**PENERAPAN ALGORITMA C4.5 UNTUK KLASIFIKASI PREDIKAT KEBERHASILAN MAHASISWA (STUDI KASUS FAKULTAS TEKNIK**

**UNIVERSITAS GRAHA NUSANTARA)**

**Oleh:**

**Alwendi**

*Dosen Fakultas Teknik Universitas Graha Nusantara Padangsidimpuan*

*email:* alwendi60@gmail.com

**Abstrak**

***Predikat keberhasilan Mahasiswa di perguruan tinggi merupakan hal utama yang menjadi acuan bagi seorang mahasiswa selama menjalani study. Data mahasiswa yang terus meningkat dari tahun ke tahun adalah bukti untuk mengembangkan Data Mining. Banyaknya jumlah Mahasiswa akan diproses Untuk mengetahui berdasarkan kriteria apa saja Mahasiswa layak mendapatkan predikat keberhasilannya dengan beberapa atribut seperti, jenis kelamin, kehadiran, sesi perkuliahan, rerata NEM dan Asal sekolah. Dalam hal ini, peneliti melakukan analisis di Amik Tunas Bangsa Jl. Sudirman no. 1,2,3 Pematangsiantar. Analisis yang digunakan adalah menggunakan Data Mining dengan metode C4.5 dan diproses menggunakan software Rapidminer untuk membuat pohon keputusan. Metode ini diharapkan dapat mempermudah pihak instansi mengklasifikasi dan memprediksi predikat keberhasilan Mahasiswa sehingga data-data yang menumpuk bisa bermanfaat untuk keperluan data mining dalam mengambil keputusan yang baik. Analisis ini diharapkan sebagai salah satu motivasi terhadap mahasiswa untuk meningkatkan IPK.***

***Kata kunci: Keberhasilan mahasiswa; metode C4.5; data mining; rapidminer; decision tree***

**BAB I. PENDAHULUAN**

Predikat keberhasilan adalah kemampuan aktual seorang mahasiswa yang berwujud penguasaan ilmu pengetahuan, sikap, dan keterampilan untuk mencapai tujuan akhir dari proses belajar yaitu Indeks Prestasi. Indeks Prestasi dihitung, baik pada setiap akhir semester dengan hasil yang disebut IP semester, maupun pada akhir program pendidikan lengkap satu jenjang, dengan hasil yang disebut IP kumulatif. Indeks Prestasi Mahasiswa merupakan data yang sangat penting di Amik Tunas Bangsa Pematangsiantar sehubung dengan terus meningkatnya jumlah Mahasiswa dari tahun ke tahun maka perlu dilakukan Data Mining dengan teknik klasifikasi bertujuan membantu suatu lembaga yang memiliki data melimpah untuk menyajikan informasi kedalam bentuk pengambilan keputusan.

Banyak teknik klasifikasi di Data Mining salah satunya adalah Algoritma C4.5 yang menyajikan klasifikasi data kedalam bentuk pohon keputusan. Kelebihan utama Algoritma C4.5 dapat membuat pohon keputusan (decision tree) yang efisien menangani atribut tipe diskrit dan tipe diskrit- numerik, mudah untuk diinterprestasikan dan memiliki tingkat akurasi yang dapat diterima (Han 2001). Kelemahan Algoritma C4.5 salah satunya terdapat di skalabilitas yaitu data training hanya dapat digunakan dan disimpan secara keseluruhan pada waktu yang bersamaan di memori (Veronica, 2007)

