

PENERAPAN *NAÏVE BAYES CLASSIFIER* UNTUK PEMILIHAN KONSENTRASI MATA KULIAH

Annisa Paramitha Fadillah¹, Bella Hardiyana²
Program Studi Sistem Informasi, Universitas Komputer Indonesia
annisa@email.unikom.ac.id¹, bella.hardiyana@email.unikom.ac.id²

ABSTRAK

UNIKOM merupakan perguruan tinggi yang berkembang pesat di Indonesia, yang memanfaatkan perkembangan teknologi berupa layanan sistem informasi. UNIKOM menggunakan sistem informasi untuk membantu berjalannya proses akademik, terutama dalam pengolahan data yang berkaitan dengan kegiatan perkuliahan. Sistem informasi merupakan salah satu program studi yang cukup besar yang ada di UNIKOM. Sebagai program studi yang cukup besar, program studi sistem informasi memiliki mahasiswa cukup banyak untuk setiap angkatan. Pada prodi sistem informasi mahasiswa semester 6 wajib melakukan pemilihan konsentrasi matakuliah, akan tetapi mahasiswa merasa cukup kesulitan dalam melakukan pemilihan konsentrasi mata kuliah, bahkan dosen wali kesulitan dalam memberikan rekomendasi pilihan konsentrasi mata kuliah kepada mahasiswanya. Oleh karena itu, akan dilakukan penelitian mengenai pemilihan konsentrasi matakuliah, dengan menggunakan *naïve bayes*. Dengan menggunakan *naïve bayes* diharapkan dapat memberikan rekomendasi pemilihan konsentrasi mata kuliah baik untuk mahasiswa maupun dosen wali. Penggunaan *naïve bayes classifier*, dikarenakan *naïve bayes classifier* merupakan metode yang mudah dipahami dan cukup sederhana.

Kata kunci: *Naïve bayes*, Data mining, Pemilihan konsentrasi mata kuliah.

I. PENDAHULUAN

1.1. Latar Belakang Masalah

Perkembangan dunia pendidikan sebagai salah satu aspek penting dalam kehidupan saat ini sangatlah pesat. Pendidikan saat ini didapatkan dengan berbagai cara, baik secara formal dan non formal. Perguruan tinggi sebagai salah satu institusi pendidikan juga ikut berkembang. Perguruan tinggi merupakan salah satu organisasi yang memberikan jasa pelayanan pendidikan kepada masyarakat. Untuk meningkatkan kualitas pendidikan harus diikuti oleh perkembangan teknologi yang ada saat ini, sehingga perkembangan pendidikan yang ada pada institusi tersebut semakin maksimal untuk mencapai tujuannya.

Sistem informasi adalah kumpulan perangkat keras dan perangkat lunak yang dirancang untuk mentransformasikan data ke dalam bentuk informasi yang berguna. Informasi adalah data yang telah diproses sedemikian rupa sehingga meningkatkan pengetahuan seseorang yang menggunakannya. [1]

UNIKOM merupakan salah satu universitas yang memanfaatkan perkembangan teknologi berupa layanan sistem informasi. UNIKOM menggunakan sistem informasi untuk membantu berjalannya proses akademik, terutama dalam pengolahan data yang berkaitan dengan kegiatan perkuliahan. Salah satu sistem informasi yang digunakan untuk mengolah data tersebut yaitu sistem informasi perwalian. Data yang diolah berupa mata kuliah, penilaian dan absensi.

Prodi sistem informasi merupakan program studi yang cukup besar yang ada di UNIKOM dengan jumlah mahasiswa >200 mahasiswa pada setiap angkatan. Sesuai dengan kurikulum yang berlaku pada prodi SI saat ini, mahasiswa yang akan melanjutkan studi di semester 6 harus melakukan pemilihan matakuliah berdasarkan minatnya.

