

**PENCARIAN POLA KONSENTRASI PEMINATAN PROGRAM STUDI TEKNIK  
INFORMATIKA DI UNIVERSITAS MUHAMMADIYAH JEMBER  
MENGUNAKAN  
ALGORITMA APRIORI DAN ASSOCIATION RULE**

**Nama Mahasiswa** : Istiana Fii Sabilillah  
**NIM** : 1210651066  
**Jurusan** : Teknik Informatika  
**Dosen Pembimbing I:** Lutfi Ali Muharrom, S. Si, M.Si

**Abstrak**

*Association rule* mining adalah sebuah teknik data mining untuk menemukan asosiasi yang sesuai atau hubungan antar data dari sejumlah besar data. Sejumlah besar data yang diproses, dapat ditemukan *rule – rule* nya menggunakan algoritma apriori dengan dua proses utama yaitu pencarian frequent itemsets berdasarkan minimum *support* (*join step* dan *prune step*) dan pembentukan *rule* berdasarkan minimum *confidence*.

Dalam Tugas Akhir ini, akan membuat sistem pencarian pola asosiasi konsentrasi peminatan program studi terhadap data akademik alumni dengan menggunakan *Association Rule* dan Algoritma Apriori yang di mana metode ini dapat digunakan lebih lanjut dalam sistem rekomendasi bidang minat pilihan mahasiswa. Pada saat ini, mahasiswa di jurusan Teknik Informatika UNMUH Jember masih belum memiliki sistem untuk membantu mereka menentukan bidang minat. Pemilihan bidang minat ini sangat penting karena berkaitan dengan pengambilan tugas akhir dari mahasiswa tersebut.

Hasil analisis menunjukkan bahwa sistem pencarian pola asosiasi konsentrasi ini di temukan bahwa semakin kecil *confidence*, maka akan memunculkan rule semakin banyak dan pilihan rule untuk sistem rekomendasi bidang minat dapat lebih beragam

Kata Kunci : association rule mining, dan Algoritma Apriori.

**Abstract**

Association rule mining is a data mining technique for finding the appropriate association or the correlation between the data among a huge amount of data. The rules can be found in the most of the processed data by using Apriori Algorithm (*join step* and *prune step*) and rule establishment based on the minimum confidence.

This thesis will make a system to find association pattern which focuses on the preference in a study program to the alumni academic data by using Association Rule and Apriori Algorithm; in which can be used more effectively in students' preference recommendation system. At the time being, Computer Science / Informatics students in UNMUH Jember haven't found the system to help them finding their preference; which is very important because this is related to their thesis subjects.

The analysis result has shown that the less confidence will show more rules and the choice of rules for the students' preference recommendation system will be more various.

*Key words : association rule mining and Apriori Algorithm.*

## 1.1 Latar Belakang

Algoritma Apriori termasuk jenis aturan asosiasi pada data mining. Selain apriori, yang termasuk pada golongan ini adalah metode *generalized rule induction* dan algoritma *hash based*. Aturan yang menyatakan asosiasi antara beberapa atribut sering disebut analisis keranjang belanja.

Analisis asosiasi atau *Association Rule* adalah teknik penggalian data untuk menemukan aturan asosiatif antara suatu kombinasi barang. Contoh aturan asosiatif dari analisa pembelian di suatu pasar swalayan adalah dapat diketahuinya berapa besar kemungkinan seorang pelanggan membeli roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik pasar swalayan dapat mengatur penempatan barangnya atau merancang kampanye pemasaran dengan memakai kupon diskon untuk kombinasi barang tertentu.

Pada saat ini, mahasiswa di Jurusan Teknik Informatika UNMUH JEMBER masih belum memiliki sistem untuk membantu mereka menentukan bidang minat. Pemilihan bidang minat ini sangat penting karena sangat berkaitan dengan pengambilan tugas akhir dari mahasiswa tersebut. Sistem ini akan dibangun berdasarkan hasil analisa terhadap sejumlah data-data sebelumnya.