Amik Tunas Bangsa Pematang Siantar adalah lembaga pendidikan yangmemiliki banyak mahasiswa. Dengan Data Mining memudahkan lembaga pendidikan mengidentifikasi berdasarkan faktor-faktor apa sajakah yang mempengaruhi mahasiswa mendapatkan Indeks Prestasi Kumulatif baik di kelas. Peneliti menggunakan Algoritma C4.5 dalam pengklasifikasian predikat keberhasilan mahasiswa berdasarkan Atribut jenis kelamin, NEM, sesi perkuliahan, asal sekolah, kehadiran, dan IPK. Data Training yang diambil peneliti adalah data alumni mahasiswa program studi Manajemen Informatika dan Komputerisasi Akuntansi 2010-2012. Dengan demikian proses pengolahan data-data akademik menggunakan teknik klasifikasi akan membantu lembaga pendidikan dalam upaya memotivasi mahasiswa untuk meningkatkan IPK, serta bermanfaat meningkatkan akreditasi program studi di Amik Tunas Bangsa.

Berdasarkan latar belakang tersebut diterapkan Algoritma C4.5 yang berguna untuk mengklasifikasi dan memprediksi Predikat Keberhasilan Mahasiswa di Amik Tunas Bangsa Pematangsiantar.

**BAB II. BAHAN DAN METODE**

Data mining merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang dapat digeneralisasi untuk masa yang akan datang, bermanfaat dan dapat dimengerti dalam suatu database yang sangat besar (massive databases). Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu yang akan datang. (Hermawati, 2013).

Istilah data mining dan knowledge discovery in databases (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah data mining. Proses KDD (Knowlegge discovery in databases) secara garis besar dapat dijelaskan sebagai berikut (Kusrini, 2009)



Gambar 1. Tahapan knowledge discovery

in databases

Pengolahan Data Mining terdiri dari beberapa metode pengolahan, yaitu (Larose, 2006):

1. Association (Asosiasi) merupakan teknik dalam data mining yang mempelajari hubungan antar data. Contoh penggunaannya seperti untuk menganalisis perilaku mahasiswa yang datang terlambat. Contohnya jika mahasiswa memiliki jadwal dengan dosen A dan B, maka mahasiswa akan datang terlambat.
2. Predictive modelling yang merupakan pengolahan data mining dengan melakukan prediksi/peramalan. Tujuan metode ini untuk membangun model prediksi suatu nilai yang mempunyai ciri-ciri tertentu.
3. Classification merupakan teknik mengklasifikasikan data. Perbedaannya dengan metode clustering terletak pada data, dimana pada clustering variabel dependen tidak ada, sedangkan pada classification diharuskan ada variabel dependen.
4. Clustering merupakan pengelompokan terhadap record, yang berfungsi memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan.

Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami, juga dapat diekspresikan dalam bentuk bahasa basis data seperti Structured Query Language untuk mencari record pada kategori tertentu. Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk masing-masing cabang sampai semua kasus pada cabang memiliki kelas yang sama.
5. Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitung gain digunakan rumus seperti yang tertera berikut (Craw, 2005):

Gain(S,A) = Entrropy(S) – \*Entropy(Si)

Keterangan :

S : Himpunan Kasus A : Atribut

N : Jumlah partisi atribut A

|Si| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

Sedangkan perhitungan nilai Entropy dapat dilihat pada rumus 2 berikut (Craw, 2005):

*Entropy(A)= *

Keterangan:

S : Himpunan Kasus A : Fitur

N : Jumlah partisi S

Pi : Proporsi dari Si terhadap S

**BAB III. HASIL DAN PEMBAHASAN**

1. Algoritma Apriori

Algoritma apriori adalah satu algoritma dasar yang diusulkan oleh Agrawal dan Srikan pada tahun 1994 untuk menemukan frequent itemsets pada aturan asosiasi Boolean. Ide utama pada algoritma apriori adalah : pertama, mencari frequent itemset (himpunan item-item yang memenuhi minimum support.) dari basis data transaksi, kedua – menghilangkan itemset degaan frekuensi yang rendah berdasarkan level minimum support yang telah ditentukan sebelumnya. Selanjutnya membangun aturan asosiasi dari itemset yang memenuhi nilai minimum confidence dalam basis data [4].