Pemilihan konsentrasi mata kuliah dilakukan mahasiswa saat akan mengambil studi di semester 6, terdapat 2 pilihan konsentrasi mata kuliah yaitu Rekayasa Sistem Informasi dan Teknologi Informasi. Jika mahasiswa sudah memilih salah satu konsentrasi mata kuliah maka mahasiswa tersebut tidak dapat memilih konsentrasi lain.

Dari 364 mahasiswa yang terdaftar terdapat 312 mahasiswa yang aktif dan hanya ada 286 mahasiswa yang telah melakukan pemilihan konsentrasi matakuliah pada semester 6 tahun ajaran 2015/2016, diketahui bahwa pemilihan mata kuliah berdasarkan minat dapat dilihat dari data pada tabel 1 dibawah ini :

Tabel 1 Data Pilihan Konsentrasi Mahasiswa

Konsentrasi Mata Kuliah	Mata Kuliah	Jumlah Mahasiswa
Rekayasa Sistem Informasi	Sistem Informasi Terdistribusi	211
Teknologi Informasi	Jaringan Komputer Lanjut	78
Total Mahasiswa		289

Dari tabel diatas dapat diketahui jumlah mahasiswa yang memilih konsentrasi mata kuliah Rekayasa Sistem Informasi dengan mata kuliah yang diajarkan Sistem Informasi Terdistribusi sebanyak 211 mahasiswa. Sedangkan untuk konsentrasi mata kuliah Teknologi Informasi dengan mata kuliah yang diajarkan Jaringan Komputer Lanjut sebanyak 78 mahasiswa. Akan tetapi dari data diatas belum diketahui model klasifikasi pemilihan mata kuliah untuk mahasiswa semester 6 tersebut, dan juga belum diketahui faktor-faktor apa saja yang mendasari pemilihan mata kuliah tersebut. Hal ini disebabkan salah satunya adalah pengolahan data yang kurang maksimal. Pengolahan data yang belum maksimal mengakibatkan informasi yang dihasilkan menjadi tidak berkualitas atau kurang bernilai.

Berdasarkan uraian tersebut, maka yang akan diteliti adalah klasifikasi pemilihan mata kuliah pada mahasiswa semester 6 program studi sistem informasi UNIKOM. Agar dapat dapat memberikan pengetahuan juga rekomendasi standar untuk pemilihan konsentrasi mata kuliah pada program studi sistem informasi, agar data yang digunakan sebagai *sample* dapat dijadikan acuan untuk pemilihan mata kuliah mahasiswa untuk angkatan berikutnya.

1.2. Identifikasi Masalah

Berdasarkan uraian latarbelakang masalah di atas maka dapat diketahui identifikasi masalah adalah sebagai berikut:

1. Belum adanya standar untuk pemilihan konsentrasi pemilihan mata kuliah di program studi sistem informasi untuk mahasiswa semester 6 UNIKOM.
2. Dosen Wali masih kesulitan dalam memeberikan rekomendasi pemilihan mata kuliah.

1.3. Batasan Masalah

1. Pemodelan menggunakan naïve bayes, pada program studi sistem informasi dengan kurikulum 2012 dengan data sample mahasiswa semester 6 tahun ajaran 2015/2016.
2. Hanya terdapat dua pilihan konsentrasi mata kuliah, yaitu Rekayasa Sistem Informasi dan Teknologi Informasi.
3. Parameter yang digunakan dalam perhitungan naïve bayes hanya berdasarkan nilai mata kuliah yang terkait dengan pilihan konsentrasi mata kuliah.

2. LANDASAN TEORI

2.1. Tinjauan Pustaka

Penelitian yang dilakukan saat ini merupakan pengembangan dari penelitian sebelumnya yang berjudul “*Classification of Subject Concentration using Algorithm C4.5*”, penelitian sebelumnya menggunakan algoritma C4.5 dalam tahapan pemodelannya, dan menghasilkan rekomendasi pola berupa pohon keputusan. [2] Sedangkan dalam penelitian ini metode yang digunakan adalah *naïve bayes classifier* yang menghasilkan rekomendasi berdasarkan perhitungan dataset.

