
3. Integrasi Data (Data Integration) Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru.
4. Transformasi Data (Data Transformation) Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining.
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 kedalam knowledge based yang ditemukan.
7. Presentasi Pengetahuan (knowledge presentation) merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna.

Flowchart Algoritma Apriori



Gambar 2 Flowchart Algoritma Apriori

algoritma apriori untuk menghasilkan kombinasi terdekat antar variabel. Teknik yang digunakan dalam aplikasi data

mining ini adalah aturan asosiasi dengan algoritma apriori. pada bidang data mining untuk penggalian aturan asosiasi atau yang lebih dikenal dengan istilah association rule mining (ARM) (Srikanti et al., 2018). Algoritma apriori ini melakukan proses iterasi untuk menghasilkan kombinasi item yang memiliki pola frekuensi tinggi, berdasarkan nilai ambang batas support dan confidence yang diberikan oleh user (Buaton et al., 2018).

METODE

Data Penelitian

Penulis mengobservasi data meliputi tabel tersedia terhadap data mahasiswa yang melekat terhadap atribut yang diperoleh dari universitas perguruan tinggi.

Tabel 1. Tabel Data Mahasiswa

Atribute	Keterangan
Tahun Angkatan	Merupakan tahun angkatan mahasiswa yang masuk ke Universitas Indraprasta
Tahun lahir	Merupakan tanggal mahasiswa yang bersangkutan dilahirkan
Tempat lahir	Merupakan Tempat dimana mahasiswa yang bersangkutan dilahirkan
Jenis Kelamin	Merupakan jenis kelamin mahasiswa yang bersangkutan
Agama	Merupakan agama yang dianut mahasiswa yang bersangkutan
Proses masuk	Merupakan jenis jalur masuk ke perguruan tinggi yang diikuti mahasiswa bersangkutan. Proses masuk universitas Indraprasta 2013-2016 merupakan nilai rapor dan beasiswa akademik
Prodi	Program studi dari mahasiswa yang bersangkutan
Asal sekolah	Merupakan asal sekolah menengah lanjutan mahasiswa yang bersangkutan
Kota asal sekolah	Merupakan kota asal sekolah menengah lanjutan dari mahasiswa yang bersangkutan
Tahun lulus	Merupakan tahun lulus dari asal sekolah menengah lanjutan mahasiswa yang bersangkutan
Jurusan sekolah	Merupakan jurusan di asal sekolah menengah lanjutan

Tanggal lulus	Merupakan tanggal mahasiswa yang bersangkutan dinyatakan lulus
IPK	Indeks Prestasi Kumulatif (IPK) adalah ukuran kemampuan mahasiswa sampai pada waktu tertentu yang dapat dihitung
Lama studi	Merupakan lama tempuh studi dihitung dimulai saat terdaftar sebagai mahasiswa sampai dinyatakan lulus
Pendidikan ayah	Merupakan pendidikan ayah dari mahasiswa yang bersangkutan
Periode wisuda	Merupakan periode wisuda yang diikuti oleh mahasiswa yang bersangkutan

Metode Pemecahan Masalah

Metode yang digunakan adalah analisis asosiasi atau association rule mining. Association Rule merupakan Sebuah aturan asosiasi, yang berupa aturan-aturan data apa saja yang memenuhi hasil dari minimum support dan minimum confident. Hasil dari association rule ini mengetahui hubungan tingkat kelulusan dengan data mahasiswa.

Tahap *processing*

Data Selection

Yaitu proses menyeleksi atribut apa yang akan diproses pada mining selanjutnya.

Reprocessing/Cleaning Data

Cleaning data merupakan proses menghilangkan noise dan data yang tidak konsisten atau data yang tidak relevan. Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen.

Data Transformation

penulis mencari keterkaitan antara tingkat kelulusan dengan data mahasiswa. Tingkat kelulusan mahasiswa dapat dilihat dari lama studi dan IPK (Indeks Prestasi Kumulatif). Dari dua atribut tersebut data diubah menjadi tipe data yang memudahkan untuk diproses. Tingkat kelulusan diukur dari lama studi dan IPK, lama studi dikategorikan berdasarkan peraturan akademik Universitas Indraprasta.

