

# IMPLEMENTASI METODE NAÏVE BAYES PADA SPK UNTUK MEMPREDIKSI POLA KELULUSAN MAHASISWA PERGURUAN TINGGI SWASTA

Rogayah<sup>1</sup>, Julia Fajaryanti<sup>2</sup>, Lintang Yuniar Banowosari<sup>3</sup>

<sup>1,2</sup>Teknik Informatika, Teknologi Industri, Universitas Gunadarma

<sup>3</sup>Sistem Informasi, Ilmu Komputer dan Teknologi Informasi, Universitas Gunadarma

Email: <sup>1</sup>rogayah@staff.gunadarma.ac.id, <sup>2</sup>julia@staff.gunadarma.ac.id, <sup>3</sup>lintang@staff.gunadarma.ac.id

**Abstrak** - Mahasiswa adalah aspek penting dalam evaluasi keberhasilan penyelenggaraan program studi pada suatu perguruan tinggi. Dalam melakukan pengambilan keputusan yang dilakukan oleh kepala program studi dibutuhkan data mahasiswa dan data jumlah kelulusan. Untuk meningkatkan pelayanan dan kinerja program studi objek penelitian dalam pengambilan keputusan dan kebijakan bagi mahasiswa dibutuhkan sistem yang dapat membantu dalam proses pengolahan data mahasiswa. Dalam mengolah data, dibutuhkan basis data akademik mahasiswa pada program studi yang terjadi dalam kurun waktu empat (4) tahun. Penggunaan data mahasiswa bertujuan untuk membangun sebuah sistem penunjang keputusan yang dapat menghasilkan klasifikasi atau pola-pola mahasiswa yang lulus tepat waktu berdasarkan hasil penambangan data menggunakan metode Naïve Bayes. Metode penelitian yang digunakan meliputi pembahasan tempat, waktu dan jadwal penelitian, alat dan bahan penelitian, metodologi penelitian, jenis data, dan teknik pengumpulan dan analisis data. Hasil penelitian ini adalah pola kelulusan mahasiswa sesuai pola yang dihasilkan oleh metode Naïve Bayes sehingga membantu program studi objek penelitian untuk melakukan langkah-langkah tepat bagi mahasiswa yang termasuk ke dalam mahasiswa tidak dapat lulus tepat waktu.

**Kata Kunci** - Data Mining, Kelulusan, Mahasiswa, Naïve Bayes

**Abstract** - Students are an important aspect in the evaluation of the success of the Program Study at a University. In conducting the decisions made by the head of the study program takes the student data and data on the number of graduation. To improve services and performance of the object of study program of research in making decision and get the policy for students needed a system that could help in the processing of student data. Program Study

needs data bases on the student's academic courses that take place within a period of four (4) years. The use of student data aims to build a decision support system that can produce a classification or patterns of students who graduate on time based on the results of data mining using Naïve Bayes method. The method used includes the discussion of the place, time and schedule of research, tools and materials research, research methodology, data types, and data collection and analysis techniques. The results of this study is graduation corresponding patterns produced by Naïve Bayes method that helps the research object of study program to perform the appropriate steps for students who belong to the student cannot graduate on time.

**Keywords:** Data Mining, Graduation, Naïve Bayes, Students

## 1. PENDAHULUAN

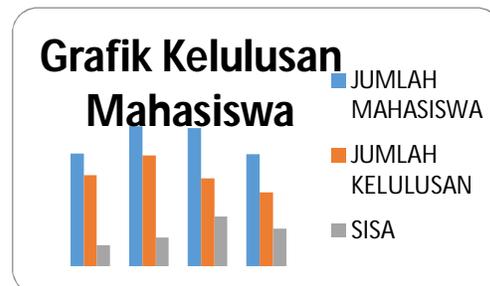
Pendidikan merupakan salah satu hal yang penting dalam kehidupan di Negara ini. Salah satu lembaga atau institusi yang menyelenggarakan kegiatan pendidikan bagi mahasiswa adalah perguruan tinggi. Lima lembaga perguruan tinggi diantaranya adalah universitas, institut, sekolah tinggi, akademi, dan politeknik. Data yang diperoleh dari Pusat Statistik Pendidikan Badan Penelitian dan Pengembangan Departemen Pendidikan Nasional Republik Indonesia ([www.pdsp.kemdikbud.go.id](http://www.pdsp.kemdikbud.go.id)) menyebutkan bahwa jumlah lembaga penyelenggara perguruan tinggi mengalami peningkatan setiap tahunnya. Sampai dengan tahun 2014 tercatat 3151 perguruan tinggi diselenggarakan di Indonesia. Perguruan tinggi yang baik adalah perguruan tinggi yang dapat mengelola semua program studi agar dapat menghasilkan mahasiswa yang dapat lulus tepat waktu dalam menyelesaikan masa studinya dalam bentuk presentasi kelulusan akademik yang nantinya

berpengaruh pada penentuan akreditasi program studi tersebut.

