

PENGEMBANGAN APLIKASI DATA MINING DENGAN ALGORITMA C4.5 DAN APRIORI DI FAKULTAS TEKNOLOGI INFORMATIKA UNIVERSITAS RESPATI INDONESIA

Yasmianti¹, Wahyudi², Andi Susilo³

^{1,2,3}Program Studi Teknik Informatika, Universitas Respati Indonesia, Jakarta

Jl. Bambu Apus I No. 3 Cipayung 13890,

*Email : yasmianti@fti.urindo.ac.id

Diterima: 4 Oktober 2016

Direvisi: 3 November 2016

Disetujui: 10 Desember 2016

ABSTRAK

Kemampuan dosen yang baik akan berdampak positif terhadap kualitas layanan terhadap mahasiswa. Salah satu cara untuk mengetahui, memperbaiki, dan menjaga mutu dosen adalah dengan mengevaluasi kinerja dosen. Evaluasi atau penilaian kinerja harus dilakukan karena memiliki pengaruh yang kuat terhadap bentuk dan sifat hubungan kerja dalam suatu lembaga pendidikan. Data nilai mahasiswa dan data hasil evaluasi belum pernah dianalisis lebih mendalam sehingga data tersebut digunakan dengan penerapan menggunakan algoritma C4.5 dan Apriori dari metode Data Mining. Algoritma C4.5 untuk melakukan klasifikasi dengan hasil pembentukan pohon keputusan (decision tree). Sedangkan Algoritma Apriori digunakan dengan cara pengkombinasian Item dengan Item yang lainnya dan menghasilkan nilai support dan confidence.

Kata kunci: Kualitas, layanan, mahasiswa, algoritmac 4.5, data mining

ABSTRACT

The ability of the good lecturer would have a positive impact on the quality of services to students. One way to find out, to improve, and to maintain the capability of lecturers is to evaluate the performance of lecturers. The performance's evaluation or assessment should be done because it has a strong influence on the shape and nature of the employment relationship in an educational institution. The students' evaluation data has not been analyzed deeply so that the data is used by the application of algorithm C4.5 and Apriori of the methods of Data Mining. C4.5 algorithms is use to classify the results of the formation by using Decision tree. While Apriori Algorithm used in a manner combining the item with other items and generate support and confidence values.

Keywords: Quality, service, student, algoritmac 4.5, data mining

PENDAHULUAN

Kesadaran pentingnya pendidikan di Indonesia dari tahun ke tahun semakin tinggi. Hal tersebut ditandai dengan naiknya jumlah pendaftar perguruan tinggi baik pendaftar perguruan tinggi negeri maupun swasta di

Indonesia seperti yang dilansir oleh UNESCO tahun 2013.

Kecenderungan pendaftar perguruan tinggi yang meningkat di Indonesia harus disikapi dengan benar. Institusi perguruan tinggi harus menyiapkan suatu sistem

pendidikan yang berkualitas tinggi agar dapat membentuk sumber daya manusia yang berkualitas. Salah satu elemen sistem pendidikan perguruan tinggi adalah dosen. Dosen merupakan ujung tombak dan motor penggerak pada perguruan tinggi untuk dapat menghasilkan mahasiswa-mahasiswa yang berkualitas. Kemampuan dosen yang baik akan berdampak positif terhadap kualitas layanan terhadap mahasiswa (Supriyanto, 2011).

Salah satu cara untuk mengetahui, memperbaiki, dan menjaga mutu dosen adalah dengan mengevaluasi kinerja dosen (Agung, 2011). Evaluasi atau penilaian kinerja harus dilakukan karena memiliki pengaruh yang kuat terhadap bentuk dan sifat hubungan kerja dalam suatu lembaga. Dengan kata lain, motivasi, sifat kerja sama, kepuasan psikologis, dan efisiensi kerja yang berpengaruh terhadap kemampuan dosen dalam proses mengajar.

Penambahan data merupakan salah satu metode analisis data yang menggabungkan statistika, mesin pembelajar (*machine learning*) yang merupakan cabang kecerdasan buatan (*artificial in*, teknologi basis data, ilmu komputasi, dan disiplin ilmu lainnya (Larose, 2005). Salah satu kegunaan penambahan data adalah digunakan untuk menemukan pola-pola tersembunyi dari data yang ada (Larose, 2005). Dari pola-pola yang terungkap, dapat dianalisis dan dievaluasi yang hasilnya dapat dijadikan basis pengetahuan untuk mendukung kegiatan-kegiatan yang dilakukan oleh pemilik data tersebut (U.S. Department of Education, 2012) yaitu:

- Memprediksi perilaku mahasiswa di masa depan dengan membuat model yang memiliki informasi mengenai pengetahuan, motivasi, metakognisi, dan sikap.
- Mengungkap dan memperbaiki karakteristik model domain untuk dipelajari agar dapat memperbaiki urutan instruksional yang optimal.
- Membantu dalam mempelajari efek-efek pendukung pedagogi yang berbeda.
- Memajukan pengetahuan dan sains tentang belajar mengajar untuk membangun model komputasi dalam bidang pendidikan.

Setiap perguruan tinggi pada umumnya memiliki data-data akademik seperti data mahasiswa, data nilai mahasiswa, data jadwal

kuliah, data dosen, data evaluasi dosen, dan data realisasi perkuliahan. Pada Universitas Respati Indonesia (URINDO), data masih diinput dan dihitung secara manual pada dokumen cetak. Dari dokumen cetak, kemudian akan diinput ke dalam basis data oleh staf di fakultas. Jumlah data mahasiswa baru dari tahun ke tahun memiliki kecenderungan meningkat.

Data evaluasi dosen masih berbentuk dokumen cetak dan diinput ke dalam berkas-berkas lembar sebar (*spreadsheet*).