Tabel 2 Predikat Kelulusan

Indeks Prestasi Kumulatif	Predikat
---------------------------	----------

2,00 – 2,75	Memuaskan
2,76 – 3,50	Sangat Memuaskan
3,51 – 4,00	Dengan Pujian (cumlaude)

Tabel 3 Transformasi Data

Kategori	Keterangan
A1	Lama studi 4 tahun kurang dari 4 tahun dan IPK 3,51 – 4,00
A2	Lama studi 4 tahun kurang dari 4 tahun dan IPK 2,76 – 3,50
A3	Lama studi 4 tahun kurang dari 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
B3	Lama studi lebih dari 4 tahun dan IPK 2,00– 2,75

1. Rancangan Pengujian

Pengujian pertama menggunakan 100 Perpustakaan Universitas Indraprasta 21 data dan kedua dengan 344 data mahasiswa angkatan 2015 fakultas teknik informatika di Universitas Indraprasta (Nurjoko & Kurniawan, 2016). Hal ini bertujuan untuk mengetahui aturan (*rules*) yang terbentuk apakah berpengaruh dengan jumlah data yang dipakai. Data yang penulis gunakan adalah data dengan format Microsoft Excel (*.xlsx). Association (asosiasi) adalah model yang penulis pakai pada tahap ini dengan menggunakan algoritme Apriori untuk mencari aturan asosiasi tingkat kelulusan dengan menggunakan data mahasiswa.

Tabel 4 Sampel Data Mahasiswa

No	Atribut
1	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A1
2	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMAN,A1
3	Nilai Rapor ≥ 7.00 , Jakarta Selatan, SMA,A2
4	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A2
5	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMKN,A2
6	Nilai Rapor ≥ 7.00 , Jakarta Barat, SMK,A2
7	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A2
8	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A2

9	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A2
10	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A1
11	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,B2
12	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMAN,B2
13	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A2
14	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,B2
15	Nilai Rapor ≥ 7.00 , Jakarta Barat, SMK,A2
16	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMAN,B2
17	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,B2
18	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A1
19	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,B2
20	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A1
21	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A2
22	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,A2
23	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMA,B2
24	Nilai Rapor ≥ 7.00 , Jakarta Selatan,MA,A1
25	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMKN,A2
26	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMA,A2
27	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,B2
28	Nilai Rapor ≥ 7.00 , Kab Tangerang, SMK,B2
29	Nilai Rapor ≥ 7.00 , Jakarta Barat, SMKN,A2
30	Nilai Rapor ≥ 7.00 , Jakarta Selatan, SMK,B2
31	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,A1
32	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMA,B2
33	Nilai Rapor ≥ 7.00 , Jakarta Selatan, SMAN,A1
34	Nilai Rapor ≥ 7.00 , Jakarta Selatan, SMK,B2
35	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMK,B2
36	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMAN,A2
37	Nilai Rapor ≥ 7.00 , Jakarta Barat, SMA,A1
38	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMA,A2
39	Nilai Rapor ≥ 7.00 , Kodya Tangerang, SMA,A2

40	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,A2
41	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A1
42	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,A2
43	Nilai Rapor >=7.00, Kodya Tangerang ,SMKN,A2
44	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,A1
45	Nilai Rapor >=7.00, Kab Tangerang ,SMK,B2
46	Nilai Rapor >=7.00, Kab Tangerang ,SMK,B2
47	Nilai Rapor >=7.00, Kab Tangerang ,MA,A2
48	Nilai Rapor >=7.00, Kab Tangerang ,SMKN,A2
49	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,A2
50	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,B2
51	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,B2
52	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,A2
53	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
54	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,B2
55	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,A2
56	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,B2
57	Nilai Rapor >=7.00, Kab Tangerang ,SMK,A2
58	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,B2
59	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,A2
60	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
61	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,A2
62	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
63	Nilai Rapor >=7.00, Kab Tangerang ,SMKN,A2
64	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
65	Nilai Rapor >=7.00, Kab Tangerang ,SMA,B2
66	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,A2
67	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,A1
68	Nilai Rapor >=7.00, Jakarta Selatan ,SMAN,A1
69	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A1
70	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,B2

