

Sistem Rekomendasi Pemilihan Mata kuliah Peminatan Menggunakan Algoritma K-means dan Apriori (studi kasus: Jurusan S1 Teknik Informatika Fakultas Informatika)

Riki Nur Afifuddin¹, Dade Nurjanah²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹twensever@students.telkomuniversity.ac.id, ²dadenurjanah@telkomuniversity.ac.id

Abstrak

Setiap mahasiswa Universitas Telkom wajib merencanakan mata kuliah pilihan yang akan diambil mulai dari semester empat. Masih banyak mahasiswa yang mengalami kesulitan untuk memilih mata kuliah pilihan sesuai dengan kelompok keahlian. Oleh karena itu, dibutuhkan sistem rekomendasi yang dapat digunakan untuk melakukan prediksi terhadap pengambilan mata kuliah pilihan yang ada di Universitas Telkom pada jurusan S1 Teknik Informatika. Tugas Akhir ini melakukan studi dan implementasi metoda K-means dan Apriori untuk memberikan rekomendasi pemilihan mata kuliah pilihan sesuai dengan kelompok keahlian. Eksperimen dilakukan dengan menggunakan data 660 Sajarana S1 Teknik Informatika tahun 2013. Hasil eksperimen menunjukkan algoritma K-means dapat mengelompokkan mahasiswa kedalam kelompok keahlian dan Apriori dapat menghasilkan *rule* yang dapat digunakan sebagai rekomendasi pemilihan mata kuliah pilihan.

Kata kunci : rekomendasi, K-means, Apriori.

Abstract

Every Telkom University student is required to plan elective courses to be taken starting from the fourth semester. There are many students who have difficulty choosing elective courses according to the expertise group. Therefore, a recommendation system that can be used to predict the selection of elective subjects at Telkom University in the S1 Informatics Department Faculty of Informatics. The final task is to study and implement the K-means and Apriori methods to provide recommendations for the selection of elective courses according to the expertise group. Experiments were carried out using 660 S1 Informatics Engineering students data in 2013. The experimental results showed that the K-means Algorithm can group students into expertise groups and Apriori can produce rules that can be used as recommendation for the choice of elective subjects.

Keywords: recommendation, K-means, Apriori.

1. Pendahuluan

Latar Belakang

Setiap mahasiswa wajib merencanakan mata kuliah pilihan yang akan diambil mulai dari semester empat. Pengambilan mata kuliah atau yang biasa disebut dengan registrasi di Universitas Telkom dilakukan secara online melalui website Igracias Universitas Telkom.

Saat melakukan registrasi, masih banyak mahasiswa yang mengalami kesulitan untuk memilih mata kuliah pilihan sesuai penjurusannya. Oleh karena itu, perlu adanya sistem rekomendasi untuk membantu mahasiswa menentukan mata kuliah pilihan yang akan diambil sesuai dengan penjurusannya. Sistem rekomendasi yang dibutuhkan adalah sistem yang dapat digunakan untuk melakukan prediksi terhadap pengambilan mata kuliah pilihan yang ada di Universitas Telkom pada jurusan S1 Teknik Informatika.

Telah ada penelitian-penelitian sebelumnya untuk memprediksi pemilihan mata kuliah dengan metode dan tujuan yang berbeda-beda. Kiratijuta Bhumichitr, Songsak Channarukul, Nattachai Saejiem, Rachsuda Jiamthaphaksin dan Kwankamol Nongpong [1], mengimplementasikan algoritma Pearson Correlation Coefficient dan ALS untuk rekomendasi pemilihan *course*, berdasarkan *course* yang telah diambil sebelumnya. Data yang digunakan berisi informasi dari 3614 siswa dengan 52 *course* dalam rentang waktu 8 tahun. Hasil pengujian yang dilakukan menunjukkan algoritma ALS mengasilkan akurasi sebesar 86%, akan tetapi hasil akurasi bagus dikarenakan data yang digunakan hanya pada object yang mirip.

Riset lainya [2] membangun sistem rekomendasi pemilihan mata kuliah dengan menggunakan metode *switching hybrid* dengan *collaborative filtering* dan *content based filtering*, dengan hasil metode *collaborative filtering* memberikan rekomendasi lebih akurat dibandingkan dengan metode *content based filtering*. Dengan akurasi *precision* 20.25% dan *recall* 67.5% lebih tinggi 3.05% *precision*, 10% *recall* dibandingkan dengan metode *content based*.