## 1.2 Tujuan

Tujuan dari Tugas Akhir ini adalah membuat sistem pencarian pola asosiasi konsentrasi peminatan dimana sistem akan menemukan aturan – aturan (*rule*) yang akan digunakan untuk membantu menentukan bidang minat pilihan.

## 2.1 Penggalian Data

Penggalian data (*data mining*) adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam basis data. Penggalian data adalah tentang memecahkan suatu masalah dengan menganalisis data yang sudah ada, dan juga didefinisikan sebagai proses pola

dalam data, dimana pola yang didapat harus memiliki beberapa keuntungan (Witten dan Frank, 2005).

## 2.2 Association Rule

*Association rule* atau aturan asosiasi sering dinamakan analisis keranjang belanja, karena awalnya berasal dari studi tentang basis data transaksi pelanggan untuk menentukan kebiasaan suatu produk dibeli bersama produk apa (Santosa, 2007). Sebagai contoh studi transaksi di supermarket, seseorang yang membeli susu bayi juga membeli sabun mandi. Disini berarti susu bayi bersama sabun mandi.

Menurut Agrawal dan Srikant dalam (Ridowati,2003), masalah dalam pencarian seluruh *association rule* dapat dibagi menjadi dua, yaitu bagaimana menemukan seluruh barang - barang (*itemset*) yang memiliki nilai penunjang (*support*) di atas nilai penunjang terkecil yang telah ditentukan. Nilai penunjang untuk sebuah barang adalah jumlah transaksi yang terdapat dalam *itemset* itu. *itemset* dengan nilai penunjang terkecil disebut *large itemset (I-itemset)*. Masalah yang kedua adalah bagaimana menggunakan *large itemset* untuk menghasilkan aturan-aturan yang diperlukan.

Aturan yang hendak dihasilkan adalah dalam bentuk  $a \rightarrow (I-a)$ , yaitu perbandingan antara *support* (I) dengan *support* (a) haruslah lebih besar dari *minconf* (minimum confidence). Misalkan  $[I]=\{I_1, I_2, \dots, I_m\}$  adalah himpunan literal yang disebut barang. Himpunan *item-item* disebut dengan *itemset*. D adalah himpunan transaksi, di mana setiap transaksi T itemset sehingga  $TE$ . I. Sebuah association rule adalah sebuah implikasi berbentuk  $X \rightarrow Y$ , di mana  $X \subseteq I$ ,  $Y \subseteq I$ , dan  $X \cap Y = \emptyset$ .

## 3.1 Algoritma Apriori

Sistem pencarian pola asosiasi bidang bidang minat ini menggunakan algoritma apriori dalam menentukan bidang minat mahasiswa. Langkah-

langkah dalam memproses dengan algoritma apriori ini adalah pertama, proses pencarian rule terlebih dahulu yaitu memasukkan data alumni yang terpilih, lalu Algoritma Apriori akan digunakan untuk menganalisis data. Saat itu pula Algoritma Apriori mulai mengerjakan 2 tahapan yaitu proses pembentukan *frequent item* dan proses pembentukan *Association Rule*, kemudian keluarlah *Output Association Rule*.

Analisis dari data keranjang belanja dapat digunakan untuk mendapatkan aturan asosiasi dari kombinasi IPK mahasiswa per KBK dengan bidang minat yang di ambil oleh mahasiswa seperti contoh tabel di bawah ini :

ID	RPL	SBC	KBJ	MINAT
0910651001	3,21	3,18	3,35	KBJ
0910651002	2,98	2,00	3,41	KBJ
0910651003	3,33	2,73	3,66	SBC

Untuk memudahkan analisis, maka IPK masing – masing KBK dibuat menjadi 3 interval :

**IPK menjadi 3 interval :**

**IPK KBK RPL**

( Minimal – 2,83 ) = RPL 1

( 2, 84 – 3,17 ) = RPL 2

( 3, 18 – Maksimal ) = RPL 3

**IPK KBK SBC**

( Minimal – 2,83 ) = SBC 1

( 2, 84 – 3,17 ) = SBC 2

( 3, 18 – Maksimal ) = SBC 3

**IPK KBK KBJ**

( Minimal – 3,17 ) = KBJ 1

( 3, 18 – 3,33 ) = KBJ 2

( 3, 34 – Maksimal ) = KBJ 3

**Support Count**

Support count adalah banyaknya itemsets yang sama muncul secara

bersamaan pada suatu data transaksi pada keranjang belanja.

