

Aplikasi Pemrediksi Masa Studi dan Predikat Kelulusan Mahasiswa Informatika Universitas Muhammadiyah Surakarta Menggunakan Metode Naive Bayes

Muh Amin Nurrohmat^{1*}, Yusuf Sulisty Nugroho¹

¹Program studi Informatika
Universitas Muhammadiyah Surakarta
Surakarta

*Amin.nurrohmat@gmail.com

Abstrak

Informatika merupakan salah satu program studi di Universitas Muhammadiyah Surakarta yang memiliki data mahasiswa cukup besar baik data mahasiswa aktif maupun mahasiswa yang sudah lulus. Setiap tahun data tersebut semakin bertambah banyak. Di sisi lain, data yang dimiliki tersebut jika tidak dikelola dengan baik, maka hanya akan menjadi tumpukan data yang tidak bermanfaat, sehingga informasi yang dihasilkan juga tidak banyak. Salah satu cara untuk mengatasi masalah ini adalah data tersebut diubah menjadi sebuah informasi yang bersifat strategis. Oleh karena itu, tujuan penelitian ini adalah membuat suatu aplikasi untuk melakukan prediksi terhadap lama studi dan predikat kelulusan mahasiswa dengan menerapkan teknik data mining. Metode aplikasi yang digunakan dalam data mining untuk prediksi adalah algoritma naive bayes. Hal ini digunakan untuk menganalisis data, terutama dalam proses pengenalan pola, memprediksi masa studi dan predikat kelulusan. Setelah pengolahan data, aplikasi akan menampilkan laporan serta laporan ringkasan. Hasil penelitian menunjukkan bahwa aplikasi dapat digunakan untuk membantu program studi informatika dalam rangka menemukan informasi strategis terkait dengan lama studi dan predikat kelulusan mahasiswa.

1. PENDAHULUAN

Dengan perkembangan teknologi informasi yang sangat pesat, informasi menjadi elemen penting dalam perkembangan kehidupan saat ini dan waktu mendatang. Data yang dihasilkan teknologi informasi sangat besar, termasuk bidang pendidikan, ekonomi, industry, dan berbagai bidang lainnya sehingga keperluan akan informasi yang tinggi tetapi tidak diimbangi dengan penyajian informasi yang memadai. Dari data yang jumlahnya sangat besar ini tentu belum digunakan secara optimal, sehingga perlu adanya penggalian ulang dari data yang jumlahnya sangat besar untuk menyajikan informasi yang dibutuhkan.

Dalam dunia pendidikan, data yang berlimpah dan berkesinambungan mengenai siswa yang dibina dan alumni terus dihasilkan. Menurut Jing dan Merceron dalam Ayub[1], data yang berlimpah membuka peluang diterapkannya *data mining* untuk pengelolaan pendidikan yang lebih baik dan *data mining* dalam pelaksanaan pembelajaran berbantuan komputer yang lebih efektif. Di sisi lain, Nugroho dan Setyawan[2] menyatakan bahwa Fakultas Komunikasi dan Informatika UMS sejak berdiri

pada tahun 2006 telah memiliki sebanyak 2358 mahasiswa termasuk yang sudah lulus sebanyak kurang lebih 600-700 mahasiswa. Dengan demikian data-data akademik mahasiswa yang ada juga cukup banyak. Jika data yang melimpah ini hanya dibiarkan menumpuk, maka hanya akan menjadi beban *database* yang dimiliki. Sementara itu, data-data yang melimpah ini sebenarnya bisa dimanfaatkan sebagai sumber informasi strategis bagi program studi untuk melakukan klasifikasi masa studi lulusan mahasiswa dengan menggunakan teknik data mining. Hal ini tentunya selain dapat memberikan informasi yang bersifat strategis bagi fakultas dan program studi, juga dapat meningkatkan upaya untuk mendorong dan mempercepat kelulusannya. Sehingga selain dapat bermanfaat bagi mahasiswa sendiri, juga dapat meningkatkan nilai akreditasi bagi program studi.