Berikut penelitian yang melandasi penggunaan algoritma K-means dan algoritma Apriori, “*Consumer Buying Pattern Analysis using Apriori Association Rule*”[3] dan “*Modified Centroid Selection Method of K-means Clustering*”[4]. Pada paper [3] menganalisis data penjualan pada toko dengan tujuan untuk mendapatkan *association rule* yang digunakan untuk mengatur penempatan barang. Hasil penelitian menunjukkan penggunaan algoritma Apriori menghasilkan *association rule* yang dapat meningkatkan omset dan penghasilan toko. Sehingga digunakan algoritma Apriori dalam dunia pendidikan untuk pencarian *association rule* pada data pengambilan matakuliah. Pada paper [4] menjelaskan bahwa pengelompokan data dengan kriteria tertentu dapat dilakukan menggunakan algoritma K-means yang telah dimodifikasi nilai awal *centroid*. Sehingga digunakan algoritma K-means untuk pengelompokan mahasiswa berdasarkan kelompok keahlian.

Topik dan Batasannya

Tugas Akhir ini menggunakan algoritma K-means untuk *clustering* mahasiswa sesuai dengan kelompok keahlian dan algoritma Apriori untuk mencari *association rule* yang ada pada mata kuliah setiap mahasiswa.

Masalah yang akan dibahas pada Tugas Akhir ini adalah:

1. Bagaimana implementasi K-means dan Apriori untuk memprediksi pengambilan mata kuliah pilihan?
2. Bagaimana peformansi algoritma K-means dan Apriori untuk prediksi pengambilan mata kuliah pilihan?

Tugas Akhir ini memiliki batasan dan ruang lingkup penilaian yang mencakup:

1. Rekomendasi mata kuliah peminatan hanya untuk jurusan S1 Teknik Informatika Universitas Telkom
2. Rekomendasi hanya berlaku untuk pemilihan mata kuliah penjurusan untuk semester lima keatas.
3. Maksimal jumlah *rule* yang terbentuk untuk setiap cluster kelompok keahlian sebanyak 20 *rule*.

Tujuan

Berdasarkan masalah yang ada, tujuan yang ingin dicapai dalam tugas akhir ini adalah membangun sistem rekomendasi untuk melakukan prediksi terhadap pengambilan mata kuliah pilihan menggunakan metode K-means *clustering* dan Apriori dengan penentuan *support* dan *confidence* yang sesuai dengan *dataset* yang digunakan.

Organisasi Tulisan

Tugas Akhir ini disusun dengan struktur sebagai berikut. Setelah dijelaskan pendahuluan pada bagian pertama, pada bagian kedua menyantumkan studi terkait, selanjutnya dijelaskan pemodelan sistem pada bagian ketiga, setelah itu, dijelaskan evaluasi terhadap sistem yang dibangun pada bagian ketiga. Pada bagian kelima dijelaskan kesimpulan dan saran untuk penelitian selanjutnya.

2. Studi Terkait

2.1 Clustering

Clustering adalah cara pengelompokan objek sedemikian rupa sehingga objek dalam satu kelompok akan mirip satu sama lain dan berbeda dari objek dalam kelompok lain. *Clustering* dapat dianggap sebagai teknik pengelompokan *unsupervised learning*[5].

2.1.1 Algoritma K-means

K-means merupakan salah satu metode data *clustering* non hirarki yang berusaha mempartisi data yang ada kedalam bentuk satu atau lebih *cluster*. Metode ini mempartisi ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama (*High intra similarity*) dikelompokkan ke dalam satu cluster yang sama dan yang memiliki karakteristik yang berbeda (*Low intra class similarity*) dikelompokkan pada kelompok yang lain[6].

Pada paper[7], membahas teknik K-means untuk pengelompokan tren topik dari data laporan Cina tentang *World No Tobacco Day* WHO, dari tahun 2011 sampai tahun 2015 yang bertujuan untuk mempelajari hubungan antar kebijakan public Cina dalam pengendalian tembakau, dengan hasil algoritma K-means dapat mengekstrasi tren topik dan dapat mengidentifikasi topik baru yang menonjol.

Pada paper[8], mendeteksi pola pengguna layanan telepon seluler dari perusahaan korea dengan menggunakan metode K-means clustering dan *quantitative association rule*, yang menghasilkan delapan *cluster* pengguna, dengan kesimpulan dari *cluster* yang didapat bahwa adanya *marketing implication*.

Fokus pembahasan pada paper[5] yaitu menganalisis algoritma apriori pada dua *cluster* yang dihasilkan oleh algoritma K-means dengan menggunakan sample data dari Moodle, *cluster* pertama menunjukkan *rule* kategori *no* dengan nilai *minimum support* 0.7 dan *minimum confidence* 0.9, pada *cluster* kedua menunjukkan

cluster yes dengan nilai *minimum support* 0.85 dan *minimum confidence* 0.9, kedua *cluster* memiliki 10 *best rule*.

2.2 Association Rule

Association rule adalah teknik data mining yang bertujuan untuk mengekstrak keterkaitan dan pola yang berada pada itemset dalam suatu data transaksi atau data *repository* lainnya[6].