71	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,A2
72	Nilai Rapor >=7.00, Jakarta Selatan ,SMAN,A1
73	Nilai Rapor >=7.00, Kab Tangerang ,SMK,A2
74	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,B2
75	Nilai Rapor >=7.00, Kab Tangerang ,SMK,B2
76	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
77	Beasiswa Akademik, Jakarta Selatan, SMKN,A1
78	Beasiswa Akademik, Kodya Tangerang, SMK,A1
79	Beasiswa Akademik, Kodya Tangerang, SMK,B2
80	Beasiswa Akademik, Kodya Tangerang, SMA,A2
81	Beasiswa Akademik, Kab Tangerang, SMKN,A1
82	Beasiswa Akademik, Kodya Tangerang, SMKN,B2
83	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,A2
84	Nilai Rapor >=7.00, Kab Tangerang ,SMK,A2
85	Nilai Rapor >=7.00, Kab Tangerang ,SMK,A2
86	Nilai Rapor >=7.00, Jakarta Selatan ,SMAN,A1
87	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,B2
88	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,A2
89	Nilai Rapor >=7.00, Kodya Tangerang ,SMA,B2
90	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,B2
91	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
92	Nilai Rapor >=7.00, Jakarta Selatan,SMK,A2
93	Nilai Rapor >=7.00, Kodya Tangerang ,SMK,A2
94	Nilai Rapor >=7.00, Kab Tangerang ,SMK,A2
95	Nilai Rapor >=7.00, Jakarta Selatan ,SMK,A2
96	Nilai Rapor >=7.00, Kodya Tangerang ,SMAN,A2
97	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,A2
98	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,A2
99	Nilai Rapor >=7.00, Kab Tangerang ,SMAN,A2
100	Nilai Rapor >=7.00, Kab Tangerang ,SMK,B2

- a. Penentuan Pola Frekuensi Tinggi (*Frequent Pattern*)
 Melakukan dengan mengkombinasi semua jenis *Itemset* dan pemasangan jika kandidat tidak memenuhi nilai *minimum support* yang diberikan. Untuk mendapatkan nilai *support* pada setiap kandidat kombinasi *itemset* dapat diperoleh dengan rumus sebagai berikut:
 - Nilai Support 1 item
 - Nilai Support 2 item
- b. Pembentukan Aturan Asosiasi data mahasiswa ditentukan pola *large item set* yang terbentuk
- c. Pembuktian kekuatan suatu aturan, dibutuhkan perhitungan *liftratio* untuk mengukur kevalidan

HASIL DAN PEMBAHASAN

1. Tahapan Pengumpulan Data

Tahap pengumpulan data merupakan tahap yang pertama kali dilakukan sebelum tahap pemrosesan data. Data pada penelitian tugas akhir ini diperoleh dari Universitas Indraprasta. Pada penelitian ini data yg digunakan adalah data mahasiswa angkatan 2013,2014,2015. Diasumsikan bahwa mahasiswa angkatan 2013-2015 akan lulus dari rentang waktu tahun 2016-2019. Adapun langkah-langkah yang akan penulis lakukan untuk proses data mining yang pertama yaitu dengan perhitungan manual menggunakan 100 data mahasiswa sebagai data sampel yang ada pada bab 3, selanjutnya menggunakan 100 dan 341 data mahasiswa sebagai data testing dengan menggunakan system aplikasi data mining ini.

2. Tahap Apriori

Setelah data mahasiswa melalui tahap preprocessing, kemudian masuk ke tahap perancangan metode Apriori. Langkah pertama pada tahap ini adalah menetapkan nilai minimum support atau threshold yaitu 20, kemudian dilakukan pencarian kombinasi semua jenis itemset dengan cara melakukan iterasi menscan

database untuk mendapatkan kandidat kombinasi itemset, kandidat yang tidak memenuhi nilai minimum support akan dihapus. Dari hasil seleksi kandidat 1-itemset (C-1) digunakan untuk menentukan kandidat dengan 2-itemset (C-2) dan seterusnya sampai tidak terdapat kombinasi lagi.

Tabel 5 Kandidat 1 Itemset

No	Item	Jumlah	Support
1	Nilai Rapor >=7.00	92	92,00
2	Kab, Tangerang	33	33,00
3	SMK	51	51,00
4	A1	20	20,00
5	Kedoya Tangerang	41	41,00
6	SMAN	22	22,00
7	Jakarta Selatan	20	20,00
8	SMA	16	16,00
9	A2	50	50,00
10	SMKN	9	9,00
11	Jakarta Barat	4	4,00
12	B2	30	30,00
13	Kab. Tanggerang	2	2,00
14	MA	2	2,00
15	Nilai Rapor < 7.00	2	2,00
16	Beasiswa Akademik	6	6,00

Dari tabel 5 dapat diketahui ada 7 itemset yang tidak memenuhi minimum support atau threshold 20, maka untuk iterasi selanjutnya 7 itemset tersebut akan dihapus. Kandidat itemset yang memenuhi nilai threshold atau minimum support 20.