Penelitian yang berkaitan dengan penelitian ini salah satunya yaitu, penelitian dengan judul *Data Mining Menggunakan Algoritma Naïve Bayes Untuk Klasifikasi Kelulusan Mahasiswa Universitas Dian Nuswantoro*. [3] Penelitian tersebut membahas klasifikasi kelulusan mahasiswa, dan peningkatan maupun penurunan kelulusan mahasiswa, sebagai salah satu bahan evaluasi dalam penentuan kebijakan.

Penelitian lain yang berkaitan yaitu penelitian dengan judul, *Prediksi Persediaan Obat Dengan Metode Naïve Bayes*. Penelitian ini membahas mengenai prediksi stok obat yang terjual. Sehingga pihak apotek bisa mengetahui jumlah obat yang akan dipasok. [4]

Penelitian berikutnya yang terkait lainnya, yaitu berjudul teknik data mining menggunakan metode bayes classifier untuk optimalisasi pencarian pada aplikasi perpustakaan (studi kasus: perpustakaan Universitas Pasundan – Bandung). Penelitian ini menggunakan *Naïve Bayes Classifier* untuk klasifikasi dokumen (di sini dokumen berupa data buku yang ada di perpustakaan) yang akan diterapkan dalam membangun perangkat lunak pencarian pada Aplikasi Perpustakaan. Fungsinya untuk mempercepat proses pencarian data buku pada Aplikasi Perpustakaan sehingga membantu meningkatkan pelayanan pada perpustakaan. [5]

2.2. Data Mining

Data Mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. [6]

Data Mining Konsep dan Aplikasi Menggunakan MATLAB' mendefinisikan *data mining* sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Istilah *data mining* kadang disebut juga *knowledge discovery*. [7]

3. METODE PENELITIAN

3.1. Pengumpulan Data

Tahap pengumpulan data dapat dilakukan dengan cara:

- a. Observasi
- b. Wawancara.
- c. Kusioner

3.2. Persiapan Data

Data – data yang telah dikumpulkan pada tahap sebelumnya, kemudian akan dilakukan proses persiapan data, dengan cara mengumpulkan data-data nilai mahasiswa secara keseluruhan.

3.3. Pemodelan

Pemodelan merupakan tahapan yang secara langsung melibatkan *data mining*. Pemilihan teknik data mining, algoritma dan menentukan parameter dengan nilai yang optimal. Pada tahapan pemodelan, ada beberapa hal yang dilakukan antara lain, memilih teknik pemodelan, membangun model, dan menilai model. [6]

3.4. Rancangan Rekomendasi

Dari hasil rancangan pola yang telah dilakukan pada tahap sebelumnya, maka pada tahap ini dapat menghasilkan rekomendasi dari hasil analisis pola yang ada, dan diharapkan rekomendasi ini dapat menjadi acuan atau standar untuk pemilihan konsentrasi mata kuliah mahasiswa, dan diharapkan dapat menjadi pendukung keputusan dalam pemilihan konsentrasi pemilihan matakuliah mahasiswa.

4. PEMBAHASAN DAN HASIL PENELITIAN

4.1. Pengumpulan Data

Dari hasil pengumpulan data didapatkan data *sample* sebanyak 40 mahasiswa. Proses pengambilan data dilakukan dengan menggunakan *sample* pada sebuah kelas mahasiswa angkatan tahun 2013 yang terdapat pada program studi System informasi. Data yang digunakan merupakan data mahasiswa yang dinyatakan aktif melakukan kegiatan perkuliahan. Dari data yang digunakan diketahui:

- a) Terdapat 22 Mahasiswa yang memilih konsentrasi mata kuliah Rekayasa Sistem Informasi
- b) Terdapat 15 Mahasiswa yang memilih konsentrasi mata kuliah Teknologi Informasi
- c) Sedangkan terdapat 3 Mahasiswa yang belum memilih konsentrasi mata kuliah.