2.2.1 Algoritma Apriori

Apriori adalah salah satu algoritma yang melakukan pencarian *frequent itemset* dengan menggunakan teknik *association rule*[6], algoritma Apriori memiliki dua tahapan utama yaitu *join* dan *prune* dengan menggunakan nilai *minimum support* dan *minimum confidence*[9]. Nilai *support* merupakan presentasi kemunculan suatu transaksi dalam *dataset*, sedangkan *confidence* adalah presentasi banyaknya Y pada transaksi yang mengandung X. formula nilai *support* (1) dan nilai *confidence* (2).

$$support(X) = \frac{\sum transaksi X}{\sum transaksi} \quad (1) \quad confidence(X|Y) = \frac{support X|Y}{occurrence X} \quad (2)$$

Pada paper[10], menerapkan algoritma Apriori untuk penentuan materi yang sesuai dengan kebutuhan mahasiswa Cina Open University, penentuan materi berdasarkan penjurusan yang tersedia, algoritma Apriori menghasilkan dua kelompok *rule* kuat yang dapat digunakan sebagai rekomendasi.

Berikut contoh pencarian *rule* menggunakan nilai *minimum support* dan *minimum confidence* pada transaksi:

Transaksi ID	Item set
1	A,B,C
2	A,C
3	A,D
4	B,E,F

Nilai yang digunakan, untuk *minimum support* = 0.5 dan *minimum confidence* = 0.5

Tahap pertama pencarian nilai *support* setiap item, item yang memiliki nilai *support* kurang dari *minimum support* akan dilakukan *prune* atau pemangkasan. Untuk item yang memiliki nilai *support* di atas *minimum support* dilakukan proses *join*.

C menunjukkan kandidat item, L menunjukkan *Large itemset*.

C1 =

Items	Support
A	3
B	2
C	2
D	1
E	1
F	1

L1=

Items	Support
A	3
B	2
C	2

. C2 =

Items	Support
{A,B}	1
{A,C}	2
{B,C}	1

L2=

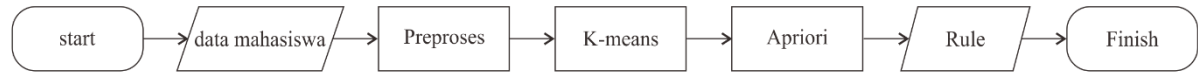
Items	Support
{A,C}	2

Setelah tidak ada lagi proses *join* yang dapat dilakukan pada *Large itemset*, selanjutnya dilakukan perhitungan nilai *confidence*.

Association rule	Support	Support %	Confidence	Confidence %
A → C	2/4	50%	2/3	66%
C → A	2/4	50%	2/2	100%

3. Rancangan Sistem Rekomendasi K-means dan Apriori

Sub-bab ini menjelaskan langkah-langkah sistem dalam memproses data sampai mendapatkan rekomendasi dalam bentuk *rule* untuk setiap kelompok keahlian yang ada di Universitas Telkom. Gambar 1 menunjukkan gambaran umum sistem yang dibangun pada tugas akhir ini.



Gambar 1. Gambaran umum sistem

Dataset yang digunakan berupa nilai mahasiswa program S1 Teknik Informatika tahun 2010-2013 yang sesuai dengan kurikulum 2012 Teknik Informatika, data didapat dari direktorat SISFO Universitas Telkom, dengan rincian NIM, nama mata kuliah, nilai dan SKS. Ukuran data keseluruhan 661x43 dengan baris menyatakan mahasiswa dan kolom menyatakan mata kuliah. Sample data dapat dilihat pada table 1.

Tabel 1. Sample data

NIM	TPB	ML	SOFT_COMP	PCD	SPPK	DAMIN	SWN
sc44f0dc04616fb6	BC		A	A			
s51f159192a543bc	AB				AB	B	
....

Data yang dipakai adalah data yang terdiri dari data matakuliah pilihan saja yang telah diurutkan sesuai dengan kelompok keahlian, karena rekomendasi yang dihasilkan hanya untuk mata kuliah penjurusan. Mata kuliah kelompok keahlian dapat di lihat pada tabel 2.