**Confidence**

Confidence adalah suatu ukuran yang menunjukkan hubungan kondisi antar dua barang ( misal seberapa sering bidang minat pilihan B diambil jika mahasiswa memiliki IPK A ) dari keseluruhan transaksi. Berikut ini adalah salah satu contoh perhitungan confidence dari salah satu transaksi yang diambil dari tabel Transaksi Belanja Akademik :

Confidence,

$$\begin{aligned}
 (x \ y) &= (x \cup y) / (x) \\
 &= (IPK KBK RPL ( 3.17- \\
 &\text{maksimal, IPK KBK KBJ ( minimal - 3.17 } \\
 &), IPK KBK KCV ( minimal - 2.83 ), RPL \\
 &)/ 4 = 1 / 4 = 0.25
 \end{aligned}$$

**Contoh Proses Algoritma Apriori**

Algoritma Apriori memiliki dua tahap penting yaitu proses Join Step dan Prune Step. Proses Join Step adalah proses untuk mencari Frequent Itemsets dan yang tidak sesuai dengan ketentuan maka akan dihilangkan melalui proses Prune Step.

**Proses pencarian Frequent Itemsets.**

Langkah awal adalah tentukan dahulu *minimum support count* yang diperlukan adalah 0.1 dan *minimum confidence* yang dibutuhkan adalah 0.9.

**Generating 1**

Itemset	Support
RPL (RPL-3)	2
SBC (SBC-3)	1
KBJ (KBJ-3)	3
(KBJ)	2
RPL (RPL-2)	1
SBC (SBC-1)	2

(SBC)	1
-------	---

### Generating 2

Itemset	Support
RPL (RPL-3) - SBC (SBC-3)	1
RPL (RPL-3) - KBJ (KBJ-3)	2
RPL (RPL-3) - (KBJ)	1
RPL (RPL-3) - RPL (RPL-2)	0
RPL (RPL-3) - SBC (SBC-1)	1
RPL (RPL-3) - (SBC)	1
SBC (SBC-3) - KBJ (KBJ-3)	1
SBC (SBC-3) - (KBJ)	1
SBC (SBC-3) - RPL (RPL-2)	0
SBC (SBC-3) - SBC (SBC-1)	0
SBC (SBC-3) - (SBC)	0
KBJ (KBJ-3) - (KBJ)	2
KBJ (KBJ-3) - RPL (RPL-2)	1
KBJ (KBJ-3) - SBC (SBC-1)	2
KBJ (KBJ-3) - (SBC)	1
(KBJ) - RPL (RPL-2)	1
(KBJ) - SBC (SBC-1)	1

(KBJ) - konsen (SBC)	0
RPL (RPL-2) - SBC (SBC-1)	1
RPL (RPL-2) - (SBC)	0
SBC (SBC-1) - (SBC)	1

Tabel di atas menunjukkan proses *joint step* dan kolom berwarna kuning, menunjukkan *prune step* artinya *support count* yang tidak memenuhi *minimum support count* harus dibuang.