Menurut buku panduan akademik Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta[4], dalam BAB II disebutkan bahwa "Beban studi mahasiswa program Strata satu (S-1) yang harus ditempuh minimal 144 satuan kredit mahasiswa (SKS) dan maksimal 146 satuan kredit mahasiswa (SKS), dengan batas waktu kurang dari 8 semester dan waktu paling

lama 14 semester”. “Predikat Kelulusan untuk mahasiswa sarjana Strata satu (S-1) yaitu : Memuaskan, Sangat Memuaskan, dan Dengan Pujian (Cumlaude)”.

Berdasarkan permasalahan tersebut, dalam penelitian ini akan dibangun sebuah aplikasi untuk membantu program studi dalam memprediksi lama studi dan predikat kelulusan mahasiswa dengan teknik data mining yaitu menggunakan metode Naive Bayes. Aplikasi tersebut diharapkan dapat digunakan untuk menemukan informasi tentang lama studi dan predikat kelulusan sehingga dapat membantu program studi Informatika dalam mencari solusi dan kebijakan untuk meningkatkan prestasi mahasiswa agar dapat menyelesaikan studinya tepat waktu.

2. METODE

2.1 PENGUMPULAN DATA DAN PENENTUAN ATRIBUT

Dalam penelitian ini diperlukan data dari seluruh mahasiswa Informatika baik yang sudah lulus maupun yang belum lulus. Semua data diperoleh dari Biro Administrasi Akademik (BAA) dan Laboratorium Fakultas Komunikasi dan Informatika. Data yang digunakan dalam penelitian ini berhubungan dengan data kelulusan dan data akademik mahasiswa.

2.2 ATRIBUT DATA KELULUSAN

Data kelulusan yang digunakan sebagai data pelatihan adalah data mahasiswa yang sudah lulus antara angkatan 2007 dan 2010. Atribut yang digunakan dapat dilihat dalam tabel 1.

Tabel 1. Atribut pada Data Kelulusan

Atribut	Keterangan
Jurusan SMA (X1)	Jurusan yang diambil saat di sekolah menengah lanjutan
Gender (X2)	Jenis kelamin mahasiswa
Daerah Asal (X3)	Daerah asal mahasiswa
Asal Sekolah (X4)	Daerah asal sekolah menengah lanjutan
Asisten (X5)	Peran mahasiswa sebagai asisten dalam proses praktikum
Lama Studi (Y1)	Masa studi yang dihitung mulai saat terdaftar sebagai mahasiswa sampai dinyatakan lulus
Predikat Kelulusan (Y2)	Predikat untuk menyatakan kemajuan hasil studi mahasiswa yang dapat dihitung berdasarkan nilai masing-masing mata kuliah dikalikan dengan jumlah satuan kredit semester (SKS) mata kuliah yang diambil pada periode tertentu dibagi dengan jumlah seluruh SKS mata kuliah

2.3 ATRIBUT DATA MAHASISWA

Data mahasiswa yang digunakan sebagai data uji adalah data mahasiswa yang masih aktif mengikuti perkuliahan hingga semester genap 2014/2015, dan

diambil sampel secara acak. Atribut yang digunakan dapat dilihat dalam tabel 2.

Tabel 2. Atribut pada Data Mahasiswa

Atribut	Keterangan
Jurusan Asal Sekolah	Jurusan yang diambil saat duduk di sekolah menengah lanjutan
Gender	Jenis kelamin mahasiswa
Daerah Asal	Daerah asal mahasiswa
Asal Sekolah	Daerah asal sekolah menengah lanjutan
Asisten Lab	Peran mahasiswa sebagai asisten dalam proses praktikum

2.4 PENENTUAN NILAI CLASS DATA

Nilai *class* data penelitian yang telah dikumpulkan kemudian diubah menyesuaikan kebutuhan data mining. Tabel 3 menunjukkan nilai *class* masing-masing atribut yang digunakan.

Nilai *class* pada atribut Lama Studi dikategorikan berdasarkan semester yang ditempuh pada saat lulus, yaitu :

- 1) Tepat waktu, jika lama studi 8 semester atau kurang dari 8 semester.
- 2) Terlambat, jika lama studi lebih dari 8 semester.