Tabel 6 Frequent 1 Itemset

No	Item	Jumlah	Support
1	Nilai Rapor >=7.00	92	92,00
2	Kab, Tangerang	33	33,00
3	SMK	51	51,00
4	A1	20	20,00
5	Kedoya Tangerang	41	41,00
6	SMAN	22	22,00
7	Jakarta Selatan	20	20,00
8	A2	50	50,00
9	B2	30	30,00

Setelah diketahui kandidat 1 itemset yang memenuhi nilai support pada seleksi pertama, selanjutnya dilakukan iterasi

dengan menscan itemset yang terdapat pada tabel 6 untuk mengetahui kandidat kombinasi 2 itemset yang memenuhi nilai minimum support atau threshold 20 dan hasilnya dapat dilihat pada tabel 7 sebagai kandidat 2 itemset dan tabel 7 sebagai kandidat yang memenuhi nilai minimum support atau threshold 20.

Tabel 7 Kandidat 2 Itemset

No	Item1	Item2	Jumlah	Support
1	Nilai Rapor >= 7.00	Kab. Tangerang	32	32,00
2	Nilai Rapor >= 7.00	SMK	48	48,00
3	Nilai Rapor >= 7.00	A1	17	17,00
4	Nilai Rapor >= 7.00	Kodya Tangerang	37	37,00
5	Nilai Rapor >= 7.00	SMAN	22	22,00
6	Nilai Rapor >= 7.00	Jakarta Selatan	18	18,00
7	Nilai Rapor >= 7.00	A2	47	47,00
8	Nilai Rapor >= 7.00	B2	28	28,00
9	Kab. Tangerang	SMK	20	20,00
10	Kab. Tangerang	A1	4	4,00
11	Kab. Tangerang	Kodya Tangerang	0	0,00
12	Kab. Tangerang	SMAN	7	7,00
13	Kab. Tangerang	Jakarta Selatan	0	0,00
14	Kab. Tangerang	A2	15	15,00
15	Kab. Tangerang	B2	14	14,00
16	SMK	A1	8	8,00
17	SMK	Kodya Tangerang	17	17,00

18	SMK	SMAN	0	0,00
19	SMK	Jakarta Selatan	12	12,00
20	SMK	A2	27	27,00
21	SMK	B2	16	16,00
22	A1	Kodya Tangerang	6	6,00
23	A1	SMAN	7	7,00
24	A1	Jakarta Selatan	8	8,00
25	A1	A2	0	0,00
26	A1	B2	0	0,00
27	Kodya Tangerang	SMAN	10	10,00
28	Kodya Tangerang	Jakarta Selatan	0	0,00
29	Kodya Tangerang	A2	22	22,00
30	Kodya Tangerang	B2	13	13,00
31	SMAN	Jakarta Selatan	4	4,00
32	SMAN	A2	9	9,00
33	SMAN	B2	6	6,00
34	Jakarta Selatan	A2	10	10,00
35	Jakarta Selatan	B2	2	2,00
36	A2	B2	0	0,00

Kandidat yang terbentuk dengan 2 itemset Dari tabel 7 ada 9 data dan yang memenuhi nilai minimum support atau threshold 20.

Tabel 8 Frequent 2 Itemset

No	Item1	Item2	Jumlah	Support
1	Nilai Rapor >= 7.00	Kab. Tangerang	32	32,00
2	Nilai Rapor >= 7.00	SMK	48	48,00
3	Nilai Rapor >= 7.00	Kodya Tangerang	37	37,00
4	Nilai Rapor >= 7.00	SMAN	22	22,00
5	Nilai Rapor >= 7.00	A2	47	47,00
6	Nilai Rapor >= 7.00	B2	28	28,00
7	Kab. Tangerang	SMK	20	20,00
8	SMK	A2	27	27,00
9	Kodya Tangerang	A2	22	22,00

Setelah diketahui kandidat yang memenuhi nilai minimum support atau threshold dari pembentukan dengan 2 itemset, maka dilakukan kembali iterasi pembentukan kandidat dengan 3 itemset.

Tabel 9 Kandidat 3 Itemset

No	Item1	Item2	Item3	Jumlah	Support
1	Nilai Rapor >= 7.00	Kab. Tangerang	SMK	20	20,00
2	Nilai Rapor >= 7.00	Kab. Tangerang	Kodya Tangerang	0	0,00
3	Nilai Rapor >= 7.00	Kab. Tangerang	SMAN	7	7,00
4	Nilai Rapor >= 7.00	Kab. Tangerang	A2	15	15,00
5	Nilai Rapor >= 7.00	Kab. Tangerang	B2	14	14,00
6	Nilai Rapor >= 7.00	SMK	A2	26	26,00
7	SMK	Kab. Tangerang	A2	8	8,00
8	Nilai Rapor >= 7.00	Kodya Tangerang	A2	21	21,00
9	Kodya Tangerang	Kab. Tangerang	A2	0	0,00
10	Nilai Rapor >= 7.00	SMK	Kodya Tangerang	15	15,00
11	Nilai Rapor >= 7.00	SMK	SMAN	0	0,00
12	Nilai Rapor >= 7.00	SMK	B2	15	15,00
13	Kodya Tangerang	SMK	A2	9	9,00
14	Nilai Rapor >= 7.00	Kodya Tangerang	SMAN	10	10,00
15	Nilai Rapor >= 7.00	Kodya Tangerang	B2	11	11,00
16	Kab. Tangerang	Kodya Tangerang	SMK	0	0,00
17	Nilai Rapor >= 7.00	SMAN	A2	9	9,00

