

ANALISA DATA MINING KEMAMPUAN LULUSAN DENGAN KEBUTUHAN STAKEHOLDER MENGGUNAKAN ALGORITMA APRIORI (STUDI KASUS UNIVERSITAS SEMARANG)

Henny Indriyawati¹, Titin Winarti²

^{1,2}Program Studi Sistem Informasi, Fakultas teknologi Informasi dan Komunikasi, Universitas Semarang
e-mail: ¹henny@usm.ac.id, ²titin@usm.ac.id

ABSTRAK

Tracer Study saat ini telah menjadi salah satu aspek yang dibutuhkan dalam akreditasi sebuah Perguruan Tinggi ataupun Jurusan, bahkan *tracer study* juga sangat dibutuhkan oleh Perguruan Tinggi dalam mengevaluasi keberhasilan pembelajaran yang telah diterapkan di kampus. Data *tracer study* juga bisa dimanfaatkan kampus untuk mengambil keputusan dan sebagai masukan dalam pengembangan kurikulum atau sarana penunjang akademik lainnya. Data mining digunakan untuk mengekstraksi informasi pada data yang berskala besar. Dengan algoritma apriori, maka akan muncul pola hubungan antar atribut yang sebelumnya telah diisi oleh para alumni ketika *tracer study*. Hasil data mining yang berupa informasi pola hubungan kemampuan alumni yang kebutuhan stakeholder dengan melihat antar atribut pada *tracer study* yang telah dilakukan. Tujuannya adalah mempermudah Perguruan Tinggi dalam menyusun strategi agar lulusan yang dihasilkan sesuai kebutuhan di dunia usaha.

Kata Kunci : *tacer study, data_mining, apriori, lulusan, stakeholder*

1. PENDAHULUAN

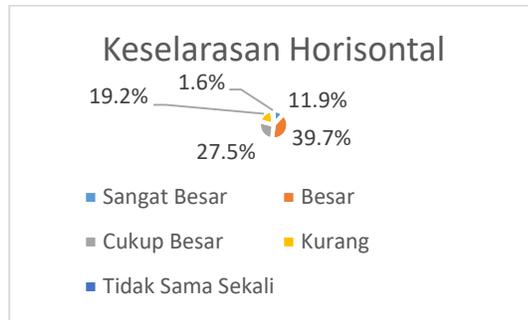
Perguruan tinggi sebagai salah satu wadah pendidikan nasional memiliki peranan penting dalam melahirkan generasi cerdas dalam membangun bangsa. Fungsi tersebut menuntut perguruan tinggi untuk mampu berperan dalam penyelenggaraan dan pengembangan pendidikan serta pemeliharaan, pembinaan dan pengembangan ilmu pengetahuan, teknologi dan/atau kesenian.

Universitas Semarang (USM) sebagai salah satu perguruan tinggi terbaik di Jawa Tengah selalu berusaha meningkatkan kualitas pendidikan, salah satunya dengan upaya menciptakan lulusan yang berkualitas dan mampu bersaing di dunia nyata. Banyaknya mahasiswa membuat USM mempunyai tanggung jawab yang besar terhadap pendidikan mahasiswa sehingga kelak menjadi lulusan yang siap kerja dan sesuai dengan kebutuhan dunia usaha atau industri. Hal ini membuat USM harus memperhatikan kualitas lulusan, salah satu upaya yang bisa dilakukan dengan cara mengumpulkan data dan melakukan evaluasi terhadap lulusan yang dihasilkan. Data yang tersimpan dan sudah tidak digunakan lagi dengan jumlah data yang dikumpulkan dan disimpan dalam basis data berukuran besar, dibutuhkan sebuah metode atau teknik yang dapat merubah tumpukan data tersebut menjadi sebuah informasi yang berharga yang bermanfaat untuk pengambilan keputusan bisnis dengan menggunakan perhitungan data dengan algoritma apriori.[1]

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Pemilihan metode yang tepat tergantung pada tujuan dan proses KDD (Knowledge Discovery in Database) secara keseluruhan. Data mining mampu menganalisis data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan.[2] Data yang dihimpun adalah data alumni dengan cara mengajukan beberapa pertanyaan yang berhubungan dengan pengetahuan dan kemampuan yang didapat selama kuliah serta kesesuaian dengan bidang pekerjaan dengan dunia industry atau perusahaan. Parameter jenis kemampuan yang dipakai untuk membentuk pola kemampuan lulusan terhadap kebutuhan dunia usaha dan industry berdasarkan perspektif lulusan.

