

ALGORITMA APRIORI UNTUK MENEMUKAN HUBUNGAN ANTARA JURUSAN SEKOLAH DENGAN TINGKAT KELULUSAN MAHASISWA

Indah Puji Astuti

Universitas Muhammadiyah Ponorogo
indahsan.0912@gmail.com

ABSTRACT

The accreditation of the university or study program, graduation rate is one of the influential factors. Educational period targeted in 4 years or 8 semesters of study period. But in reality there are still many students who pass beyond the study period. In this case the university or course of study can utilize students' self data to predict the student's graduation rate. One of them by using the concept of data mining. In this research the authors used an a priori algorithm to find the relationship between departments taken at high school level with the level of graduation students. The student's graduation rate is measured by length of study and GPA. The calculation is using 2 ways that is by manual calculation and by using Tanagra software. Based on the results of the analysis from regular class A students of 2012/2013 which amounted to 23 data it can be found rule, if the majors taken at high school level is SMK, then the possibility of the student will graduate on time with a period of 4 years or less. The GPA that the student will achieve between 3.1 - 3.5. The results of this analysis can be an alternative for study programs to make decisions.

Keywords: *Algoritma Apriori, Data Mining, Student Graduation, Tanagra*

ABSTRAK

Pada saat universitas atau program studi mengajukan akreditasi, tingkat kelulusan merupakan salah satu faktor yang berpengaruh. Masa tempuh pendidikan yang umum ditargetkan adalah 4 tahun atau 8 semester masa studi. Namun dalam kenyataannya masih banyak 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. Salah satunya dengan menggunakan konsep *data mining*. Pada penelitian ini penulis menggunakan algoritma apriori untuk menemukan hubungan antara jurusan yang diambil waktu sekolah tingkat SLTA dengan tingkat kelulusan mahasiswa. Tingkat kelulusan mahasiswa diukur berdasarkan lama studi dan IPK. Perhitungan dilakukan menggunakan 2 cara yaitu dengan perhitungan manual dan dengan menggunakan *software* Tanagra. Berdasarkan hasil analisis dari data mahasiswa kelas reguler A angkatan 2012/2013 yang berjumlah 23 data maka dapat ditemukan rule atau aturan bahwa, jika jurusan yang diambil waktu sekolah tingkat SLTA adalah SMK maka, kemungkinan mahasiswa tersebut akan lulus tepat waktu dengan masa studi 4 tahun atau kurang. IPK yang akan dicapai mahasiswa tersebut antara *rank* 3.1 – 3.5. Hasil analisis ini dapat menjadi alternatif bagi program studi untuk mengambil keputusan.

Kata Kunci: *Algoritma Apriori, Data Mining, Kelulusan Mahasiswa, Tanagra*

I. PENDAHULUAN

Dalam suatu universitas atau program studi tingkat kelulusan mahasiswa merupakan salah satu unsur yang diutamakan. Hal ini sangat berpengaruh pada saat universitas atau program studi tersebut mengajukan akreditasi. Masa tempuh pendidikan yang umum ditargetkan adalah 4 tahun atau 8 semester masa studi. Namun masih banyak mahasiswa yang nyatanya lulus melebihi masa studi yang ditargetkan. Dalam hal ini universitas atau program studi dapat memanfaatkan data diri mahasiswa untuk memprediksi tingkat kelulusannya. Salah satunya dengan menggunakan konsep *data mining*.

Salah satu contoh penggunaan *data mining* adalah dalam hal peningkatan fasilitas kartu kredit. Dari data pelanggan yang ada, maka akan dicari pola dari data tersebut untuk menemukan informasi pelanggan mana yang kira-kira berpotensi dan mana yang tidak. Dengan diketahuinya prediksi ini maka akan bermanfaat sekali bagi pihak perusahaan dalam menentukan pelanggan yang akan diberi izin untuk memiliki kartu kredit. Tentunya perusahaan akan berfikir dua kali untuk memberikan kartu kredit ke pelanggan yang tidak berpotensi [1]. Salah satu algoritma yang ada dalam konsep *data mining* adalah algoritma apriori. Algoritma *apriori* merupakan salah satu algoritma *data mining* yang menggunakan aturan asosiasi kombinasi *item* [2].

Fajri tahun 2016 melakukan penelitian yang mengimplementasikan algoritma *data mining* yaitu algoritma *apriori* untuk menentukan program studi yang diambil mahasiswa. Pemilihan program studi terkadang memang dirasa cukup membingungkan untuk calon mahasiswa baru. Pada penelitiannya Fajri menggunakan *item* jurusan SLTA, predikat lulus UN, dan nilai tes masuk untuk penentuan program studi yang akan diambil [3].