### Generating 3

Itemset	Support
RPL (RPL-3) - SBC (SBC-3) - KBJ (KBJ-3)	1
RPL (RPL-3) - SBC (SBC-3) - (KBJ)	1
RPL (RPL-3) - KBJ (KBJ-3) - (KBJ)	1
RPL (RPL-3) - KBJ (KBJ-3) - SBC (SBC-1)	1
RPL (RPL-3) - KBJ (KBJ-3) - (SBC)	1
RPL (RPL-3) - SBC (SBC-1) - (SBC)	1
SBC (SBC-3) - KBJ (KBJ-3) - (KBJ)	1
KBJ (KBJ-3) - (KBJ) - RPL (RPL-2)	1
KBJ (KBJ-3) - (KBJ) - SBC (SBC-1)	1
KBJ (KBJ-3) - RPL (RPL-2)	1

2) - SBC (SBC-1)	
KBJ (KBJ-3) - SBC (SBC-1) - (SBC)	1
(KBJ) - RPL (RPL-2) - SBC (SBC-1)	1

#### Generating 4

Itemset	Support
RPL (RPL-3) - SBC (SBC-3) - KBJ (KBJ-3) - (KBJ)	1
RPL (RPL-3) - KBJ (KBJ-3) - SBC (SBC-1) - (SBC)	1
KBJ (KBJ-3) - (KBJ) - RPL (RPL-2) - SBC (SBC-1)	1

#### Proses Pecarian Association Rules

Proses pencarian *Association Rules* ini memanfaatkan *minimum confidence* yang telah ditentukan sebelumnya yaitu 0,9. Contoh :

**R1 : IPK KBK RPL = (RPL 3) → IPK KBK SBC = (SBC 3)**

$Confidence = Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) ; IPK\ KBK\ SBC = (SBC\ 3) \} / Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) \} = 1 / 2 = 0,5$

**R2 : IPK KBK RPL = (RPL 3) → IPK KBK KBJ = (KBJ 3)**

$Confidence = Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) ; IPK\ KBK\ KBJ = (KBJ\ 3) \} / Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) \} = 2 / 2 = 1$

**R3 : IPK KBK RPL = (RPL 3) → KBJ**

$Confidence = Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) ; KBJ \} / Support\ Count \{ IPK\ KBK\ RPL = (RPL\ 3) \} = 1 / 2 = 0,5$

**R4 : IPK KBK SBC = (SBC 3) IPK → KBK KBJ = (KBJ 3)**

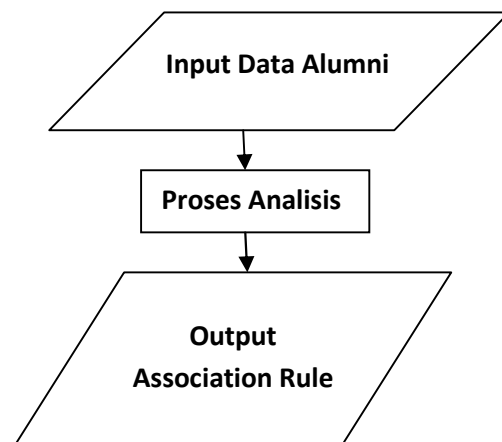
$Confidence = Support\ Count \{ IPK\ KBK\ SBC = (SBC\ 3) ; IPK\ KBK\ KBJ = (KBJ\ 3) \} / Support\ Count \{ IPK\ KBK\ SBC = (SBC\ 3) \} = 1 / 2 = 1$

**R5 : IPK KBK SBC = (SBC 3) KBJ →**

$Confidence = Support\ Count \{ IPK\ KBK\ SBC = (SBC\ 3) ; KBJ \} / Support\ Count \{ IPK\ KBK\ SBC = (SBC\ 3) \} = 1 / 2 = 0,5$

Contoh *rule* yang ditolak adalah *rule* nomor 1 dan 3, ditandai dengan cetakan tebal. *Rule* tersebut ditolak karena tidak memenuhi *minimum confidence*.