Pada penelitian sebelumnya yang berjudul “Penerapan Data Mining menggunakan Association Rule Untuk Mendukung Strategi Promosi Universitas Nusantara PGRI Kediri”. Penelitian ini mengimplementasikan metode Association Rule Apriori dalam menganalisis hubungan daerah dan minat mahasiswa terhadap suatu prodi pada Universitas Nusantara PGRI Kediri, hasil dari penelitian adalah memberikan rekomendasi wilayah promosi.[3].

Berdasarkan data *tracer* USM tahun 2019 pada gambar 1, menunjukkan keselarasan horizontal yaitu keselarasan seberapa erat hubungan antara bidang studi dengan pekerjaan alumni, tampak bahwa masih ada ketidaksesuaian (tidak sama sekali= 1,6%, kurang 19,2%, dan cukup besar 27,5%) kemampuan lulusan dengan *stakeholder*. Hal ini menjadi perhatian khusus perguruan tinggi untuk membenahi/ mengatur strategi agar prosentase data tersebut berkurang.



Gambar 1 Keselarasan Horisontal USM

Berdasarkan permasalahan diatas maka peneliti bermaksud melakukan penelitian tentang kemampuan metode association rule berbasis algoritma apriori dalam permodelan pola kelayakan lulusan dengan kebutuhan stakeholder.

2. TINJAUAN PUSTAKA

2.1 Data Mining

Data mining mampu menganalisis data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pendukung keputusan. Data mining adalah proses yang memperkerjakan satu atau lebih teknik pembelajaran komputer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis. [4]. Ada 3 jenis data set yaitu record(matriks data, data transaksi,data dokumen), graph (www, struktur molekul) dan ordered data set (data spasial, data temporal, data sekuensial, data urutan genetic) [5]

2.2 Algoritma Apriori

Algoritma apriori adalah metode dengan aturan yang menyatakan hubungan antara beberapa atribut atau sering disebut analisis afinitas atau analisis keranjang pasar. Cara kerja algoritma apriori adalah menghasilkan kandidat baru dari k-itemset pada frekuensi itemset langkah sebelumnya. Langkah akan berhenti ketika tidak ada frekuensi item set baru yang dihasilkan. Setelah itemset ditentukan maka selanjutnya adalah menghitung minconf dengan rumus yang sudah ditentukan[6]. Metode dalam analisis Apriori terbagi menjadi 2 tahap yaitu:

- a. Analisis pola frekwensi tertinggi, tahap ini mencari kombinasi item yang memenuhi syarat minimum dari nilai support dalam data base. Nilai support sebuah item diperoleh dengan rumus berikut

$$\text{Support}(A)=\frac{\text{Jumlah Transaksi mengandung } A}{\text{Total transaksi}} \times 100\% \tag{1}$$

Nilai dukungan 2 item diperoleh dari rumus

$$\text{Support}(A,B)=\frac{\text{Jumlah Transaksi mengandung } A \text{ dan } B}{\text{Total transaksi}} \times 100\% \tag{2}$$

- b. Pembentukan aturan assosiatif

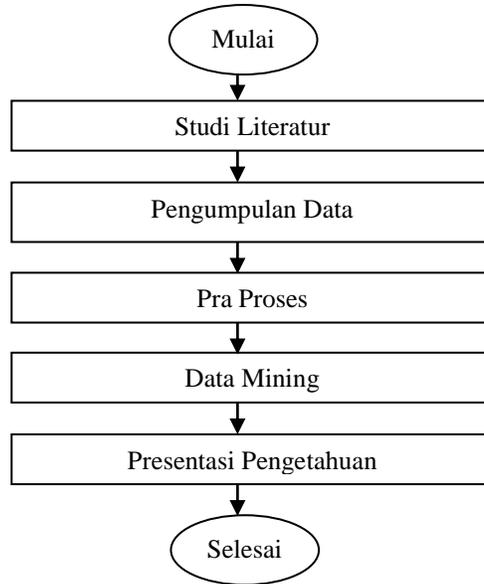
$$\text{Confident} = P(A|B)=\frac{\text{Jumlah Transaksi mengandung } A \text{ dan } B}{\text{Total transaksi } A} \times 100\% \tag{3}$$

2.3 Assosiation Rule Mining

Aturan asosisasi adalah teknik data mining untuk menemukan aturan assosiasi antar suatu kombinasi item [8].