Pada tahun 2015 Anofrozen dan Harton telah merancang aplikasi *data mining* untuk menampilkan informasi tingkat kelulusan mahasiswa dengan algoritma *apriori*. Mereka memining data induk dan data kelulusan mahasiswa untuk menemukan informasi tentang tingkat kelulusan. Dimana kategori tingkat kelulusan diukur berdasarkan lama studi dan IPK [4].

[5] mengimplementasikan *data mining* untuk menemukan hubungan antara kota kelahiran mahasiswa dengan tingkat kelulusan

mahasiswa pada Fakultas Teknologi Informasi UNISBANK. Hasil penelitian digunakan sebagai bahan pertimbangan ketika universitas tersebut akan melakukan promosi [5].

Data mahasiswa yang terdiri dari data jurusan yang diambil waktu sekolah tingkat SLTA, data lama studi, data IPK dan data kelulusan (tepat waktu dan tidak tepat waktu), data-data tersebut sifatnya masih mentah belum diketahui keterhubungannya sehingga dibutuhkan sebuah metode untuk mengetahui hubungan antar data-data tersebut. Salah satunya dengan menggunakan konsep *data mining*. Pemilihan metode *data mining* harus disesuaikan dengan tipe data yang telah terkumpul, sehingga tujuan untuk menemukan hubungan antara data-data mahasiswa tersebut dapat tercapai [6].

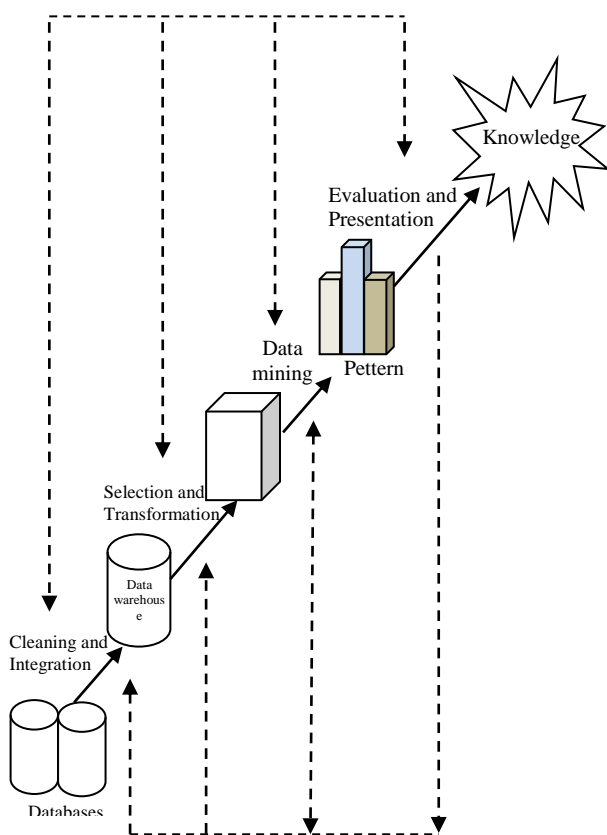
Pada penelitian-penelitian sebelumnya penulis melakukan penambangan data dengan algoritma apriori untuk menemukan keterkaitan antar atribut dari data mahasiswa dan pada penelitian ini penulis menggunakan algoritma apriori untuk menemukan hubungan antara jurusan yang diambil waktu Sekolah Lanjutan Tingkat Atas (SLTA) dengan tingkat kelulusan mahasiswa. Tingkat kelulusan mahasiswa diukur berdasarkan lama studi dan IPK.

II. TINJAUAN PUSTAKA

2.1 Data Mining

Data mining merupakan proses penambangan atau penggalian makna dari data yang jumlahnya begitu besar, dengan cara mengekstrak data tersebut untuk menemukan pola-pola tertentu dan menganalisisnya untuk memperoleh suatu pengetahuan atau informasi [7]. Suatu metode yang digunakan untuk mengekstrak pola dari suatu data atau biasa disebut dengan *Knowledge Discovery in Database* (KDD) terdiri dari langkah-langkah sebagai berikut [8]:

1. *Data cleaning*
2. *Data integration*
3. *Data selection*
4. *Data transformation*
5. *Data mining*
6. *Pattern evaluation*
7. *Knowledge presentation*



Gambar 1. Tahapan Knowledge Discovery in Database (KDD) [8]

Tabel 1. Penelitian sebelumnya

| No | Judul | Hasil |
|----|--|--|
| 1 | Implementasi Algoritma Apriori dalam Menentukan Program Studi yang diambil Mahasiswa oleh Ahmad Fikri Fajri (2016) [3] | Dengan menggunakan algoritma <i>apriori</i> telah didapat <i>rule</i> yang dapat digunakan mahasiswa dalam memilih program studi dengan nilai <i>confidence</i> paling tinggi adalah 75% |
| 2 | Rancang Bangun Aplikasi <i>Data mining</i> untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa dengan Algoritma Apriori oleh Harton Rohul Meisa Tambun dan | Dengan menggunakan algoritma <i>apriori</i> dapat menampilkan informasi berupa nilai <i>support</i> dan nilai <i>confidence</i> hubungan dari tingkat kelulusan |

Anofrizen [4]

dan data induk mahasiswa. Didapat nilai *confidence* dan *support* antara jalur UMJM dan SMPTN adalah sama yaitu 50% dan 18%.