Algoritma apriori dibagi menjadi beberapa tahapan yang disebut iterasi. Pada iterasi pertama dihasilkan pola frekuensi tinggi dengan panjang 1-itemset. Pada iterasi pertama pula, nilai support dari setiap item dihitung dengan melihat database yang ada. Setelah didapat nilai support, item yang memiliki support di atas minimum support (minsup) dipilih sebagai acuan untuk iterasi selanjutnya. Iterasi kedua menghasilkan 2-itemset yang tiap setnya memiliki dua item. Tiap kandidat 2-itemset dihitung nilai supportnya dengan melihat database sebelumnya. Support disini dimaksudkan untuk jumlah transaksi dalam database yang mengandung kedua item dalam kandidat 2-itemset. Sama seperti 1-itemset, setelah didapat nilai supportnya, maka kandidat 2-itemset yang memenuhi minsup dapat ditetapkan sebagai 2-itemset yang merupakan pola frekuensi tinggi.

Data yang diperoleh dari sitem informasi mahasiswa F-KIP Universitas Graha Nusantara berupa data sekunder berupa nama, jalur masuk Universitas Sam Ratulangi, IPK, lama studi, jenis kelamin, dan jurusan serta wawancara (data primer) berupa asal sekolah diberikan simbol sebagai berikut :

Tabel 1. Analisa data berdasarkan asal sekolah

|  |  |  |  |
| --- | --- | --- | --- |
| **Sim****bol** | **Kete****rangan** | **Sim****bol** | **Keterangan** |
| IPK | IPK | K | Jalur Masuk Kemitraan |
| A |  IPK > 3,5 | J.K | Jenis Kelamin |
| B | 3 ≤ IPK ≤ 3,5 | L | Jenis kelamin laki-laki |
| C | 2,5 ≤ IPK < 3 | P | Jenis kelaminprempuan |
| D |  IPK < 2,5 | Jurusan | Jurusan |
| LL |  Lama Lulus | Q | Jurusan biologi |
| E | Lama lulus ≤ 4 tahun | R | Jurusan Matematika |
| F | 4 tahun < lama lulus ≤ 5 tahun | S | Jurusan Kimia |
| G | Lama lulus > 5 tahun | T | Jurusan Fisika |
| JM | Jalur masuk | U | Jurusan Farmasi |
| H | Jalur masuk SPMB | Sekolah | Asal Sekolah |
| I | Jalur masuk Sumikola | X | Asal sekolahKabupaten |
| J | Jalur masuk T2 | Y | Asal sekolahKecamatan |

 Dengan menggunakan softwareOrange data tersebut diolah dan menghasilkan kombinasi itemset sampai kombinasi6-itemset. Berikut merupakan beberapa kombinasi itemset seperti pada tabel 2 dan tabel 3.

Tabel 2. Beberapa Kombinasi 2- Itemset

|  |  |  |
| --- | --- | --- |
| **2-Itemset** | **Support** | **Confidence** |
| J.K=P -> LL=F | 25,2% | 36,9% |
| J.K=L -> LL=G | 13,8% | 43,5% |
| SEKOLAH=X -> LL=G | 21,8% | 38,3% |
| SEKOLAH=Y -> LL=G | 16% | 37,2% |
| J. M=J -> LL=G | 33% | 51,2% |
| J. M=I -> LL=F | 4,2% | 53,2% |
| J. M=H -> LL=F | 3,2% | 42,2% |
| J. M=K -> LL=E | 18,3% | 90,9% |
| JURUSAN=S -> LL=G | 14,3% | 44,8% |
| JURUSAN=Q -> LL=G | 13,3% | 43,2% |
| JURUSAN=R -> LL=G | 7,5% | 38,5% |
| JURUSAN=U -> LL=F | 5% | 52,6% |
| JURUSAN=T -> LL=E | 3,8% | 46,9% |
| IPK=C -> LL=G | 18% | 48,2% |
| IPK=B -> LL=F | 17,5% | 43,4% |
| IPK=D -> LL=G | 6,5% | 65% |
| IPK=A -> LL=E | 5% | 40,5% |

