

# PREDIKSI TINGKAT KELULUSAN MAHASISWA PROGRAM STUDI SISTEM INFORMASI DI UNIVERSITAS XXX MENGGUNAKAN ALGORITMA NAÏVE BAYES

Mega Wahyuningsih<sup>1</sup>, Hari Purwanto<sup>2</sup>, Muryan Awaludin<sup>3</sup>  
<sup>1</sup>mega\_w@gmail.com, <sup>2</sup>raldy08@gmail.com, <sup>3</sup>muryanawaludin1@gmail.com  
<sup>1,2</sup> Universitas Dirgantara Marsekal Suryadarma

## ABSTRACTION

*The Information System Study Program at UNIVERSITAS XXX is a new major and the graduation of students is still few. Based on data obtained from graduates of the 2018/2019 academic year, 41 students graduated, including 26 students who experienced delays in taking their studies. The need for a system that can predict student graduation in order to major in Information Systems can produce more student graduations than ever before. Data mining using the Naïve Bayes algorithm method is one of the methods that utilizes probability calculations and statistics to predict future probabilities based on previous experience. Naïve Bayes is able to be applied in predicting graduation by utilizing previous student graduation data, The attributes used are Gender, Age, SKS, GPA, and Student Status. The results of research testing using RapidMiner 9.8 with 100 training data and 60 testing data, result accuracy 90%, recall 93.94%, and precision 88.57%. A computerized system built on a web basis using the PHP programming language.*

**Keywords:** *Graduation, Data Mining, Naïve Bayes, and Rapid Miner*

## PENDAHULUAN

Sistem prediksi kelulusan mahasiswa menggunakan data mining *Naïve Bayes*, pemanfaatan data historis kelulusan mahasiswa menjadi lebih optimal untuk melakukan prediksi kelulusan mahasiswa (**Diqy Fakhrun Shiddieq, S.T. & Patricia, 2020**). Pengujian akurasi sistem dilakukan dengan mencocokkan hasil prediksi dengan data yang sesungguhnya dengan algoritma *Naïve Bayes* menggunakan metode pengujian *confusion matrix*, hasil uji menunjukkan *accuracy* sebesar 80% (**Putu Sainanda Cahyani Moonallika, Ketut Queena Fredlina, 2020**). Hasil pengujian data latih dan data uji dengan atribut SKS semester 1 sampai 4 dan IP semester 1 sampai 4 dan skala perbandingan 60% data latih dan 40% data uji didapatkan akurasi 91,86% dengan menggunakan algoritma *Naïve*

*Bayes* (**Firman Azhar Riyadi, 2020**). Metode data mining *Naïve Bayes* dapat membuat suatu pre-diksi mengenai kelulusan mahasiswa tepat waktu dengan memperhitungkan atribut IPK (Indeks Prestasi Kumulatif), sehingga tingkat akurasi literatur meng-hasilkan akurasi di atas 90% (**Setiyani, Wahidin, Awaludin, & Purwani, 2020**).

Hasil penelitian membuktikan algoritma *Naïve Bayes* berhasil diimplementasikan untuk prediksi kelulusan mahasiswa dan mampu menghasilkan *accuracy* 73,725% dari 4000 data *instance* yang diperoleh (**Sutoyo & Almaarif, 2020**).

Terhitung jurusan Sistem Informasi sudah meluluskan 185 mahasiswa, berdasarkan data yang didapat untuk lulusan tahun ajaran 2018/2019 maha-

siswa lulus sejumlah 41 mahasiswa, diantara mahasiswa yang lulus terdapat 26 mahasiswa yang lulus terlambat dan 15 mahasiswa yang lulus tepat waktu. Sejumlah 22 mahasiswa kelas karyawan dengan status bekerja mengalami keterlambatan lulus. Hal ini merupakan masalah yang harus dihadapi jurusan Sistem Informasi karena banyaknya mahasiswa yang mengalami keterlambatan dalam menempuh masa studinya terutama dari kelas karyawan. Mahasiswa Program Studi Sistem Informasi memiliki kelulusan mahasiswa yang masih sedikit sehingga akan mempengaruhi akreditasi jurusan, dan belum tersedianya aplikasi prediksi kelulusan mahasiswa di Universitas XXX membuat pihak Kaprodi tidak bisa memprediksi mahasiswa yang dikhawatirkan akan terlambat masa studinya. Sehingga akan terjadi masalah dikemudian hari bagi mahasiswa yang terlambat tidak dapat lulus dengan tepat waktu, dikarenakan tidak adanya persiapan bagi mahasiswa tersebut dalam mengejar ketertinggalannya. Semakin banyak mahasiswa jurusan Sistem Informasi yang berhasil lulus maka semakin banyak lulusan yang bisa melanjutkan ke jenjang dunia kerja, sehingga pihak Kaprodi dapat membawa Sistem Informasi menaikan poin untuk akreditasi jurusan. Hal ini menjadikan tolak ukur keunggulan kualitas dan potensi mahasiswa, serta instansi terkait dapat menerima dan memberikan bantuan kepada mahasiswa dan lulusannya.