3. METODE PENELITIAN

Metodologi penelitian merupakan langkah-langkah yang harus dilakukan agar penelitian.Penelitian merupakan rangkaian kegiatan ilmiah[7].



Gambar 2 Metode Penelitian

Tahapan proses penelitian pada gambar 2 dapat dijelaskan sebagai berikut:

- a. Studi literatur, mengumpulkan teori-teori yang berfungsi sebagai panduan menyelesaikan masalah berhubungan dengan data mining.
- b. Mengumpulkan data, pengumpulan data dilakukan dengan observasi dokumen dari laporan yang ada tracer study pada tahun 2021.
- c. Pra Proses, Tahap pre-processing merupakan tahap mentransformasikan data mentah ke format yang sesuai untuk analisis, meliputi data cleaning, data integrasi, data transformasi, data reduksi.
- d. Data Mining, data yang sudah diperoleh diolah dengan proses data mining. Pengolahan data ini menggunakan algoritma apriori. Hasil dari proses ini adalah data dikelompokkan berdasarkan kemiripan karakteristik setiap data sehingga di dapatkan informasi yang tersembunyi.
- e. Presentasi Pengetahuan, dari hasil olah data sebelumnya dilakukan knowledge presentation. Knowledge presentation yaitu menyajikan pola atau informasi yang ditemukan yang sesuai fakta atau hipotesa yang ada sebelumnya. Pola ini nantinya adalah pola yang berguna sebagai output penelitian ini.

4. HASIL DAN PEMBAHASAN

Pada tahap ini, analisa data dan membuat kelompok data sesuai dengan variable yang dibutuhkan. Tabel 1 menunjukkan data penelitian tracer.

Tabel 1 Kelompok/ Kategori Item

Kode	Kategori
A	Etika
B	Keahlian Bidang Ilmu
C	Kemampuan Bahasa Asing
D	Penggunaan Teknologi Informasi
E	Kemampuan Berkomunikasi
F	Kerja Sama
G	Pengembangan Diri

Tahapan selanjutnya adalah seleksi data yang memiliki atribut yang relevan dengan melakukan seleksi data.

Gambar 3 Data Kombinasi

Dataset terdiri dari 1316 data yang diolah menggunakan Microsoft Exel, Nilai 1 mempunyai arti alumni memilih variabel tersebut, nilai 0 mempunyai arti mahasiswa tidak memilih variabel tersebut.

Nilai minimum sport adalah 50%, dari hasil pengolahan data di peroleh beberapa kombinasi itemset yaitu: Kombinasi 2 itemset

1. Kombinasi 2 itemset

Kombinasi 2 itemset dipilih yang mempunyai nilai itemset 50% keatas, sehingga di dapat kombinasi yang tampak pada tabel 2

Tabel 2 *Minimum support 2 itemset*

Item	Σ	Nilai Support
A4,B4	638	48%
A4,C4	635	48%
A4,D4	571	43%
A4,E4	670	51%
A4,G4	647	49%
B4,C4	660	50%
B4,D4	567	43%
B4,E4	600	46%
B4,G4	579	44%
C4,D4	508	39%
D4,E4	578	44%
E4,C4	595	45%
D4,G4	559	42%
(E4,G4)	710	54%
G4,C4	579	44%

2. Kombinasi 3 itemset

Kombinasi 3 itemset dipilih yang mempunyai nilai itemset 50% keatas, sehingga di dapat 2 kombinasi yang tampak pada tabel 3

