

Prediksi Tingkat Kelulusan Mahasiswa dengan Algoritma Naive Bayes

¹Neni Purwati, ²Agnes Dwi Januanti

^{1,2} Fakultas Ilmu Komputer Prodi Sistem Informasi
Institut Informatika dan Bisnis Darmajaya
Jl. ZA. Pagar Alam No.93, Gedong Meneng, Kec. Rajabasa, Bandar Lampung
Kode pos. 35142 Telpn. 0721-787214
e-mail: nenipurwati87@darmajaya.ac.id, agnesdwijanuanti@gmail.com

Abstract — Stacks of data from company or agency transactional systems are assets that are currently not utilized for analysis to produce new knowledge or information that is very valuable in the future. The purpose of this study is to generate new knowledge about graduation rates, by utilizing student graduation data in 5 years, which can be used to determine whether students graduate on time or not. The Data Mining method used is Naïve Bayes which is a simple probability classification method that calculates probabilities by adding up the frequency and combination of values from existing data sets. The research method used is RUP (Rational Unified Process) with stages including: Inception, Elaboration, Construction, and Transition. The data used is the graduation data of students who graduated from 2016-2019 at XYZ college. The attributes used are, Gender, Age, Origin, Class, Major, GPA. As a preventive measure by using a web application, which can be accessed anywhere and anytime by students, the prediction results can motivate students in their completion, make it easier for the campus to predict graduation rates and assist in designing remedial steps to improve campus quality, as well as campus accreditation scores, so that the level of stakeholder trust can increase in the future.

Keywords: Data Mining; Kelulusan Mahasiswa; Naïve Bayes.

1. PENDAHULUAN

Sumber daya terdiri dari sumber daya sarana, prasarana dan manusia, selain itu aplikasi komputerisasi adalah salah satu sumber daya yang dapat digunakan untuk meningkatkan keunggulan bersaing. Sistem informasi dapat digunakan untuk mendapatkan, mengolah, dan menyebarkan informasi yang menunjang kegiatan pengambilan keputusan serta menunjang kegiatan transaksional sehari-hari. Perguruan Tinggi adalah program pendidikan tinggi yang memiliki beban studi sekurang-kurangnya 144 sks (satuan kredit semester) yang dijadwalkan untuk 8 semester dan dapat ditempuh minimal 7 semester dan paling lama 14 semester. Masih banyak mahasiswa Program Sarjana (S1) reguler di Perguruan Tinggi XYZ yang menempuh lama studi lebih dari 8 semester. Melihat kondisi tersebut diperlukan cara untuk menggali data yang dimiliki oleh Perguruan Tinggi XYZ tersebut, karena jika masa studi mahasiswa dapat diketahui lebih dini, maka pihak perguruan tinggi dapat melakukan tindakan-tindakan preventif agar mahasiswa dapat lulus tepat waktu. Dibutuhkan suatu teknik klasifikasi yang merupakan salah satu teknik dari data mining untuk menganalisis data tersebut.

1.1 Data mining

Data Mining mampu membantu perusahaan untuk mendapatkan pola dari data-data yang tersimpan di dalam basis data perusahaan. “Pengetahuan yang diperoleh tersebut akan menjadi pedoman dalam mengambil tindakan – tindakan bisnis sebagai upaya pemeliharaan dan peningkatan fungsi – fungsi berikut : deskripsi, estimasi, prediksi, klasifikasi, *clustering*, dan asosiasi[1].

1.2 Knowledge Discovery Database

Pada proses Knowledge Discovery Database (KDD) terdapat beberapa tahapan yaitu sebagai berikut:

1. Seleksi Data (*Data Selection*) merupakan sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *Knowledge Discovery Database (KDD)* dimulai. Data seleksi yang akan digunakan untuk proses *Data Mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.
2. Pemilihan Data (*Preprocessing / Cleaning*) proses *preprocessing* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten dan memperbaiki kesalahan pada data, seperti kesalahan cetak (*tipografi*). Juga dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk *KDD*, seperti data atau informasi *eksternal*.
3. Transformasi (*Transformation*) pada fase ini yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas ke dalam bentuk data yang *valid* atau siap untuk dilakukan proses *Data Mining*.
4. *Data Mining* pada fase ini yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan. Ini adalah langkah penting di mana teknik kecerdasan diterapkan untuk mengekstrak pola informasi yang berpotensi berguna dari data yang dipilih.
5. Interpretasi/Evaluasi (*Interpretation/Evaluation*) pada fase terakhir ini yang dilakukan adalah proses pembentukan keluaran yang bersumber pada proses *Data Mining* Pola informasi [2][3].

1.3 Kelulusan Mahasiswa

Kelulusan mahasiswa tanda berakhirnya mahasiswa dalam menyelesaikan pendidikan pada jenjang sarjana. Kelulusan juga merupakan hasil akhir pencapaian yang membanggakan dalam menempuh suatu pendidikan pada jenjang tertentu. Beban mahasiswa program pendidikan strata satu (S-1) beban studi harus ditempuh minimal 144 satuan kredit semester (SKS) dengan batas waktu kurang dari 8 semester dan waktu paling lama 14 semester. Mahasiswa dinyatakan lulus apabila ketentuan tersebut dipenuhi[7].