Data nilai mahasiswa dan data hasil evaluasi belum pernah dianalisis lebih mendalam, biasanya hanya dibuatkan grafik yang menunjukkan tren atau kecenderungan Indeks Prestasi Kumulatif dari mahasiswa atau dari dosen.

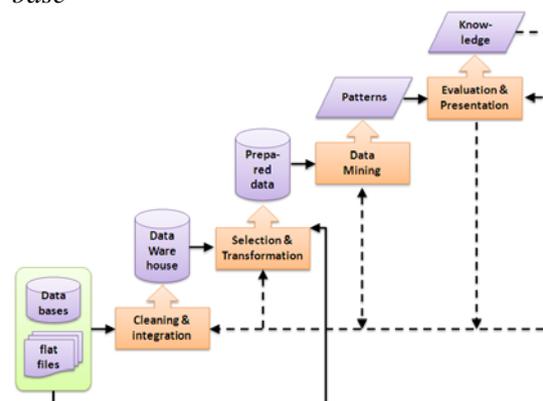
Berdasarkan latar belakang tersebut, maka penelitian ini akan meneliti bagaimana hubungan data nilai mahasiswa dengan data hasil evaluasi dosen yang telah dilakukan oleh mahasiswa dengan menggunakan teknik-teknik penambahan data yaitu algoritma Apriori dan C4.5 pada URINDO.

Pertanyaan penelitian

Dari masalah yang telah diidentifikasi dan ruang lingkup yang telah ditentukan, maka rumusan masalahnya yaitu bagaimana membangun aplikasi untuk mengklasifikasikan dan menentukan asosiasi antara nilai mahasiswa dengan nilai evaluasi dosen oleh mahasiswa dengan menggunakan algoritma C4.5 dan algoritma Apriori?

Kajian literatur

Sebagai suatu rangkaian proses, *data mining* dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 1 Tahap-tahap tersebut bersifat interaktif, pemakai terlibat langsung atau dengan perantaraan *knowledge base*



Gambar 1. Tahap-Tahap Data mining (Han, 2006).

- Tahap-tahap *data mining* ada 6 yaitu:
1. Pembersihan data (*data cleaning*)
 2. Integrasi data (*data integration*)
 3. Seleksi data (Data Selection)
 4. Loadasi data (Data Loadation)
 5. Proses *mining*
 6. Evaluasi pola (*pattern evaluation*)
 7. Presentasi pengetahuan (*knowledge presentation*)

Teknik pengumpulan data

Data nilai mahasiswa diperoleh dari dosen pengampu mata kuliah yang bersangkutan. Data nilai yang digunakan di sini adalah data nilai Kuis/Tugas Mahasiswa, Ujian Tengah Semester, dan Ujian Akhir Semester yang berbentuk dokumen cetak yang diisi oleh dosen.

Data hasil evaluasi dosen yang dilakukan oleh mahasiswa didapatkan dari hasil penyebaran angket kepada mahasiswa oleh BAAK, kemudian mahasiswa mengisi angket tersebut dan dikumpulkan lagi ke petugas BAAK. Kemudian BAAK akan menginput data-data hasil evaluasi dosen tersebut ke suatu file Excel. Data yang ada di file Excel tersebut yang akan digunakan untuk proses penambangan data.

Mata Kuliah :								
Kode Mata Kuliah :								
SKS :								
Dosen :								
NIDN :								
Hari / Tanggal :								
Kelas / Smt. :								
No	Nama	NPM	Nilai					
			Kuis & Tugas 20%	UTS 30%	UAS 50%	Angka	Huruf	
1						0.00	E	
2						0.00	E	
3						0.00	E	