Tabel 3 *Minimum support 3 itemset*

Item	Σ	Nilai Support
(A4,B4,C4)	576	44%
(A4,B4,E4)	594	45%
(A4,B4,G4)	576	44%
(A4,C4,E4)	589	45%
(A4,C4,G4)	576	44%
(E4,A4,G4)	715	54%
(B4,C4,E4)	542	41%
(B4,C4,G4)	526	40%
(B4,E4,G4)	579	44%
(C4,E4,G4)	670	51%

3. Kombinasi 4 itemset

Kombinasi 4 itemset dipilih yang mempunyai nilai itemset 50% keatas, sehingga di dapat 1 kombinasi yang tampak pada tabel 4

Tabel 4 *Minimum support 4 itemset*

Item	Σ	Nilai Support
(A4,B4,C4,E4)	536	41%
(A4,B4,C4,G4)	523	40%
(A4,B4,E4,G4)	576	44%
(C4,E4,A4,G4)	668	51%
(B4,C4,E4,G4)	526	40%

4.2 Association Rule

Tahap association rule mining merupakan teknik data mining untuk menemukan aturan asosiatif antara suatu kombinasi item. Aturan asosiatif antara suatu kombinasi item akan diperoleh dari etika, keahlian bidang ilmu, kemampuan bahasa asing, penggunaan teknologi informasi, kemampuan berkomunikasi, kerja sama, pengembangan diri. Setelah polah frekuensi tinggi ditemukan, barulah dicari aturan asosiasi yang memenuhi syarat minimum untuk confidence dengan menghitung confidence aturan asosiatif $A \rightarrow B$. Minimum Confidence = 90%.

Tabel 5 *Minimum confidence*

Item	Σ	Nilai Support	Confidance
(E4,G4)	710	0,539513678	100%
(E4,A4,G4)	715	0,54331307	100%
(C4,E4,G4)	670	0,509118541	100%
(C4,E4,A4,G4)	670	0,507598784	100%

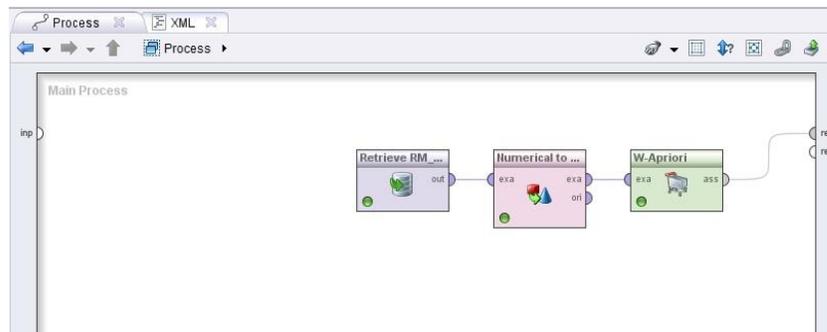
4.3 Algoritma Apriori pada Association Rule Mining

Kaidah asosiasi apriori digunakan untuk menggambarkan hubungan antar item pada tabel data transaksional ataupun data relasional. Association Rule Mining terlihat pada tabel 5.8

Tabel 6 Aturan asosiasi

Item	Σ	Nilai Support	Confidance
Jika Etika Baik maka Pengembangan Diri Baik	656	0,498480243	92%
Jika Kemampuan Berkomunikasi Baik dan Etika Baik maka Pengembangan diri Baik	670	0,509118541	94%
Jika Kemampuan Bahasa Asing Baik dan Kemampuan Berkomunikasi Baik dan Kemampuan Berkomunikasi Baik maka Pengembangan diri Baik	660	0,501519757	99%
Jika Kemampuan Bahasa Asing dan Kemampuan Berkomunikasi Baik dan Etika Baik maka Pengembangan diri Baik	656	0,498480243	98%

Dataset yang terdiri dari 1316, selanjutnya dilakukan pengujian menggunakan Rapid Miner, hal pertama yang perlu didesain yaitu terdiri dari import *database* dalam bentuk exel dalam Retrieve, kemudian dihubungkan dengan Operator Numerical to Binominal yang berfungsi untuk mengubah nilai atribut yang berada pada table menjadi binominal. Operator ketiga adalah W-Apriori, operator ini berfungsi untuk perhitungan algoritma apriori. Desain dari ketiga operator ini dapat dilihat pada gambar 4