Mahasiswa adalah aspek penting dalam evaluasi keberhasilan penyelenggaraan program studi pada suatu perguruan tinggi, yang setiap tahunnya data mahasiswa dan data jumlah kelulusan mahasiswa dapat menghasilkan informasi yang berlimpah berupa jumlah kelulusan dan hasil akademik mahasiswa selama menempuh proses kegiatan belajar mengajar di perguruan tinggi, sehingga sangat bermanfaat jika tidak dibiarkan menumpuk dan dilakukannya proses penggalan data yang lebih serius untuk memperoleh pola atau pengetahuan untuk mengambil keputusan yang tepat bagi mahasiswa dalam sebuah program studi diploma pada perguruan tinggi swasta di kota Depok yang selanjutnya disebut program studi objek penelitian. Begitu pula hal yang terjadi pada program studi objek penelitian, yang telah menggunakan Sistem Informasi Akademik (SIA) tetapi data yang diperoleh hanya dapat digunakan untuk beberapa hal saja, misalnya untuk keperluan mendaftar sidang penulisan ilmiah/tugas akhir, memverifikasi nilai akademik, dan konsultasi akademik. Pengambilan keputusan yang berhubungan dengan mahasiswa belum sepenuhnya terbantu dengan pengetahuan data yang jumlahnya semakin hari semakin bertambah. Serangkaian proses mendapatkan pengetahuan atau pola dari kumpulan data disebut dengan *data mining* [1]. *Data mining* memecahkan masalah dengan menganalisis data yang telah ada dalam basis data.

Program studi objek penelitian perlu melakukan pemetaan perilaku mahasiswa untuk mencegah secara dini kegagalan akademik mahasiswa. Faktor-faktor yang menyebabkan kegagalan adalah rendahnya kemampuan akademik, pembiayaan, domisili, dan faktor lainnya [2]. Hal inilah yang sekarang ini menjadi perhatian penting bagi program studi untuk meningkatkan kelulusan mahasiswa dengan cara memberi motivasi, kegiatan, dan arahan yang sesuai dengan klasifikasi atau pola kelulusan mahasiswa. Beberapa metode klasifikasi *data mining* yang banyak digunakan adalah Metode *Decision Tree*, *K-Nearest Neighbour*, Metode *C45*, *Naïve Bayes*, dan *Neural Network*. Rujukan [2] menyatakan bahwa Metode *Naïve Bayes* dapat digunakan untuk mengetahui faktor-faktor atau karakteristik penting dari data akademik dan non akademik mahasiswa yang dapat membangkitkan pola untuk memprediksi kemajuan studi mahasiswa berupa kategori IPK, sehingga dapat dijadikan sebagai sebuah sistem

peringatan dini (*early warning*) bagi mahasiswa yang masuk ke dalam klasifikasi mahasiswa yang berpotensi untuk lulus dengan IPK kategori rendah yang bisa menurunkan mutu lulusan sebuah Perguruan Tinggi. Program studi objek penelitian memiliki data mengenai jumlah mahasiswa yang terdaftar pada program studi, jumlah mahasiswa yang sudah lulus, dan juga jumlah mahasiswa belum lulus kuliah digambarkan dalam sebuah grafik kelulusan mahasiswa yang dapat dilihat pada gambar 1 di bawah ini:



Gambar 1. Grafik Kelulusan Mahasiswa Dari Th. 2008 s/d 2011 Per Tanggal 31 Agustus 2014  
(Sumber: Program Studi Objek Penelitian)

Grafik kelulusan pada gambar 1 menunjukkan bahwa antara jumlah mahasiswa dari tahun 2008 sampai dengan 2011, yang terdaftar dikurangi dengan jumlah mahasiswa yang lulus masih terdapat sisa mahasiswa yang belum lulus. Apalagi jika dilihat pada tahun 2010, jumlah mahasiswa yang belum lulus sebanyak 295 orang dari jumlah mahasiswa sebanyak 866 orang sehingga program studi wajib bekerja keras untuk menghasilkan mahasiswa yang lulus pada tahun 2012 mencapai lebih dari 90% terhadap jumlah mahasiswa yang terdaftar.