3 Implementasi *Data mining* untuk Menemukan Hubungan antar Kota Kelahiran Mahasiswa dengan Tingkat Kelulusan Mahasiswa pada Fakultas Teknologi Informasi UNISBANK oleh Pipin Retnosari dan Arief Jananto [5]

Dengan menggunakan algoritma *apriori* dapat menampilkan informasi berupa nilai *confidence* sampai 75%. Hasil *rule* menunjukkan bahwa kota Semarang dan Kendal yang terpilih menjadi aturan asosiasinya. Hal ini dapat digunakan sebagai bahan pertimbangan ketika universitas tersebut akan melakukan promosi.

4 *Research and Application of Data mining in College Students' Employment Based on Association Rules* oleh Liu Lijuan dan Zhai Changliang [9]

Dengan menggunakan algoritma *apriori* dapat menampilkan informasi berupa nilai *confidence* sampai 75%. Pada penelitian ini hasil *rule* digunakan untuk memprediksi pekerjaan mahasiswa setelah mereka lulus dari perguruan tinggi tersebut.

Dari beberapa penelitian sebelumnya yang disajikan dalam Tabel 1, perbedaan yang mendasar dengan penelitian ini adalah terletak pada keterkaitan hubungan antara atributnya.

Pada penelitian ini atribut yang digunakan untuk menemukan hubungan adalah jurusan sekolah yang diambil siswa ketika masih duduk di bangku SLTA dengan tingkat kelulusan yang dilihat berdasarkan IPK dan lama studi.

2.2 Algoritma Apriori

Algoritma *apriori* adalah jenis aturan asosiasi pada *data mining*. Analisis asosiasi ini merupakan teknik *data mining* untuk menemukan aturan asosiatif antara suatu kombinasi *item*/atribut [10].

Contoh aturan asosiatif dari analisa pembelian di suatu mini market adalah dapat diketahui berapa besar kemungkinan seorang pelanggan membeli minuman dingin bersamaan dengan membeli makanan ringan. Dengan pengetahuan tersebut pemilik mini market dapat mengatur penempatan rak kedua produk tersebut secara berdekatan.

Metodologi dasar analisis asosiasi terbagi menjadi dua tahap [1]:

a. Analisa pola frekuensi tinggi

Tahap ini mencari kombinasi *item* yang memenuhi syarat minimum dari nilai *support* dalam *database*. Nilai *support* sebuah *item* diperoleh dengan rumus berikut:

$$\text{Support (A)} = \frac{\text{Jumlah Transaksi mengandung A}}{\text{total transaksi}} \times 100\% \quad \dots(1)$$

Sedangkan nilai *support* dari 2 *item* diperoleh dari rumus berikut:

$$\text{Support (A n B)} = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{total transaksi}} \times 100\% \quad \dots(2)$$

b. Pembentukan aturan asosiatif

Setelah semua pola frekuensi tinggi ditemukan, barulah dicari aturan asosiatif yang memenuhi syarat minimum untuk *confidence* dengan menghitung *confidence* aturan *asosiatif* A_B

Nilai *confidence* dari aturan A_B diperoleh dari rumus berikut:

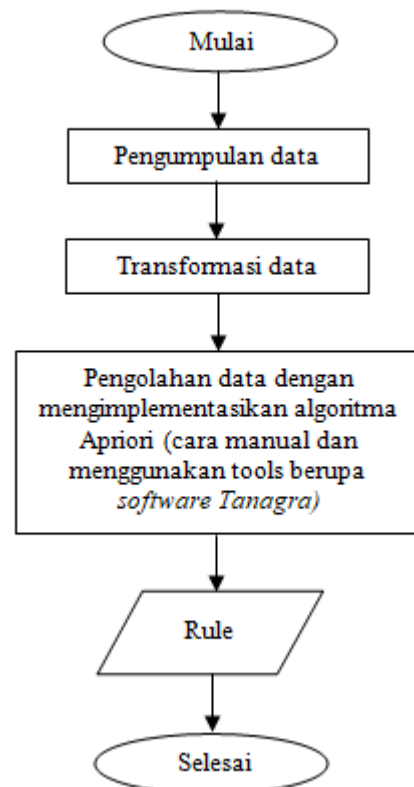
$$\text{Confidence} = P(B | A) = \frac{\text{Jumlah Transaksi mengandung A dan B}}{\text{Jumlah transaksi yang mengandung A}} \times 100\% \quad \dots(3)$$

III. METODE PENELITIAN

Dalam melakukan analisis penulis menggunakan dua cara yaitu dengan cara manual dengan membuat beberapa tabel untuk menemukan *rule* sesuai dengan aturan yang ada pada algoritma apriori. Cara yang kedua yaitu dengan menggunakan *software* Tanagra. Data yang digunakan dalam penelitian ini adalah data diri mahasiswa Universitas Muhammadiyah Ponorogo Jurusan Teknik Informatika yang ada di Sistem Informasi Akademik (SIMTIK). Data yang digunakan adalah data mahasiswa kelas reguler A angkatan 2012/2013.