4.2. Persiapan Data

Data yang digunakan untuk penelitian merupakan data nilai / transkrip nilai sementara mahasiswa. Contoh data yang digunakan adalah sebagai berikut:

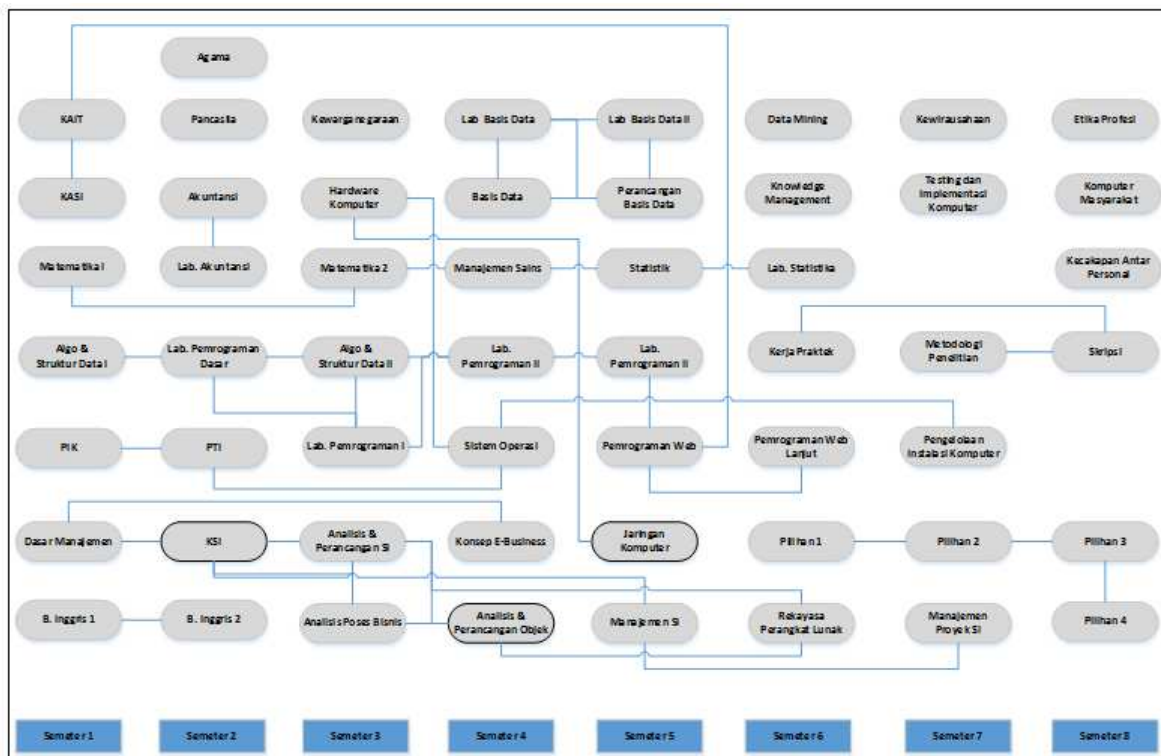
Tabel 3 Contoh Data Nilai Transkrip Mahasiswa

Tahun Akademik 2013/2014 Semester Ganjil			
Kode Mata Kuliah	Nama Mata Kuliah	SKS	Nilai
IS31271	Bahasa Inggris I	2	B
IS31272	Pengantar Ilmu Komputer	3	A
IS31273	Matematika I	3	A
IS31274	Dasar Manajemen & Bisnis	3	A
IS31275	Algoritma & Struktur Data I	3	B
IS31371L	Komputer Aplikasi SI	2	A
IS31372L	Komputer Aplikasi IT	2	B