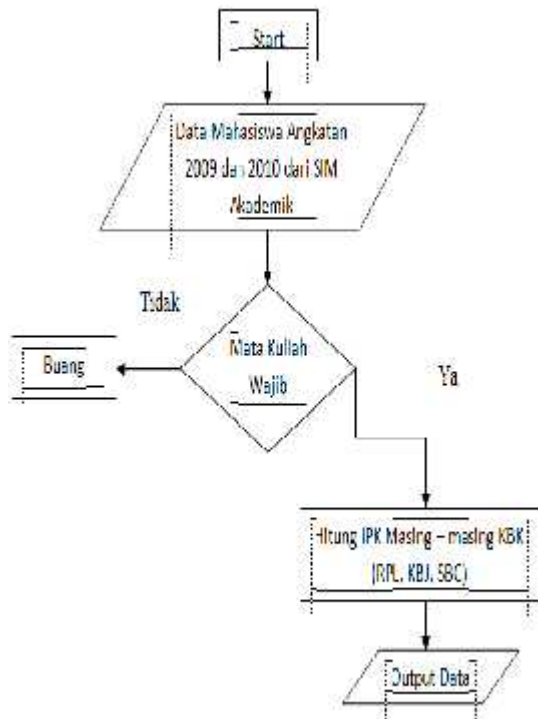
#### 3.4 Flowchart Sistem



Sistem pencarian pola asosiasi bidang bidang minat ini menggunakan algoritma apriori dalam menentukan bidang minat mahasiswa. Langkah-langkah dalam memproses dengan algoritma apriori ini adalah pertama, proses pencarian rule terlebih dahulu yaitu memasukkan data alumni yang terpilih, lalu Algoritma Apriori akan digunakan untuk menganalisis data. Saat itu pula Algoritma Apriori mulai mengerjakan 2

tahapan yaitu proses pembentukan *frequent item* dan proses pembentukan *Association Rule*, kemudian keluarlah *Output Association Rule*.

### 3.5 Perancangan Data



Langkah – langkah perancangan data di atas akan dijelaskan sebagai berikut:

1. Pertama ambil data mahasiswa angkatan 2009 dan 2010 dari SIM akademik, data tersebut tidak bisa langsung dipakai karena hanya memerlukan matakuliah wajib saja yang akan di ambil.
2. Setelah selesai memilih matakuliah wajib lalu langkah selanjutnya menghitung IPK masing – masing KBK (RPL, KBJ, dan SBC).
3. Dan output yang dihasilkan adalah ID Mahasiswa 1910651001 mendapat IPK KBK RPL (3,57), IPK KBK SBC (3,77), dan IPK KBK KBJ (3,65).

#### 4.1. Lingkungan Pengujian

Pada bagian ini akan dijelaskan spesifikasi perangkat yang akan digunakan dalam pengujian aplikasi baik itu perangkat lunak maupun perangkat keras.

Berikut spesifikasi perangkat yang digunakan untuk implementasi dan pengujian aplikasi.

ASUS X450L	
Computer Name	DEKSTOP – 14UOUDH
Operating System	Windows 10 Pro 32 – bit ( 10.0 , Build 10240 )
Language	English ( Regional Setting English )
System Manufacturer	ASUSTeK COMPUTER INC.
System Model	X450LCP
BIOS	BIOS Date: 10/23/13 10:19:13 Ver: 04.06.05
Processor	Intel ® Core™ i5 – 4200U CPU @ 1.60 GHz (4CPUs), ~2.3 GHz
Memory	4096MB RAM
Page File	3051MB Used, 449MB Available
DirectX Version	11.3
Database	MySQL
Program Language	PHP ( <i>Hypertext Preprocessor</i> )

#### 4.2 Uji Coba Penemuan Rule

Uji coba ini dilakukan untuk menemukan rule-rule yang nantinya dapat digunakan untuk sistem rekomendasi bidang minat mahasiswa. Uji coba dilakukan dengan rentang minsupport 1 hingga minsupport 10 dan confidence 0.9 hingga confidence 0.1. Kemudian dari hasil pengamatan akan ditemukan bahwa

semakin kecil confidence, maka akan memunculkan rule semakin banyak dan pilihan rule untuk sistem rekomendasi bidang minat dapat lebih beragam. Hal ini terjadi karena confidence adalah syarat dipilih atau ditolaknya sebuah rule, jika nilai confidence tersebut diturunkan, maka akan memberikan kesempatan lebih banyak rule untuk ditemukan.