Hasilnya adalah kedua cara telah menampilkan *rule* yang sama dimana tahapan pada cara manual dapat dilihat pada Tabel 2 sampai Tabel 14 dan cara kedua dapat dilihat pada Gambar 3 dan Gambar 4.

Flowchart tahapan penelitian dapat dilihat pada Gambar 2.



Gambar 2. Metodologi penelitian

IV. HASIL DAN PEMBAHASAN

4.1 Kebutuhan Data

Data diambil dari data diri mahasiswa Universitas Muhammadiyah Ponorogo Jurusan

Teknik Informatika yang ada di SIMTIK. Data yang digunakan adalah data mahasiswa kelas reguler A angkatan 2012/2013. Tidak semua atribut dalam data diri mahasiswa digunakan dalam prediksi menggunakan algoritma apriori ini, tetapi yang digunakan hanya IPK dan jurusan sekolah yang dahulu diambil waktu sekolah. Penentuan kelulusan tepat waktu dan tidak tepat waktu berdasarkan IPK dibagi menjadi 2 kategori yaitu, lama studi 4 tahun atau kurang (berarti lulus tepat waktu) dan lama studi lebih dari 4 tahun (berarti lulus tidak tepat waktu). Sedangkan jurusan sekolah terdiri dari 3 kategori yaitu IPA, IPS, dan SMK.

4.2 Transformasi Data

Pengkategorian data kelulusan berdasarkan lama studi yaitu, lulus tepat waktu, bila lama studi 4 tahun atau kurang dari 4 tahun dan lulus tepat waktu, bila lama studi lebih dari 4 tahun. Dari dua pengkategorian tersebut dapat dibuat kategori berdasarkan kombinasi keduanya, seperti yang dapat dilihat pada Tabel 2.

Tabel 2. Transformasi data

| Kategori | Keterangan |
|----------|---|
| X1 | Lama studi 4 tahun / kurang dan IPK 3.51 – 4.00 |
| X2 | Lama studi 4 tahun / kurang dan IPK 3.1 – 3.5 |
| X3 | Lama studi 4 tahun / kurang dan IPK 2.51 – 3.00 |
| Y1 | Lama studi lebih dari 4 tahun dan IPK 3.51 – 4.00 |
| Y2 | Lama studi lebih dari 4 tahun dan IPK 3.1 – 3.5 |
| Y3 | Lama studi lebih dari 4 tahun dan IPK 2.51 – 3.00 |

Adapun tabel atribut dapat dilihat pada Tabel 3.

Tabel 3. Atribut / item yang digunakan

| Kategori |
|----------|
| X1 |
| X2 |
| X3 |
| Y1 |
| Y2 |
| Y3 |
| IPA |
| IPS |
| SMK |

4.3 Pengolahan Data

Pengolahan data dilakukan dengan perhitungan menggunakan algoritma apriori. Data mahasiswa yang digunakan berjumlah 23 data. Data mentah dapat dilihat pada Tabel 4.

Tabel 4. Data mentah

| No | NIM | Jurusan | IPK | Kategori kelulusan |
|----|----------|---------|-----|--------------------|
| 1 | 12531376 | IPS | 2.7 | Y3 |
| 2 | 12531407 | IPA | 2.6 | Y3 |
| 3 | 12531641 | IPA | 3.2 | X2 |
| 4 | 12531391 | IPA | 3.5 | X2 |
| 5 | 12531553 | IPA | 2.9 | Y3 |
| 6 | 12531612 | SMK | 3.3 | X2 |
| 7 | 12531579 | IPA | 2.8 | Y3 |
| 8 | 12531468 | SMK | 3.3 | X2 |
| 9 | 12531426 | SMK | 3.1 | X2 |
| 10 | 12531406 | SMK | 3.4 | X2 |
| 11 | 12531457 | IPA | 3.5 | X2 |
| 12 | 12531386 | SMK | 3.5 | X2 |
| 13 | 12531455 | IPA | 3.5 | X2 |
| 14 | 12531383 | IPA | 3 | Y3 |
| 15 | 12531393 | IPA | 3.8 | X1 |
| 16 | 12531421 | IPS | 3.5 | X2 |
| 17 | 12531374 | IPA | 3.4 | X2 |
| 18 | 12531458 | SMK | 2.9 | Y3 |
| 19 | 12531446 | IPA | 3.2 | X2 |
| 20 | 12531394 | SMK | 3.1 | Y2 |
| 21 | 12531389 | SMK | 3.4 | X2 |
| 22 | 12531424 | IPS | 2.9 | Y3 |
| 23 | 12531554 | IPS | 2.9 | Y3 |

Data mentah ditransformasikan dalam bentuk tabel seperti pada Tabel 5.