Tahun Akademik 2013/2014 Semester Genap			
---	--	--	--

4.3. Pemodelan

Tahapan awal yang dilakukan sebelum memodelkan data dengan menggunakan data mining, yaitu mengetahui pola kurikulum yang digunakan sehingga dapat diketahui mata kuliah yang berhubungan dengan pemilihan konsentrasi pada mahasiswa semester 6. Pola dari kurikulum yang digunakan tersebut dapat dilihat pada gambar dibawah ini;



Gambar 1 Pola Kurikulum Mata Kuliah

Dari pola tersebut dapat diketahui mata kuliah yang menjadi standar awal untuk mata kuliah pilihan. Dapat dilihat pada tabel 4 berikut ini:

Tabel 4 mata kuliah yang terkait dengan pilihan konsentrasi

Teknologi Informasi	Rekayasa Sistem Informasi
1. Lab pemrograman I	a. Konsep Sistem informasi
2. Lab pemrograman II	b. Analisis Proses Bisnis
3. Lab pemrograman III	c. Konsep E-Business
4. Pemrograman WEB	d. Manajemen Sistem Informasi
5. Jaringan Komputer	

Dari data-data yang telah dikumpulkan diatas maka tahap berikutnya adalah melakukan pemisahan data *sample* nilai mahasiswa berdasarkan konsentrasi mata kuliah, agar perhitungan dengan algoritma data mining dapat dilakukan. Data nilai yang dikumpulkan kemudian di konveriskan kedalam bentuk predikat, sesuai dengan aturan yang berlaku di universitas. Konversi data *sample* berdasarkan aturan berikut:

Tabel 5 Indeks Nilai Mahasiswa

NILAI	INDEKS	PREDIKAT
$80 \leq NA \leq 100$	A	LULUS, Sangat Baik
$68 \leq NA \leq 79$	B	LULUS, Baik
$56 \leq NA \leq 67$	C	LULUS, Cukup
$45 \leq NA \leq 55$	D	LULUS, Kurang
$0 \leq NA \leq 44$	E	TIDAK LULUS

Pada tahap ini data *sample* yang digunakan kemudian dihitung dengan menggunakan *naive bayes*. Diketahui jumlah data yang digunakan sebanyak 37, dengan konsentrasi pemilihan Teknologi Informasi (TI) sebanyak 15 dan pemilihan konsentrasi Rekayasa Sistem Informasi (RSI) sebanyak 22. Perhitungan yang pertama kali dilakukan adalah menghitung $P(X)$, probabilitas untuk konsentrasi mata kuliah Teknologi Informasi dan Rekayasa Sistem Informasi.

1. $P(\text{RSI}) = \frac{22}{37} = 0,595$
2. $P(\text{TI}) = \frac{15}{37} = 0,405$

Kemudian tahap berikutnya adalah melakukan perhitungan probabilitas X berdasar kondisi pada hipotesis H, untuk setiap atribut yang ada pada matakuliah. Berikut perhitungan menggunakan *naive bayes*:

Tabel 6 Hasil perhitungan Naïve Bayes

1) $P(\text{Lab. I} = \text{Sangat Baik} \text{TI}) = \frac{11}{15} = \mathbf{0,733}$ $P(\text{Lab. I} = \text{Baik} \text{TI}) = \frac{4}{15} = \mathbf{0,267}$ $P(\text{Lab. I} = \text{Cukup} \text{TI}) = \frac{0}{15} = \mathbf{0}$	2) $P(\text{Lab. I} = \text{Sangat Baik} \text{RSI}) = \frac{18}{22} = \mathbf{0,818}$ $P(\text{Lab. I} = \text{Baik} \text{RSI}) = \frac{3}{22} = \mathbf{0,136}$ $P(\text{Lab. I} = \text{Cukup} \text{RSI}) = \frac{1}{22} = \mathbf{0,045}$
3) $P(\text{Lab. II} = \text{Sangat Baik} \text{TI}) = \frac{9}{15} = \mathbf{0,600}$ $P(\text{Lab. II} = \text{Baik} \text{TI}) = \frac{6}{15} = \mathbf{0,400}$ $P(\text{Lab. II} = \text{Cukup} \text{TI}) = \frac{0}{15} = \mathbf{0}$	4) $P(\text{Lab. II} = \text{Sangat Baik} \text{RSI}) = \frac{3}{22} = \mathbf{0,136}$ $P(\text{Lab. II} = \text{Baik} \text{RSI}) = \frac{14}{22} = \mathbf{0,636}$ $P(\text{Lab. II} = \text{Cukup} \text{RSI}) = \frac{5}{22} = \mathbf{0,227}$
5) $P(\text{Lab. III} = \text{Sangat Baik} \text{TI}) = \frac{10}{15} = \mathbf{0,667}$	6) $P(\text{Lab. III} = \text{Sangat Baik} \text{RSI}) = \frac{8}{22} = \mathbf{0,364}$

P(Lab. III= Baik TI)	= 4/15 = 0,267	P(Lab. III= Baik RSI)	= 13/22 = 0,591
P(Lab. III = Cukup TI)	= 1/15 = 0,067	P(Lab. III= Cukup RSI)	= 1/22 = 0,045
7) P(Pem.Web = Sangat Baik TI)	= 9/15 = 0,600	8) P(Pem.Web = Sangat Baik RSI)	= 8/22 = 0,364
P(Pem.Web = Baik TI)	= 6/15 = 0,400	P(Pem.Web = Baik RSI)	= 13/22 = 0,591
P(Pem.Web = Cukup TI)	= 0/15 = 0	P(Pem.Web = Cukup RSI)	= 1/22 = 0,045
9) P(Jarkom = Sangat Baik TI)	= 11/15 = 0,733	10) P(Jarkom = Sangat Baik RSI)	= 18/22 = 0,818
P(Jarkom = Baik TI)	= 3/15 = 0,200	P(Jarkom = Baik RSI)	= 3/22 = 0,136
P(Jarkom = Cukup TI)	= 1/15 = 0,067	P(Jarkom = Cukup RSI)	= 1/22 = 0,045
11)P(KSI = Sangat Baik TI)	= 7/15 = 0,467	12)P(KSI = Sangat Baik RSI)	= 10/22 = 0,455
P(KSI = Baik TI)	= 8/15 = 0,533	P(KSI = Baik RSI)	= 7/22 = 0,318
P(KSI = Cukup TI)	= 0/15 = 0	P(KSI = Cukup RSI)	= 5/22 = 0,227
13)P(APB = Sangat Baik TI)	= 5/15 = 0,333	14) P(APB = Sangat Baik RSI)	= 8/22 = 0,364
P(APB = Baik TI)	= 10/15 = 0,667	P(APB = Baik RSI)	= 14/22 = 0,636
P(APB = Cukup TI)	= 0/15 = 0	P(APB = Cukup RSI)	= 0/22 = 0
15)P(Konsep = Sangat Baik TI)	= 10/15 = 0,667	16) P(Konsep = Sangat Baik RSI)	= 16/22 = 0,727
P(Konsep = Baik TI)	= 5/15 = 0,333	P(Konsep = Baik RSI)	= 6/22 = 0,273
P(Konsep = Cukup TI)	= 0/15 = 0	P(Konsep = Cukup RSI)	= 0/22 = 0
17)P(MSI = Sangat Baik TI)	= 12/15 = 0,800	18) P(MSI = Sangat Baik RSI)	= 19/22 = 0,864
P(MSI = Baik TI)	= 2/15 = 0,133	P(MSI = Baik RSI)	= 3/22 = 0,136
P(MSI = Cukup TI)	= 1/15 = 0,067	P(MSI = Cukup RSI)	= 0/22 = 0