Tabel 2. Kelompok Keahlian

ICM		SIDE		TELE	
Kode MK	Nama MK	Kode MK	Nama MK	Kode MK	Nama MK
CIG4A3	Pembelajaran Mesin (Machine Learning)	CDG4A3	Sistem Informasi Lanjut (Advanced Information Systems)	CNG4A3	Jaringan Komputer Lanjut (Advanced Computer Network)
CIG4B3	Soft Computing (Soft Computing)	CDG4B3	Perancangan Strategis Sistem Informasi (Information System Strategic Planning)	CNG4B3	Sistem Nirkabel (Wireless Systems)
CIG4C3	Evolutionary Computation (Evolutionary Computation)	CDG4D3	Rekayasa Pengembangan Web (Web Engineering)	CNG4C3	Rekayasa Protokol (Protocol Engineering)
CIG4D3	Information Retrieval (Information Retrieval)	CDG4E3	Data Warehousing (Data Warehousing)	CNG4D3	Sistem dan Kinerja Jaringan (System And Network Performance)
CIG4E3	Pengolahan Citra Digital (Digital Image Processing)	CDG4F3	Teknologi Manajemen Pengetahuan (Teknologi Manajemen Pengetahuan)	CNG4E3	Sensor Jaringan Nirkabel (Wireless Sensor Network)
CIG4F3	Metode Formal (Formal Methods)	CDG4G3	Sistem Pendukung Pengambilan Keputusan	CNG4L3	Forensik Komputer dan Jaringan Network Forensik (Computer and Network Forensic)
CIG4G3	Rekayasa Aplikasi Internet (Web Application And Technology)	CDG4H3	Konsep Pengembangan E-Learning (E-learning Development Concept)	CNG4M3	Sistem Waktu Nyata (Real Time Systems)
CIG4I3	Sistem Rekognisi (Recognition System)	CDG4I3	Audit Sistem Informasi (Information System Audit)	CNG4N3	Sistem Embedded (Embedded System)
CIG4K3	Sistem Berbasis Pengetahuan (Knowledge Based System)	CDG4J3	Penjaminan Mutu Perangkat Lunak (Software Quality Assurance)	CNG4O3	Keamanan Sistem (Systems Security)
CIG4L3	Kriptografi (Cryptography)	CDG4K3	Data Mining (Data Mining)	CNG4P3	Disaster Recovery (Disaster Recovery)
CIG4M3	Kompresi Data (Data Compression)		Teks Mining	CNG4R3	Perancangan Kartu Cerdas (Smart-Card Design)
				CNG4S3	Mikrokontroler : Assembly Language (Mikrokontroler : Assembly Language)

Selanjutnya dilakukan *clustering* menggunakan algoritma K-means untuk mengelompokkan mahasiswa pada kelompok keahlian masing masing. Terdapat tiga *cluster* yang dihasilkan dimana setiap *cluster* mewakili kelompok keahlian yang ada di Fakultas Informatika yaitu *cluster* 1 ICM, *cluster* 2 SIDE dan *cluster* 3 TELE. Algoritma K-means menggunakan perhitungan jarak untuk proses *clustering*. Untuk itu, sebelumnya perlu dilakukan konversi data nilai, dengan skala E=1, D=2, C=3, BC=4, B=5, AB=6 dan A=7, untuk mendapatkan *cluster* sesuai dengan kelompok keahlian dilakukan penempatan nilai awal *centroid* pada masing masing kelompok keahlian. Dalam implementasinya, perhitungan jarak menggunakan rumus *Euclidean* (3):

$$Distance(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (3)$$

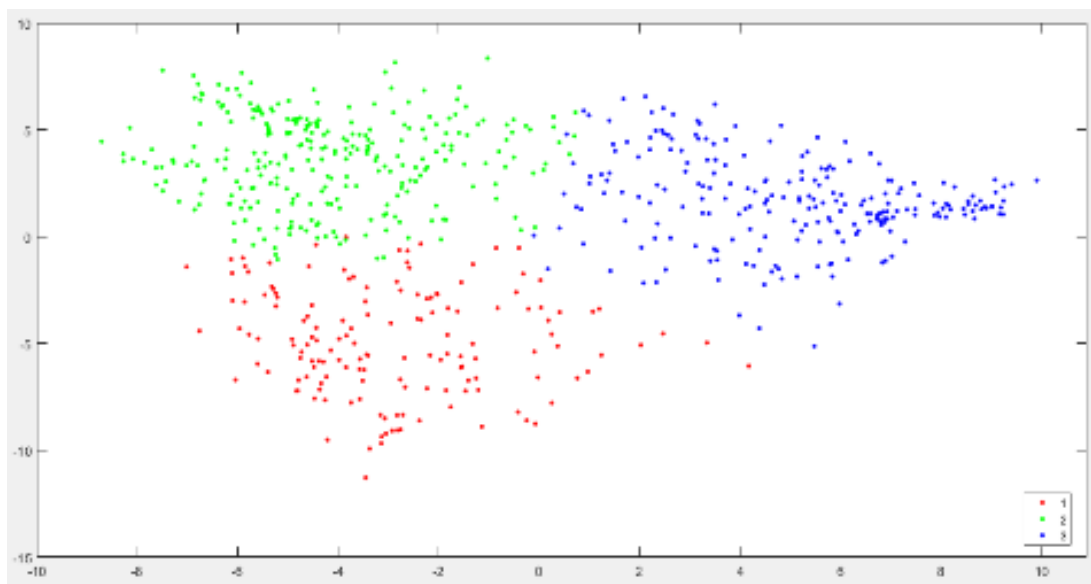
Setelah mendapatkan tiga *cluster*, dilakukan proses pencarian *rule* untuk setiap *cluster*. Sebelum dilakukan pencarian *rule* dengan menggunakan algoritma Apriori, perlu dilakukan konversi data nilai dimana nilai 1, 2 dan kosong menjadi 0, dan 3, 4, 5, 6 dan 7 menjadi 1, dimana nilai 1 menyatakan mahasiswa lulus mata kuliah tersebut dan 0 menyatakan mahasiswa tidak mengambil atau tidak lulus mata kuliah tersebut. Selain dilakukan konversi nilai, perlu juga dilakukan penentuan nilai *minimum support* dan *minimum confidence*. Setelah semua tahapan selesai, akan mendapatkan *rule* untuk setiap kelompok keahlian yang dapat digunakan sebagai rekomendasi pemilihan mata kuliah.