1.4 Penelitian Terdahulu

Tabel 1.1 Penelitian Terdahulu

No.	Author	Judul	Masalah	Hasil
1	Agus Budiyantra, Irwansyah, Egi Prengki, Pandi Ahmad Pratama, Ninuk Wiliani	Komparasi Algoritma <i>Decision Tree</i> , <i>Naive Bayes</i> Dan <i>K-Nearest Neighbor</i> Untuk Memprediksi Mahasiswa Lulus Tepat Waktu	Prediksi mahasiswa lulus tepat waktu dibutuhkan oleh manajemen Perguruan Tinggi dalam menentukan kebijakan preventif terkait pencegahan dini kasus <i>Drop Out</i> (DO)	Hasil evaluasi dan validasi yang telah dilakukan menggunakan <i>tools</i> RapidMiner diperoleh hasil akurasi dari Metode <i>Decision Tree</i> (C4.5) sebesar 98.04% pada pengujian ke 3. akurasi Metode <i>Naive Bayes</i> sebesar 96.00% pada pengujian ke 4. Dan akurasi Metode <i>K-Nearest Neighbor</i> (K-NN) sebesar 90.00% pada pengujian ke 2.
2	Azahari, Yulindawati, Dewi Rosita, dan Syamsuddin Mallala	Komparasi <i>Data Mining Naive Bayes</i> Dan <i>Neural Network</i> Memprediksi Masa Studi Mahasiswa S1	Prediksi kelulusan dibutuhkan oleh manajemen perguruan tinggi dalam menentukan kebijakan preventif terkait pencegahan dini kasus drop out.	Data <i>testing</i> diperoleh tingkat akurasi 57,63%. Kelemahan dari hasil prediksi <i>naive bayes</i> dikarenakan tingkat akurasi kevalidannya tidak terlalu tinggi. Sedangkan akurasi prediksi <i>neural network</i> adalah 72,58%, sehingga menjadi metode alternatif yang lebih baik.

No.	Author	Judul	Masalah	Hasil
3		Implementasi Data Mining Menggunakan Algoritme <i>Naive Bayes Classifier</i> dan <i>C4.5</i> untuk Memprediksi Kelulusan Mahasiswa	Ketidakmampuan mahasiswa untuk menyelesaikan studi tepat waktu dialami oleh STMIK Bina Nusantara Jaya Lubuklinggau. Dalam banyak kasus para mahasiswa menyelesaikan studi mereka lebih lama dari rentang waktu yang diharapkan.	Hasil penelitian menunjukkan bahwa metode Algoritme <i>C4.5</i> dapat digunakan untuk memprediksi status kelulusan mahasiswa dengan tingkat akurasi 79,08% sedangkan metode <i>Naive Bayes Classifier</i> hanya 78,46%.
4	Neni Purwati, Rini Nurlistiani, Oscar Devinsen	Data Mining Dengan Algoritma <i>Neural Network</i> Dan Visualisasi Data Untuk Prediksi Kelulusan Mahasiswa	Status kelulusan mahasiswa jika tidak diprediksi secara dini, akan mengakibatkan banyak mahasiswa lulus tidak tepat waktu, dan hal tersebut akan sangat merugikan mahasiswa dan instansi pendidikan.	<i>Training</i> dan <i>testing</i> data dengan mengubah parameter yaitu: <i>Hidden Layer Size: 3, Training Time: 500, Learning Rate: 0.3, Momentum: 0.2</i> menghasilkan klasifikasi dengan <i>Precision</i> 87.80%, <i>Recall</i> 86.90% dan <i>Accuracy</i> (tingkat akurasi) sebesar 92,83%. Visualisasi datanya dapat membantu kelulusan mahasiswa untuk dapat lulus tepat waktu, dan merekomendasikan kepada manajemen atau pihak yang berwenang dalam mengambil keputusan.
5	Edi Sutoyo, Ahmad Almaarif	<i>Educational Data Mining</i> untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme <i>Naive Bayes Classifier</i>	Tingginya jumlah mahasiswa yang tidak lulus tepat waktu atau DO dapat diminimalkan dengan melakukan deteksi mahasiswa yang berisiko pada tahap awal pendidikan dan ditunjang dengan membuat kebijakan yang dapat mengarahkan mahasiswa agar dapat menyelesaikan pendidikannya.	Hasil penelitian ini membuktikan bahwa NBC berhasil diimplementasikan untuk memprediksi kelulusan mahasiswa. Prediksi kelulusan mahasiswa tersebut mampu menghasilkan <i>accuracy</i> sebesar 73.725%, <i>precision</i> 0.742, <i>recall</i> 0.736 dan <i>F-measure</i> sebesar 0.735.

Tabel 1.1 menjelaskan bahwa tingkat akurasi Metode *Naive Bayesian* berada di atas *K-Nearest Neighbor*, walaupun berada di bawah akurasi *Decision Tree(C4.5)* dan *Neural Network*. Selain tingkat akurasi metode *Naive Bayesian* juga cocok untuk data yang ada, sehingga penelitian ini dilakukan menggunakan metode *Naive Bayesian*.

2. METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian ini terdiri dari 2, yaitu:

2.1 Data Mining dan Teorema *Bayesian Classification*

Bayesian Classification atau metode *Naive Bayes* adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu *class*. *Bayesian Classification* didasarkan pada teorema Bayes yang memiliki kemampuan klasifikasi serupa dengan *decision tree* dan *neural network*. *Bayesian*

Classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [4][5][6].