18	Nilai Rapor >= 7.00	SMAN	B2	6	6,00
19	Kab. Tangerang	SMAN	SMK	0	0,00
20	SMK	SMAN	A2	0	0,00
21	Kodya Tangerang	SMAN	A2	6	6,00
22	Nilai Rapor >= 7.00	A2	B2	0	0,00
23	Kab. Tangerang	B2	SMK	9	9,00
24	SMK	B2	A2	0	0,00
25	Kodya Tangerang	B2	A2	0	0,00

Dari iterasi pembentukan dengan 3 itemset terdapat 25 kandidat yang terbentuk, namun hanya ada 3 kandidat yang memenuhi nilai minimum support atau threshold.

Tabel 10 Frequent 3 Itemset

No	Item1	Item2	Item3	Jumlah	Support
1	Nilai Rapor >= 7.00	Kab. Tangerang	SMK	20	20,00
2	Nilai Rapor >= 7.00	SMK	A2	26	26,00
3	Nilai Rapor >= 7.00	Kodya Tangerang	A2	21	21,00

Aturan asosiatif dapat dilakukan setelah semua pola frekuensi tinggi didapatkan. Untuk menentukan aturan asosiasi minimal itemset pada proses pencarian pola frekuensi tinggi harus memiliki 2 kandidat A dan B. Pada kandidat berlaku hukum asosiasi $A \rightarrow B$ dan tidak berlaku $B \rightarrow A$. Sebelum menentukan aturan $A \rightarrow B$ terlebih dahulu menentukan nilai minimum confidence yaitu sebesar 80%.

Tabel 11 Data Pencarian Nilai Confidence

No	$X \Rightarrow Y$	Support X U Y	Support X	Confidence
----	-------------------	---------------	-----------	------------

1	Nilai Rapor ≥ 7.00 , SMK \Rightarrow A2	26	48	54,17
2	SMK \Rightarrow Kab. Tangerang, Nilai Rapor ≥ 7.00	20	51	39,22
3	Kab. Tangerang \Rightarrow Nilai Rapor ≥ 7.00 , SMK	20	33	60,61
4	Nilai Rapor $\geq 7.00 \Rightarrow$ SMK, Kab. Tangerang	20	92	21,74
5	SMK, Nilai Rapor $\geq 7.00 \Rightarrow$ Kab. Tangerang	20	48	41,67
6	Kab. Tangerang, SMK \Rightarrow Nilai Rapor ≥ 7.00	20	20	100
7	Nilai Rapor ≥ 7.00 , Kab. Tangerang \Rightarrow SMK	20	32	62,5
8	SMK, A2 \Rightarrow Nilai Rapor ≥ 7.00	26	27	96,3
9	A2, Nilai Rapor $\geq 7.00 \Rightarrow$ SMK	26	47	55,32
10	Nilai Rapor $\geq 7.00 \Rightarrow$ A2, SMK	26	92	28,26
11	SMK \Rightarrow Nilai Rapor ≥ 7.00 , A2	26	51	50,98
12	A2, SMK, Nilai Rapor ≥ 7.00	26	50	52,00
13	Nilai Rapor ≥ 7.00 , Kodya Tangerang \Rightarrow A2	21	37	56,76
14	Kodya Tangerang, A2 \Rightarrow Nilai Rapor ≥ 7.00	21	22	95,45
15	A2, Nilai Rapor $\geq 7.00 \Rightarrow$ Kodya Tangerang	21	47	44,68
16	Nilai Rapor $\geq 7.00 \Rightarrow$ A2, Kodya Tangerang	21	92	22,83
17	Kodya Tangerang \Rightarrow Nilai Rapor ≥ 7.00 , A2	21	41	51,22
18	A2 \Rightarrow Kodya Tangerang, Nilai Rapor ≥ 7.00	21	50	42,00

Setelah di bentuk aturan asosiasi maka di lakukan pencarian kombinasi itemset yang memenuhi minimum nilai confidence yang telah di tentukan di awal tadi sebesar 80%. Untuk membuktikan keakuratan kombinasi

yang terbentuk yaitu dengan menggunakan lift ratio. Nilai Lift ratio yang lebih besar dari 1 memiliki aturan yang kuat untuk item tersebut.