Sedangkan nilai *class* pada atribut Predikat Kelulusan ditentukan berdasarkan IPK yang dibagi menjadi tiga, yaitu :

- 1) Memuaskan, jika IPK antara 2,00 dan 2,75
- 2) Sangat Memuaskan, jika IPK antara 2,76 dan 3,50
- 3) Cumlaude, jika IPK antara 3,51 dan 4,00

Tabel 3. Atribut dan Nilai Class dalam Data

Atribut	Class dalam Data
Jurusan SMA	IPA, IPS dan Lain
Gender	Pria dan Wanita
Daerah asal	WIB, WITA, dan WIT
Asal sekolah	WIB, WITA, dan WIT
Asisten Lab	Ya dan Tidak
Lama studi	Tepat dan Terlambat
Predikat Kelulusan	Cumlaude, Sangat Memuaskan, Memuaskan

2.5 ANALISIS DATA

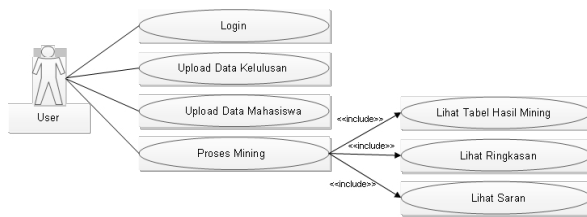
Tahap analisis data dilakukan perhitungan menggunakan algoritma *Naive Bayes*. Algoritma *naive bayes* merupakan pengklasifikasian dengan metode probabilitas dan *statistic* untuk memprediksi peluang di masa depan[4]. Adapun algoritma ini dapat dihitung menggunakan persamaan 1.

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

2.6 PERANCANGAN USE CASE DIAGRAM

Aplikasi yang dibangun memiliki tujuan untuk melakukan analisis data mahasiswa terkait dengan

kelulusannya. Perancangan aplikasi dimulai dari tahap pembuatan *use case diagram* seperti yang ditunjukkan pada gambar 1 dan tabel 4.



Gambar 1. Use Case Diagram dengan User sebagai Aktor

Tabel 4. Keterangan use case diagram

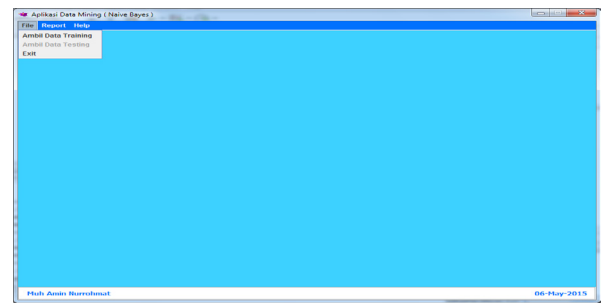
Nama Use Case	Aktor	Deskripsi
Login	User	Menggambarkan kegiatan input username dan password
Ambil data training	User	Menggambarkan kegiatan mengambil data kelulusan
Ambil data testing	User	Menggambarkan kegiatan mengambil data mahasiswa aktif
Proses Mining	User	Menggambarkan kegiatan proses mining atribut lama studi dan IPK
Laporan hasil Mining	User	Menggambarkan kegiatan menampilkan hasil mining
Laporan Rangkuman	User	Menggambarkan kegiatan menampilkan rangkuman dari proses mining
Laporan saran	User	Menggambarkan kegiatan melihat saran yang diberikan untuk prodi

3. HASIL

Hasil dari penelitian ini adalah sebuah aplikasi *data mining* untuk memprediksi lama studi dan predikat kelulusan mahasiswa program studi Informatika menggunakan metode *Naive Bayes*. Aplikasi ini memiliki form-form yang bisa diakses oleh *user* untuk melakukan prediksi, antara lain:

3.1 FORM UTAMA

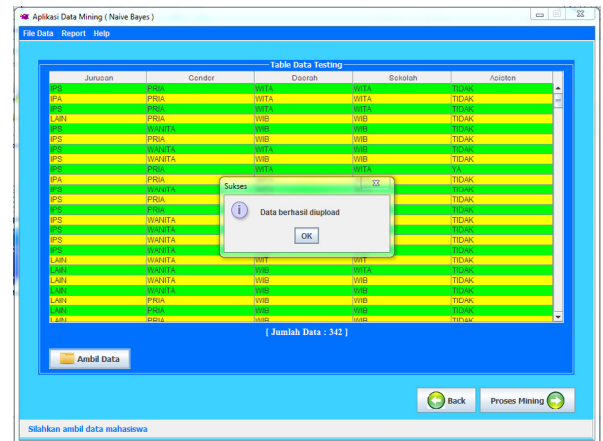
Form utama akan muncul jika *user* sukses melakukan login. Form yang ditunjukkan seperti gambar 2 ini berfungsi untuk mengakses menu-menu yang terdapat dalam aplikasi antara lain Menu File Data yang berisi submenu Upload data training, Upload data testing, Exit, Menu Report dan Menu Help berisi sub menu Petunjuk dan About.



Gambar 2. Form Utama Aplikasi Prediksi Masa Studi dan Predikat Kelulusan

3.2 FORM UPLOAD DATA TRAINING

Form *Upload Data Training* yang ditunjukkan pada gambar 3 akan ditampilkan setelah *user* memilih submenu *Upload Data Training* pada Menu File. *User* dapat mengambil data kelulusan kemudian menyimpannya dalam tabel. Jika berhasil, *user* dapat melanjutkan proses mining dan dilanjutkan ke form *Upload Data Testing*.



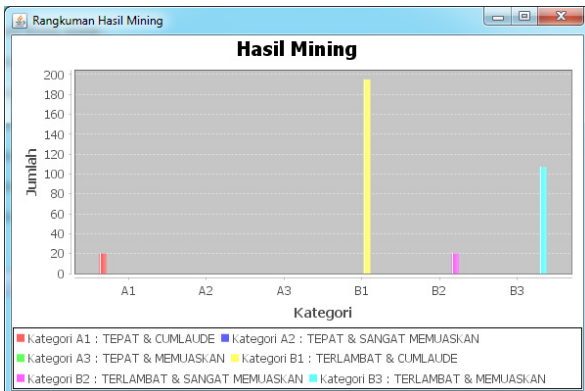
Gambar 3. Form Upload Data Training

3.3 FORM HASIL MINING

Form *Upload Data Testing* akan ditampilkan setelah *user* mengambil data kelulusan pada form *Upload Data Training*. *Data training* dan *data testing* siap diproses *mining* menggunakan algoritma *Naive Bayes* dan hasil proses mining tersebut akan muncul pada form Hasil Mining seperti yang ditunjukkan pada gambar 4 dan bisa dilihat dalam bentuk grafik seperti pada gambar 5.

Jurusan	Gender	Daerah	Sekolah	Asisten	Lama	IPK
IPA	PRIA	WITA	WITA	TIDAK	TERLAMBAT	CUMLAUDE
IPS	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
LAIN	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
IPS	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
IPS	WANITA	WIB	WIB	TIDAK	TERLAMBAT	CUMLAUDE
IPS	PRIA	WITA	WITA	TIDAK	TERLAMBAT	CUMLAUDE
IPS	PRIA	WITA	WITA	TIDAK	TERLAMBAT	MEMUASKAN
IPS	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
IPS	WANITA	WIB	WITA	TIDAK	TERLAMBAT	CUMLAUDE
IPS	PRIA	WITA	WITA	TIDAK	TERLAMBAT	MEMUASKAN
IPS	WANITA	WIT	WIT	TIDAK	TERLAMBAT	MEMUASKAN
LAIN	WANITA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
LAIN	WANITA	WIB	WIB	TIDAK	TERLAMBAT	CUMLAUDE
LAIN	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
LAIN	PRIA	WIB	WIB	TIDAK	TERLAMBAT	MEMUASKAN
LAIN	PRIA	WIB	WIT	TIDAK	TERLAMBAT	MEMUASKAN