1. Pembentukan Aturan Asosiasi

Untuk pembentukan aturan asosiasi, frequent itemset yang telah didapat akan dibentuk menjadi pola “jika A,

maka B (A ⟹ B)” dan akan dicari nilai confidence masing- masing pola tersebut dengan rumus :

Support dan Confidence dari kombinasi 2-itemset diatas diperoleh menggunakan rumus seperti berikut :

$$Support \left(X∩G\right)\frac{Jumlah lulusan mengandung X dan G}{Total Kelulusan}$$

$Support \left(X∩G\right)\frac{131}{600}$ = 0,218

 = 21,8 %

$$Confidence=P \left(G|X\right)$$

$$\frac{Jumlah kelulusan mengandung X dan G}{Jumlah Kelulusan Mengandung X}$$

$$Confidence =P \left(G|X\right)\frac{131}{342}$$

 = 0,383 = 38,3 %

Pada rumus yang sama digunakan untuk mencari nilai Supportdan confidencedari kombinasi 2- itemset lainnya.

1. Kombinasi 2-itemset berdasarkan asal sekolah
2. Jika X (asal sekolah kebupaten ) maka G (lama lulus > 5 tahun) dengan nilai support 21,8% dan confidence 38,3% artinya sebanyak 21,8% lulusan yang memiliki asal sekolah kabupaten menempuh studi selama lebih dari 5 tahun dengan nilai confidence 38,3%.
3. Jika Y (asal sekolah kecamatan ) maka G (lama lulus > 5 tahun) dengan nilai support 16% dan confidence 37,2% artinya sebanyak 16% lulusan yang memiliki asal sekolah kecamatanmenempuh studi selama lebih dari 5 tahun dengan nilaiconfidence 37,2%.
4. Berdasarkan pola diatas dapat dilihat bahwa alumni F-KIP Universitas Graha Nusantara yang berasal dari kabupaten dan kecamatan cenderung memiliki lama studi yang sama yaitu lebih dari 5 tahun.
5. Kombinasi 2-itemset berdasarkan Jenis Kelamin
6. Jika P (jenis kelamin perempuan) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan nilai support25,2% dan confidence 36,9% artinya sebanyak 25,2% lulusan yang berjenis kelamin perempuan menempuh studi selama lebih dari 4 tahun dengan nilaiconfidence 36,9%.
7. Jika L (jenis kelamin laki-laki) maka G (lama lulus > 5 tahun) dengan nilai support 13,8% dan confidence 43,5% artinya sebanyak 13,8% lulusan yang berjenis kelamin laki-laki menempuh studi selama lebih dari 5 tahun dengan nilaiconfidence 43,5%
8. Berdasarkan pola diatas dapat dilihat bahwa alumni F-KIP Universitas Graha Nusantara yang berjenis kelamin perempuan cunderung lebih cepat lulus di bandingkan dengan alumni F-KIP Universitas Graha Nusantara yang berjenis kelamin laki-laki.