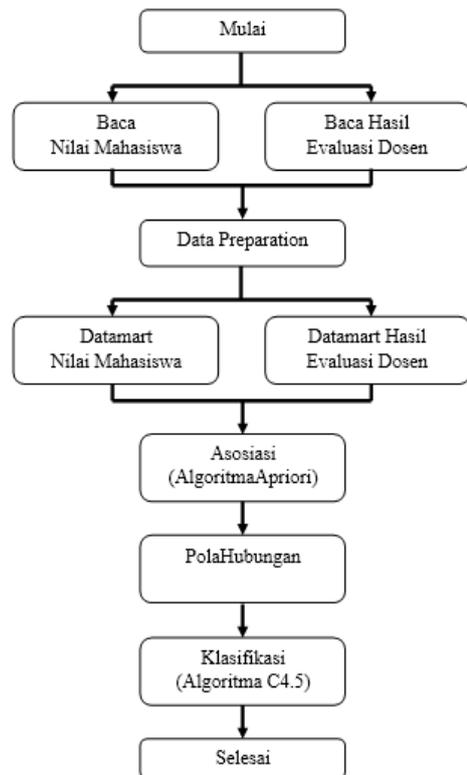
Gambar 2. Formulir Data Nilai Mahasiswa

Nama Dosen :		
Mata Kuliah :		155
Kelas :		155
Kategori Pembagian di Ruang Kuliah :		
1	1	75%
2	2	75%
3	3	75%
4	4	80%
5	5	80%
6	6	75%
7	7	80%
8	8	75%
9	9	75%
10	10	80%
11	11	75%
12	12	75%
13	13	75%
14	14	75%
15	15	75%
16	16	75%
17	17	75%
18	18	75%
19	19	75%
20	20	75%
21	21	75%
22	22	75%
23	23	75%
24	24	75%
25	25	75%
26	26	75%
27	27	75%
28	28	75%
29	29	75%
30	30	75%
31	31	75%
32	32	75%
33	33	75%
34	34	75%
35	35	75%
36	36	75%
37	37	75%
38	38	75%
39	39	75%
40	40	75%
41	41	75%
42	42	75%
43	43	75%
44	44	75%
45	45	75%
46	46	75%
47	47	75%
48	48	75%
49	49	75%
50	50	75%
51	51	75%
52	52	75%
53	53	75%
54	54	75%
55	55	75%
56	56	75%
57	57	75%
58	58	75%
59	59	75%
60	60	75%
61	61	75%
62	62	75%
63	63	75%
64	64	75%
65	65	75%
66	66	75%
67	67	75%
68	68	75%
69	69	75%
70	70	75%
71	71	75%
72	72	75%
73	73	75%
74	74	75%
75	75	75%
76	76	75%
77	77	75%
78	78	75%
79	79	75%
80	80	75%
81	81	75%
82	82	75%
83	83	75%
84	84	75%
85	85	75%
86	86	75%
87	87	75%
88	88	75%
89	89	75%
90	90	75%
91	91	75%
92	92	75%
93	93	75%
94	94	75%
95	95	75%
96	96	75%
97	97	75%
98	98	75%
99	99	75%
100	100	75%

Gambar 3. Hasil Evaluasi Dosen

Metodologi penelitian

Bentuk penelitian yang dipergunakan adalah terapan yang bersifat kuantitatif deskriptif (Walliman, 2011). Penelitian akan menerapkan algoritma-algoritma yang berkembang di bidang penambangan data. Data yang diperoleh adalah data sekunder dan akan diekplorasi untuk menggambarkan objek-objek yang diteliti yang tujuannya untuk memperkirakan gejala-gejala yang mungkin dapat terjadi disebabkan data dari lapangan.



Gambar 4. Alur Proses Penelitian

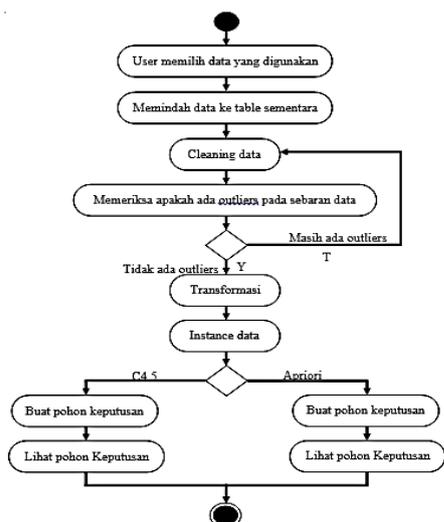
Fase analisis

Dari hasil Penelitian yang dilakukan akan dicari C4.5 (pohon kehidupan) dan Apriori (nilai *support* dan *confidence*) dari hubungan penilaian dosen dengan data nilai mahasiswa. Tidak semua data nilai mahasiswa akan dicari hubungannya dengan data penilaian dosen, hanya beberapa atribut yang kira-kira berguna dan tidak terlalu acak. Karena data yang terlalu acak akan membuat proses *mining* memakan waktu lama dan tingkat hubungannya pun rendah. Analisis sistem saat ini dilihat berdasarkan hasil angket penilaian dosen dan nilai mahasiswa pada proses belajar mengajar

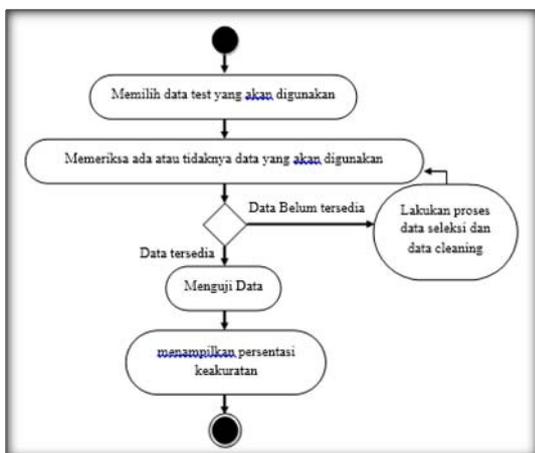
yang berjalan pada universitas, juga sebagai dasar untuk membuat atau memperbaiki serta mempermudah proses belajar mengajar yang sedang berjalan di universitas agar menjadi lebih baik lagi.

Activity Diagram

Activity diagram memodelkan alur kerja(work flow) sebuah proses bisnis dan urutan aktivitas pada suatu proses. Diagram ini sangat mirip dengan flowchart karena dapat memodelkan prosedur logika, proses bisnis dan alur kerja. Perbedaan utamanya adalah flowchart dibuatkan untuk menggambarkan alur kerja dari sebuah sistem, sedangkan activity diagram dibuat untuk menggambarkan aktivitas dari aktor. Perancangan activity diagram pada aplikasi data mining adalah sebagai berikut :



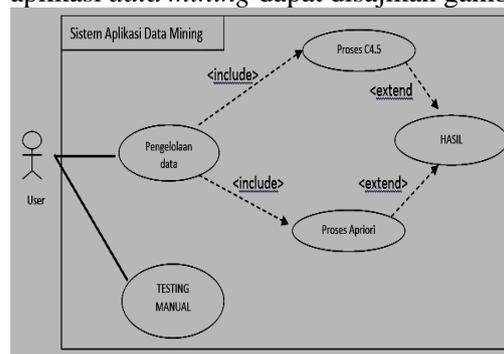
Gambar 5. Activity Diagram Pembentukan Pohon Keputusan dan Apriori



Gambar 6. Activity Diagram Pengujian Pohon Keputusan Apriori

Use case Diagram

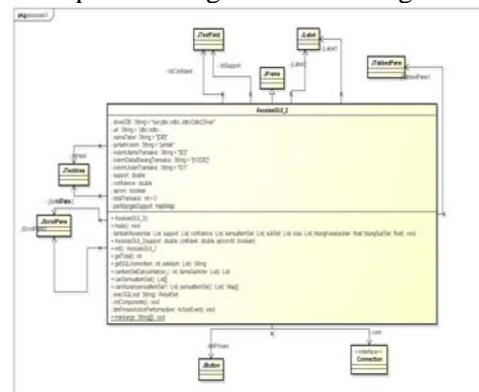
Dalam perancangan Aplikasi ini secara garis besar diagram use case digunakan untuk mendeskripsikan apa yang seharusnya dilakukan oleh sistem. Diagram use case menyediakan pandangan dari luar terhadap sistem, sementara model rancangan adalah pandangan dari dalam. Use case diagram pada aplikasi data mining dapat disajikan gambar 7.