4. Evaluasi

4.1 Hasil Clustering K-means

Langkah pertama pada *clustering* yaitu penentuan banyak *cluster* yang akan dibentuk dari data 660 mahasiswa, berdasarkan kelompok keahlian, *Cluster* yang dibentuk sebanyak tiga *cluster*, *cluster* 1 mewakili kelompok keahlian ICM, *cluster* 2 kelompok keahlian SIDE dan *cluster* 3 kelompok keahlian TELE. Untuk mendapatkan hasil *cluster* sesuai dengan kelompok keahlian mahasiswa, dilakukan penempatan nilai awal *centroid* pada setiap kelompok keahlian.

Dengan menggunakan algoritma K-means yang telah dimodifikasi nilai awal *centroid* menghasilkan *cluster* ICM sebanyak 152 mahasiswa, *cluster* SIDE 279 mahasiswa dan *cluster* TELE 229 mahasiswa. Visualisasi hasil *cluster* dapat di lihat pada Gambar 2. Dengan menggunakan *principal component analysis* (PCA)[9], dimensi mata kuliah dikurangi menjadi dua dimensi sehingga dapat divisualisasikan. PCA dalam hal ini hanya digunakan sebagai pembantu visualisasi tidak digunakan untuk perhitungan *cluster*.



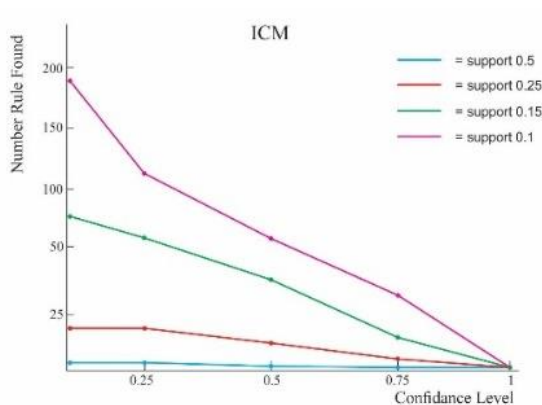
Gambar2. Cluster mahasiswa berdasarkan KK

4.2 Pengujian Nilai Minimum Support dan Minimum Confidence

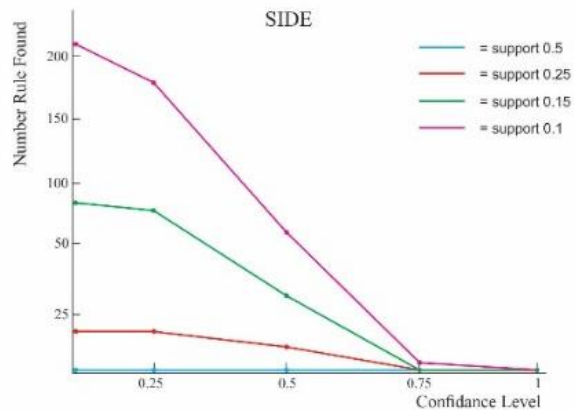
Setelah didapatkan tiga *cluster*, tahap selanjutnya pembentukan *rule* untuk setiap *cluster*. Sebelum algoritma Apriori dapat berjalan perlu ditentukan nilai *minimum support* dan *minimum confidence*, nilai *minimum support* dan *minimum confidence* yang digunakan akan berpengaruh pada hasil akhir *rule*, dikarenakan setiap *dataset* memiliki nilai *support* dan *confidence* yang sesuai berbeda-beda, maka perlu dilakukan pengujian dengan menggunakan kombinasi nilai *support* dan *confidence* yang berbeda-beda. Berikut hasil pengujian nilai *support* dan *confidence*:

Tabel 3. Hasil pengujian nilai minimum support dan minimum confidence

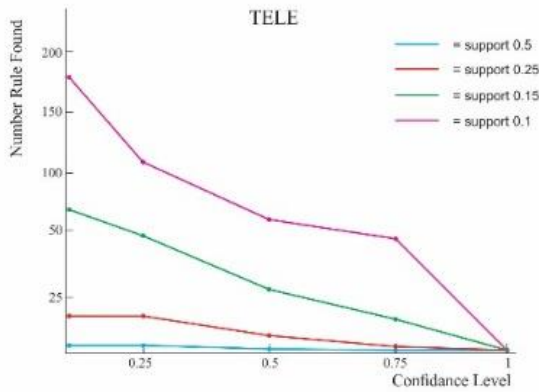
Support 50%	Confidence: 0%	Rule yang dibentuk: SIDE =0; TELE=2; ICM=2;
	Confidence: 25%	Rule yang dibentuk: SIDE =0; TELE=2; ICM=2;
	Confidence: 50%	Rule yang dibentuk: SIDE =0; TELE=2; ICM=2;
	Confidence: 75%	Rule yang dibentuk: SIDE =0; TELE=1; ICM=1;
	Confidence: 100%	Rule yang dibentuk: SIDE =0; TELE=0; ICM=0;
Support 25%	Confidence: 0%	Rule yang dibentuk: SIDE =18; TELE=16; ICM=20;
	Confidence: 25%	Rule yang dibentuk: SIDE =18; TELE=16; ICM=20;
	Confidence: 50%	Rule yang dibentuk: SIDE =12; TELE=9; ICM=14;
	Confidence: 75%	Rule yang dibentuk: SIDE =0; TELE=4; ICM=6;
	Confidence: 100%	Rule yang dibentuk: SIDE =0; TELE=0; ICM=0;
Support 15%	Confidence: 0%	Rule yang dibentuk: SIDE =82; TELE=60; ICM=68;
	Confidence: 25%	Rule yang dibentuk: SIDE =80; TELE=48; ICM=58;
	Confidence: 50%	Rule yang dibentuk: SIDE =34; TELE=28; ICM=35;
	Confidence: 75%	Rule yang dibentuk: SIDE =0; TELE=16; ICM=15;
	Confidence: 100%	Rule yang dibentuk: SIDE =0; TELE=0; ICM=0;
Support 10%	Confidence: 0%	Rule yang dibentuk: SIDE =210; TELE=176; ICM=188;
	Confidence: 25%	Rule yang dibentuk: SIDE =170; TELE=117; ICM=129;
	Confidence: 50%	Rule yang dibentuk: SIDE =65; TELE=77; ICM=74;
	Confidence: 75%	Rule yang dibentuk: SIDE =4; TELE=48; ICM=74;
	Confidence: 100%	Rule yang dibentuk: SIDE =0; TELE=0; ICM=0;



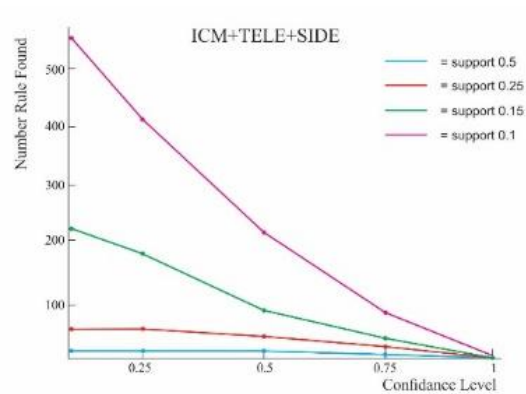
Gambar 3. Jumlah rule cluster ICM



Gambar 4. Jumlah rule cluster SIDE



Gambar 5. Jumlah rule cluster TELE



Gambar 6. Jumlah rule cluster ICM, SIDE, TELE

Pengamatan nilai *support* dan *confidence* pada *dataset* terhadap jumlah *rule* yang terbentuk pada *cluster* ICM, SIDE dan TELE:

- *Minimum Support* 10% dan *minimum confidence* 75%: *rule* yang terbentuk sebanyak 82, dengan detail 4 *rule* SIDE, 48 *rule* TELE dan 30 *rule* ICM. Dikarenakan *rule* yang terbentuk tidak seimbang dan terlalu banyak *rule* yang terbentuk sehingga nilai *minimum support* 10% dan *minimum confidence* 75% tidak dapat digunakan.
- *Minimum support* 10% dan *minimum confidence* 50%: *rule* yang terbentuk sebanyak 216, jumlah tersebut terlalu banyak untuk diamati sehingga nilai *minimum support* 10% dan *minimum confidence* 50% tidak dapat digunakan.
- *Minimum support* 15% dan *minimum confidence* 75%: *rule* yang terbentuk sebanyak 31, dengan detail 0 *rule* SIDE, 16 *rule* TELE dan 15 *rule* ICM, dikarenakan tidak ada *rule* yang terbentuk pada *cluster* SIDE, nilai *minimum support* 15% dan *minimum confidence* 75% tidak dapat digunakan.
- *Minimum support* 15% dan *minimum confidence* 50%: *rule* yang terbentuk sebanyak 97, terlalu banyak *rule* untuk diamati. Sehingga nilai *minimum support* 15% dan *minimum confidence* 50% tidak dapat digunakan.
- *Minimum support* 25% dan *minimum confidence* 75%: *rule* yang terbentuk sebanyak 10, dengan detail 0 *rule* SIDE, 4 *rule* TELE dan 6 *rule* ICM, nilai tersebut tidak dapat digunakan karena tidak ada *rule* yang terbentuk pada *cluster* SIDE.
- *Minimum support* 25% dan *minimum confidence* 50%: *rule* yang terbentuk sebanyak 35, dengan detail 12 *rule* SIDE, 9 *rule* TELE dan 14 *rule* ICM, penggunaan nilai tersebut dapat menghasilkan *rule* pada setiap *cluster* dengan tingkat *confidence* diatas 50%.
- *Minimum support* 50% dan *minimum confidence* 75%: hanya mengidentifikasi beberapa *rule* dengan tingkat *confidence* yang rendah, berarti tidak ada asosiasi yang sering dari data. Dengan kesimpulan nilai *minimum support* 0.5 tidak dapat dipilih, *rule* yang dihasilkan tidak dapat mewakili data.

