

IMPLEMENTASI ALGORITMA APRIORI DALAM MENENTUKAN PROGRAM STUDI YANG DIAMBIL MAHASISWA

Ahmad Fikri Fajri

Sistem Informasi, STMIK Jayanusa, Padang, Sumatera Barat, Kode Pos : 25116
email: fajri.bayang@gmail.com

Submitted: 16-05-2016, Reviewed:16-05-2016, Accepted:24-05-2016

<http://dx.doi.org/10.22216/jit.2016.v10i2.402>

Abstract

The decision is something that is very influential in the process of facing the alternative chosen, various constraints in selecting the concentration of study programs that fit the criteria is quite confusing. The purpose of this study to determine the concentration of courses that will be taken by students with the technique of Data Mining using Apriori Algorithm. The system requires training data on whether the student fits the Prodi X. A match prediction can be obtained based on the comparison with other students who have similar student demographics data with A. By using Apriori algorithm is obtained in the form of rules a collection of frequent itemset with high confidence value.

Keywords: Apriori Algorithm, CollegeStudent, Recommendation.

Abstrak

Keputusan merupakan suatu hal yang sangat berpengaruh dalam proses menghadapi alternative yang dipilih, Berbagai kendala dalam memilih konsentrasi prodi yang sesuai dengan kriteria memang cukup membingungkan. Tujuan penelitian ini untuk menentukan konsentrasi program studi yang akan diambil oleh mahasiswa dengan teknik Data Mining menggunakan metode Algoritma Apriori. Sistem membutuhkan data latih agar dapat membandingkan apakah mahasiswa A cocok dengan prodi X. Prediksi kecocokan dapat diperoleh berdasarkan hasil perbandingan dengan mahasiswa lain yang memiliki kemiripan data demografi dengan mahasiswa A. Dengan menggunakan Algoritma Apriori didapatkan hasil berupa aturan yang merupakan kumpulan frequent item set dengan nilai confidence yang tinggi.

Kata Kunci: AlgoritmaApriori, Mahasiswa, Rekomendasi.

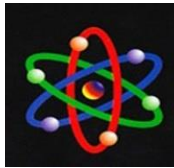
PENDAHULUAN

Keputusan merupakan suatu hal yang sangat berpengaruh dalam proses menghadapi alternatif yang dipilih. Untuk masuk kedalam prodi yang diinginkan merupakan suatu hal yang sangat diharapkan setiap mahasiswa, tentunya harus sesuai dengan kriteria.

Berawal dari criteria tersebut diperlukan sebuah sistem untuk menjembatani mahasiswa dalam pemilihan konsentrasi program studi. Sistem yang dimaksud adalah sebuah sistem yang dapat memberikan rekomendasi pemilihan program studi.

Sebagai algoritma klasifikasi, *Apriori* memiliki beberapa keunggulan dalam teknik pengklasifikasian, untuk pengklasifikasian diperlukan data dari calon mahasiswa seperti jurusan waktu SLTA/SMA atau sederajat, nilai kelulusan, predikat lulus dan nilai tes masuk pada perguruan tinggi

Kennedi Tampubolon, dkk (2013) menyimpulkan penerapan Algoritma *Apriori* pada teknik *Data Mining* sangat efisien dan dapat mempercepat proses pembentukan kecenderungan pola kombinasi itemset hasil penjualan alat-alat kesehatan di Apotek Kelambir-2 Medan, yaitu dengan *support* dan *confidence* tertinggi adalah *Stick Asam*



Urut - Stick Gula dan Stick Colestrol-Stick Gula.

Prinsip dari Apriori ini adalah bila *itemset* digolongkan sebagai *frequent itemset*, yang memiliki *support* lebih dari yang ditetapkan sebelumnya, maka semua subsetnya juga termasuk golongan *frequent itemset*, dan sebaliknya.

METODE PENELITIAN

Penelitian ini dilakukan dengan cara *interview* untuk mengumpulkan data yang diperlukan serta melakukan kajian pustak untuk mendukung dalam pengambilan hasil penelitian

Masing-masing langkah dalam penelitian dapat diuraikan seperti berikut ini :

Mendefinisikan Ruang Lingkup Masalah

Ruang lingkup masalah yang akan diteliti harus ditentukan terlebih dahulu, karena tanpa mampu menentukan serta mendefinisikan batasan masalah yang akan diteliti, maka tidak akan pernah didapat suatu solusi yang terbaik dari masalah tersebut.