1. Kombinasi 2-itemset berdasarkan Jalur masuk
2. Jika J (jalur masuk T2) maka G (lama lulus > 5 tahun) dengan nilai support 33% dan confidence 51,2% artinya sebanyak 33% lulusan yang masuk melalui jalur masuk T2 menempuh studi selama lebih dari 5 tahun dengan nilaiconfidence 51,2%.
3. Jika K (jalur masuk kemitraan) maka E (lama lulus ≤ 4 tahun) dengan nilai support 18,3% dan confidence 90,9% artinya sebanyak 18,3% lulusan yang masuk melalui jalur masuk kemitraan menempuh studi selama kurang dari 4 tahun dengan nilai confidence 90,9%.
4. Jika I (jalur masuk sumikola) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan nilai support 4,2% dan confidence 53,2% artinya sebanyak 4,2% lulusan yang masuk melalui jalur masuk sumikola menempuh studi selama lebih dari 4 tahun dengan nilai confidence 53,2%.
5. Jika H (jalur masuk SPMB) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan nilai support 3,2% dan confidence 42,2% artinya sebanyak 3,2% lulusan yang masuk melalui jalur masuk SPMB menempuh studi selama lebih dari 4 tahun dengan nilai confidence 42,2%
6. Berdasarkan pola diatas dapat dilihat bahwa alumni F-KIP Universitas Graha Nusantarayang masuk melalui jalur masuk kemitraan cenderung lebih cepat lulus di ikuti dengan jalur masuk sumikola dan jalur masuk SPMB, kemudian di ikuti oleh jalur masuk T2.
7. Kombinasi 2-itemset berdasarkan Jurusan
8. Jika S (jurusan kimia) maka G (lama lulus > 5 tahun) dengan support14,3% dan confidence 44,8% artinya sebanyak 14,3% lulusan yang berasal dari jurusan kimia menempuh studi selama lebih dari 5 tahun dengan nilai confidence 44,8%.
9. Jika Q (jurusan biologi) maka G (lama lulus > 5 tahun) dengan support13,3% dan confidence 43,2% artinya sebanyak 13,3% lulusan yang berasal dari jurusan biologi menempuh studi selama lebih dari 5 tahun dengan nilai confidence 43,2%
10. Jika R (jurusan matematika) maka G (lama lulus > 5 tahun) dengan support7,5% dan confidence 38,5% artinya sebanyak 7,5% lulusan yang berasal dari jurusan matematika menempuh studi selama lebih dari 5 tahun dengan nilai confidence 38,5%.
11. Jika U (jurusan farmasi) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan support5 dan confidence 52,6% artinya sebanyak 5% lulusan yang berasal dari jurusan farmasi menempuh studi selama lebih dari 4 tahun dengan nilai confidence 52,6%.
12. Jika T (jurusan fisika) maka E (lama lulus ≤ 4 tahun) dengan support 3,8% dan confidence 46,9% artinya sebanyak 3,8% lulusan yang berasal dari jurusan fisika menempuh studi selama kurang dari 4 tahun dengan nilai confidence 46,9%.
13. Berdasarkan pola diatas dapat dilihat bahwa alumni F-KIP Universitas Graha Nusantara yang berasal dari jurusan fisika cenderung lebih cepat lulus diikuti jurusan farmasi, kemudian jurusan matematika, jurusan kimia dan jurusan biologi.
14. Kombinasi 2-itemset berdasarkan IPK
15. Jika C (2,5 ≤ IPK < 3) maka G (lama lulus > 5 tahun) dengan support 18% dan confidence 48,2% artinya sebanyak 18% lulusan yang memiliki IPK diantara 2,5 sampai kurang dari 3 menempuh studi selama lebih dari 5 tahun dengan nilai confidence 48,2%.
16. Jika B (3 ≤ IPK ≤ 3,5) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan support17,5% dan confidence 43,4% artinya sebanyak 17,5% lulusan yang memiliki IPK diantara 3 sampai 3,5 menempuh studi selama lebih dari 4 tahun dengan nilai confidence 43,4%.
17. Jika D (IPK < 2,5) maka G (lama lulus > 5 tahun) dengan support 6,5% dan confidence 65% artinya sebanyak 6,5% lulusan yang memiliki IPK kurang dari 2,5 menempuh studi selama lebih dari 5 tahun dengan nilai confidence 65%.
18. Jika A (IPK > 3,5) maka E (lama lulus ≤ 4 tahun) dengan support 5% dan confidence 40,5% artinya sebanyak 5% lulusan yang memiliki IPK lebih dari 3,5 menempuh studi selama kurang dari 4 tahun dengan nilai confidence 40,5%.
19. Berdasarkan pola diatas dapat dilihat bahwa alumni F-MIPA yang memiliki IPK lebih dari 3,5 cenderung lebih cepat lulus diikuti IPK diantara 2,5 sampai kurang dari 3, kemudian alumniyang memiliki IPK diantara 3 sampai 3,5 dan alumni yang memiliki IPK kurang dari 2,5.