Gambar 4. Form Hasil Mining



Gambar 5. Grafik Hasil Mining

Tabel 5. Potongan Data Training

Jurusan SMA	Gender	Daerah Asal	Asal Sekolah	Asisten	Lama Studi	Predikat Kelulusan (IPK)
IPA	Pria	WIB	WIB	Ya	Tepat	Cumlaude
IPS	Wanita	WIB	WIB	Ya	Tepat	Cumlaude
IPA	Pria	WIB	WIB	Ya	Tepat	S a n g a t Memuaskan
IPS	Pria	WIT	WIT	Ya	Terlambat	Memuaskan
IPA	Pria	WITA	WITA	Ya	Terlambat	Memuaskan
IPA	Pria	WIB	WIB	Tidak	Tepat	Cumlaude
IPS	Pria	WITA	WIB	Tidak	Terlambat	Memuaskan
IPS	Wanita	WIB	WIB	Tidak	Terlambat	S a n g a t Memuaskan
IPA	Pria	WITA	WITA	Ya	Terlambat	Memuaskan
Lain	Pria	WITA	WITA	Ya	Terlambat	Memuaskan

4. DISKUSI

Algoritma Naive Bayes digunakan untuk melakukan perhitungan terhadap probabilitas nilai class data dalam data testing untuk setiap variabel dependen (Y) berdasarkan data training. Data mahasiswa yang telah lulus diatur sebagai data training dan data mahasiswa aktif sebagai data testing yang dipilih secara acak. Tabel 5 adalah contoh potongan data training yang diambil secara random, sedangkan tabel 6 merupakan contoh data testing yang digunakan sebagai hipotesis.

Untuk mencari nilai class variabel Lama Studi, diperlukan perhitungan probabilitas prior terhadap data training, dengan tahapan sebagai berikut:

4.1 MENGHITUNG PROBABILITAS PRIOR MASING-MASING CLASS VARIABEL LAMA STUDI (Y1).

Pada tahap ini dilakukan perhitungan, persamaan (2) dan (3), untuk mencari nilai probabilitas masing-masing class yaitu “Tepat” dan “Terlambat” pada variabel Y1.

$$P(Y1 = \text{Tepat}) = \frac{\text{Jumlah class "Tepat"}}{\text{Jumlah data}} = \frac{4}{10} \tag{2}$$

Tabel 6. Contoh Data *Testing*

Jurusan SMA	Gender	Daerah Asal	Asal Sekolah	Asisten	Lama Studi	IPK
IPA	Pria	WIB	WIB	Ya	?	?

$$P(Y1 = Terlambat) = \frac{\text{Jumlah class "Terlambat"}}{\text{Jumlah data}} = \frac{6}{10} \quad (3)$$

4.2 MENGHITUNG PROBABILITAS BERSYARAT SESUAI NILAI CLASS YANG DIAJUKAN DALAM DATA TESTING.

Pada tahap ini dilakukan perhitungan, persamaan (4)-(8), untuk mencari nilai probabilitas bersyarat masing-masing *class* dalam variabel X terhadap variabel Y1.

$$P(X1 = IPA|Y1 = Tepat) = \frac{3}{4}; P(X1 = IPA|Y1 = Terlambat) = \frac{2}{6} \quad (4)$$

$$P(X2 = Pria|Y1 = Tepat) = \frac{3}{4}; P(X2 = Pria|Y1 = Terlambat) = \frac{5}{6} \quad (5)$$

$$P(X3 = WIB|Y1 = Tepat) = \frac{4}{4}; P(X3 = WIB|Y1 = Terlambat) = \frac{1}{6} \quad (6)$$

$$P(X4 = WIB|Y1 = Tepat) = \frac{4}{4}; P(X4 = WIB|Y1 = Terlambat) = \frac{2}{6} \quad (7)$$

$$P(X5 = Ya|Y1 = Tepat) = \frac{3}{4}; P(X5 = Ya|Y1 = Terlambat) = \frac{3}{6} \quad (8)$$

4.3 PERHITUNGAN NAÏVE BAYES.

Algoritma naïve bayes digunakan untuk mencari nilai *confidence* atau peluang terjadinya masing-masing *class* variabel Y1 berdasarkan variabel X yang diajukan, persamaan (9) dan (10), Nilai *confidence* masing-masing *class* tersebut kemudian dibandingkan untuk mengetahui peluang yang lebih besar untuk memprediksi nilai *class* yang muncul pada variabel Y1.