Proses testing yang terdapat pada *data mining* dapat digunakan untuk memprediksi data siswa tentang tingkat kelulusan yang diperoleh melalui pemetaan pola yang tersusun atas atribut yang digunakan, yaitu tempat lahir, dan SMP, dan jurusan studi. Atribut dipilih karena memiliki nilai perkalian *support confidence* yang tinggi dibandingkan atribut yang lain. Nilai masing-masing atribut tersebut berdasarkan perhitungan antara lain tempat lahir = 39,083, dan SMP = 23,828, dan jurusan studi = 31,464 [3].

Penelitian ini dilakukan untuk membangun sebuah sistem penunjang keputusan yang dapat digunakan sebagai sarana untuk menentukan langkah dan kebijakan bagi mahasiswa yang ditargetkan untuk lulus tepat waktu (sesuai dengan periode sidang yang ditetapkan oleh perguruan tinggi) sesuai dengan tujuan program studi objek penelitian.

## 2. METODE PENELITIAN

### 2.1 Objek dan Tahapan Penelitian

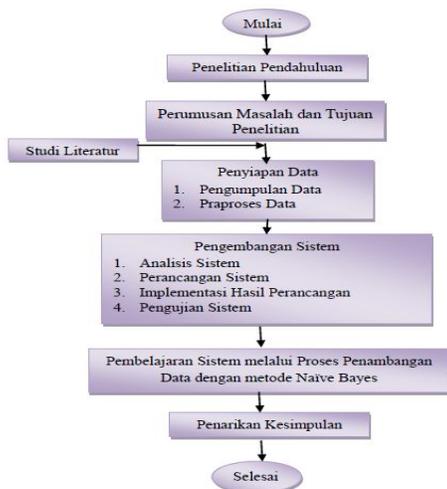
Penelitian dilakukan di sebuah program studi objek penelitian di sebuah Perguruan Tinggi Swasta, kota Depok, dengan alamat Jalan Margonda Raya No. 100, Pondok Cina, Depok 16424. Tahapan proses pelaksanaan penelitian mulai indentifikasi masalah sampai penarikan kesimpulan yang disajikan dalam diagram alir tahapan penelitian pada gambar 2 berikut ini:

Gambar 2. Tahapan Pada Metodologi Penelitian

### 2.2 Analisa Sistem

#### 2.2.1 Analisa Kebutuhan Sistem

Pada sistem yang dibuat, data yang digunakan tersedia pada Sistem Informasi Akademik dalam bentuk tabel-tabel. Adapun tabel-tabel yang tersedia yaitu tabel data master, tabel rangkuman nilai lokal, tabel rangkuman



nilai utama, tabel mahasiswa, dan tabel kelulusan.

#### 2.2.2 Analisa Sistem yang Diusulkan

Variabel yang digunakan dalam penelitian ini meliputi variabel *input* dan variabel *output* atau target. Hubungan yang ditemukan dari sejumlah variabel input dengan variabel target dinyatakan sebagai aturan (*rule*) untuk memprediksi apakah mahasiswa lulus tepat waktu atau tidak tepat waktu. Variabel input dalam penelitian ini adalah: (1) IPK; (2) SKS; (3) Kota Domisili; dan (4) Jenis Sekolah Asal.

### 2.3 Teknik Klasifikasi Data Mining

#### 2.3.1 Metode Naïve Bayes

Metode Naïve Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes* yang menciptakan metode Bayes persamaan 1 berikut:

$$P(X_k | Y) = \frac{P(Y | X_k)}{\sum_i P(Y | X_i)} \quad (1)$$

Dimana, keadaan Posterior (Probabilitas  $X_k$  di dalam  $Y$ ) dapat dihitung dari keadaan prior (Probabilitas  $Y$  di dalam  $X_k$  dibagi dengan jumlah dari semua probabilitas  $Y$  di dalam semua  $X_i$ ).