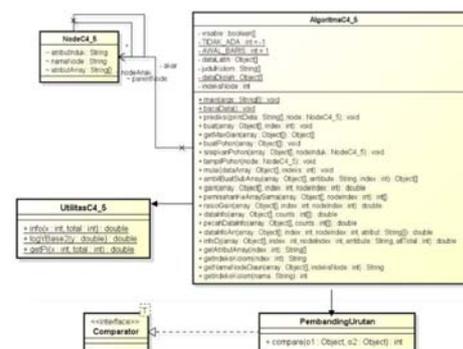
Gambar 7. Use case diagram

Class Diagram

Setelah memodelkan sistem utama dengan use case diagram, langkah selanjutnya menganalisa kelas-kelas yang dapat disimpulkan dengan use case diagram.



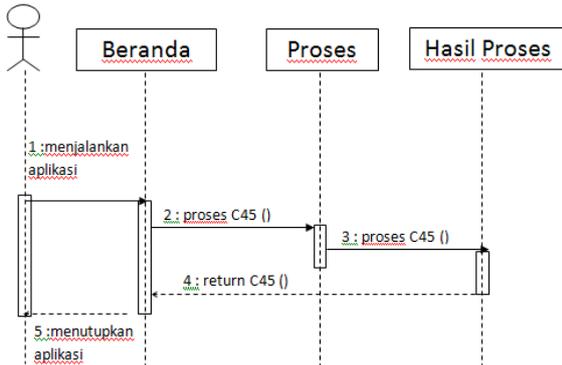
Gambar 8. Class diagram Apriori



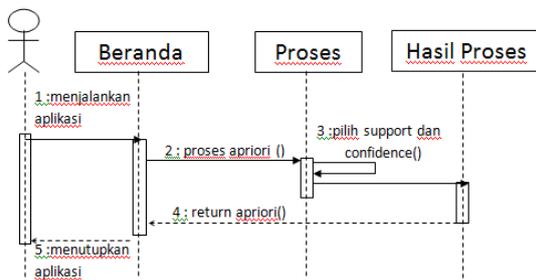
Gambar 9. Class diagram C4.5

Sequence Diagram

Sequence diagram digunakan untuk menjabarkan aktifitas yang ada pada use case kepada level yang lebih detail untuk menggambarkan interaksi antar objek di dalam dan di sekitar sistem yang berupa message yang digambarkan terhadap waktu.



Gambar 10. Sequence diagram C4.5



Gambar 11. Sequence diagram Apriori

Fase perancangan

Tahapan ini meliputi mengkonfigurasi komponen-komponen perangkat lunak dan perangkat keras dari suatu sistem sehingga setelah instalasi sistem akan benar-benar memuaskan berdasarkan rancang bangun yang telah ditetapkan pada tahap akhir analisis.

Dalam penelitian ini terdiri dari dua Sumber data, yaitu data nilai mahasiswa dan data evaluasi dosen. Kedua data tersebut dicari hubungan beberapa atribut, oleh karena itu tidak semua tabel digunakan, maka perlu dilakukan pembersihan data agar data yang akan diolah benar-benar relevan dengan yang dibutuhkan. Cara pembersihan dengan menghapus atribut yang tidak terpakai dan menghapus data-data yang tidak lengkap isianya. atribut yang digunakan terdiri dari

atribut pada data nilai mahasiswa dan pada Data Evaluasi Dosen adalah sebagai berikut:

1. Data nilai Mahasiswa

Data nilai mahasiswa adalah nilai mahasiswa yang didapat dari dosen untuk setiap matakuliah. Atribut yang ada dapat dilihat dalam tabel 1.

Tabel 1. Nilai mahasiswa

FIELD NAME	DATA TYPE	LENG TH	DESCRIPTION
DOS	C	3	Nama Dosen
MK	C	3	Nama Matakuliah
TM	C	6	Tugas Mahasiswa
UTS	C	6	Nilai Ujian Tengah Semester
UAS	C	6	Nilai Ujian Akhir Semester
Hsl	C	6	Hasil
Ket	C	6	Keterangan

2. Data Evaluasi Dosen

Data Evaluasi Dosen adalah lembar angket yang diberikan kepada mahasiswa untuk diisi/dinilai pada matakuliah yang telah dipelajari. Data yang dicatat dapat dilihat pada tabel 2.

Tabel 2. Evaluasi Dosen

FIELD NAME	DATA TYPE	LENG TH	DESCRIPTION
DOS	C	3	Nama Dosen
MK	C	3	Nama Matakuliah
N1	C	6	Jawaban Pertanyaan 1
N2	C	6	Jawaban Pertanyaan 2
N3	C	6	Jawaban Pertanyaan 3
N4	C	6	Jawaban Pertanyaan 4
N5	C	6	Jawaban Pertanyaan 5
N6	C	6	Jawaban Pertanyaan 6
N7	C	6	Jawaban Pertanyaan 7
N8	C	6	Jawaban Pertanyaan 8
N9	C	6	Jawaban Pertanyaan 9
N10	C	6	Jawaban Pertanyaan 10
N11	C	6	Jawaban Pertanyaan 11
N12	C	6	Jawaban Pertanyaan 12
N13	C	6	Jawaban Pertanyaan 13
N14	C	6	Jawaban Pertanyaan 14
N15	C	6	Jawaban Pertanyaan 15
N16	C	6	Jawaban Pertanyaan 16
N17	C	6	Jawaban Pertanyaan 17

Data nilai mahasiswa yang diambil dalam sampel adalah data mahasiswa Angkatan 2012. Hal ini didasarkan pada kebutuhan data yang akan dihubungkan dengan Data Evaluasi Dosen, dengan asumsi bahwa mahasiswa angkatan 2012-2013 yang mengikuti mata kuliah pada tahun 2012-2013.