Tabel 3. Beberapa Kombinasi 3- Itemset

|  |  |  |
| --- | --- | --- |
| **3-Itemset** | **Support** | **Confi-****dence** |
| JURUSAN=R J.K=L -> LL=G | 3,5% | 48,8% |
| JURUSAN=R J.K=P -> LL=E | 5,3% | 43,2% |
| JURUSAN=R SEKOLAH=Y -> LL=E | 3,2% | 61,3% |
| JURUSAN=R SEKOLAH=X -> LL=G | 6,5% | 45,3% |
| JURUSAN=R J. M=H -> LL=G | 1,3% | 57,1% |
| JURUSAN=R J. M=J -> LL=G | 5,8% | 58,3% |
| JURUSAN=R J. M=K -> LL=E | 5,2% | 96,9% |
| JURUSAN=R J. M=I -> LL=F | 1% | 54,5% |
| JURUSAN=R IPK=A -> LL=E | 1,2 % | 58,3% |
| JURUSAN=R IPK=B -> LL=G | 4,3% | 51% |
| JURUSAN=R -> LL=G | 7,5% | 38,5% |
| JURUSAN=U -> LL=F | 5% | 52,6% |
| JURUSAN=T -> LL=E | 3,8% | 46,9% |
| JURUSAN=R IPK=C -> LL=E | 4,5% | 52,9% |

Support dan Confidence dari kombinasi 3-itemset diperoleh menggunakan rumus seperti berikut :

$$Support \left(R∩K∩G\right)==\frac{Jumlah kelulusan mengandung R, K dan G}{total kelulusan}$$

$Support \left(R∩K∩G\right)\frac{21}{600}$ = 0,035 = 3,5 %

$$Confidence=P \left(G|R,K\right)$$

$$\frac{Jumlah kelulusan mengandung R,K dan G}{Jumlah Kelulusan Mengandung R K}$$

$$Confidence =P \left(G| J X\right)\frac{21}{43}$$

 = 0,488 = 48,8 %

Rumus yang sama digunakan untuk mencari nilai Support dan confidence dari kombinasi3-itemset lainnya.