Untuk itu penulis tertarik melakukan penelitian dan menjadikan sebagai skripsi dengan judul **“IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK MEMPREDIKSI TINGKAT KELULUSAN MAHASISWA DI UNIVERITAS XXX”**.  
**KAJIAN LITERATUR.**

## **2.1 Kelulusan Mahasiswa**

Kelulusan mahasiswa merupakan hasil akhir dari pencapaian dalam menempuh suatu pendidikan pada jenjang Perguruan Tinggi. Sesuai dengan ketentuan Program Studi Sistem Informasi, ada beberapa hal yang akan menjadi pertimbangan untuk mahasiswa dalam syarat skripsinya yaitu SKS yang ditempuh minimal sudah menempuh beban 143 SKS, tidak ada nilai D, sudah mengikuti orientasi studi mahasiswa di Universitas XXX, dan mengikuti minimal dua kali *Workshop* dari Program Studi Sistem Informasi. Lama studi paling lambat di Program Studi Sistem Informasi Universitas XXX adalah 14 Semester (7 Tahun), maka lebih dari 7 tahun mahasiswa akan *drop out* atau dikeluarkan dari Perguruan Tinggi.

## **2.2 Data Mining**

Data mining adalah teknologi untuk menemukan struktur dan pola dalam kumpulan data yang besar. Data mining mengacu pada metode analisis bisnis yang melampaui hitungan, teknik deskriptif, pelaporan, dan metode berdasarkan aturan bisnis. Secara khusus, ini mencakup metode statistik dan pembelajaran mesin yang menginformasikan pengambilan keputusan, seringkali secara otomatis (Shmueli, Bruce, Yahav, Patel, & Lichtendahl, 2018). Menurut (Febie Elfaladonna, 2019) data mining merupakan gabungan sejumlah ilmu komputer, yang didefinisikan sebagai proses penemuan pola-pola baru dari kumpulan-kumpulan data yang sangat besar, yaitu meliputi metode-metode yang merupakan irisan dari *artificial intelligence, machine learning, statistics, dan database systems*. Data mining merupakan bagian dari tahapan proses *Knowledge Discovery in Database (KDD)*, dengan data mining kita dapat melakukan pengklasifikasian, memprediksi, memperkirakan dan mendapatkan

informasi lain yang bermanfaat dari kumpulan data dalam jumlah yang besar (Mardi, 2017).

### 2.2.1 Klasifikasi dan Prediksi

Prediksi biasanya merupakan komponen penting, seringkali pada tingkat individu (Shmueli et al., 2018). Klasifikasi adalah suatu teknik pembentukan model dari data yang belum terklasifikasi, untuk digunakan mengklasifikasi data baru (Sartika & Indra, 2017). Tujuan dari klasifikasi yaitu menemukan model dari data *training* yang membedakan dalam setiap atributnya ke dalam kelas/label yang sesuai, model tersebut lalu digunakan untuk mengklasifikasikan atribut yang kelas/labelnya belum diketahui (Harmianty, 2017). Ada banyak metode lain untuk membangun model klasifikasi salah satunya adalah *Naïve Bayes*, dapat memprediksi yang mencakup identifikasi terdistribusi berdasarkan data yang telah tersedia. Klasifikasi merupakan sebuah metode dalam Data Mining yang sering digunakan untuk menyelesaikan masalah-masalah di dunia nyata, salah satu yang terpopuler yaitu menggunakan teknik *machine learning*, klasifikasi mempelajari pola-pola yang ada dari data historis (Rahman & Firdaus, 2016).

Klasifikasi data terdiri dari dua tahap yaitu tahap pembelajaran dan tahap pengklasifikasian. Tahap pembelajaran adalah tahapan dalam pembentukan model klasifikasi, sedangkan tahap pengklasifikasian adalah tahapan penggunaan model klasifikasi untuk memprediksi kelas/label dari suatu data (Sartika & Indra, 2017). Komponen-komponen utama dari proses klasifikasi, yaitu;

1. Kelas adalah suatu variabel tidak bebas yang merupakan label dari hasil klasifikasi.
2. Prediktor adalah variabel bebas dari suatu model berdasarkan dari

karakteristik atribut data yang sudah diklasifikasi.

3. Dataset latih adalah sekumpulan data lengkap yang berisi kelas dan prediktor untuk dilatih agar model dapat mengelompokkan ke dalam kelas yang tepat.
4. Dataset uji yaitu berisi data-data baru yang akan dikelompokkan oleh model agar mengetahui akurasi dari model yang dibuat.