Menganalisa Masalah

Dengan menganalisa masalah yang telah ditentukan tersebut, maka diharapkan masalah dapat dipahami dengan baik. Pada analisa masalah ini digambarkan proses pengambilan keputusan dalam pemilihankonsentrasi program studi. Metode analisa yang dipakai untuk menganalisis kebutuhan dalam mengambil keputusan yang optimal telah ditetapkan dengan menggunakan kriteria-kriteria yang telah ditentukan.

Menentukan Tujuan

Berdasarkan pemahaman dari masalah, maka ditentukan tujuan yang akan dicapai dari penulisan ini. Pada tujuan ini ditentukan target yang dicapai, terutama yang dapat mengatasi masalah-masalah yang ada.

Mempelajari Literatur

Untuk mencapai tujuan, maka dipelajari beberapa literatur-literatur yang diperkirakan dapat digunakan. Kemudian literatur-literatur yang dipelajari tersebut diseleksi untuk dapat ditentukan literatur-literatur mana yang akan digunakan dalam penelitian. Literatur diambil dari internet, yang berupa artikel dan jurnal ilmiah tentang *Data Mining Algoritma Apriori* serta bahan bacaan lain yang mendukung penelitian.

Mengumpulkan Data

Dalam pengumpulan data dilakukan observasi yaitu pengamatan secara langsung di tempat penelitian sehingga permasalahan yang ada dapat diketahui secara jelas. Kemudian dilakukan *interview* yang bertujuan untuk mendapatkan informasi atau data yang dibutuhkan. Selain itu juga dilakukan studi kepustakaan yaitu dengan membaca buku-buku yang menunjang dalam melakukan penganalisaan terhadap data dan informasi yang didapat. Adapun data-data yang diperlukan dalam penelitian ini adalah data pemilihan konsentrasi program studi oleh mahasiswa.

Analisis Data

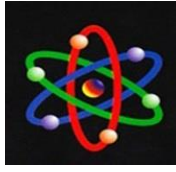
Data yang diperoleh dari hasil penelitian selanjutnya dilakukan pengolahan dengan *Algoritma Apriori*

Implementasi Data Menggunakan Algoritma Apriori

Tahapan proses dimana data yang sudah diproses ditentukan *Market Basket Analisis* dengan mencari *Frequent Itemset*. Bagian ini bertujuan untuk menganalisis dan memahami teknik yang digunakan dalam pencarian data sering muncul atau *Frequent Itemset* dan pembangunan aturan asosiasi.

Pengujian Hasil Pengolahan Data

Pengujian dilakukan dalam konsep pada *Apriori Clustering* untuk menentukan jenis program studi yang sering di ambil oleh mahasiswa. Mekanisme pengujian dilakukan



dengan menggunakan variabel yang ada berupa data matakuliah. Berdasarkan variabel-variabel tersebut maka dapat ditentukan program studi yang efektif untuk mahasiswa tersebut.

HASIL DAN PEMBAHASAN

Dalam analisa ini penulis menganalisa dan mengelompok kelompok data untuk memudahkan dalam mengerjakan perancangan sistem yang telah direncanakan sebelumnya sesuai dengan *variable-variabel* yang dibutuhkan, guna untuk menganalisa data yang diperlukan dalam perancangan system ini.

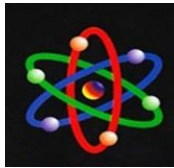
Tahap dalam menganalisa data dengan *Algoritma Apriori* pada penentuan program studi dimulai dengan menyeleksi dan membersihkan data-data yang akan dianalisis, kemudian mencari semua jenis *item* yang ada didalam *list* data program studi, selanjutnya mencari jumlah setiap item yang ada. Sesuai dengan *support* yang telah ditentukan maka terseleksilah beberapa *item* data yang memenuhi minimal *support* sesuai dengan jumlah *item* didalam program studi, itu disebut pembentukan kombinasi satu *item*. Tahap selanjutnya dari *item-item* yang telah terseleksi dibentuklah kombinasi dua *item*, maka terbentuklah beberapa *item* data dengan kombinasi 2 *item* yang berbeda, dengan *support* yang ditentukan maka terseleksilah beberapa data dua *item* ini disebut pembentukan kombinasi dua *item*. Demikian seterusnya sampai kombinasi batas maksimal *item* program studi. Apabila nilai *Support* item \geq minimal *support* maka proses pencarian calon kombinasi *item set* berlanjut, jika tidak proses pencarian selesai.