Kemudian dari hasil perhitungan diatas, akan diuji menggunakan data sample baru, sehingga dapat diketahui apakah data sample awal sudah sesuai. Contoh pengetesan pada data berikut:

Tabel 7 Contoh Data Uji

NIM	10513122
Lab. Pemrograman I	Sangat Baik
Lab. Pemrograman II	Cukup
Lab. Pemrograman III	Baik
Pem. Web	Sangat Baik
Jarkom	Sangat Baik
Konsep Sistem Informasi	Sangat Baik

Analisis Proses Bisnis	Sangat Baik
Konsep E-Business	Baik
Manajemen Sistem Informasi	Baik
Pilihan Konsentrasi	Dicari?

Maka setelah dihitung kemudian menghasilkan data sebagai berikut:

Tabel 8 Hasil perhitungan

Lab-Pemrograman 1	Sangat Baik	Rekayasa Sistem Informasi	0.888888888888889
		Teknologi Informasi	0.76923076923077
Lab-Pemrograman 2	Cukup	Rekayasa Sistem Informasi	0.111111111111111
		Teknologi Informasi	0
Lab-Pemrograman 3	Baik	Rekayasa Sistem Informasi	0.555555555555556
		Teknologi Informasi	0.15384615384615
Pemrograman Web	Sangat Baik	Rekayasa Sistem Informasi	0.388888888888889
		Teknologi Informasi	0.61538461538462
Jaringan Komputer	Sangat Baik	Rekayasa Sistem Informasi	0.888888888888889
		Teknologi Informasi	0.76923076923077
Konsep Sistem Informasi	Sangat Baik	Rekayasa Sistem Informasi	0.388888888888889
		Teknologi Informasi	0.38461538461538
Analisis Proses Bisnis	Sangat Baik	Rekayasa Sistem Informasi	0.388888888888889
		Teknologi Informasi	0.23076923076923
Konsep E-Business	Baik	Rekayasa Sistem Informasi	0.222222222222222
		Teknologi Informasi	0.30769230769231
Manajemen Sistem Informasi	Baik	Rekayasa Sistem Informasi	0.166666666666667
		Teknologi Informasi	0.23076923076923
Nilai Atribut	Total Nilai	Hasil Akhir	
Rekayasa Sistem Informasi	0.00010624115442692	0.0019123407796845	
Teknologi Informasi	0	0	

Dari hasil perhitungan dengan menggunakan data uji dapat dilihat pada tabel 8, sehingga dapat diperoleh kesimpulan dibawah ini:

Kesimpulan	
Mahasiswa dengan nilai :	Lab-Pemrograman 1 : Sangat Baik Lab-Pemrograman 2 : Cukup Lab-Pemrograman 3 : Baik Pemrograman Web : Sangat Baik

Kesimpulan	
	Jaringan Komputer : Sangat Baik Konsep Sistem Informasi : Sangat Baik Analisis Proses Bisnis : Sangat Baik Konsep E-Business : Baik Manajemen Sistem Informasi : Baik Direkomendasikan Untuk Memilih Pilihan Konsentrasi : Rekayasa Sistem Informasi dengan Nilai Terbesar = 0.0019123407796845

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan, dapat diambil kesimpulan sebagai berikut:

1. Analisis pemilihan konsentrasi pemilihan mata kuliah di program studi sistem informasi UNIKOM, dilakukan dengan menggunakan naïve bayes classifier.
2. Hasil pemodelan dengan menggunakan naïve bayes classifier menghasilkan rekomendasi yang dapat membantu dosen wali maupun mahasiswa memilih konsentrasi mata kuliah, berdasarkan data set (*data sample* dari angkatan sebelumnya) yang telah dijadikan acuan perhitungan.