Teorema Bayes memiliki bentuk umum sebagai berikut,

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

Dengan :

X : Data dengan kelas yang belum diketahui;

H : Hipotesis data X merupakan suatu label kelas tertentu;

$P(H|X)$: Probabilistik hipotesis H berdasarkan kondisi X (*posteriori probability*);

$P(H)$: Probabilistik hipotesis H (*prior probability*);

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H;

$P(X)$: Probabilistik X

Merupakan proses menghitung data yang sudah diseleksi, dibersihkan dan ditransform. Sebelum dihitung dengan metode *Naïve Bayes*, data dibuat menjadi data training dengan atribut yang sudah disesuaikan[8][9]. Berikut total data (*dataset*) terdapat 2349 record, namun 15 record ini hanya sebagai Data training dapat dilihat pada tabel 2.1 berikut:

Tabel 2.1 Data Training

No	Jenis Kelamin	Umur	Asal	Kelas	Jurusan	IPK	Class Hasil
1	Laki-Laki	24..29	Dalam Kota	Ekstensi	Sistem Komputer	Sangat Memuaskan	Terlambat
2	Laki-Laki	30..40	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	Terlambat
3	Laki-Laki	> 40	Luar Kota	Reguler	Sistem Komputer	Memuaskan	Terlambat
4	Perempuan	24..29	Luar Kota	Ekstensi	Teknik Informatika	Sangat Memuaskan	Tepat
5	Laki-Laki	24..29	Luar Kota	Ekstensi	Teknik Informatika	Sangat Memuaskan	Terlambat
6	Laki-Laki	24..29	Dalam Kota	Ekstensi	Teknik Informatika	Sangat Memuaskan	Terlambat
7	Laki-Laki	24..29	Luar Kota	Reguler	Sistem Informasi	Sangat Memuaskan	Terlambat
8	Perempuan	24..29	Dalam Kota	Ekstensi	Sistem Informasi	Sangat Memuaskan	Terlambat
9	Laki-Laki	18..23	Dalam Kota	Ekstensi	Sistem Informasi	Sangat Memuaskan	Tepat
10	Perempuan	24..29	Luar Kota	Ekstensi	Manajemen	Dengan Pujian	Tepat
11	Perempuan	24..29	Dalam Kota	Ekstensi	Manajemen	Sangat Memuaskan	Tepat
12	Perempuan	24..29	Luar Kota	Ekstensi	Manajemen	Dengan Pujian	Tepat
13	Perempuan	24..29	Dalam Kota	Reguler	Akuntansi	Dengan Pujian	Tepat
14	Laki-Laki	24..29	Dalam Kota	Ekstensi	Akuntansi	Sangat Memuaskan	Tepat
15	Perempuan	24..29	Dalam Kota	Ekstensi	Akuntansi	Sangat Memuaskan	Tepat

Hitung jumlah class.

Jumlah dari masing-masing class hasil dibagi dengan total data yang terdapat pada *data training*.

- $H1: P(\text{Class Hasil} = \text{"tepat"}) = 8/15 = 0.533$
- $H2: P(\text{Class Hasil} = \text{"terlambat"}) = 7/15 = 0.467$

Hitung $P(X|Ci)$, yaitu probabilitas dari setiap atribut pada data X, kemudian dibagi dengan banyaknya jumlah class tepat dan terlambat:

Data X= Jenis Kelamin = "Laki-laki", Umur= "24..29", Asal= "Dalam Kota", Kelas= "Ekstensi", Jurusan= "Teknik Informatika", IPK= "Memuaskan".

- $P(\text{Jenis Kelamin} = \text{"Laki-laki"} | \text{class_hasil} = \text{"tepat"}) = 2/8 = 0,250$
- $P(\text{Jenis Kelamin} = \text{"Laki-laki"} | \text{class_hasil} = \text{"terlambat"}) = 6/7 = 0,857$
- $P(\text{Umur} = \text{"24..29"} | \text{class_hasil} = \text{"tepat"}) = 7/8 = 0,875$
- $P(\text{Umur} = \text{"24..29"} | \text{class_hasil} = \text{"terlambat"}) = 5/7 = 0,714$
- $P(\text{Asal} = \text{"Dalam Kota"} | \text{class_hasil} = \text{"tepat"}) = 5/8 = 0,625$
- $P(\text{Asal} = \text{"Dalam Kota"} | \text{class_hasil} = \text{"terlambat"}) = 4/7 = 0,571$
- $P(\text{Kelas} = \text{"Ekstensi"} | \text{class_hasil} = \text{"tepat"}) = 7/8 = 0,875$
- $P(\text{Kelas} = \text{"Ekstensi"} | \text{class_hasil} = \text{"terlambat"}) = 4/7 = 0,571$
- $P(\text{Jurusan} = \text{"Teknik Informatika"} | \text{class_hasil} = \text{"tepat"}) = 1/8 = 0,125$
- $P(\text{Jurusan} = \text{"Teknik Informatika"} | \text{class_hasil} = \text{"terlambat"}) = 2/7 = 0,286$
- $P(\text{IPK} = \text{"Memuaskan"} | \text{class_hasil} = \text{"tepat"}) = 0/8 = 0$
- $P(\text{IPK} = \text{"Memuaskan"} | \text{class_hasil} = \text{"terlambat"}) = 1/7 = 0,143$

Hitung total masing-masing nilai pada setiap atribut pada data X :