Tabel 12 Rule Asosiasi yang terbentuk

No	X \Rightarrow Y	Support X U Y	Support X	Confidence
1	Kab. Tangerang, SMK \Rightarrow Nilai Rapor ≥ 7.00	20	20	100
2	SMK, A2 \Rightarrow Nilai Rapor ≥ 7.00	26	27	96,3
3	Kodya Tangerang, A2 \Rightarrow Nilai Rapor ≥ 7.00	21	22	95,45

Setelah proses mining akan disajikan hasil dari data mining berupa tabel hubungan kekuatan dengan nilai support dan confidence masing- masing atribut serta threshold yang digunakan. Semakin tinggi nilai confidence dan support maka semakin kuat nilai hubungan antar atribut.

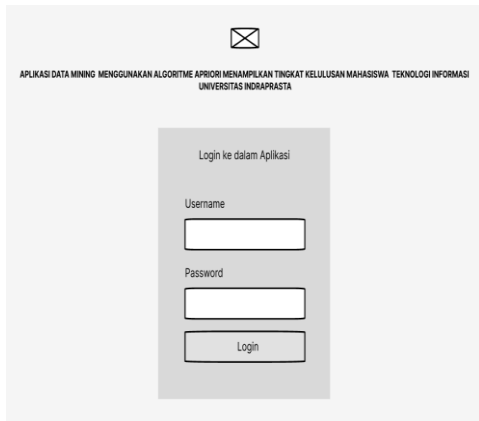
3. Hasil Pengujian

Sistem yang sudah diimplementasikan selanjutnya diuji, pengujian terhadap aplikasi yang dibuat ini bertujuan apakah aplikasi data mining ini berhasil iimplementasikan atau tidak untuk menghitung tingkat kelulusan mahasiswa. Pengujian pertama menggunakan 100 data dan kedua dengan 341 data mahasiswa angkatan 2015 fakultas teknik informatika di Universitas Indraprasta, hal ini bertujuan untuk mengetahui aturan (rules) yang terbentuk apakah berpengaruh dengan jumlah data yang dipakai

4. Rancangan Layar

a) Rancangan Layar Login

rancangan layar form login dimana pada menu ini user diharuskan melakukan login terlebih dahulu agar bisa masuk ke menu Halaman Utama dengan mengisi username dan password, lalu klik tombol Login.



Gambar 3 Rancangan layar login

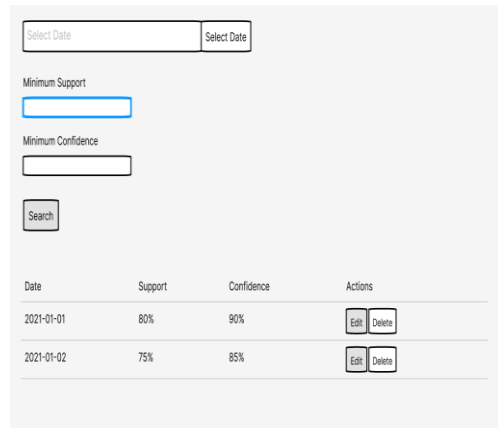
- b) Rancangan Layar Data Mahasiswa rancangan layar form Data Mahasiswa. Rancangan ini berfungsi untuk menambah data, menghapus, dan menampilkan data mahasiswa

DATA MAHASISWA

#	Tanggal	Attribut
1	2019-XX-XX	Nilai Rapor >=7,00, Kab Tangerang, SMK,A1
2	2019-XX-XX	Nilai Rapor >=7,00, Kodya Tangerang, SMAN,A1
3	2019-XX-XX	Nilai Rapor >=7,00, Jakarta Selatan, SMA,A2
4	2019-XX-XX	Nilai Rapor >=7,00, Kodya Tangerang, SMK,A2
5	2019-XX-XX	Nilai Rapor >=7,00, Kab Tangerang, SMKN,A2

Gambar 4 Rancangan layar data mahasiswa

- c) Rancangan Layar Proses Apriori rancangan layar proses Apriori. Rancangan ini berfungsi untuk menginput support dan confidencenya agar data bisa dihitung. dan rancangan ini menampilkan perhitungan support dan confidence dari data yang sudah di upload di form data mahasiswa