Tabel 3. ID Mata Kuliah dan Dosen

MK	DSN	ID	MK	DSN	ID
AAM	PAG	01	JNH	TBD	20
ABS	FIS	02	RDW	PAG	21
ABS	MTD	03	SHS	TBA	22
AS	ECM	04	SJI	ALP	23
AS	MTN	05	SJI	PAP	24
AS	PD2	06	SLM	ALL	25
DDK	PTB	07	SLM	MTD	26
DDK	TBD	08	SON	A3D	27
DJA	IMK	09	SON	PBW	28
DSO	PTP	10	SON	PGF	29
EMN	BII	11	SON	PGM	30
EMN	ERP	12	SRS	FIS	31
EMN	KSD	13	SRS	KSD	32
EMN	PKN	14	SUZ	DWM	33
HRT	ALP	15	SUZ	IMK	34
IND	PSG	16	SUZ	RPL	35
IND	SIG	17	TTS	ETP	36
JNH	PAP	18	TTS	KAP	37
JNH	PTB	19	TWI	BII	38

Atribut yang digunakan dalam data nilai mahasiswa meliputi:

- ID digunakan sebagai penggabungan nama dosen dengan matakuliah.
- Atribut nilai tugas mahasiswa (TM) digunakan untuk proses *mining* guna mengetahui nilai tugas mahasiswa pada setiap matakuliah.
- Atribut nilai Ujian Tengah Semester (UTS) digunakan untuk proses *mining* guna mengetahui nilai UTS pada setiap matakuliah.
- Atribut nilai Ujian Akhir Semester (UAS) digunakan untuk proses *mining* guna mengetahui nilai UAS pada setiap matakuliah.
- Atribut Hasil (Hsl) digunakan untuk proses *mining* guna mengetahui keberhasilan mahasiswa pada setiap matakuliah .

Atribut yang digunakan dalam Data Evaluasi Dosen meliputi:

- ID digunakan sebagai penggabungan nama dosen dengan matakuliah.

- Atribut Jawaban Pertanyaan (Nilai a) digunakan untuk proses data *mining* guna mengetahui jawaban pertanyaan No 1 – 13 pada daftar penilaian dosen.
- Atribut Jawaban Pertanyaan (Nilai b) digunakan untuk proses data *mining* guna mengetahui jawaban pertanyaan No 14 – 17 pada daftar penilaian dosen.
- Atribut Hasil (Hsl) digunakan untuk proses *mining* guna mengetahui keberhasilan mahasiswa pada setiap mata kuliah.

Tabel 4. Predikat Nilai

Indeks Prestasi Kumulatif	Predikat
0,00 - 57,00	Nilai Kurang
57,01 - 79,00	Nilai Cukup
78,00 - 100,00	Nilai Baik

Tabel 5. *Learning Dataset*

No	ID	TM	UTS	UAS	HSL
1	15	kurang	Cukup	kurang	Kurang
2	15	Baik	Cukup	Baik	Baik
3	15	Baik	Cukup	Cukup	Baik
4	15	Cukup	Baik	Cukup	Cukup
5	15	kurang	Cukup	Baik	Cukup
6	15	kurang	Cukup	Baik	Baik
7	15	Baik	Baik	Baik	Baik
8	15	Cukup	Baik	Cukup	Cukup
9	15	kurang	Cukup	Baik	Cukup
10	15	Baik	Cukup	Cukup	Baik
11	15	kurang	Baik	Cukup	Cukup
12	15	Baik	Cukup	Baik	Baik
13	15	Cukup	Baik	Baik	Baik
14	15	Baik	Cukup	Baik	Baik
15	15	kurang	Cukup	Baik	Cukup
16	15	Cukup	Cukup	Baik	Cukup
17	15	kurang	Cukup	kurang	Kurang
18	15	Baik	Baik	Baik	Baik

Jadi

$$\begin{aligned}
 Entropi(S) &= \left(-\left(\frac{9}{18}\right) \times \log_2\left(\frac{9}{18}\right)\right) + \left(-\left(\frac{7}{18}\right) \times \log_2\left(\frac{7}{18}\right)\right) + \left(-\left(\frac{2}{18}\right) \times \log_2\left(\frac{2}{18}\right)\right) \\
 &= 1.3821
 \end{aligned}$$

Tabel 6. Hasil Perhitungan pada Dataset

Total Kasus	Sum (baik)	Sum (cukup)	Sum (kurang)	Entropi Total
18	9	7	2	1.3821

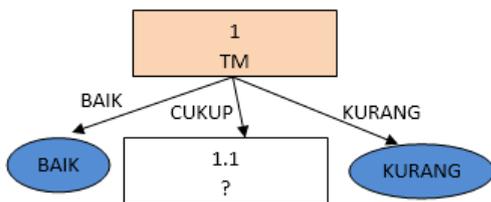
Tabel 7. Analisis atribut, nilai, banyaknya kejadian nilai, Entropi, dan Gain