- $P(X | \text{class_hasil} = \text{"tepat"})$
 $= 0,250 \times 0,875 \times 0,625 \times 0,875 \times 0,125 \times 0 \times 0,533$
 $= 0$
- $P(X | \text{class_hasil} = \text{"terlambat"})$
 $= 0,857 \times 0,714 \times 0,571 \times 0,571 \times 0,286 \times 0,143 \times 0,467$
 $= 0,0038$

2.1.1 Evaluation

Berdasarkan data uji yang belum diketahui class hasilnya, kemudian dihitung dengan metode Naïve Bayes maka data X = (Jenis Kelamin="Laki-laki", Umur="24..29", Asal="Dalam Kota", Kelas="Ekstensi", Jurusan="Teknik Informatika", IPK="Memuaskan") menghasilkan nilai dari class hasil "**terlambat**" yaitu 0,0038 lebih besar daripada class hasil "**tepat**".

2.2 Metode Pengembangan Perangkat Lunak

Metode yang digunakan dalam penelitian ini adalah RUP (*Rational Unified Process*) dengan tahapan[10]: *Inception, Elaboration, Construction, dan Transition*.

3. HASIL DAN PEMBAHASAN

Metode RUP (*Rational Unified Process*) dengan tahapannya, dapat dijelaskan sebagai berikut[10]:

2.2.1 Inception

Berawal dari menentukan tempat penelitian, selanjutnya melakukan analisa permasalahan. Permasalahan yang terjadi adalah penentuan prediksi tingkat kelulusan kurang efektif karena kriteria yang digunakan hanya

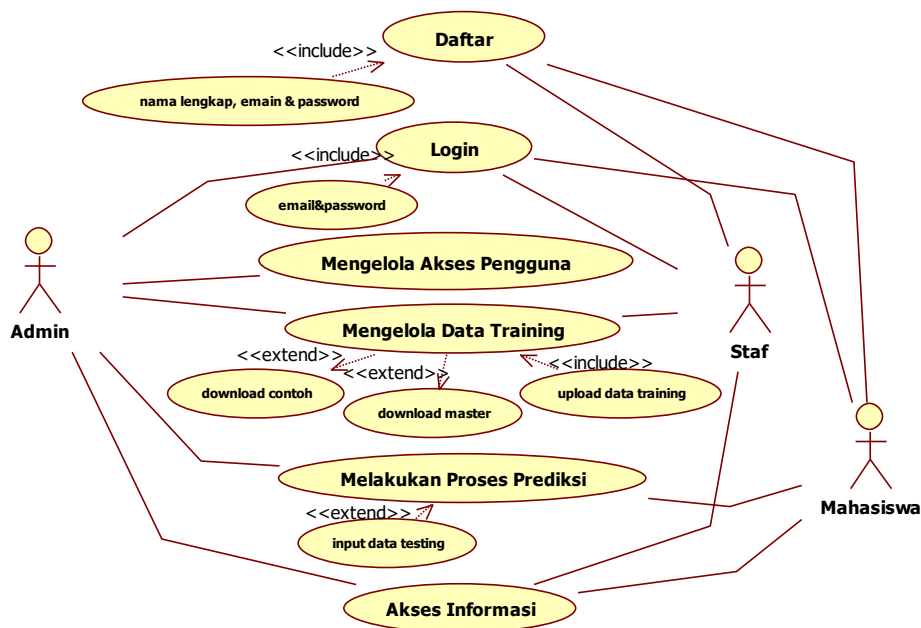
mencakup lama studi mahasiswa yang lulus dan tools yang digunakan masih bersifat konvensional belum ada database.

2.2.2 Elaboration

Pada tahapan ini dilakukan proses desain secara lengkap berdasarkan hasil analisis pada tahap inception. Aktivitas yang dilakukan yaitu melakukan perancangan aplikasi dengan *use case diagram* dan *class diagram*[11].

2.2.2.1 Use Case Diagram

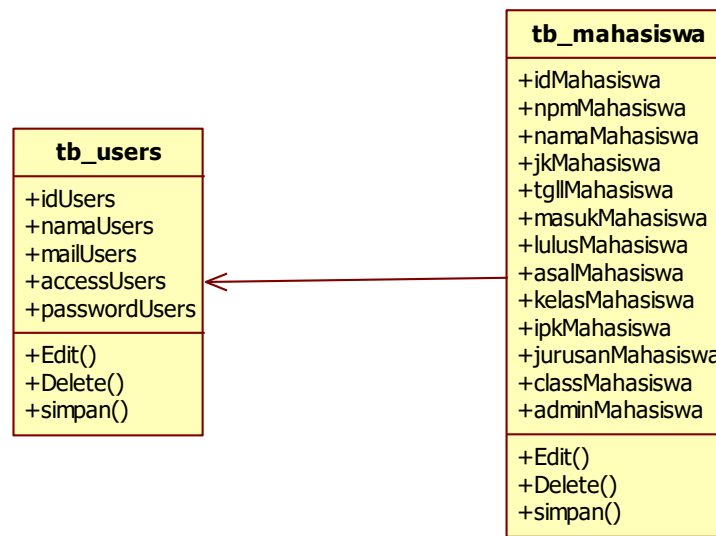
Perancangan *use case diagram* aplikasi prediksi tingkat kelulusan di perguruan tinggi yang diusulkan adalah seperti pada Gambar 2.1 berikut:



Gambar 2.1 Use Case Diagram Aplikasi Baru

2.2.2.1 Class Diagram

Class diagram menggambarkan struktur sistem dari segi pendefinisian kelas – kelas yang akan dibuat untuk membangun sistem penerapan teknik *data mining* untuk memprediksi tingkat kelulusan dengan metode *Naïve Bayes* di Perguruan Tinggi XYZ seperti Gambar 2.2 berikut:



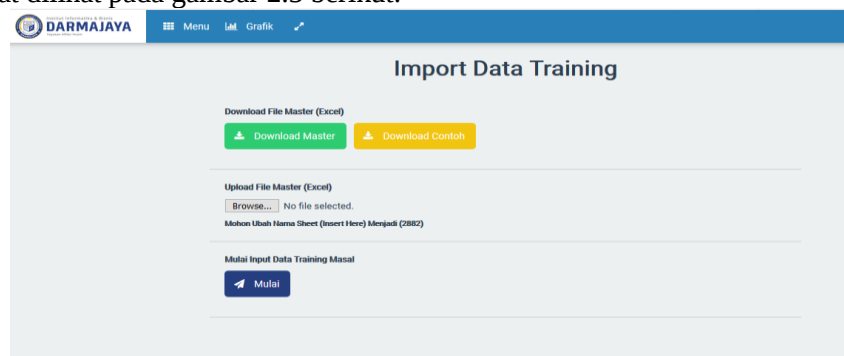
Gambar 2.2 Class Diagram

2.2.3 Construction

Tahap melakukan penulisan coding program untuk aplikasi[12] prediksi tingkat kelulusan mahasiswa berbasis web[13][14], dengan menggunakan bahasa pemrograman PHP(*Hypertext Preprocessor*) dan databasenya dengan MySQL[15].

1. Menu Import Data Training

Tampilan menu ini digunakan oleh user untuk *mendownload* contoh data dan master data, serta *upload* data *training*, dapat dilihat pada gambar 2.3 berikut:



Gambar 2.3 Tampilan Menu Import Data Training

Kemudian tampilan menu berisi data *training* yang telah *terupload* ke *website* dapat dilihat pada gambar 2.4 berikut:

No	Jenis Kelamin	Umur	Asal	Kelas	Jurusan	IPK	Hasil
1	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Memuaskan	TERLAMBAT
2	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
3	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
4	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
5	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Memuaskan	TERLAMBAT
6	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
7	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
8	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
9	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Memuaskan	TERLAMBAT
10	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
11	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TERLAMBAT
12	Laki-Laki	< 30 Tahun	Luar Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TEPAT
13	Perempuan	< 30 Tahun	Luar Kota	Reguler	Sistem Komputer	Dengan Pujian	TEPAT
14	Laki-Laki	< 30 Tahun	Luar Kota	Reguler	Sistem Komputer	Sangat Memuaskan	TEPAT
15	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Dengan Pujian	TEPAT
16	Laki-Laki	< 30 Tahun	Dalam Kota	Reguler	Sistem Komputer	Dengan Pujian	TEPAT

Gambar 2.4 Tampilan Menu Data Training

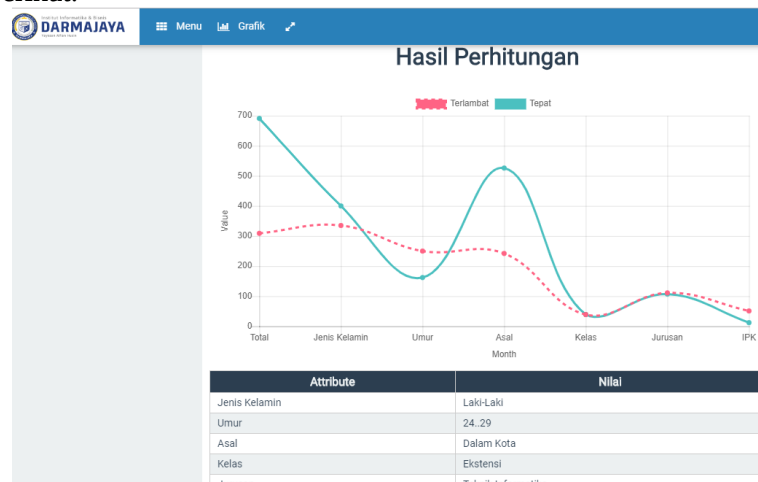
2. Menu *Input Data Testing*

Menu ini digunakan untuk simulasi oleh mahasiswa dengan menginput data *testing* dan memulai proses perhitungan. Untuk lebih jelasnya dapat dilihat pada gambar 2.5 berikut:

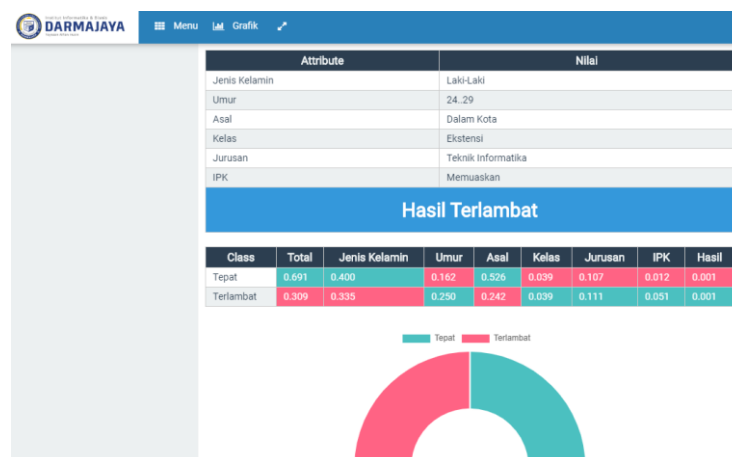
Gambar 2.5 Tampilan Menu *Input Data Testing*