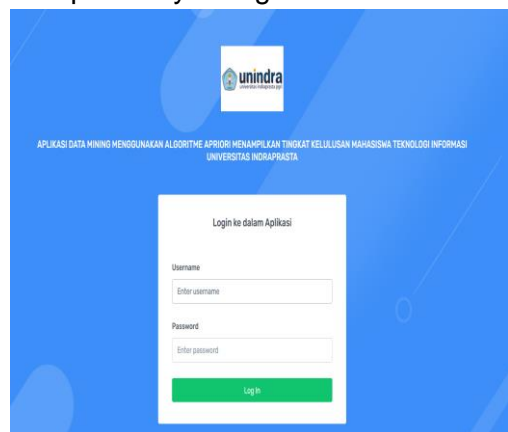
Gambar 5 Rancangan layar proses apriori

- d) Rancangan Layar Hasil Rancangan layar form hasil. Rancangan ini berfungsi untuk menampilkan hasil-hasil data yang sudah dihitung sesuai support dan confidence yang telah diinput oleh users

Tabel Hasil			
Nama Penguji	Waktu Pengujian	Support	Confidence
John Doe	2021-01-01	0.8	0.9
Jane Smith	2021-02-01	0.7	0.8
Mike Johnson	2021-03-01	0.6	0.7

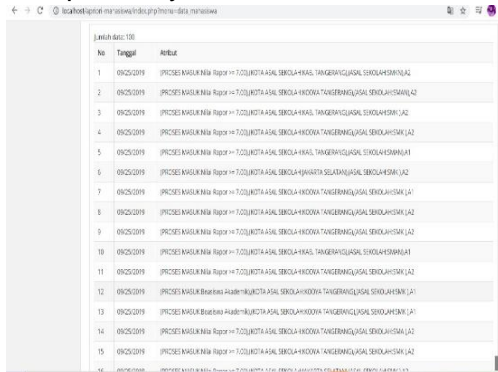
Gambar 6 Rancangan layar hasil

5. Tampilan Layar
 a) Tampilan Layar Login



Gambar 7 Tampilan Layar Login

b) Tampilan Layar Data Mahasiswa



Gambar 8 Tampilan Layar Data

c) Tampilan Layar Proses Apriori



Gambar 9 Tampilan Layar Proses Apriori

d) Tampilan Layar Hasil



Gambar 10 Tampilan Layar Hasil

KESIMPULAN

Setelah hasil diketahui dari pengujian dengan aplikasi data mining, maka akan dilakukan analisis dari hasil tersebut dengan cara membandingkan hasil pengujian 341 dan 100 data mahasiswa untuk mengetahui hasil yang didapat apakah berpengaruh terhadap jumlah data yang digunakan. Ada dua hasil yang akan dibandingkan yaitu hasil Frequent itemset (pola frekuensi tinggi) an hasil Association rule (aturan asosiasi).

100 data mahasiswa didapatkan 4 Itemset yang terdiri dari kombinasi 3 item, itemset dengan nilai tertinggi terdapat pada kombinasi (PROSES MASUK : Nilai Rapor >= 7.00),(Asal Sekolah:SMK),A2

a) hasil pengujian 341 data mahasiswa didapatkan 16 itemset dengan kombinasi 3 item, dan nilai tertinggi nya juga terdapat pada kombinasi (PROSES MASUK : Nilai Rapor >= 7.00),(Asal Sekolah:SMK),A2.

b) Sedangkan hasil pengujian 341 data mahasiswa dengan minimum support 50 didapatkan 2itemset dengan nilai tertinggi terdapat pada kombinasi (PROSES MASUK : Nilai Rapor >= 7.00),(Asal Sekolah:SMK),A2 juga

Maka dapat di ambil kesimpulan sebagai berikut:

Meskipun dari jumlah *itemset* yang didapatkan berbeda namun untuk hasil *itemset* dengan nilai tertinggi adalah sama yang terdapat pada kombinasi (PROSES MASUK : Nilai Rapor >= 7.00),(Asal Sekolah:SMK),A2.

Semakin banyak data yang diuji maka akan semakin banyak *itemset* yang terbentuk. Hal ini terbukti dari 100 data dihasilkan 4 *itemset* dan 341 data dihasilkan 16 *itemset*

Semakin tinggi nilai *minimum support* yang *diinput* maka semakin sedikit *itemset* yang terbentuk. Hal ini terbukti dari 341 data mahasiswa yang *diinput* dengan *minimum support* 20 dihasilkan 16 itemset dan 341 data *itemset*.