Berdasarkan pengamatan diatas, nilai *minimum support* yang digunakan 0.25 dan *minum confidence* 0.5, setelah didapat nilai *minimum support* dan *minimum confidence* dilakukan proses pencarian *rule* setiap *cluster*. Hasil *rule* setiap *cluster* kelompok keahlian dapat dilihat pada tabel 2.

Tabel 2. Hasil rule setiap cluster kelompok keahlian

<p>Minimum support: 0.25 Minimum confidence: 0.5</p>	<p>Hasil cluster 1 ICM Rule:</p> <ol style="list-style-type: none"> 1. SISTEM_REKOGNISI (ICM)→ PCD (ICM) 2. RAI (ICM),SISTEM_REKOGNISI (ICM)→ PCD (ICM) 3. DAMIN (SIDE)→ PCD (ICM) 4. RAI (ICM)→ PCD (ICM) 5. SPPK (SIDE)→ PCD (ICM) 6. SOFT_COMPUTING (ICM)→ PCD (IMC) 7. PCD (ICM)→ RAI (ICM) 8. SISTEM_REKOGNISI (ICM)→ RAI (ICM) 9. PCD (ICM), SISTEM REKOGNISI (ICM)→ RAI (ICM) 10. SOFT_COMPUTING (ICM)→ RAI (ICM) 11. SISTEM_REKOGNISI (ICM)→ PCD,RAI (ICM) 12. PCD (ICM)→ SISTEM_REKOGNISI (ICM) 13. PCD (ICM),RAI (ICM)→ SISTEM REKOGNISI (ICM) 14. PCD (ICM)→ SOFT COMPUTING (ICM)
<p>Minimum support: 0.25 Minimum confidence: 0.5</p>	<p>Hasil cluster 2 SIDE Rule:</p> <ol style="list-style-type: none"> 1. KPE_LEARNING (SIDE)→ SPPK (SIDE) 2. DAMIN (SIDE)→ AUDIT_SISTEM (SIDE) 3. PMPL (SIDE)→ SPPK (SIDE) 4. PSSI (SIDE)→ SPPK (SIDE) 5. AUDIT_SISTEM (SIDE)→ SPPK (SIDE) 6. DAMIN (SIDE)→ SPPK (SIDE) 7. PMPL (SIDE)→ AUDIT_SISTEM (SIDE) 8. SPPK (SIDE)→ AUDIT_SISTEM (SIDE) 9. KPE_LEARNING (SIDE)→ AUDIT SISTEM (SIDE) 10. AUDIT_SISTEM (SIDE)→ DAMIN (SIDE) 11. KPE_LEARNING (SIDE)→ PMPL (SIDE) 12. SPPK (SIDE)→ PMPL (SIDE)
<p>Minimum support: 0.25 Minimum confidence: 0.5</p>	<p>Hasil cluster 3 TELE Rule:</p> <ol style="list-style-type: none"> 1. SISTEM_NIRKABEL (TELE),FORENSIK (TELE)→ JARKOMLAN (TELE) 2. FORENSIK (TELE)→ JARKOMLAN (TELE) 3. SISTEM_NIRKABEL (TELE)→ JARKOMLAN (TELE) 4. SENSOR_NIRKABEL (TELE)→ JARKOMLAN (TELE) 5. JARKOMLAN (TELE)→ SISTEM_NIRKABEL (TELE) 6. JARKOMLAN (TELE), FORENSIK (TELE)→ SISTEM NIRKABEL (TELE) 7. FORENSIK (TELE)→ SISTEM_NIRKABEL (TELE) 8. SENSOR_NIRKABEL (TELE)→ SISTEM NIRKABEL (TELE) 9. FORENSIK (TELE)→ JARKOMLAN (TELE),SISTEM_NIRKABEL (TELE)

4.2 Analisis Hasil Pengujian

Berdasarkan hasil pengujian menggunakan data 660 Sarjana S1 Teknik Informatika, penempatan nilai awal *centroid* pada algoritma K-means dapat menghasilkan *cluster* sesuai dengan kelompok keahlian mahasiswa, dapat dibuktikan dengan hasil *rule* yang terbentuk pada proses Apriori dimana setiap *cluster* menghasilkan *rule* sesuai dengan mata kuliah kelompok keahlian. Dapat dilihat pada table 2, pada *cluster* ICM terdapat 14 rule dengan 1 *rule* lintas kelompok keahlian, maka didapat 92,86% *rule* dengan mata kuliah ICM, pada *cluster* SIDE terdapat 12 *rule* yang terbentuk dimana semua *rule* merupakan kelompok keahlian SIDE, maka didapat 100% *rule* dengan mata kuliah SIDE dan pada *cluster* TELE terbentuk 9 rule dimana semua *rule* merupakan kelompok keahlian TELE, maka didapat 100% *rule* dengan mata kuliah TELE.