3. Menu Hasil Perhitungan

Menu ini berisi hasil dari proses perhitungan data *testing*, yang dapat dilihat pada gambar 2.6, gambar 2.7, gambar 2.8, gambar 2.9 berikut:



Gambar 2.6 Tampilan Hasil Perhitungan



Gambar 2.7 Tampilan Kesimpulan

Data Traning		
Hasil Class	Jumlah Class	Total Class
Tepat	1624	0.691
Terlambat	725	0.309
Total	2349	

Perhitungan		
Class Jenis Kelamin	Jumlah	Hasil
Tepat	650	0.4
Terlambat	544	0.335
Total	1194	

Perhitungan: untuk attribute (Jenis Kelamin) didapatkan data persentase yang berpeluang "TEPAT" (54%) untuk nilai (Laki-Laki) dalam data traning

Class Umur	Jumlah	Hasil
Tepat	363	0.162
Terlambat	406	0.25
Total	669	

Perhitungan: untuk attribute (Umur) didapatkan data persentase yang berpeluang "TEPAT" (39%) untuk nilai (24-29) dalam data traning

Gambar 2.8 Tampilan Perhitungan

Class Asal	Jumlah	Hasil
Tepat	854	0.526
Terlambat	393	0.242
Total	1247	

Perhitungan: untuk attribute (Asal) didapatkan data persentase yang berpeluang "TEPAT" (68%) untuk nilai (Dalam Kota) dalam data traning

Class Kelas	Jumlah	Hasil
Tepat	63	0.039
Terlambat	64	0.039
Total	127	

Perhitungan: untuk attribute (Kelas) didapatkan data persentase yang berpeluang "TEPAT" (49%) untuk nilai (Ekstensi) dalam data traning

Class Jurusan	Jumlah	Hasil
Tepat	173	0.107
Terlambat	181	0.111
Total	354	

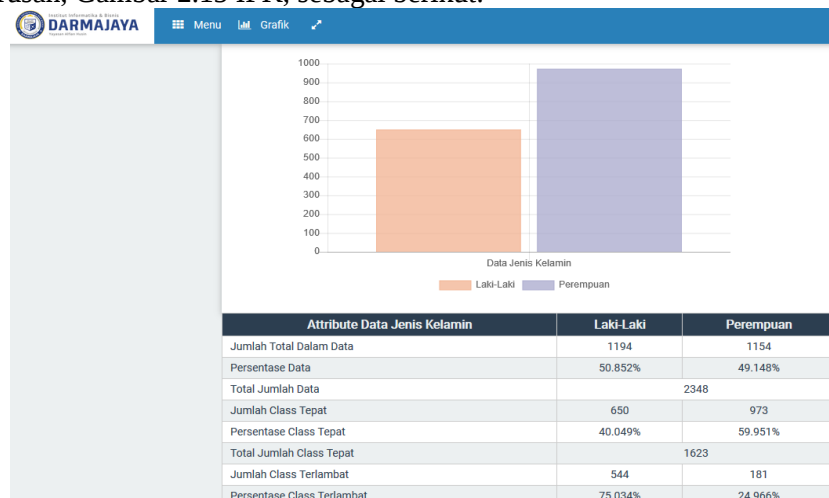
Perhitungan: untuk attribute (Jurusan) didapatkan data persentase yang berpeluang "TEPAT" (48%) untuk nilai (Teknik Informatika) dalam data traning

Class IPK	Jumlah	Hasil
Tepat	20	0.012
Terlambat	60	0.004
Total	80	

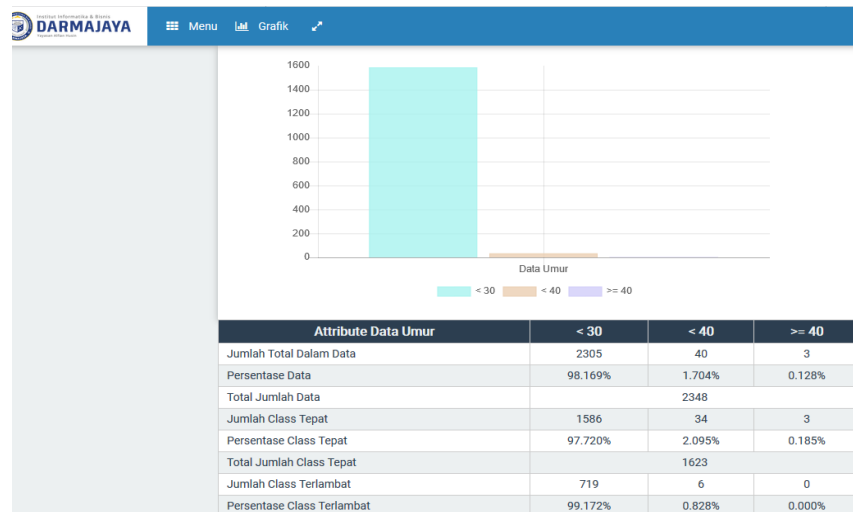
Gambar 2.9 Tampilan Perhitungan (lanjutan)