#### 2.3.2 HMAP (Hypothesis Maximum Appropri Probability) Pada Naive Bayes

HMAP menyatakan hipotesa yang diambil berdasarkan nilai probabilitas berdasarkan kondisi prior yang diketahui.

$$P(S | X) = \frac{\text{Argmax}_{x \in X}}{x \in X} \quad (2)$$

Dengan perhitungan sebagai berikut:

$$\begin{aligned} \frac{\text{Argmax}_{x \in X}}{x \in X} &= \frac{P(Y|X) P(X)}{P(X)} \quad (3) \\ &= P(Y | X) P(X) \end{aligned}$$

Dimana,  $P(S|X)$  adalah sebuah himpunan semesta yang menghasilkan  $\text{Argmax}_{x \in X}$  sebagai hasil kelas prediksi,  $P(X)$  adalah jumlah data kelas,  $P(Y|X)$  adalah probabilitas atribut. HMAP inilah yang digunakan di dalam mesin pembelajaran (*machine learning*) sebagai metode untuk mendapatkan hipotesis untuk suatu keputusan. HMAP dapat diartikan untuk mencari probabilitas terbesar dari semua instansi pada atribut target atau semua kemungkinan keputusan.

#### 2.3.3 Tahap Penyiapan Data

Proses penyiapan data dimulai dengan tahapan pengumpulan data akademik mahasiswa dan aktifitasnya dalam kurun waktu empat (4) tahun yaitu angkatan tahun 2008 sampai dengan 2011. Tahapan selanjutnya adalah melakukan *preprocessing* (praproses) data. Proses yang terjadi dalam tahap ini adalah sebagai berikut:

- a. *Cleaning* atau pembersihan data (untuk membuang data yang tidak konsisten dan *noise*).
- b. Integrasi data (penggabungan data dari beberapa sumber)
- c. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di *mining*). Beberapa teknik *data mining* membutuhkan format data yang khusus sebelum bisa diaplikasikan. Pada sistem penunjang keputusan ini, untuk menentukan inisialisasi tahun kelulusan dapat dilakukan dengan cara:
  - a. Mahasiswa angkatan 2008 dengan tahun kelulusan 2011 dinyatakan lulus tepat waktu, dan diberi inisial "LTW".
  - b. Mahasiswa angkatan 2008 dengan tahun kelulusan lebih dari tahun 2011 dinyatakan tidak lulus tepat waktu, dan diberi inisial "TTW".
  - c. Mahasiswa angkatan 2009 dengan tahun kelulusan 2012 dinyatakan lulus tepat waktu, dan diberi inisial "LTW".
  - d. Mahasiswa angkatan 2009 dengan tahun kelulusan lebih dari tahun 2012 dinyatakan tidak lulus tepat waktu, dan diberi inisial "TTW".
  - e. Mahasiswa angkatan 2010 dengan tahun kelulusan 2013 dinyatakan lulus tepat waktu, dan diberi inisial "LTW".
  - f. Mahasiswa angkatan 2010 dengan tahun kelulusan lebih dari tahun 2013 dinyatakan tidak lulus tepat waktu, dan diberi inisial "TTW".
  - g. Mahasiswa angkatan 2011 dengan tahun kelulusan 2014 dinyatakan lulus tepat waktu, dan diberi inisial "LTW".
  - h. Mahasiswa angkatan 2011 dengan tahun kelulusan lebih dari tahun 2014 dinyatakan tidak lulus tepat waktu, dan diberi inisial "TTW".
- d. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai).
- e. Presentasi pola yang ditemukan untuk menghasilkan aksi.

#### 2.3.4 Tahap Pengolahan Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa data akademik mahasiswa. Proses konversi data data yang telah diperoleh dari tahap persiapan data adalah sebagai berikut :

- a. Atribut IPK disederhanakan menjadi atribut *Range* IPK yang dapat dilihat pada tabel 1 sebagai berikut:

Tabel 1 *Range* IPK

Kelompok	IPK
1	0.01 s/d 1.00
2	1.01 s/d 2.00
3	2.01 s/d 3.00
4	3.01 s/d 4.00

- b. Atribut SKS disederhanakan menjadi atribut *Range* SKS yang dapat dilihat pada tabel 2 sebagai berikut:

Tabel 2 *Range* SKS

Kelompok	SKS
1	0 s/d 54
2	55 s/d 75
3	> 76

- c. Atribut Kota Domisili Tinggal disederhanakan menjadi atribut Kota Domisili yang dapat dilihat pada tabel 3 sebagai berikut:

Tabel 3 Kota Domisili

Kelompok	Kota Domisili
1	Jakarta, Depok, atau Bekasi
2	Lainnya

- d. Atribut Jenis Sekolah menengah disederhanakan menjadi atribut Jenis Sekolah Asal yang dapat dilihat pada tabel 4 sebagai berikut:

Tabel 4 Jenis Sekolah Asal

Kelompok	Sekolah Asal
1	SMA/SMU/SLTA
2	SMK/SMEA/STM (Kejuruan)

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Implementasi Dengan Perhitungan *Naïve Bayes*

Implementasi metode *Naïve Bayes* pada tahap ini menghasilkan prediksi kelulusan berdasarkan pola yang telah ditentukan. Berikut langkah-langkah implementasi metode *Naïve Bayes*:

- a. Membuat Dataset

Berikut ini adalah contoh *dataset* berdasarkan kelulusan yang memiliki atribut NPM, IPK Total, SKS, Alamat, Sekolah Asal, Tanggal\_Lulus dan atribut Lulus yang memiliki nilai data "LTW" (Lulus Tepat Waktu) atau "TTW" (Tidak Tepat Waktu) yang merupakan inisialisasi untuk kelulusan mahasiswa

berdasarkan atribut Tanggal\_Lulus seperti yang ditunjukkan pada tabel 5 berikut ini:

Tabel 5 Contoh input Dataset mahasiswa

NPM	IPK TOTAL	SKS	ALAMAT	ASAL SMU	TGL LULUS	LULUS
31108260	3.17	116	BEKASI	MARTIA BHAKTI	01-10-2011	LTW
30108665	2.00	116	TANGERANG	MUHAMMADIYAH 25		TTW
34109104	3.09	116	JAKARTA TIMUR	BINA DHARMA	20-10-2012	LTW
33109207	3.07	110	BEKASI BARAT	BPS&K II	29-12-2012	LTW
30110986	2.66	110	BEKASI	SMAN 67	05-07-2014	TTW
34110753	2.59	114	JAKARTA PUSAT	SMK 1	16-10-2014	TTW
31110792	2.70	110	BEKASI	SMA ANGKASA 1	28-12-2013	LTW
33111385	2.84	116	JAKARTA SELATAN	BAKTI IDHATA	25-09-2014	LTW
38111787	3.36	116	DEPOK	MAN 7 JAKARTA	25-09-2014	LTW
33111720	3.05	110	JAKARTA TIMUR	PKP 2 DKI JAKARTA	23-05-2015	TTW

b. Membuat *Data Train*

Setelah tabel *dataset* dibuat, selanjutnya adalah proses pembuatan *data train* yang sebelumnya telah diinisialisasi seperti yang ditunjukkan pada tabel 6 berikut ini:

Tabel 6 Contoh input *Data Train* mahasiswa yang telah dilakukan inisialisasi

NPM	IPK TOTAL	Ran ge IPK Total	SKS	SKS RANG E	ALAMAT	KOTA	SEKOLAH ASAL	JENIS SEKOLAH	TGL LULUS	LULUS
31108260	3.17	4	116	3	BEKASI	2	MARTIA BHAKTI	1	01-10-2011	1
30108665	2.00	2	116	3	TANGERA NG	2	MUHAMMADIYA H 25	1		2
34109104	3.09	4	116	3	JAKARTA TIMUR	1	BINA DHARMA	1	20-10-2012	1
33109207	3.07	4	110	3	BEKASI BARAT	2	BPS&K II	1	29-12-2012	1
30110986	2.66	3	110	3	BEKASI	2	SMAN 67	1	05-07-2014	2
34110753	2.59	3	114	3	JAKARTA PUSAT	1	SMK 1	2	16-10-2014	2
31110792	2.70	3	110	3	BEKASI	2	SMA ANGKASA 1	1	28-12-2013	1
33111385	2.84	3	116	3	JAKARTA SELATAN	1	BAKTI IDHATA	1	25-09-2014	1
38111787	3.36	4	116	3	DEPOK	1	MAN 7 JAKARTA	1	25-09-2014	1
33111720	3.05	4	110	3	JAKARTA TIMUR	1	PKP 2 DKI JAKARTA	1	23-05-2015	2

Pada tabel *data training* di atas, terdapat pendefinisian atribut atau kelas target yang dinamakan atribut “Lulus”, yang memiliki 2 nilai data yaitu “LTW” atau yang diinisialisasi dengan angka 1, dan “TTW” atau yang diinisialisasi dengan angka 2.