5.2. Saran

Berdasarkan kesimpulan yang dipaparkan, hal yang diharapkan adalah sebagai berikut:

1. Penggunaan Algoritma lain untuk membuat pola atau standar dalam pemilihan konsentrasi mata kuliah untuk membandingkan hasil yang didapatkan.
2. Hasil pemodelan dapat diterapkan untuk kurikulum baru (KKNI) yang digunakan pada prodi saat ini.
3. Membangun sistem pendukung keputusan, untuk mempermudah pengolahan data untuk pemilihan konsentrasi mata kuliah.

UCAPAN TERIMA KASIH

Ucapan terimakasih ditujukan kepada Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jendral Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi yang telah mendanai penelitian ini.

DAFTAR PUSTAKA

- [1] A. Kadir, Pengenalan Sistem Informasi Edisi Revisi, Yogyakarta: Penerbit Andi, 2014.
- [2] A. P. Fadillah and B. Hardiyana, "Classification of Subject Concentration using Algorithm C4.5," in *INCITEST- IOP conference series : Science & Engineering*, Bandung, 2018.
- [3] Y. S. Nugroho, "DATA MINING MENGGUNAKAN ALGORITMA NAÏVE BAYES UNTUK KLASIFIKASI KELULUSAN MAHASISWA UNIVERSITAS DIAN NUSWANTORO," 2014.
- [4] I. C. Gumilang, D. S. M. and A. R. S. , "PREDIKSI PERSEDIAAN OBAT DENGAN METODE NAÏVE BAYES (STUDI KASUS : APOTEK SAPUTRA)," PROGRAM STUDI TEKNIK INFORMATIKA, FAKULTAS KOMUNIKASI DAN INFORMATIKA UNIVERSITAS MUHAMMADIYAH SURAKARTA, Surakarta, 2014.
- [5] S. L. B. Ginting and R. P. Trinanda, "TEKNIK DATA MINING MENGGUNAKAN METODE BAYES CLASSIFIER UNTUK OPTIMALISASI Pencarian pada Aplikasi Perpustakaan (STUDI KASUS : PERPUSTAKAAN UNIVERSITAS PASUNDAN – BANDUNG)," *JATI : Jurnal Teknologi dan Informasi UNIKOM*, vol. Volume 1 No 6, 2014.
- [6] Kusriani, Algoritma Data Mining, Yogyakarta: Penerbit Andi, 2009.

- [7] E. Prasetyo, *Data Mining konsep dan Aplikasi menggunakan MATLAB*, Yogyakarta: Penerbit Andi, 2012.
- [8] Budiman, Irawan; Prahasto, Toni; Christyono, Yuli;, “Data Clustering menggunakan metodologi CRISP –DM untuk pengenalan pola proporsi pelaksanaan Tridharma,” *Jurnal Sistem Informasi Bisnis*, vol. Volume 1 Nomor 3, pp. 129-134, 2011.
- [9] A. Prajitno, “Prediksi Kinerja Penjualan Karya Musik Menggunakan Framework CRISP-DM (studi kasus: X Music Indonesia),” 2012.
- [10] Suyanto, *Buku Data Mining untuk Klasifikasi dan Klasterisasi Data*, Bandung: Penerbit Informatika, 2017.
- [11] Kamagi, David Hartanto; Hansun, Seng;, “library.umn.ac.id,” 2014. [Online]. Available: <http://library.umn.ac.id/jurnal/public/uploads/papers/pdf/be20fbb34fbc7c5d2fe06044edf6d37e.pdf>. [Accessed Mei 2015].
- [12] K. Ahmed and T. Jesmin, “Comparative Analysis of Data Mining Classification Algorithms in Type-2 Diabetes Prediction Data Using WEKA Approach,” *International Journal of Science and Engineering (IJSE)*, vol. Vol. 7 (2), pp. 155-160, 2014.
- [13] E. d. Turban, *Decicion Support Systems and Intelligent Systems*, Andi Offset, 2005.
- [14] A. P. Fadillah, “Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ),” *JUTISI*, vol. VOL. 1, pp. 260-270, 3 Desember 2015.