- a) Nilai *confidence* pada variabel Y1 untuk *class* = "Tepat"

$$P(X1 = IPA, X2 = Pria, X3 = WIB, X4 = WIB, X5 = Ya|Y1 = Tepat) = \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{4}{4} \cdot \frac{4}{4} \cdot \frac{3}{10} = 0,168 \quad (9)$$

- b) Nilai *confidence* pada variabel Y1 untuk *class* = "Terlambat"

$$P(X1 = IPA, X2 = Pria, X3 = WIB, X4 = WIB, X5 = Ya|Y1 = Terlambat) = \frac{2}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{2}{6} \cdot \frac{3}{10} = 0,00432 \quad (10)$$

Berdasarkan nilai *confidence* masing-masing *class* pada variabel Y1 dapat dilihat bahwa nilai *confidence* untuk *class* = "Tepat" lebih besar daripada nilai *confidence* pada *class* = "Terlambat". Dengan demikian, prediksi *class* variabel Y1 pada data *testing* adalah "Tepat".

Sedangkan untuk mencari nilai *class* variabel Predikat Kelulusan, diperlukan perhitungan probabilitas *prior* terhadap data *training*, dengan tahapan sebagai berikut:

4.4 MENGHITUNG PROBABILITAS PRIOR MASING-MASING CLASS VARIABEL PREDIKAT KELULUSAN (Y2).

Pada tahap ini dilakukan perhitungan, persamaan (11)-(13), untuk mencari nilai probabilitas masing-masing *class* yaitu "Cumlaude", "Sangat Memuaskan" dan "Memuaskan" pada variabel Y2.

$$P(Y2 = Cumlaude) = \frac{\text{Jumlah class "Cumlaude"}}{\text{Jumlah data}} = \frac{3}{10} \quad (11)$$

$$P(Y2 = Sangat Memuaskan) = \frac{\text{Jumlah class "Sangat Memuaskan"}}{\text{Jumlah data}} = \frac{2}{10} \quad (12)$$

$$P(Y2 = Memuaskan) = \frac{\text{Jumlah class "Memuaskan"}}{\text{Jumlah data}} = \frac{5}{10} \quad (13)$$

4.5 MENGHITUNG PROBABILITAS BERSYARAT SESUAI NILAI CLASS YANG DIAJUKAN DALAM DATA TESTING.

Pada tahap ini dilakukan perhitungan, persamaan (14)-(21), untuk mencari nilai probabilitas bersyarat masing-masing *class* dalam variabel X terhadap variabel Y2.

$$P(X1 = IPA|Y2 = Cumlaude) = \frac{2}{3}; P(X1 = IPA|Y2 = Sangat Memuaskan) = \frac{1}{2} \quad (14)$$

$$P(X1 = IPA|Y2 = Memuaskan) = \frac{2}{5}; P(X2 = Pria|Y2 = Cumlaude) = \frac{2}{3} \quad (15)$$

$$P(X2 = Pria|Y2 = Sangat Memuaskan) = \frac{1}{2}; P(X2 = Pria|Y2 = Memuaskan) = \frac{5}{5} \quad (16)$$

$$P(X3 = WIB|Y2 = Cumlaude) = \frac{3}{3}; P(X3 = WIB|Y2 = Sangat Memuaskan) = \frac{2}{7} \quad (17)$$

$$P(X3 = WIB|Y2 = Memuaskan) = \frac{0}{5}; P(X4 = WIB|Y2 = Cumlaude) = \frac{3}{3} \quad (18)$$

$$P(X4 = WIB|Y2 = Sangat Memuaskan) = \frac{2}{2}; P(X4 = WIB|Y2 = Memuaskan) = \frac{1}{5} \quad (19)$$

$$P(X5 = Ya|Y2 = Cumlaude) = \frac{2}{3}; P(X5 = Ya|Y2 = Sangat Memuaskan) = \frac{1}{2} \quad (20)$$

$$P(X5 = Ya|Y2 = Memuaskan) = \frac{4}{5} \quad (21)$$

4.6 PERHITUNGAN NAÏVE BAYES.

Nilai *confidence* masing-masing *class* dalam variabel Y2, persamaan (22)-(24), kemudian dibandingkan untuk mengetahui peluang yang lebih besar untuk memprediksi nilai *class* yang muncul pada variabel Y2 tersebut.