Atrbt	Nilai	Sum (Nilai)	Sum (Baik)	Sum (Ckp)	Sum (Krg)	Entrop	Gain
TM	Baik	7	7	0	0	0.00	
	Cukup	4	1	3	0	0.81	
	Kurang	7	1	4	2	1.38	0.666
UTS	Baik	6	3	3	0	0.00	
	Cukup	12	6	4	2	1.46	
	Kurang	0	0	0	0	0.00	0.076
UAS	Baik	11	7	4	0	0.95	
	Cukup	5	2	3	0	0.97	
	Kurang	2	0	0	2	0.00	0.534

Untuk menghitung gain adalah:

$$Gain(TM) = 0.764205 - \left(\left(\frac{7}{18} \right) \times 0 + \left(\frac{4}{18} \right) \times 0.81 + \left(\frac{7}{18} \right) \times 1.3788 \right)$$

$$Gain(TM) = 0,665625$$



Gambar 12. Pohon Keputusan Node 1 (root node)

Tabel 8. Data yang memiliki TM = Kurang

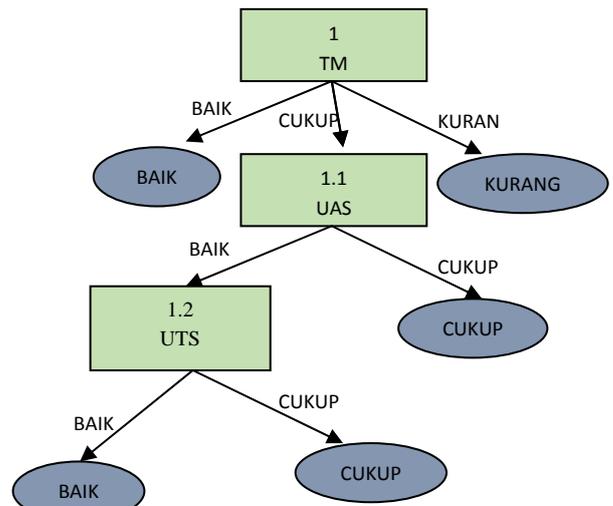
No	ID	TM	UTS	UAS	HSL
4	15	CUKUP	BAIK	CUKUP	CUKUP
8	15	CUKUP	BAIK	CUKUP	CUKUP
13	15	CUKUP	BAIK	BAIK	BAIK
16	15	CUKUP	CUKUP	BAIK	CUKUP

Tabel 9. Hasil Analisis Node 1.1

TM CUKUP	Sum (baik)	Sum (cukup)	Sum (cukup)	Entropi
7	1	3	0	0.811278

Tabel 10. Analisis atribut, nilai, banyaknya kejadian nilai, Entropi, dan Gain

Atrbt	Nilai	Sum (Nilai)	Sum (Baik)	Sum (Ckp)	Sum (Krg)	Entropi	Gain
UTS	BAIK	3	1	2	0	0.9183	
	CUKUP	1	0	1	0	0	
	KURANG	0	0	0	0	0	0.123
UAS	BAIK	2	1	1	0	1	
	CUKUP	2	0	2	0	0	
	KURANG	0	0	0	0	0	0.311



Gambar 13. Pohon Keputusan Analisis Node 1.1 dan Node 1.2

Tabel 11. Learning Dataset

No	ID	Nilai a	Nilai b	Hasil
1	15	Baik	Kurang	Cukup
2	15	Baik	Baik	Baik
3	15	Baik	Baik	Baik
4	15	Baik	Baik	Baik
5	15	Baik	Baik	Baik
6	15	baik	Baik	Baik
7	15	baik	Baik	Baik
8	15	baik	Baik	Baik
9	15	baik	Baik	Baik
10	15	baik	Baik	Baik
11	15	baik	Baik	baik
12	15	baik	Baik	baik
13	15	baik	Baik	baik
14	15	baik	Baik	baik
15	15	baik	Baik	baik
16	15	baik	Baik	baik

17	15	baik	Kurang	cukup
18	15	baik	Baik	Baik
19	15	baik	Baik	Baik
20	15	baik	Baik	Baik
21	15	baik	Baik	Baik
22	15	baik	Kurang	Cukup
23	15	baik	Baik	Baik
24	15	baik	Kurang	Cukup
25	15	baik	Baik	Baik
26	15	baik	Baik	Baik
27	15	kurang	Kurang	Kurang
28	15	baik	Baik	Baik
29	15	baik	Baik	Baik
30	15	baik	Kurang	Cukup
31	15	baik	Baik	Baik
32	15	baik	Baik	Baik
33	15	baik	Baik	Baik
34	15	baik	Baik	Baik
35	15	baik	Baik	Baik
36	15	baik	Baik	Baik
37	15	baik	Baik	Baik
38	15	baik	Baik	Baik
39	15	baik	Kurang	Cukup
40	15	baik	Baik	Baik
41	15	baik	Kurang	Kurang
42	15	baik	Baik	Baik
43	15	baik	Baik	Baik

Jadi

$$Entropi (S) = \left(-\left(\frac{35}{43}\right) \times \log_2 \left(\frac{35}{43}\right)\right) + \left(-\left(\frac{6}{43}\right) \times \log_2 \left(\frac{6}{43}\right)\right) + \left(-\left(\frac{2}{43}\right) \times \log_2 \left(\frac{2}{43}\right)\right) = 0.844069$$

Tabel 12. Hasil Perhitungan pada Dataset

Total Kasus	Sum (Baik)	Sum (Cukup)	Sum (Kurang)	Entropi Total
43	35	6	2	0.844

Gambar 14. Pohon Keputusan Node 1 (root node)

Tabel 13. Analisis Atribut, Nilai, Banyaknya Kejadian Nilai, Entropi dan Gain

Atrbt	Nilai	Sum (Nilai)	Sum (Baik)	Sum (Ckp)	Sum (Krg)	Entropi	Gain
nilai a	baik	42	35	6	1	0.749	
	cukup	0	0	0	0	0	
	kurang	1	0	0	1	0	0.113
nilai b	baik	35	35	0	0	0	
	cukup	0	0	0	0	0	
	kurang	8	0	6	2	0,811	0.844