4. Menu Grafik

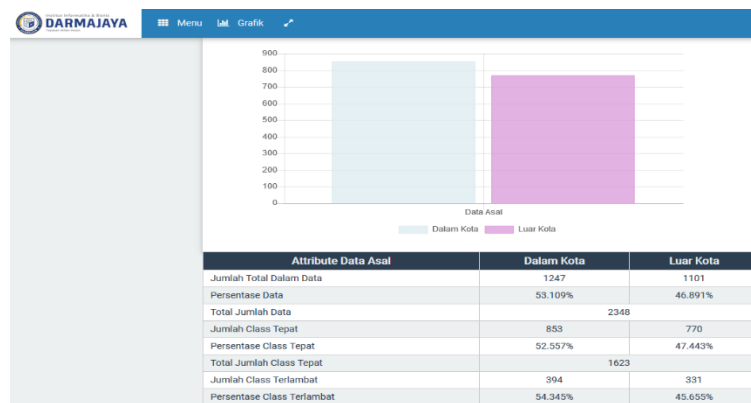
Menu ini berisi grafik-grafik tingkat kelulusan mahasiswa berdasarkan atribut yang ada. Adapun grafik tersebut dapat dilihat pada gambar 2.10 Jenis Kelamin, gambar 2.11 Umur, gambar 2.12 Asal, gambar 2.13 Kelas, gambar 2.14 Jurusan, Gambar 2.15 IPK, sebagai berikut:



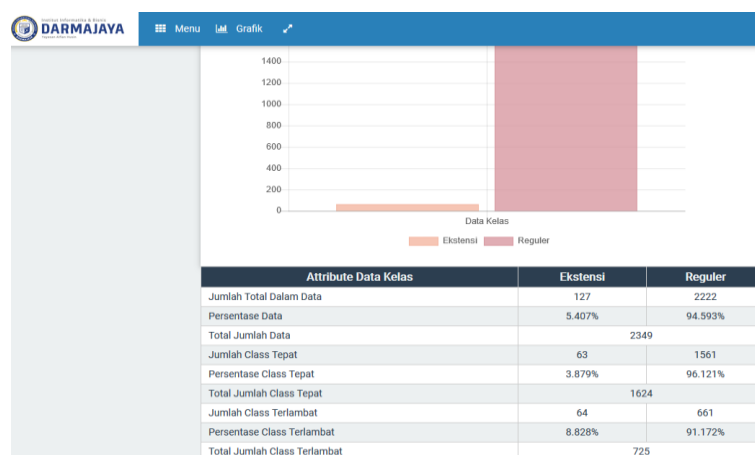
Gambar 2.10 Tampilan Grafik Jenis Kelamin



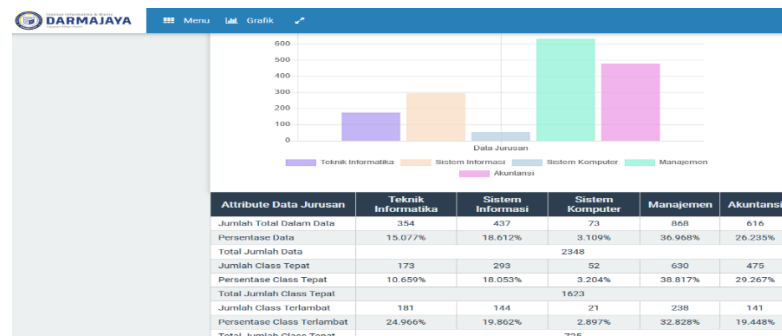
Gambar 2.11 Tampilan Grafik Umur



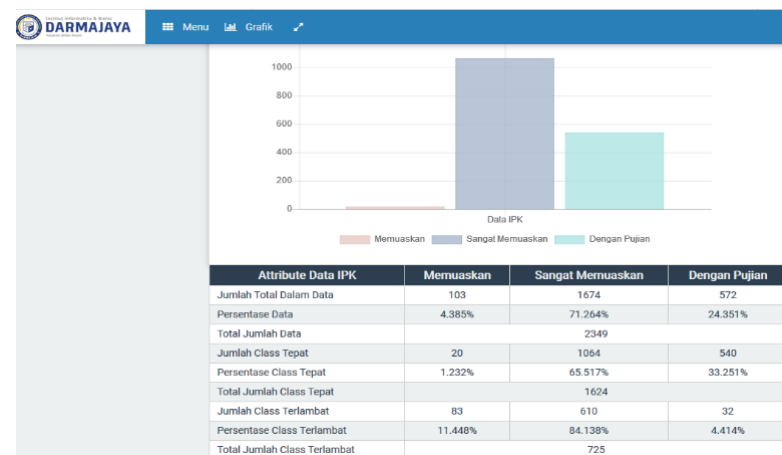
Gambar 2.12 Tampilan Grafik Asal



Gambar 2.13 Tampilan Grafik Kelas



Gambar 2.14 Tampilan Grafik Jurusan



Gambar 2.15 Tampilan Grafik IPK

2.2.4 Transition

Dilakukan pelatihan kepada staf sebagai user yang diberi kewenangan mengakses aplikasi, setelah user mengetahui dan memahami cara penggunaannya maka dilakukan hosting dan sehingga aplikasi berbasis web ini siap digunakan. Kemudian aplikasi tersebut diterapkan oleh Perguruan Tinggi XYZ, sehingga apabila ditemukan *error* dapat dilakukan perbaikan sesegera mungkin.