- a) Nilai *confidence* pada variabel Y2 untuk *class* = "Cumlaude"

$$P(X1 = IPA, X2 = Pria, X3 = WIB, X4 = WIB, X5 = Ya|Y2 = Cumlaude)$$

$$= \frac{2}{3} \cdot \frac{2}{3} \cdot \frac{3}{3} \cdot \frac{3}{3} \cdot \frac{2}{3} \cdot \frac{3}{10} = 0,097746 \quad (22)$$

- b) Nilai *confidence* pada variabel Y2 untuk *class* = ‘Sangat Memuaskan’

$$P(X1 = IPA, X2 = Pria, X3 = WIB, X4 = WIB, X5 = Ya | Y2 = Sangat Memuaskan)$$

$$= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{2} \cdot \frac{2}{2} \cdot \frac{1}{2} \cdot \frac{2}{10} = 0,025 \quad (23)$$

- c) Nilai *confidence* pada variabel Y2 untuk *class* = ‘Memuaskan’

$$P(X1 = IPA, X2 = Pria, X3 = WIB, X4 = WIB, X5 = Ya | Y2 = Memuaskan)$$

$$= \frac{2}{5} \cdot \frac{5}{5} \cdot \frac{0}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{5}{10} = 0 \quad (24)$$

Berdasarkan nilai *confidence* masing-masing *class* pada variabel Y2 dapat dilihat bahwa nilai *confidence* untuk *class* = ‘Cumlaude’ lebih besar daripada nilai *confidence* pada *class* = ‘Sangat Memuaskan’ dan *class* = ‘Memuaskan’. Dengan demikian, prediksi *class* variabel Y2 pada data *testing* adalah “Cumlaude”.

5. KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan maka dapat disimpulkan bahwa aplikasi *Data Mining* ini dapat digunakan untuk menampilkan informasi hasil prediksi

lama studi dan predikat kelulusan dalam bentuk tabel dengan menggunakan data mahasiswa yang telah lulus sebagai data *training* dan data mahasiswa yang masih aktif sebagai data *testing*. Variabel bebas (X) yang terdiri dari jurusan asal sekolah, gender, daerah asal, asal sekolah, dan asisten dapat digunakan sebagai variabel untuk memprediksi lama studi dan predikat kelulusan mahasiswa.

Algoritma naïve bayes dapat digunakan untuk memprediksi suatu nilai *class* variabel dalam data *testing* berdasarkan perbandingan nilai probabilitas masing-masing *class* variabel pada data *training* terhadap variabel X yang diajukan dalam data *testing*.

DAFTAR PUSTAKA

- [1] Ayub, Mewati, “Proses Data Mining dalam Sistem Pembelajaran Berbantuan Komputer”, Jurnal Sistem Informasi Vol. 2 No. 1 Maret 2007 : 21-30
- [2] Nugroho, Yusuf Sulisty, dan Setyawan., “Klasifikasi Masa Studi Mahasiswa Fakultas Komunikasi dan Informatika”, Jurnal Komunikasi dan Teknologi Informasi (KomuniTi), Volume VI No. I Maret 2014.
- [3] UMS, “Panduan Akademik Fakultas Komunikasi dan Informatika UMS 2014/2015”, Universitas Muhammadiyah Surakarta., Surakarta, 2014
- [4] Bustami, “Penerapan Algoritma Naive Bayes untuk Mengklasifikasi Data Nasabah Asuransi”, Jurnal Penelitian Teknik Informatika, 2013