Gambar 4 Desain Rapid Miner

Setelah diproses menggunakan Rapid Miner menghasilkan rule-rule, yaitu Jika Etika Baik maka Pengembangan Diri Baik dengan nilai *confidence* 92%, Jika Kemampuan Berkomunikasi Baik dan Etika Baik maka Pengembangan diri Baik dengan nilai *confidence* 94%, Jika Kemampuan Bahasa Asing Baik dan Kemampuan Berkomunikasi Baik dan Kemampuan Berkomunikasi Baik maka Pengembangan diri Baik dengan nilai *confidence* 99%, Jika Kemampuan Bahasa Asing dan Kemampuan Berkomunikasi Baik dan Etika Baik maka Pengembangan diri Baik dengan nilai *confidence* 98% yang ditunjukkan pada gambar 5

W-Apriori

Apriori

```

Minimum support: 0.45 (592 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 11

Generated sets of large itemsets:

Size of set of large itemsets L(1): 7
Size of set of large itemsets L(2): 1
Size of set of large itemsets L(3): 2
Size of set of large itemsets L(4): 1

Best rules found:

1. E4=false 658 ==> G4=false 658   conf:(92)
2. E4=false A4=false 605 ==> G4=false 658   conf:(94)
3. C4=false E4=false 605 ==> G4=false 658   conf:(99)
4. C4=false E4=false A4=false 605 ==> G4=false 658   conf:(98)
    
```

Gambar 5 Assosiasi Rule

5. KESIMPULAN

Dari hasil perhitungan data mining menggunakan algoritma apriori, data tracer dengan batasan *minimum support* 50% dan *minimum confidence* 90%, membentuk 4 *rules*. Salah satu *rules* yang terbentuk adalah Jika Kemampuan Bahasa Asing Baik dan Kemampuan Berkomunikasi Baik dan Kemampuan Berkomunikasi Baik maka Pengembangan diri Baik dengan nilai *support* 50% dan nilai *confidence* 98%, sehingga informasi ini dapat memberi rekomendasi kepada pihak universitas untuk mengatur pengetahuan yang diberikan mahasiswa.

DAFTAR PUSTAKA

- [1] Sari, Dewi Puspita. (2015). Data Mining Perkiraan Produksi Spanduk dengan Algoritma Apriori (Studi Kasus: CV. Mentati Persada Medan): 33-41.
- [2] Sunjaya. 2010. Aplikasi Mining Data Mahasiswa dengan Metode Klasifikasi DecisionTree. Seminar Nasional Aplikasi Teknologi Informasi 2010. Yogyakarta.
- [3] Risky, H. &. (2017). Penerapan Data Mining menggunakan Association Rule Untuk Mendukung Strategi Promosi Universitas Nusantara PGRI Kediri. Seminar Nasional Teknologi Informasi, Komunikasi dan Aplikasinya Volume 04, Tahun 2017 (pp. 138-142). kediri: SEKOLAH TINGGI INFORMATIKA DAN KOMPUTER INDONESIA (STIKI) MALANG.
- [4] Hermawati, Fajar Astuti.(2013)Data Mining. Yogyakarta: Andi
- [5]. Prasetyo, Eko. 2014. DATA MINING Mengolah Data Menjadi Informasi menggunakan Matlab. Yogyakarta: CV. Andi Offset.
- [6] P. Bagus, I. Sukadiana, N. Putu, S. Merta, and S. Aryani, "Analysis of Apriori Algorithm on Sales Transactions to Arrange Placement of Goods on Minimarket," Int. J. Eng. Emerg. Technol., vol. 3, no. 1, pp. 13–17, 2018.
- [7] J. Han and M. Kamber, *Data Mining Concepts and Techniques*. San Diego: Morgan Kaufmann, 2001.
- [8] Ardani, N.R., & Fitriana, N. "Sistem Rekomendasi Pemesanan Sparepart Dengan Algoritma FP-Growth". Seminar Nasional Aplikasi Teknologi Informasi Dan Multimedia, 6–7. 2016.