4. KESIMPULAN

Berdasarkan penerapan aplikasi data *mining* dengan metode *Naïve Bayes*, maka dapat disimpulkan bahwa: Metode *Naïve Bayes* dapat diterapkan untuk prediksi tingkat kelulusan mahasiswa karena algoritmanya sangat mudah, penerapan metode *Naïve Bayes* dalam bentuk *website* memberikan kemudahan bagi mahasiswa untuk memprediksi kelulusan tepat/terlambat sejak dini, yang bisa diakses dimana saja dan kapan saja, dan menyampaikan hasil prediksinya sangat rahasia karena melalui akun masing-masing, sehingga hasil prediksi tersebut dapat memotivasi mahasiswa dalam penyelesaian kuliahnya, dan dengan semakin banyaknya mahasiswa lulus tepat waktu dapat meningkatkan kepercayaan stakeholder, serta dapat membantu pihak kampus untuk meningkatkan kualitas kampus. Kelebihan sistem ini adalah: dapat memproses data *training*, dapat menampilkan informasi prediksi tingkat kelulusan mahasiswa, menggunakan grafik untuk informasi tingkat kelulusan mahasiswa sehingga mudah dimengerti dan mudah digunakan oleh pengguna. Kekurangan sistem ini adalah atribut atau kriteria untuk menghasilkan pola-pola informasi baru hanya sedikit. Saran untuk penelitian selanjutnya agar menggunakan atribut yang lain, dan menggunakan algoritma atau metode lainnya, seperti *Decicion Tree*, *Neural Network*, dan lain-lain.

UCAPAN TERIMA KASIH

Puji syukur kami panjatkan atas terselesaikannya penelitian ini, dan kami ucapkan terimakasih kepada Biro Administrasi Akademik dan Kemahasiswaan yang telah memberikan data, dan kepada Ketua Program Studi Sistem Informasi Perguruan Tinggi XYZ yang sangat membantu dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] Nurjoko and A. Darmawan, "Penerapan Data Mining Menggunakan Association Rules untuk Mendukung Strategi Pemasaran Calon Mahasiswa baru (Studi Kasus IBI Darmajaya)," *Jurnal TIM Darmajaya*, vol. 1, no. 1, pp. 17-32, 2015.
- [2] D. Nofriansyah, *Konsep Data Mining Vs Sistem Pendukung Keputusan*, Yogyakarta: Deepublish, 2014.
- [3] N. Purwati and R. Nurlistiani, "DATA MINING DENGAN ALGORITMA NEURAL NETWORK DAN VISUALISASI DATA UNTUK PREDIKSI KELULUSAN MAHASISWA," *Jurnal Informatika*, vol. 20, no. 2, pp. 156-163, 2020.
- [4] H. Al Fatta, "Analisis dan Perancangan Sistem Informasi untuk Keunggulan Bersaing Perusahaan dan Organisasi Modern" Yogyakarta: Andi Offset. 2007.
- [5] A. Budiyantra, I. Irwansyah, E. Prengki, P. A. Pratama, and N. Wiliani, "KOMPARASI ALGORITMA DECISION TREE, NAIVE BAYES DAN K-NEAREST NEIGHBOR UNTUK MEMREDIKSI MAHASISWA LULUS TEPAT WAKTU," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 5, no. 2, pp. 265-270, 2020.
- [6] A. Azahari, Y. Yulindawati, D. Rosita, and S. Mallala, "Komparasi Data Mining Naive Bayes dan Neural Network memprediksi Masa Studi Mahasiswa S1," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 3, pp. 443-452, 2020.
- [7] Kemenristekdikti, *Peraturan Menteri Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia Nomor 44 Tahun 2015 tentang Standar Nasional Perguruan Tinggi*. 2015.
- [8] E. Etriyanti, D. Syamsuar, and Y. N. Kunang "Implementasi Data Mining Menggunakan Algoritme Naive Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa," *Telematika*, vol. 13, no. 1, pp. 56-67, 2020.
- [9] E. Sutoyo and A. Almaarif, "Educational Data Mining untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme Naive Bayes Classifier," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 1, pp. 95-101, 2020.
- [10] R. A. Sukanto and M. Shalahuddin, *Rekayasa Perangkat Lunak (Terstruktur dan berbasis objek)*, Bandung: Informatika, 2014.
- [11] A. R. Pratama, "Belajar UML - Use Case Diagram," *Codepolitan*, 2019 .
- [12] A. Mariyus, N. Puwarti, and A. Aziz, "Aplikasi Pengolahan Data Puskesmas (Pusat Kesehatan Masyarakat) Desa Margodadi Kab. Tulang Bawang Barat," *Jurnal Simada*, vol. 2, no. 1, pp. 15-25, 2019.
- [13] R. Abdulloh, "7 in 1 Pemrograman Web Untuk Pemula," Jakarta: PT Elex Media Komputindo, 2018.
- [14] N. Purwati, H. Halimah, and A. Rahardi, "Perancangan Website Program Studi Sistem Informasi Institut Informatika Dan Bisnis Darmajaya Bandar Lampung," *SIMADA (Jurnal Sist. Inf. Manaj. Basis Data)*, vol. 1, no. 1, pp. 71-80, 2018.
- [15] L. MADCOMS, *Pemrograman PHP dan MySQL untuk pemula*, Bandung: ANDI 2016.