Untuk menghitung gain adalah :

$$Gain (TM) = 0.881291 - \left(\left(\frac{42}{43}\right) \times 0.749\right) - \left(\left(\frac{0}{43}\right) \times 0\right) - \left(\left(\frac{1}{43}\right) \times 0\right)$$

$$Gain (TM) = 0,112838$$

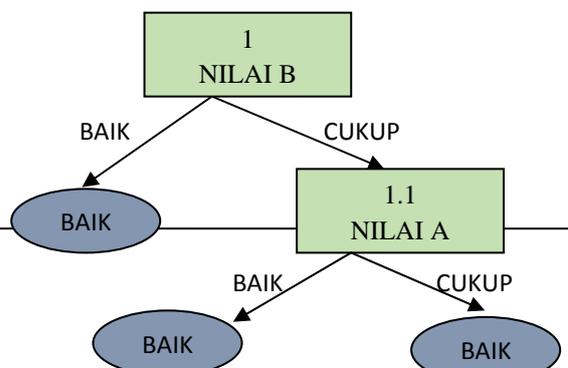
Tabel 14. Transformasi Data

Kategori	Keterangan
DB	Ket hasil baik nilai hasil mahasiswa
DC	Ket hasil cukup nilai hasil mahasiswa
DK	Ket hasil kurang nilai hasil mahasiswa
MB	Ket hasil baik nilai hasil penilaian dosen
MC	Ket hasil cukup nilai hasil penilaian dosen
MK	Ket hasil kurang nilai hasil penilaian dosen

Dari kombinasi yang terdapat di tabel 14 terdapat enam tingkatan untuk mengukur penilaian dosen dan mahasiswa.

Tabel 15. Hasil Penggabungan Nilai Mahasiswa dengan Penilaian Dosen

No	ID	DK	DC	DB	MK	MC	MB
1	01	5	2	3	0	0	22
2	02	0	0	16	0	21	13
3	03	2	3	13	0	0	12
4	04	4	4	14	0	2	67



5	05	2	2	10	11	10	13
6	06	5	15	22	0	4	68
7	07	2	0	18	0	4	29
8	08	0	0	16	0	4	29
9	09	5	17	22	0	0	20
10	10	5	15	36	5	38	28
11	11	2	2	13	1	53	1
12	12	5	12	24	0	45	13
13	13	0	2	19	0	5	13
14	14	9	8	20	4	22	12
15	15	2	6	35	2	7	9
16	16	3	6	16	0	0	60
17	17	6	8	15	0	0	61
18	18	4	6	41	0	8	13
19	19	1	2	19	0	13	21
20	20	2	2	12	0	13	23
21	21	5	2	3	0	0	17
22	22	6	7	33	9	13	17
23	23	2	6	35	6	23	16
24	24	4	6	41	0	26	16
25	25	10	6	54	5	21	37
26	26	5	4	13	0	12	22
27	27	4	2	17	0	10	53
28	28	5	7	33	0	2	56
29	29	4	3	37	0	1	40
30	30	3	7	51	0	1	60
31	31	0	1	14	0	13	22
32	32	5	12	36	0	0	37
33	33	6	6	15	24	18	18
34	34	5	1	10	1	7	14
35	35	2	4	14	1	12	7
36	36	4	5	6	26	11	21
37	37	5	9	29	7	12	17
38	38	4	2	17	0	1	11

Untuk mempermudah, nama-nama item di Tabel 15, disingkat dengan diambil huruf awalnya saja, sebagai contoh:

- MB = Baik (Hasil Nilai Mahasiswa)
- MC = Cukup (Hasil Nilai Mahasiswa)
- MK = Kurang (Hasil Nilai Mahasiswa)
- DB = Baik (Hasil Penilaian Dosen)
- DC = Cukup (Hasil Penilaian Dosen)
- MK = Kurang (Hasil Nilai Mahasiswa)

Sehingga tabel 15 menjadi seperti tabel 16 Transaksi hasil:

Tabel 16. Transaksi hasil

Transaksi	HASIL					
	DK	DC	DB	0	0	MB
T01	DK	DC	DB	0	0	MB
T02	0	0	DB	0	MC	MB
T03	DK	DC	DB	0	0	MB
T04	DK	DC	DB	0	MC	MB
T05	DK	DC	DB	MK	MC	MB
T06	DK	DC	DB	0	MC	MB
T07	DK	0	DB	0	MC	MB
T08	0	0	DB	0	MC	MB
T09	DK	DC	DB	0	0	MB
T10	DK	DC	DB	MK	MC	MB
T11	DK	DC	DB	MK	MC	MB
T12	DK	DC	DB	0	MC	MB
T13	0	DC	DB	0	MC	MB
T14	DK	DC	DB	MK	MC	MB
T15	DK	DC	DB	MK	MC	MB
T16	DK	DC	DB	0	0	MB
T17	DK	DC	DB	0	0	MB
T18	DK	DC	DB	0	MC	MB
T19	DK	DC	DB	0	MC	MB
T20	DK	DC	DB	0	MC	MB
T21	DK	DC	DB	0	0	MB
T22	DK	DC	DB	MK	MC	MB
T23	DK	DC	DB	MK	MC	MB
T24	DK	DC	DB	0	MC	MB
T25	DK	DC	DB	MK	MC	MB
T26	DK	DC	DB	0	MC	MB
T27	DK	DC	DB	0	MC	MB
T28	DK	DC	DB	0	MC	MB
T29	DK	DC	DB	0	MC	MB
T30	DK	DC	DB	0	MC	MB
T31	0	DC	DB	0	MC	MB
T32	DK	DC	DB	0	0	MB
T33	DK	DC	DB	MK	MC	MB
T34	DK	DC	DB	MK	MC	MB
T35	DK	DC	DB	MK	MC	MB
T36	DK	DC	DB	MK	MC	MB
T37	DK	DC	DB	MK	MC	MB
T38	DK	DC	DB	0	MC	MB

Langkah ke-1: Hitung banyaknya transaksi untuk setiap item.