Persiapan Data

Tabel1 : Data Untuk Proses Mining

NO PENDAFTARAN	JURUSAN SLTA	PREDIKAT LULUS UN	NILAI TES MASUK	PRODI
01894	IPS	SANGAT BAIK	B	SISTEM INFORMASI
01895	KEJURUAN	BAIK	B	SISTEM INFORMASI
01896	IPS	SANGAT BAIK	A	SISTEM INFORMASI
01897	IPA	BAIK	B	SISTEM INFORMASI
01898	IPS	SANGAT BAIK	A	SISTEM INFORMASI
01899	KEJURUAN	SANGAT BAIK	A	SISTEM KOMPUTER
01900	KEJURUAN	BAIK	B	SISTEM INFORMASI
01901	KEJURUAN	BAIK	A	SISTEM INFORMASI
01902	KEJURUAN	BAIK	A	SISTEM INFORMASI
01903	KEJURUAN	SANGAT BAIK	A	SISTEM KOMPUTER
01904	KEJURUAN	SANGAT BAIK	A	SISTEM KOMPUTER
01905	IPA	SANGAT BAIK	A	SISTEM INFORMASI
01906	KEJURUAN	BAIK	B	SISTEM INFORMASI
01907	IPS	BAIK	A	SISTEM KOMPUTER
01908	KEJURUAN	SANGAT BAIK	A	SISTEM KOMPUTER
01909	IPA	SANGAT BAIK	A	SISTEM INFORMASI
01910	KEJURUAN	SANGAT BAIK	B	SISTEM KOMPUTER
01911	IPS	SANGAT BAIK	A	SISTEM KOMPUTER
01912	IPS	SANGAT BAIK	A	SISTEM INFORMASI
01913	KEJURUAN	SANGAT BAIK	B	SISTEM KOMPUTER

Tahapan awal yang dilakukan dalam penelitian ini adalah menyiapkan data, dimana data diperoleh dari bagian Administrasi Akademik (BAA). Data yang diperoleh dan akan digunakan dalam penelitian ini berupa data berkaitan dengan data daftar ulang yaitu Asal Sekolah Lanjutan Tingkat Atas (SLTA), Nilai Ujian Nasional (UN), Nilai Tes Masuk, dan Jurusan yang Diambil. Data yang diambil yaitu data Daftar Ulang Mahasiswa Ganjil Tahun Ajaran 2014/2015 di STMIK Jayanusa Padang.



Berikut ini table 2 jumlah *item* yang sering muncul untuk masing-masing *item set*.

Tabel 2 : Jumlah Kemunculan Data

ITEM	JUMLAH
IPS	6
IPA	3
KEJURUAN	11
SANGAT BAIK	13
BAIK	7
A	13
B	7
SI	12
SK	8

Tahapan Perhitungan *Algoritma Apriori*

Untuk perhitungan *Algoritma Apriori* dimulai dengan mendefinisikan masing-masing *itemset*. Tahapan dalam perhitungan *Algoritma Apriori* untuk menemukan *association rules* yaitu menuliskan data transaksi yang terjadi dalam periode tertentu.

Mendefinisikan Masing – Masing *Item set*

Langkah Pertama adalah memuat kombinasi 1 *item sets* pada setiap criteria dan frekuensi masing-masing Produk dihitung *Support* sesuai dengan data tabular pada tabel.

Nilai Support (IPS) = (jumlah data mengandung IPS / Total Data) x 100%
= $6/20 \times 100\% = 30$

Nilai Support (IPA) = (jumlah data mengandung IPA / Total Data) x 100%
= $3/20 \times 100\% = 15$ Dst

Langkah berikutnya adalah menentukan *Minimum Support* Dengan Nilai ≥ 25 .

Tentukan *item set* yang memenuhi *minimum support*.