1. Kombinasi 3-itemset berdasarkan jenis kelamin.
2. Jika R (jurusan matematika)P (jenis kelamin perempuan) maka E (lama lulus ≤ 4 tahun) dengan nilai support 5,3% dan confidence 43,2% artinya sebanyak 5,3%lulusan yang berasal dari jurusan matematika dan berjenis kelamin perempuan makamenempuh studi selama kurang dari 4 tahun dengan nilai confidence 43,2%
3. Jika R (jurusan matematika)L (jenis kelamin laki-laki) maka G (lama lulus > 5 tahun) dengan nilai support 3,5% dan confidence 48,8% artinya sebanyak 3,5%lulusan yang berasal dari jurusan matematika dan berjenis kelamin laki-laki maka menempuh studi selama lebih dari 5 tahun dengan nilai confidence 48,8%.
4. Berdasarkan pola diatas dapat dilihat bahwa alumni jurusan matematika yang berjenis kelamin perempuan cenderung lebih cepat lulus di bandingkan dengan laki-laki.
5. Kombinasi 3-itemset berdasarkan jenis kelamin.
6. Kombinasi 3-itemset berdasarkan Asal sekolah
7. Jika R (jurusan matematika) Y (asal sekolah kecamatan ) maka E (lama lulus ≤ 4 tahun dengan nilai support 3,2% dan confidence 61,3% artinya sebanyak 3,2%lulusan yang berasal dari jurusan matematika dan memiliki asal sekolah kecamatan makamenempuh studi selama kurang dari 4 tahun dengan nilai confidence 61,3%.
8. Jika R (jurusan matematika) X (asal sekolah kabupaten) maka G (lama lulus > 5 tahun) dengan nilai support 6,5% dan confidence 45,3% artinya sebanyak 6,5%lulusan yang berasal dari jurusan matematika dan memiliki asal sekolah kabupaten makamenempuh studi selama lebih dari 5 tahun dengan nilai confidence 45,3%
9. Berdasarkan pola diatas dapat dilihat bahwa alumni jurusan matematika yang asal sekolah kecamatan cenderung lebih cepat lulus di bandingkan dengan asal sekolah kebupaten
10. Kombinasi 3-itemset berdasarkan Jalur Masuk
11. Jika R (jurusan matematika)K (jalur masuk kemitraan) maka E (lama lulus ≤ 4 tahun) dengan nilai support 5,2% dan confidence 96,9% artinya sebanyak 5,2%lulusan yang berasal dari jurusan matematika dan masuk melalui jalur masuk kemitraan maka menempuh studi selama kurang dari 4 tahun dengan nilai confidence 96,9%.
12. Jika R (jurusan matematika)I (jalur masuk sumikola) maka F (4 tahun < lama lulus ≤ 5 tahun) dengan nilai support 1% dan confidence 54,5% artinya sebanyak 1%lulusan yang berasal dari jurusan matematika dan masuk melalui jalur masuk sumikola makamenempuh studi selama lebih dari 4 tahun dengan nilai confidence 54,5%
13. Jika R (jurusan matematika) J (jalur masuk T2) maka G (lama lulus > 5 tahun) dengan nilai support 5,8% dan confidence 58.3% artinya sebanyak 5,8%lulusan yang berasal dari jurusan matematika dan masuk melalui jalur masuk T2 makamenempuh studi selama lebih dari 5 tahun dengan nilai confidence 58,3%.
14. Jika R (jurusan matematika) H (jalur masuk SPMB) maka G (lama lulus > 5 tahun) dengan nilai support 1,3% dan confidence 57,1% artinya sebanyak 1,3%lulusan yang berasal dari jurusan matematika dan masuk melalui jalur masuk SPMB maka menempuh studi selama lebih dari 5 tahun dengan nilai confidence 57.1%
15. Berdasarkan pola diatas dapat dilihat bahwa alumni jurusan matematika yang masuk melalui jalur masuk kemitraan cenderung lebih cepat lulus di ikuti dengan jalur masuk sumikola kemudian di ikuti oleh jalur masuk T2 dan jalur masuk SPMB.
16. Kombinasi 3-itemset berdasarkan IPK
17. Jika R (jurusan matematika)A (IPK > 3,5) maka E (lama lulus ≤ 4 tahun) dengan nilai support 1,2% dan confidence 58,3% artinya sebanyak 1,2%lulusan yang berasal dari jurusan matematika dan memiliki IPK lebih dari 3,5 makamenempuh studi selama kurang dari 4 tahun dengan nilai confidence 58,3%
18. Jika R (jurusan matematika) C (2,5 ≤ IPK < 3) maka E (lama lulus ≤ 4 tahun) dengan nilai support 4,5% dan confidence 52,9% artinya sebanyak 4,5% lulusan yang berasal dari jurusan matematika dan memiliki IPK diantara 2,5 sampai kurang dari 3 maka menempuh studi selama kurang dari 4 tahun dengan nilai confidence 52,9%.
19. Jika R (jurusan matematika) B (3 ≤ IPK ≤ 3,5) maka G (lama lulus >5 tahun) dengan nilai support 4,3% dan confidence 51% artinya sebanyak 4,3%lulusan yang berasal dari jurusan matematika dan memiliki IPK diantara 3 sampai 3,5 makamenempuh studi selama lebih dari 5 tahun dengan nilai confidence 51%.
20. Berdasarkan pola diatas dapat dilihat bahwa alumni jurusan matematika yang memiliki IPK lebih dari 3,5 dan IPK diantara 2,5 sampai kurang dari 3 cenderung lebih cepat lulus dibandingkan alumniyang memiliki IPK diantara 3 sampai 3,5.