Tabel 17. Banyaknya transaksi per Item

Item	Banyaknya Transaksi
DK	34
DC	35
DB	38
MK	13
MC	31
MB	38

Langkah ke-2: Berdasarkan *golden rule* yang telah disebutkan di atas, disaring data menjadi 2-item, 3-item, 4-item, 5-item, dan 6-item.

Langkah ke-3: Buat pasangan item dimulai dari item pertama sampai dengan item keenam lihat tabel 18 pasangan item.

Tabel 18. Pasangan Item

Jml item	Keterangan
2-item	AB, AC, AX, AY, AZ, BC, BX, BY, BZ, CX, CY, CZ, XY, XZ, YZ
3-item	ABC, ABX, ABY, ABZ, ACX, ACY, ACZ, AXY, AXZ, AYZ, BCX, BCY, BCZ, BXY, BXZ, BYZ, CXY, CXZ, CYZ, XYZ,
4-item	ABCX, ABCY, ABCZ, ABXY, ABXZ, ABYZ, ACXY, ACXZ, ACYZ, AXYZ, BCXY, BCXZ, BCYZ, BXYZ, CXYZ
5-item	ABCXY, ABCXZ, ABCYZ, ACXYZ, BCXYZ
6-item	ABCXYZ

Langkah ke-4: Hitung berapa kali suatu pasangan item bersamaan.

Langkah ke-5: Cari asosiasi pada semua himpunan bagian yang telah dibuat.

Simpulan

Aplikasi *Data Mining* yang dibangun bukan untuk menggantikan sistem yang lama, akan tetapi pengembangan aplikasi yaitu menggunakan bahasa *Java*.

Berdasarkan hasil dari analisis dan perancangan sistem yang dibangun, maka dapat diambil kesimpulan:

1. Algoritma Apriori telah berhasil diterapkan untuk mendapatkan asosiasi data dengan nilai *confidence* dan *support* sesuai dengan yang keinginan. Dari hasil

asosiasi ini, diharapkan dapat membantu memperbaiki dan menjaga mutu dosen.

2. Algoritma C4.5 telah berhasil diterapkan untuk mendapatkan klasifikasi data dengan membentuk pohon keputusan dari seluruh data maupun sebagian data. Dari hasil klasifikasi berupa pohon keputusan tersebut dapat digunakan untuk memprediksi hasil dari data nilai mahasiswa maupun data evaluasi dosen untuk ke depannya.

Saran

Dari pembuatan Aplikasi yang dibangun ini masih memiliki kekurangan, baik dari segi fungsionalitas maupun data yang dimiliki untuk meningkatkan kualitas aplikasi dimasa yang akan datang yaitu :

1. Aplikasi *Algoritma* C4.5 ini berupa model pohon keputusan yang belum sempurna/baik karena masih menampilkan output berupa tulisan sehingga representasi data perlu dalam bentuk grafik.
2. Bagi peneliti lain yang ingin menggunakan kasus yang sama dapat membandingkan keakurasian algoritma C4.5 dan Apriori dengan algoritma yang lain selain sebagai metode pembandingan dalam kasus ini.

DAFTAR PUSTAKA

- Dennis. Wixom. dan Tegarden, 2009. *Third Edition Systems Analisis & Design with UML Version 2.0 An Object-Oriented Approach*. USA : John Wiley & Sons, Inc.
- Fakhrurriqfi, M. dan Wardoyo, Retantyo. 2013. *Perbandingan Algoritma Nearest, C4.5 dan LVQ untuk Klasifikasi Kemampuan Mahasiswa, Vol.7, N0.2, July 2013, pp 145-154*
- Hernawati, Astuti. 2013. *Pengenalan Data Mining*. Jakarta : Media Kita.
- Kusrini. Dan Taufiq Lutfi, Emha. 2009. *Algoritma Data Mining*. Penerbit ANDI, Yogyakarta
- Munawar. 2005. *Pemodelan Visual*. Penerbit Graha Ilmu. Yogyakarta
- Susanto, S. Suryadi, D. 2010; *Pengantar DATA MINING Menggali Pengetahuan dari Bongkahan Data*. Penerbit ANDI, Yogyakarta
- Ayu Nursela, Dwi. 2009. *Penerapan Algoritma C4.5 untuk Klasifikasi Tingkat Keganasan Kanker Payudara*.

Cahyono P, Heruandika. Bettiza, Martaleli, dan Matulandan, Tekad. *Penerapan Algoritma Apriori Dalam Menemukan Hubungan Data Awal Masuk Mahasiswa Dengan Prestasi Akademik*

Masykur Huda, Nuqson. 2010. *Aplikasi Data Mining Untuk menampilkan Tingkat Kelulusan*. Tugas Akhir. Semarang: MIPA Universitas Diponegoro.

Raditya, Angga. 2011. *Implementasi Data Mining Classification untuk mencari Pola Prediksi Hujan Dengan Menggunakan Algoritma C4.5*.

Zakaria, Oliver. 2012. *Implementasi Data Mining Menggunakan Aturan Asosiasi Dengan Algoritma Apriori Terhadap Penyusunan Layout Makanan Pada Rumah Makan Padang "Murah Meriah"*. Naskah Publikasi. Yogyakarta. STMIK AMIKOM.