Tabel 5. Data mahasiswa (K = 1) terdiri dari 1 unsur

| Data ke | X ₁ | X ₂ | X ₃ | Y ₁ | Y ₂ | Y ₃ | IPA | IPS | SMK |
|---------|----------------|----------------|----------------|----------------|----------------|----------------|-----|-----|-----|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 7 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 9 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 10 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 11 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

| | | | | | | | | | |
|----|---|---|---|---|---|---|---|---|---|
| 12 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 15 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 16 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 17 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 19 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 21 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 22 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |

Tabel 6. Jumlah banyaknya mahasiswa berdasarkan kategori (K=1) terdiri dari 1 unsur

| Data ke | X 1 | X 2 | X 3 | Y 1 | Y 2 | Y 3 | IPA | IPS | SMK |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 7 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 8 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 9 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 10 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 11 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 12 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 13 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 15 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 16 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 17 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 19 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 21 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 22 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| Σ | 1 | 13 | 0 | 0 | 1 | 8 | 11 | 4 | 8 |

Threshold (Φ) = 3, maka dapat ditentukan frekuensi itemset dari tabel di atas, yaitu:

$$F1 = \{ \{X2\}, \{Y3\}, \{IPA\}, \{IPS\}, \{SMK\} \}$$

Untuk k = 2 (2 unsur). Himpunan yang mungkin terbentuk adalah: {X2,IPA}, {X2,IPS}, {X2,SMK}, {Y3,IPA}, {Y3,IPS}, {Y3,SMK}. Maka tabel-tabel untuk calon 2 item set adalah seperti di bawah ini.

Tabel 7. Tabel 2 item set untuk X2 dan IPA

| Data ke | X2 | IPA | F |
|---------|----|-----|---|
| 1 | 0 | 0 | S |
| 2 | 0 | 1 | S |
| 3 | 1 | 1 | P |
| 4 | 1 | 1 | P |
| 5 | 0 | 1 | S |
| 6 | 1 | 0 | S |
| 7 | 0 | 1 | S |
| 8 | 1 | 0 | S |
| 9 | 1 | 0 | S |
| 10 | 1 | 0 | S |
| 11 | 1 | 1 | P |
| 12 | 1 | 0 | S |
| 13 | 1 | 1 | P |
| 14 | 0 | 1 | S |
| 15 | 0 | 1 | S |
| 16 | 1 | 0 | S |
| 17 | 1 | 1 | P |
| 18 | 0 | 0 | S |
| 19 | 1 | 1 | P |
| 20 | 0 | 0 | S |
| 21 | 1 | 0 | S |
| 22 | 0 | 0 | S |
| 23 | 0 | 0 | S |
| Σ | | | 6 |

Tabel 8. Tabel 2 item set untuk X2 dan IPS

| Data ke | X2 | IPS | F |
|---------|----|-----|---|
| 1 | 0 | 1 | S |
| 2 | 0 | 0 | S |
| 3 | 1 | 0 | S |
| 4 | 1 | 0 | S |
| 5 | 0 | 0 | S |
| 6 | 1 | 0 | S |
| 7 | 0 | 0 | S |
| 8 | 1 | 0 | S |
| 9 | 1 | 0 | S |
| 10 | 1 | 0 | S |
| 11 | 1 | 0 | S |
| 12 | 1 | 0 | S |
| 13 | 1 | 0 | S |
| 14 | 0 | 0 | S |
| 15 | 0 | 0 | S |
| 16 | 1 | 1 | P |
| 17 | 1 | 0 | S |
| 18 | 0 | 0 | S |
| 19 | 1 | 0 | S |
| 20 | 0 | 0 | S |
| 21 | 1 | 0 | S |
| 22 | 0 | 1 | S |
| 23 | 0 | 1 | S |
| Σ | | | 1 |