b. Menghitung jumlah *data training*

Pada *data training* digunakan data sebanyak 3651 data mahasiswa, yang diambil dari 4 tahun angkatan (TA). Jumlah *data training* yang digunakan data sebanyak 10 saja. *Data testing* yang digunakan pada pengujian data adalah seluruh data dari jumlah data mahasiswa tahun angkatan (TA) 2012 sebanyak 458 data mahasiswa.

c. Menghitung jumlah data untuk masing-masing kelas (PX)

Pada atribut target/kelas target, data dibagi menjadi dua bagian yaitu data yang termasuk pada kelas  $P(X) \rightarrow \text{LTW}$  dan  $P(X) \rightarrow \text{TTW}$ . Sehingga didapat jumlah data LTW sebanyak 1685 data dan TTW sebanyak 1966 data.

d. Menghitung probabilitas untuk masing-masing kelas P(X)

Setelah menghitung berapa banyak jumlah data untuk masing-masing kelas P(X), selanjutnya adalah menghitung probabilitas atau kemungkinan pada masing-masing kelas baik kelas  $P(X) \rightarrow \text{LTW}$  ataupun  $P(X) \rightarrow \text{TTW}$ . Sehingga didapat probabilitas untuk  $\text{LTW} = 0,462$  yang berasal dari hasil  $1685/3651$ , dan probabilitas untuk  $\text{TTW} = 0,538$  dari hasil bagi antara  $1966/3651$ .

e. Menghitung probabilitas untuk masing-masing atribut terhadap kelas P(Y|X)

Perhitungan probabilitas untuk setiap atribut, dapat disederhanakan menggunakan metode HMAP dengan variable-variabel berikut ini:

Y = Lulus

X1 = IPK

X2 = SKS

X3 = Kota Domisili

X4 = Jenis Sekolah Asal

$$P(Y=\text{LTW}) = \frac{7}{10}$$

$$P(Y=\text{TTW}) = \frac{3}{10}$$

Fakta lain adalah:

Untuk Probabilitas  $P(X1) = \text{IPK}$

$$P(X1=1|Y=\text{LTW}) = 0,$$

$$P(X1=2|Y=\text{LTW}) = 0,$$

$$P(X1=3|Y=\text{LTW}) = \frac{4}{7},$$

$$P(X1=4|Y=\text{LTW}) = \frac{3}{7},$$

Untuk Probabilitas  $P(X2) = \text{SKS}$

$$P(X2=1|Y=\text{LTW}) = 0,$$

$$P(X2=2|Y=\text{LTW}) = 0,$$

$$P(X2=3|Y=\text{LTW}) = 1,$$

Untuk Probabilitas  $P(X3) = \text{Kota Domisili}$

$$P(X3=1|Y=\text{LTW}) = \frac{4}{7},$$

$$P(X3=2|Y=\text{LTW}) = \frac{3}{7},$$

Untuk Probabilitas  $P(X4) = \text{Jenis Sekolah Asal}$

$$P(X4=1|Y=\text{LTW}) = \frac{3}{7},$$

$$P(X4=2|Y=\text{LTW}) = \frac{4}{7},$$

Misal, Untuk  $P(X1=2, X2=1, X3=1, X4=1 | Y=\text{TTW})$

$$= \{ P(X1=2|Y=\text{TTW}) * P(X2=1|Y=\text{TTW}) * P(X3=1|Y=\text{TTW}) * P(X4=1|Y=\text{TTW}) \} * P(Y=\text{TTW})$$

$$= \left\{ \left(\frac{1}{3}\right) * \left(\frac{2}{3}\right) * (1) * \left(\frac{1}{3}\right) * \left(\frac{3}{10}\right) \right\} = \frac{1}{45}$$

Metode Naïve Bayes dengan HMAP mencari nilai probabilitas terbesar pada tiap atribut, sehingga dari perhitungan dengan HMAP dapat disimpulkan bahwa  $P(Y|\text{LTW})$  memiliki nilai lebih besar dibandingkan  $P(Y|\text{TTW})$ .

f. Menghitung total probabilitas dari masing-masing atribut U { P(Y|X) }

Menghitung total probabilitas untuk masing-masing atribut telah dilakukan pada tahap sebelumnya.