Kombinasi 2 Item sets

Langkah berikutnya adalah membuat kombinasi 2 *item sets* pada setiap frekuensi masing-masing kombinasi dihitung sesuai dengan data tabular

Kombinasi 3 Item sets

Berikutnya adalah kombinasi 3 *item sets* pada setiap frekuensi masing-masing kombinasi dihitung sesuai dengan data tabular

Kombinasi 4 Item sets

Berikutnya adalah kombinasi 4 *item sets* pada setiap frekuensi masing-masing kombinasi dihitung sesuai dengan data tabular.

Karena pada 4 item set nilai *Support* tidak Melebihi dari *Minimum Support* maka proses berhenti di sini, untuk hasil *Rule* diambil dari Kombinasi 2 *item set* dan 3 *Item set* dan akan dihitung *confidencenya*.

Contoh perhitungan Confidence :

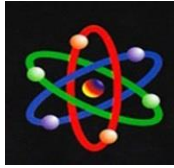
$$\text{Confidence (ips,sangat baik)} = s(\text{ips,sangat baik}) / s(\{\text{ips}\}) = 25/30 = 0,83$$

$$\text{Confidence (ips,a)} = s(\text{ips,a}) / s(\{\text{ips}\}) = 25/30 = 0,83 \quad \dots\dots\text{dst}$$

Rule Asosiasi *Algoritma Apriori*

Untuk mencari aturan asosiasi diperlukan juga minimal Confidence Misal minimal confidence :0.75, aturan asosiasi yang mungkin terbentuk dapat ditunjukkan pada Tabel 3.

Tabel 3 : Hasil Kombinasi 2 Item set dan 3 Item set Yang Memenuhi Minimum Confidence



ITEM	JUMLAH	SUPPORT (%)	CONFIDENCE (%)
IPS,SANGATBAIK	5	25	0,83
IPS,A	5	25	0,83
BAIK,SI	6	30	0,86
SANGATBAIK,A	10	50	0,77
KEJURUAN,SANGATBAIK,SK	6	30	1,00
KEJURUAN,BAIK,SI	5	25	1,00

Dilihat dari hasil di atas, ternyata item set yang tersisa, adalah (kejuruan, IPS, sangatbaik, baik, A, yang selanjutnya SI, SK) dapat dinyatakan sebagai frequent itemset. Langkah selanjutnya adalah dengan membentuk association rule yang kuat berdasarkan frequent item set yang dihasilkan tersebut berdasarkan nilai confidence.

- **Rule 1 : {SangatBaik} → {A}**
 $C = \frac{\{SangatBaik, A\}}{\{SangatBaik\}} = \frac{10}{13} = 0,77$
- **Rule 2 : {Baik} → {SI}**
 $C = \frac{\{Baik, SI\}}{\{Baik\}} = \frac{6}{7} = 0,86$
- **Rule 3 : {IPS} → {SangatBaik}**
 $C = \frac{\{IPS, SangatBaik\}}{\{IPS\}} = \frac{5}{6} = 0,83$
- **Rule 4 : {IPS} → {A}**
 $C = \frac{\{IPS, A\}}{\{IPS\}} = \frac{5}{6} = 0,83$
- **Rule 5 : {Kejuruan, SangatBaik} → {SK}**
 $C = \frac{\{Kejuruan, SangatBaik, SK\}}{\{Kejuruan, SangatBaik\}} = \frac{6}{6} = 100 \%$
- **Rule 6 : {Kejuruan, Baik} → {SI}**
 $C = \frac{\{Kejuruan, Baik, SI\}}{\{Kejuruan, Baik\}} = \frac{5}{5} = 100 \%$

SIMPULAN

Dari uraian yang sudah dibahas sebelumnya dapat ditarik kesimpulan :

Dengan menggunakan *Algoritma Apriori* didapatkan hasil berupa aturan (*rules*) yang merupakan kumpulan *frequent item set* dengan nilai *confidence* yang tinggi.

Dengan didapatkannya *rules* ini maka mahasiswa dapat menggunakan *rules* tersebut dalam memilih program studi yang diinginkan.

Hasil implementasi telah dilakukan dengan menghasilkan nilai *strong confidence* paling tinggi 75 %

DAFTAR PUSTAKA

- (Erwin (Jurusan Teknik Informatika, Fakultas Ilmu Komputer, 2009)
(Tampubolon et al., 2013)
(Fithri et al., 2014)
(Wandi, Hendrawan, & Mukhlason, 2012)
(Wahyono, 2012)