| | |
|----------|---|
| Σ | 4 |
|----------|---|

Tabel 9. Tabel 2 item set untuk X2 dan SMK

| Data ke | X2 | SMK | F |
|---------|----|----------|---|
| 1 | 0 | 0 | S |
| 2 | 0 | 0 | S |
| 3 | 1 | 0 | S |
| 4 | 1 | 0 | S |
| 5 | 0 | 0 | S |
| 6 | 1 | 1 | P |
| 7 | 0 | 0 | S |
| 8 | 1 | 1 | P |
| 9 | 1 | 1 | P |
| 10 | 1 | 1 | P |
| 11 | 1 | 0 | S |
| 12 | 1 | 1 | P |
| 13 | 1 | 0 | S |
| 14 | 0 | 0 | S |
| 15 | 0 | 0 | S |
| 16 | 1 | 0 | S |
| 17 | 1 | 0 | S |
| 18 | 0 | 1 | S |
| 19 | 1 | 0 | S |
| 20 | 0 | 1 | S |
| 21 | 1 | 1 | P |
| 22 | 0 | 0 | S |
| 23 | 0 | 0 | S |
| | | Σ | 6 |

Tabel 11. Tabel 2 item set untuk Y3 dan IPS

| Data ke | Y3 | IPS | F |
|---------|----|----------|---|
| 1 | 1 | 1 | P |
| 2 | 1 | 0 | S |
| 3 | 0 | 0 | S |
| 4 | 0 | 0 | S |
| 5 | 1 | 0 | S |
| 6 | 0 | 0 | S |
| 7 | 1 | 0 | S |
| 8 | 0 | 0 | S |
| 9 | 0 | 0 | S |
| 10 | 0 | 0 | S |
| 11 | 0 | 0 | S |
| 12 | 0 | 0 | S |
| 13 | 0 | 0 | S |
| 14 | 1 | 0 | S |
| 15 | 0 | 0 | S |
| 16 | 0 | 1 | S |
| 17 | 0 | 0 | S |
| 18 | 1 | 0 | S |
| 19 | 0 | 0 | S |
| 20 | 0 | 0 | S |
| 21 | 0 | 0 | S |
| 22 | 1 | 1 | P |
| 23 | 1 | 1 | P |
| | | Σ | 3 |

Tabel 10. Tabel 2 item set untuk Y3 dan IPA

| Data ke | Y3 | IPA | F |
|---------|----|-----|---|
| 1 | 1 | 0 | S |
| 2 | 1 | 1 | P |
| 3 | 0 | 1 | S |
| 4 | 0 | 1 | S |
| 5 | 1 | 1 | P |
| 6 | 0 | 0 | S |
| 7 | 1 | 1 | P |
| 8 | 0 | 0 | S |
| 9 | 0 | 0 | S |
| 10 | 0 | 0 | S |
| 11 | 0 | 1 | S |
| 12 | 0 | 0 | S |
| 13 | 0 | 1 | S |
| 14 | 1 | 1 | P |
| 15 | 0 | 1 | S |
| 16 | 0 | 0 | S |
| 17 | 0 | 1 | S |
| 18 | 1 | 0 | S |
| 19 | 0 | 1 | S |
| 20 | 0 | 0 | S |
| 21 | 0 | 0 | S |
| 22 | 1 | 0 | S |
| 23 | 1 | 0 | S |

Tabel 12. Tabel 2 item set untuk Y3 dan SMK

| Data ke | Y3 | SMK | F |
|---------|----|-----|---|
| 1 | 1 | 0 | S |
| 2 | 1 | 0 | S |
| 3 | 0 | 0 | S |
| 4 | 0 | 0 | S |
| 5 | 1 | 0 | S |
| 6 | 0 | 1 | S |
| 7 | 1 | 0 | S |
| 8 | 0 | 1 | S |
| 9 | 0 | 1 | S |
| 10 | 0 | 1 | S |
| 11 | 0 | 0 | S |
| 12 | 0 | 1 | S |
| 13 | 0 | 0 | S |
| 14 | 1 | 0 | S |
| 15 | 0 | 0 | S |
| 16 | 0 | 0 | S |
| 17 | 0 | 0 | S |
| 18 | 1 | 1 | P |
| 19 | 0 | 0 | S |
| 20 | 0 | 1 | S |
| 21 | 0 | 1 | S |
| 22 | 1 | 0 | S |

| | | | |
|----|---|----------|---|
| 23 | 1 | 0 | S |
| | | Σ | 1 |

Keterangan:

P artinya kategori yang terjadi bersamaan

S artinya tidak ada kategori yang terjadi bersamaan

Jumlah frekuensi *item set* harus lebih besar atau sama dengan jumlah Frekuensi *item set* ($S \geq \Phi$). Dari tabel di atas, maka didapat:

$F2 = \{\{X2, IPA\}, \{X2, SMK\}, \{Y3, IPA\}, \{Y3, IPS\}\}$

4.4 Proses Pembentukan Rule

Rule yang dipakai adalah jika x maka y, dimana x adalah *antecedent* dan y adalah *consequent*. Berdasarkan rule tersebut, maka dibutuhkan 2 buah *item* yang mana salah satunya sebagai *antecedent* dan satunya lagi sebagai *consequent*.