g. Mengalikan total probabilitas atribut dengan masing-masing atribut kelas U { P(Y|X) } \* P(X)

Mengalikan total probabilitas atribut dengan masing-masing atribut kelas, dapat ditunjukkan dengan contoh perhitungan di bawah ini:

Untuk  $P(X1=4, X2=3, X3=2, X4=2 | Y=\text{LTW})$

$$= \{ P(X1=4|Y=\text{LTW}) * P(X2=3|Y=\text{LTW}) * P(X3=2|Y=\text{LTW}) * P(X4=2|Y=\text{LTW}) \} * P(Y=\text{LTW})$$

$$= \left\{ \left(\frac{3}{7}\right) * (0) * \left(\frac{3}{7}\right) * \left(\frac{4}{7}\right) * \left(\frac{7}{10}\right) \right\} = \frac{18}{245}$$

Pada perhitungan tersebut dapat diungkapkan bahwa, probabilitas Y terhadap atribut X1, X2, X3, dan X4 adalah sebesar  $\frac{18}{245}$  atau senilai **0.07**

h. Mencari nilai probabilitas maksimum

Langkah terakhir adalah mencari nilai probabilitas tertinggi (maksimum) dari hasil perkalian masing-masing kelas prediksi, sehingga dapat diambil contoh dari perhitungan HMAP pada point-point di atas adalah untuk probabilitas  $P(Y|TTW)$  adalah sebesar  $1/45$  (0,02) sedangkan  $P(Y|LTW)$  adalah sebesar  $32/245$  (0,13) atau dapat dikatakan nilai Maksimum =  $P(Y|LTW)$ .

### 3.2 Implementasi Dengan Sistem

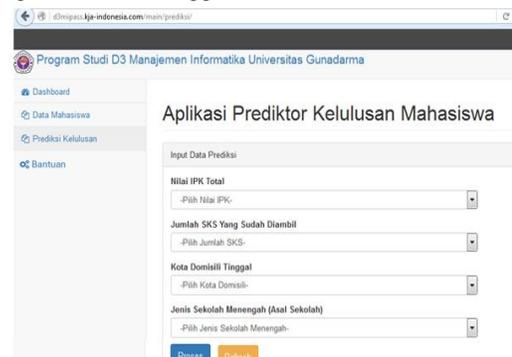
Setelah melalui tahapan perhitungan menggunakan metode Naïve Bayes selanjutnya adalah implementasi sistem. Implementasi sistem merupakan bagian akhirdaripada perancangan sistem yang telah dibangun dimana tahapan ini juga merupakan pengujian program dimana pada sistem ini, program akan menguji kebenaran dari perhitungan yang telah dilakukan dengan serangkaian algoritma yang dituangkan dalam sintaks program berikut ini untuk mencari probabilitas masing-masing pola kelulusan.

```
//Finding Probability of each class P(X)
foreach($allclass as $c=>$p)
{
    $Pc[$c] = round($p/$C,4);
    $argmax[$c] = 1;
}
/*echo"<pre>";
print_r($Pc);
exit;*/

//Finding the prediction
foreach($allclass as $c=>$p)
{
    foreach($args['X'] as
    $x=>$y)
    {
        $i = $this-
        >prediksi_model-
        >get_count_allclass(array('n'=>$args['n'],
        'table'=>$args['table'], 'c'=>$c, 'x'=>$x,
        'y'=>$y))->result_array();
        $j = count($i);
```

```
// calculate probability each attributes
P(Y|X)
$P[$c][$x] =
round($j/$allclass[$c],2);
// U{P(Y|X)}
$argmax[$c] *= $P[$c][$x];
}
//U{P(Y|X)} * P(X)
$argmax[$c] *= $Pc[$c];
}
//argmax U{P(Y|X)} * P(X)
$result
array_keys($argmax,max($argmax));
return $result[0];
}
}>
```

Adapun Langkah-langkah dalam mengelola data menggunakan SPK metode



Naïve Bayes yaitu:

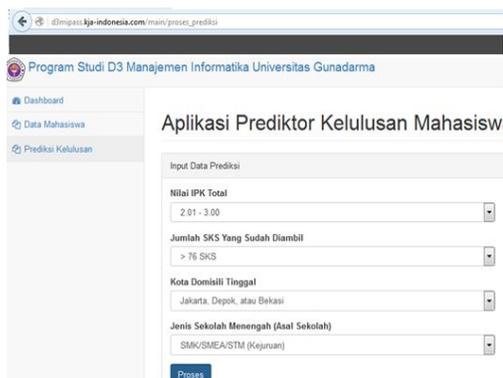
#### a. Halaman Prediksi Kelulusan

Merupakan halaman yang digunakan untuk menguji data *training* yang telah ditambah menggunakan metode Naïve Bayes dan menjadikan beberapa aturan atau yang dapat disebut dengan pola. Pada halaman ini pengguna dapat memasukkan data yang ingin diprediksi kelulusannya.