Selain itu, penentuan *nilai minimum support* dan *minimum confidence* pada algoritma Apriori sangat berpengaruh pada hasil akhir *rule* yang terbentuk dikarenakan setiap *dataset* memiliki karakter yang berbeda-beda, dengan pemilihan *nilai minimum support* dan *minimum confidence* yang tepat dapat menghasilkan *rule* yang sesuai dengan kebutuhan rekomendasi.

5. Kesimpulan

Berdasarkan hasil penelitian dan analisis yang dilakukan, maka ditarik kesimpulan sebagai berikut:

1. Implementasi algoritma K-means dan Apriori untuk rekomendasi mata kuliah pilihan diperlukan beberapa langkah, pada algoritma K-means perlu dilakukan penempatan nilai awal *centroid* pada data setiap kelompok keahlian agar K-means dapat mengelompokkan mahasiswa sesuai dengan kelompok keahlian. Kemudian pada Apriori perlu dilakukan pengamatan nilai support dan confidence yang sesuai dengan data, sehingga menghasilkan *rule* yang dapat digunakan sebagai rekomendasi pemilihan mata kuliah penjurusan.
2. Dengan penggunaan nilai *minimum support* 0.25 dan *minimum confidence* 0.5 didapatkan 35 *rule* yang dapat digunakan sebagai rekomendasi pemilihan mata kuliah pilihan, dengan penggunaan nilai tersebut setiap *rule* menggambarkan pola pengambilan mata kuliah tahun 2010 – 2013 sesuai dengan kelompok keahlian, dari 35 *rule* terdapat 1 *rule* yang memiliki rekomendasi lintas kelompok keahlian.

Daftar Pustaka

- [1] Bhumichitr, K., Channarukul, S., Saejiem, N., Jiamthaphaksin, R., & Nongpong, K. 2017. Recommender Systems for university elective course recommendation. In Computer Science and Software Engineering (JCSSE), 2017 14th International Joint Conference on (pp. 1-5). IEEE.
- [2] Khasanah, I. N. Analisis dan Implementasi Switching Hybrid Recommender System Menggunakan Support Vector Machines Classifier dan Collaborative Filtering Switching Hybrid Recommender System Analysis and Implementation with Support Vector Machines Classifier and Collaborative Filtering. 2012.
- [3] Kumar, V. S., Renganathan, R., VijayaBanu, C., & Ramya, I. Consumer Buying Pattern Analysis using Apriori Association Rule. 2018. International Journal of Pure and Applied Mathematics.
- [4] Mawati, R., Sumertajaya, I. M., & Afendi, F. M. Modified Centroid Selection Method of K-means Clustering. 2014. IOSR Journal of Mathematics (IOSR-JM) on (pp. 49-53).
- [5] Aher, S. B., & Lobo, L. M. R. J. 2012. Prediction of Course Selection by Student using Combination of Data Mining Algorithms in E-learning. International Journal of Computer Applications, 40(15), 1-7.
- [6] P.-N. Tan, M. Steinbach, and V. Kumar. 2005. Introduction to Data Mining, (First Edition). Boston, MA. USA Addison-Wesley Longman Publishing Co.Inc.
- [7] Tao, H., Li, J., Luo, T., & Wang, C. 2017. Research on topics trends based on weighted K-means. In Electronics Information and Emergency Communication (ICEIEC) on (pp. 457-460). IEEE.
- [8] Sohn, S. Y., & Kim, Y. 2008. Searching customer patterns of mobile service using clustering and quantitative association rule. Expert systems with Applications Expert Systems with Applications: An International Journal.
- [9] Patil, S. D., Deshmukh, R. R., & Kirange, D. K. 2016. Adaptive Apriori Algorithm for frequent itemset mining. In System Modeling & Advancement in Research Trends (SMART), International Conference (pp. 7-13). IEEE.
- [10] Liu, F., Zhang, S., Ge, J., Lu, F., & Zou, J. 2016. Agricultural Major Courses Recommendation Using Apriori Algorithm Applied in China Open University System. In Computational Intelligence and Design (ISCID), 9th International Symposium on (Vol. 1, pp. 442-446). IEEE.
- [11] O'Mahony, M. P., & Smyth, B. 2007. A recommender system for online course enrolment: an initial study. In Proceedings of the 2007 ACM conference on Recommender systems (pp. 133-136). ACM.