**BAB IV. KESIMPULAN**

Berdasarkan hasil penelitian didapatkan kesimpulan dari Penggunaan Association Rule Data Mining Untuk Menentukan pola lama Studi Mahasiswa.

1. Dengan menggunakan association rule data mining dihasilkan pola lama studi mahasiswa F-KIP Universitas Graha Nusantara. Beberapa pola lama studi mahasiswa pada F-KIP Universitas Graha Nusantara antara lain :
2. Kombinasi 2-itemset antara lain sebanyak 18,3% mahasiswa yang masuk F-KIP Universitas Graha Nusantara melalui jalur masuk kemitraan menempuh studi F-KIP Universitas Graha Nusantara selama kurang dari 4 tahun dengan nilai confidance sebesar 90,9%.
3. Kombinasi 2-itemset antara lain sebanyak 7,5% lulusan yang berasal dari jurusan matematika menempuh studi selama lebih dari 5 tahun dengan nilai confidence 38,5%.
4. Kombinasi 2-itemset antara lain sebanyak 5% lulusan yang memiliki IPK lebih dari 3,5 menempuh studi selama kurang dari 4 tahun dengan nilai confidence 40,5%.
5. Kombinasi 3-itemset antara lain sebanyak 1,2% lulusan yang berasal dari jurusan matematika dan memiliki IPK lebih dari 3,5 maka menempuh studi selama kurang dari 4 tahun dengan nilai confidence 58,3%.
6. Kombinasi 3-itemset antara lain sebanyak 5,2% lulusan yang berasal dari jurusan matematika dan masuk melalui jalur masuk kemitraan maka menempuh studi selama kurang dari 4 tahun dengan nilai confidence 96,9%.
7. Dengan Pembangkitan 2-itemset pada algoritma hash-based menggunakan tabel hash sedangkan pada algoritma apriori menggunakan minimum support. Namun jumlah 2-itemset yang dihasilkan tidak memiliki perbedaan yaitu sebanyak 49 kombinasi 2-itemset.

DAFTAR PUSTAKA

1. E. Aribowo dan O. SAD, "Analisis Perbandingan Algoritma Apriori Dan Algoritma Hash Based Pada Market Basket Analysis Di Apotek Universitas Ahmad Dahlan," *Jurnal Sarjana Teknik Informatika volume 3 Nomor 1, Februari 2015.*
2. D. Fitriati, "Implementasi Data Mining untuk Menentukan Kombinasi Media Promosi Barang Berdasarkan Perilaku Pembelian Pelanggan Menggunakan Algoritma Apriori,"*Annual Research Seminar volume 2 Nomor 1, Desember 2016.*
3. W. Weku, " Implementasi Data Mining Untuk Menentukan Studi Mahasiswa Menggunakan Algoritma Apriori," *Jurnal Matematika Dan Aplikasi Volume 3 Nomor 1 Tahun 2014.*
4. D. Listriani dan H.Setyaningrum, "Penerapan Metode Asosiasi Menggunakan Algoritma Apriori Pada Aplikasi Analisa Pola Belanja Konsumen,"*Jurnal Teknik Informatika Volume 9 Nomor. 2, Oktober 2016.*
5. T. Pradana, "Penggalian Kaidah Multilevel Association Rule Dari Data Mart Swalayan ASGAP," *Jurnal SPIRIT Volume. 7 Nomor. 2 Nopember 2015.*