x sebagai *antecedent* dan y sebagai *consequent* dari Fk. Dari F2 maka dapat disusun aturan sebagai berikut:

- Untuk {X2,IPA}:
 - Jika (x) = X2, Jika (y) = IPA,
 - jika X2 maka IPA
 - Jika (x) = IPA, Jika (y) = X2,
 - Jika IPA maka X2
- Untuk {X2,SMK}:
 - Jika (x) = X2, Jika (y) = SMK,
 - jika X2 maka SMK
 - Jika (x) = SMK, Jika (y) = X2,
 - Jika SMK maka X2
- Untuk {Y3,IPA}:
 - Jika (x) = Y3, Jika (y) = IPA,
 - jika Y3 maka IPA
 - Jika (x) = IPA, Jika (y) = Y3,
 - Jika IPA maka Y3
- Untuk {Y3,IPS}:
 - Jika (x) = Y3, Jika (y) = IPS,
 - jika Y3 maka IPS
 - Jika (y) = IPS, Jika (y) = Y3,
 - Jika IPS maka Y3

Dari aturan di atas didapat 8 rule sebagai berikut:

- Jika lama studi 4 tahun / kurang dan IPK 3.1 – 3.5 maka jurusan yang dulu diambil adalah IPA
- Jika jurusan yang dulu diambil adalah IPA maka lama studinya adalah 4 tahun / kurang dan IPK 3.1 – 3.5

- Jika lama studi 4 tahun / kurang dan IPK 3.1 – 3.5 maka jurusan yang dulu diambil adalah SMK
- Jika jurusan yang dulu diambil adalah SMK maka lama studinya adalah 4 tahun/kurang dan IPK 3.1 – 3.5
- Jika lama studi lebih dari 4 tahun dan IPK 2.51 – 3.00 maka jurusan yang dulu diambil adalah IPA
- Jika jurusan yang dulu diambil adalah IPA maka lama studinya adalah lebih dari 4 tahun dan IPK 2.51 – 3.00
- Jika lama studi lebih dari 4 tahun dan IPK 2.51 – 3.00 maka jurusan yang dulu diambil adalah IPS
- Jika jurusan yang dulu diambil adalah IPS maka lama studinya adalah lebih dari 4 tahun dan IPK 2.51 – 3.00

Tabel 13 adalah tabel untuk mencari nilai *support* dan *confidence*.

Tabel 13. Tabel kandidat asosiasi rule untuk 1 antecedent

| If antecedent then consequent | Support | Confidence |
|-------------------------------------|-----------------------------------|----------------------------------|
| Jika X2 maka IPA | $(6/23) \times 100\%$ = 26.087 | $(6/13) \times 100\%$ = 46.15 |
| Jika IPA maka X2 | $(6/23) \times 100\%$ = 26.087 | $(6/11) \times 100\%$ = 54.54 |
| Jika X2 maka SMK | $(6/23) \times 100\%$ = 26.087 | $(6/13) \times 100\%$ = 46.15 |
| Jika SMK maka X2 | $(6/23) \times 100\%$ = 26.087 | $(6/8) \times 100\%$ = 75 |
| Jika Y3 maka IPA | $(4/23) \times 100\%$ = 17.39 | $(4/8) \times 100\%$ = 50 |
| Jika IPA maka Y3 | $(4/23) \times 100\%$ = 17.39 | $(4/11) \times 100\%$ = 36.36 |
| Jika Y3 maka IPS | $(3/23) \times 100\%$ = 13.04 | $(3/8) \times 100\%$ = 37.5 |
| Jika IPS maka Y3 | $(3/23) \times 100\%$ = 13.04 | $(3/4) \times 100\%$ = 75 |

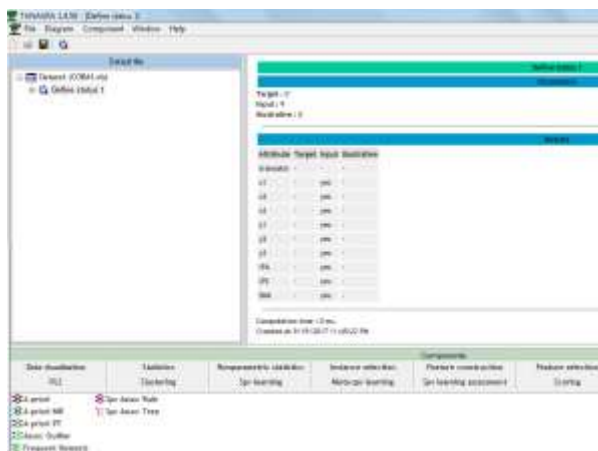
Setelah di dapat *support* dan *confidence* untuk masing-masing kandidat, dilakukan perkalian antara *support* dan *confidence*, dimana *confidence*-nya diambil 70% ke atas, sehingga di dapat tabel sebagai berikut:

Tabel 14. Rule prediksi ketepatan kelulusan

| If antecedent | Support | Confidence | Support x Confidence |
|------------------|------------------------|-------------------|----------------------|
| Jika SMK maka X2 | (6/23) x 100% = 26.087 | (6/8) x 100% = 75 | 0.1956 |
| Jika IPS maka Y3 | (3/23) x 100% = 13.04 | (3/4) x 100% = 75 | 0.0978 |

Setelah didapat hasil perkalian antara *support* dan *confidence*, maka hasil perkaliannya paling besar dari perkalian perkalian tersebut merupakan *rule* yang dipakai pada saat menentukan prediksi kelulusan tepat waktu berdasar IPK dan jurusan asal sekolah waktu SLTA.

Rule yang digunakan adalah → Jika jurusan yang dahulu diambil adalah SMK maka lama studinya adalah 4 tahun/kurang dan IPK 3.1 – 3.5 (lulus tepat waktu). Penghitungan dengan menggunakan Tanagra dapat dilihat pada Gambar 2 dan Gambar 3.



Gambar 3. Item yang digunakan



Gambar 4. Rule yang dihasilkan

V. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan di atas, maka penulis dapat mengambil kesimpulan dari data mahasiswa kelas reguler A angkatan 2012/2013 yang berjumlah 23 data maka dapat ditemukan *rule* atau aturan bahwa, jika jurusan yang diambil waktu sekolah adalah SMK maka, kemungkinan mahasiswa tersebut akan lulus tepat waktu dengan masa studi 4 tahun atau kurang dari 4 tahun. IPK yang akan dicapai mahasiswa tersebut antara *rank* 3.1–3.5. Algoritma *apriori* menemukan *rule* prediksi ketepatan kelulusan dengan nilai *support* sebesar 26. 087 dan nilai *confidence* sebesar 75%. Dibandingkan dengan hasil penelitian sebelumnya, pada penelitian ini telah memiliki nilai *confidence* yang sama yaitu sebesar 75%.

5.2 Saran

Beberapa saran untuk penelitian lebih lanjut adalah:

1. Menggunakan algoritma lain untuk menganalisa data mahasiswa yang nantinya digunakan untuk bahan program studi mengambil keputusan.
2. Menggunakan atribut lainnya untuk menemukan hubungan antar atribut.
3. Membandingkan beberapa algoritma sehingga dapat dipilih algoritma mana yang paling optimal untuk diimplementasikan dalam studi kasus seperti ini.

DAFTAR PUSTAKA

[1] Kusriani, Luthfi E .T. 2009. Algoritma Data mining. Penerbit Andi. Yogyakarta.

[2] Listiani, Dewi. et. al. 2016. Penerapan Metode Asosiasi Menggunakan algoritma Apriori pada Aplikasi Analisa Pola Belanja Konsumen. Jurnal Teknik Informatika. Vol. 9 (2): 120 - 127

[3] Fajri Ahmad Fikri. 2016. Implementasi Algoritma Apriori dalam Menentukan Program Studi yang Diambil Mahasiswa. Jurnal Ipteks Terapan. Research of Applied Science and Education. Vol. 10 (2): 81 – 85

- [4] Anofrizen, Harton, R. M. T. 2015. Rancang Bangun Aplikasi *Data mining* untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa dengan Algoritma Apriori. *Jurnal Rekayasa dan Manajemen Sistem Informasi*. Vol. 1(1): 31 – 38
- [5] Jananto, A., Retnosari, P. 2013. Implementasi *Data mining* untuk Menemukan Hubungan Antara Kota Kelahiran Mahasiswa Dengan Tingkat Kelulusan Mahasiswa Pada Fakultas Teknologi Informasi UNISBANK. *Dinamika informatika*. Vol. 5 (2): 112 – 121
- [6] Grand. 2018. Penerapan Algoritma Apriori Untuk Menentukan Hubungan Data Murid Dengan Nilai Sekolah. *IKRAITH-INFORMATIKA*. 2(1): 7 – 12.
- [7] Dennis Aprilla C, *et. al.* Belajar *Data mining* dengan Rapidminer.
- [8] Jiawei Han, Micheline Kamber, Jian Pei. *Data mining: Concepts And Techniques*. 3rd Ed
- [9] Liu Lijuan dan Zhai Changliang. 2017. Research and Application of *Data mining* in College Students' Employment Based on Association Rules. *Computer and Information Science*. Published by Canadian Center of Science and Education .Vol. 10 (3): 54 - 59
- [10] Nofriansyah, Dicky. 2014. Konsep *Data mining* Vs Sistem Pendukung Keputusan. Deepublish Publisher. Yogyakarta