Oleh karena itu dapat diambil kesimpulan yang hampir sama dengan kesimpulan analisis Frequent itemset sebagai berikut:

Semakin banyak data yang diuji maka semakin banyak juga aturan yang terbentuk hal ini terbukti dari 100 data mahasiswa dihasilkan 4 aturan dan 341 data mahasiswa dihasilkan 12 aturan.

Dari jumlah hasil yang didapatkan berbeda, untuk hasil akurasi terbaik juga akan berbeda.

Semakin tinggi nilai minimum confidence diinput maka semakin sedikit hasil aturan yang didapatkan. hal ini terbukti dari 341 data mahasiswa yang diinput dengan minimum confidence 80% dihasilkan 12 aturan, dan 341 data mahasiswa yang diinput dengan minimum confidence 90% dihasilkan 2 aturan

Setelah analisis perbandingan hasil Frequent itemset maupun Association rules dilakukan, meskipun hasil pengujian dari 341 yang diinput nilai minimum support 20 dan minimum confidence 80% data lebih banyak. Berdasarkan akurasi, pola yang dijadikan sebagai hasil akhir dari penelitian ini adalah pola yang didapatkan dari pengujian 341 data mahasiswa yang diinput nilai minimum support nya 50 dan minimum confidence 90% Karena akurasi lebih optimal. Adapun pola yang didapatkan adalah sebagai berikut:

Jika Mahasiswa kota asal sekolah nya di Kodya Tangerang dan kategori kelulusannya A2 (Lama studi 4 tahun atau kurang dari 4 tahun dan IPK 2,76-3,50) maka Mahasiswa proses masuk nya dengan Nilai Rapor ≥ 7.00 dengan nilai support 19,35% , confidence 91,67%, dan lift ratio 1,06

Jika Mahasiswa asal sekolahnya dari SMK dan kategori kelulusannya A2 (Lama studi 4 tahun atau kurang dari 4 tahun dan IPK 2,76-3,50) maka Mahasiswa tersebut proses masuk nya dengan Nilai Rapor ≥ 7.00 dengan nilai support 16,13%, confidence 92,96% dan lift ratio 1,07.

REFERENSI

Jurnal/Prosiding/Disertasi/Tesis/Skripsi

[1] Buaton, R., Maulita, Y., & Kristiawan,

- A. (2018). Korelasi Faktor Penyebab Tindak Kekerasan dalam Rumah Tangga Menggunakan Data Mining Algoritma A Priori. *Jurnal Media Infotama*, 14(1).
<https://doi.org/10.37676/jmi.v14i1.468>
- [2] Indah Puji Astuti. (2019). Algoritma Apriori Untuk Menemukan Hubungan Antara Jurusan Sekolah Dengan Tingkat Kelulusan Mahasiswa. *Jurnal Teknik Informatika*, 12(1), 69–78.
- [3] Kusumo, H., Sedyono, E., & Marwata, M. (2019). Analisis Algoritma Apriori untuk Mendukung Strategi Promosi Perguruan Tinggi. *Walisongo Journal of Information Technology*, 1(1), 49.
<https://doi.org/10.21580/wjit.2019.1.1.4000>
- [4] Mulyani, H. (2018). Implementasi Algoritma Apriori Pada Transaksi Penjualan Barang (Studi Kasus: Toko Isna Kabupaten Bintan). *Jurnal Bangkit Indonesia*, 7(1), 15.
<https://doi.org/10.52771/bangkitindonesia.v7i1.82>
- [5] Nurjoko, & Kurniawan, H. (2016). Aplikasi Data Mining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Apriori di IBI Darmajaya, Bandar Lampung. *Jurnal TIM Darmajaya*, 02(01), 79–93.
- [6] Saefudin, S., & DN, S. (2019). Penerapan Data Mining Dengan Metode Algoritma Apriori Untuk Menentukan Pola Pembelian Ikan. *JSil (Jurnal Sistem Informasi)*, 6(2), 36.
<https://doi.org/10.30656/jsii.v6i2.1587>
- [7] Saputro, R. A. (2015). Penerapan Association Rule Dengan Algoritma Apriori Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa Teknik Informatika S1 Fakultas Ilmu.
- [8] Srikanti, E., Yansi, R. F., Norhavina, Permana, I., & Salisah, F. N. (2018). Penerapan Algoritma Apriori untuk Mencari Aturan Asosiasi pada Data Peminjaman Buku di Perpustakaan. *Jurnal Ilmiah Rekayasa Dan Manajemen Sistem Informasi*, 4(1), 77–80.