1. Uji coba MinSupport : 0,1 dan Confidence : 0.9 dengan parameter IPK. Memunculkan 15 rules terpilih.
2. Uji coba MinSupport : 0.1 dan Confidence : 0.8 dengan parameter IPK. Memunculkan 20 rules terpilih.
3. Uji coba MinSupport : 0.1 dan Confidence : 0.7 dengan parameter IPK. Memunculkan 24 rules terpilih.
4. Uji coba MinSupport : 0.1 dan Confidence : 0.6 dengan parameter IPK. Memunculkan 31 rules terpilih.
5. Uji coba MinSupport : 0.1 dan Confidence : 0.5 dengan parameter IPK. Memunculkan 37 rules terpilih.
6. Uji coba MinSupport : 0.1 dan Confidence : 0.4 dengan parameter IPK. Memunculkan 56 rules terpilih.
7. Uji coba MinSupport : 0.1 dan Confidence : 0.3 dengan parameter IPK. Memunculkan 77 rules terpilih.
8. Uji coba MinSupport : 0.1 dan Confidence : 0.2 dengan parameter IPK. Memunculkan 84 rules terpilih.
9. Uji coba MinSupport : 0.1 dan Confidence : 0.1 dengan parameter IPK. Memunculkan 88 rules terpilih.

### 4.3 Pengujian

#### 1. Antarmuka Menu Utama

Pada gambar 4.1 tampilan antarmuka menu utama menampilkan form upload data pada mahasiswa dari data mentah excel.



#### 2. Antarmuka Upload Data.

Tampilan antarmuka *upload* data berupa data mentah excel sebelum di lakukan *preprocessing* data dan menentukan *min support* dan *confidence*.



#### 3. Antarmuka Menu Data Mahasiswa

Terdapat tampilan hasil *preprocessing* data yang meliputi hasil IPK per KBK dan nilai IPK yang sudah di kategorikan.

No	Nama	IPK	Kategori	IPK	MHS
1	1202010001	80	B	80	1202010001
2	1202010002	75	B	75	1202010002
3	1202010003	70	B	70	1202010003
4	1202010004	65	B	65	1202010004
5	1202010005	60	B	60	1202010005





## 5.2 Saran

Saran - saran untuk pengembangan selanjutnya yaitu, untuk menemukan rule terbaik dari yang terbaik yang nantinya dapat digunakan oleh algoritma atau metode yang lain sebagai bahan pertimbangan sistem pertimbangan rekomendasi mahasiswa.

### DAFTAR PUSTAKA

Cristobal Romero, Sebastian Ventura, Enrique Garcia, “ Data Mining in Course Management Systems : Moodle Study and Tutorial “, 2007

Irina Tudor,” Association Rule Mining as a Data Mining Technique “, 2009  
<http://vokasi.uho.ac.id/statistika/assets/download/15121204230717.%20Jurnal%20Kenendy.pdf>

Kusrini, Emha Taufiq, “ Algoritma Data Mining “, Yogyakarta : ANDI, 2009

<http://digilib.its.ac.id/public/ITS-paper-22013-5105100070-Paper.pdf>

Witten dan Frank, “ Data Mining “ , 2005

[https://www.academia.edu/7847134/Penerapan\\_Metode\\_Association\\_Rule\\_Menggunakan\\_Algoritma\\_Apriori\\_untuk\\_Analisa\\_Pola\\_Data\\_Hasil\\_Tangkapan\\_Ikan](https://www.academia.edu/7847134/Penerapan_Metode_Association_Rule_Menggunakan_Algoritma_Apriori_untuk_Analisa_Pola_Data_Hasil_Tangkapan_Ikan)

Erwin. 2009. Analisis Market Basket dengan Algoritma Apriori dan FP – Growth. Jurnal Generic 26 – 30.

<http://ejournal.its.ac.id/index.php/teknik/article/viewFile/1293/578>

Han, Jiawei, kamber, Micheline. 2006. Data Mining : Concepts and Technique. Morgan Kaufirman : San Francisco.