Gambar 3. Halaman Prediksi Kelulusan

#### b. Input Data Kelulusan

Pada halaman untuk menguji data mahasiswa, maka halaman ini akan membantu staf untuk mengetahui prediksi kelulusan mahasiswa dengan menginputkan Nilai IPK, Jumlah SKS, Kota Domisili tinggal, dan Jenis Sekolah Asal.



Gambar 4. Input Prediksi Kelulusan

Pada pengujian Sistem Penunjang Keputusan ini dan perhitungan probabilitas yang telah dilakukan pada tahap sebelumnya dapat diketahui bahwa terdapat 48 pola kelulusan mahasiswa yang digunakan untuk menghasilkan data akurat pada saat pengujian data, dengan perincian 6 pola menghasilkan pola kelulusan Lulus Tepat Waktu (LTW) dan 42 pola kelulusan Tidak Tepat Waktu. Perincian pola dengan pola kelulusan yang dihasilkan adalah Lulus Tepat Waktu sebanyak enam (6) pola

#### 4. SIMPULAN

Berdasarkan hasil pengujian yang dilakukan untuk memprediksi pola kelulusan mahasiswa dengan metode Naïve Bayes, maka dapat disimpulkan bahwa:

1. Perwakilan dari pola kelulusan mahasiswa lulus tepat waktu (LTW) adalah dengan pola sebagai berikut:  
 $P(X_1 = \text{IPK } 2,01 - 3,00; X_2 = \text{SKS} > 76; X_3 = \text{Kota} = \text{Jakarta Depok Bekasi}; X_4 = \text{Jenis Sekolah Asal} = \text{SMK/STM/Kejuruan}), \text{LTW}$ .  
Sedangkan kesimpulan lain adalah bahwa perwakilan dari pola kelulusan mahasiswa lulus tidak tepat waktu (TTW) adalah dengan pola sebagai berikut:  
 $P(X_1 = \text{IPK } 1,01 - 2,00; X_2 = \text{SKS} = 55 - 75; X_3 = \text{Kota} = \text{Jakarta Depok Bekasi}; X_4 = \text{Jenis Sekolah Asal} = \text{SMA/SMU/SLTA}), \text{TTW}$ .
2. Pengukuran akurasi prediksi dilakukan dengan mengujikan data kelulusan mahasiswa pada sistem telah bermanfaat bagi pengguna.

#### 5. SARAN

Adapun saran yang dapat diberikan adalah Aplikasi sistem penunjang keputusan ini diharapkan dapat segera digunakan tidak hanya oleh staf program studi objek penelitian, tetapi juga oleh mahasiswa yang membutuhkan informasi mengenai berbagai pertanyaan seputar masalah akademik tanpa harus menunggu waktu pelayanan oleh staf.

#### DAFTAR PUSTAKA

- [1] Han, J. dan Kamber, M. 2006. *Data Mining Concepts and Techniques Second Edition*. Morgan Kauffman. San Francisco.
- [2] Hastuti, Khaviizh. *Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif*. Jurnal Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012 (Semantik 2012), Vol 2, no.2, Juni 2012, Hlm. 241-249.[3] Dwi Meilani, Budanis, dkk. *Aplikasi Data Mining Untuk Menghasilkan Pola Kelulusan Siswa Dengan Metode Naïve Bayes*. Jurnal LINK, vol 21, no. 2, september 2014 Hlm. 1-6.
- [4] Nugroho, Bunafit 2004. *Aplikasi Pemrograman Web Dinamis dengan PHP dan MySQL*. Gava Media. Yogyakarta.
- [5] Peranginangin, Kasiman. 2006. *Aplikasi Web Dengan PHP dan MySQL*. Andi Offset, Yogyakarta.
- [6] Ridwan, Mujib, dkk. 2013. *Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier*. Jurnal EECCIS Vol.7, No. 1, Juni 2013 Hlm. 59-64