### 2.3 Algoritma *Naïve Bayes*

Algoritma *Naïve Bayes* memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya, teori ini dikemukakan oleh **Teorema Bayes** (Awaludin & Mantik, 2023). Metode yang perhitungannya berupa pembagian perbedaan kedalam kelas-kelas, lalu penghitungan digunakan oleh statistik akan dihitung dengan cara probabilitas dan akan dibagi kedalam kelas. *Naïve Bayes* adalah salah satu algoritma pembelajaran induktif yang paling efektif dan efisien untuk *machine learning* dan data mining (Syarli & Muin, 2016). *Naïve Bayes* merupakan pengklasifikasian sederhana yang menghitung probabilitas dengan menjumlahkan frekuensi dan nilai dari data yang ada (Sutoyo & Almaarif, 2020). *Naïve Bayes* merupakan sebuah model klasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas (Dewi, 2016).

Algoritma *Naïve Bayes* memiliki tahapan-tahapan dalam prosesnya, yaitu:

1. Menghitung jumlah kelas/label pada dataset.
2. Menghitung jumlah kasus per-kelas dari data yang ada.
3. Mengkalikan semua variabel kelas.
4. Membandingkan hasil per-kelas, guna mendapatkan kesimpulan dari perbandingan tersebut.

**Teorema Bayes** mempermudah pemahaman dari algoritma *Naïve Bayes*, disederhanakan pada persamaan berikut:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)}$$

Keterangan:

- X : Data dengan kelas yang belum diketahui.
- C : Hipotesis data X merupakan suatu kelas spesifik.
- P(C|X) : Probabilitas hipotesis C berdasarkan kondisi X.
- P(X|C) : Probabilitas X berdasarkan kondisi pada hipotesis C.
- P(C) : Probabilitas hipotesis C (probabilitas *prior*).
- P(X) : Probabilitas X.

Menurut (Sartika & Indra, 2017) yang mengutip I. Yoo, P. Alafaireet, M. Marinov, K. Pena-Hernandez, R. Gopidi, J.-F. Chang, dan L. Hua (2012) *Naive Bayes* adalah perhitungan yang paling sederhana, karena mampu mengurangi kompleksitas komputasi menjadi multiplikasi sederhana dari probabilitas. Ciri utama dari *Naïve Bayes* ini adalah asumsi yang sangat kuat akan independensi dari masing-masing kondisi atau kejadian yang terjadi dan algoritma ini mengasumsikan bahwa atribut objek adalah independen. (Firman Azhar Riyadi, 2020). Selain itu, algoritma *Naive Bayes* juga mampu menangani dataset yang memiliki banyak atribut. Berikut adalah kelebihan dan kekurangan yang dimiliki algoritma *Naïve Bayes*.

#### 2.4 Data Training dan Data Testing

Data *testing* merupakan data uji dalam mengimplementasikan algoritma *Naive Bayes* pada prediksi tingkat kelulusan mahasiswa di Universitas XXX, dalam data ini digunakan sebagai

pengujian akurasi pada perhitungan *Naïve Bayes*, dalam studi kasus ini peneliti mengambil data *testing* sebanyak 60% dari data *training* pada mahasiswa lulusan Program Studi Sistem Informasi untuk dijadikan data uji.

#### 2.5 Rapid Miner

*Rapid miner* dibuat oleh Ralf Klinkenberg, Ingo Mierswa, dan Simon Fischer pada tahun 2001 di Artificial Intelligence Unit dari University of Dortmund, dibuat menggunakan Bahasa pemrograman Java berhasil meraih penghargaan sebagai *software* data mining dan data analisis terbaik oleh KDnuggets, yaitu sebuah portal data mining pada tahun 2010-2011.

*Rapid miner* adalah suatu alat data mining yang digunakan untuk menganalisis suatu informasi web yang diakses, *rapid miner* juga digunakan untuk penelitian, pendidikan, *rapid prototyping*, pengembangan aplikasi dan aplikasi industry (Mulya, 2019). *Rapid miner* merupakan perangkat lunak untuk pengolahan data dengan menggunakan prinsip dan algoritma data mining, *rapid miner* mengekstrak pola-pola dari dataset yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database (Novianti, 2019). *Rapid miner* adalah platform perangkat lunak data ilmu pengetahuan yang dikembangkan oleh perusahaan dengan nama yang sama menyediakan lingkungan terpadu untuk *machine learning* (pembelajaran mesin), *deep learning* (pembeajaran mendalam), *text mining* (penambangan text), dan *predictive analytics* (analisis prediktif) (Nofitri & Irawati, 2019).

#### 2.6 Bahasa Pemrograman dan Perangkat Lunak Pendukung

##### 2.6.1 XAMPP

Pada jurnal (Fitri Ayu and Nia

**Permatasari, 2018)** yang mengutip **MADCOMS (2016)** *XAMPP* merupakan sebuah paket kumpulan software yang terdiri dari *Apache, MySQL, Phpmyadmin, PHP, Perl, Filezilla*, dan lain-lain. *XAMPP* berfungsi untuk memudahkan instalasi lingkungan PHP, di mana lingkungan pengembangan *web* memerlukan PHP (*server side scripting*), *Apache (web server)*, *MySQL (database)* dan *Phpmyadmin*.

### 2.6.2 PHP

Menurut **MADCOMS (2016)** PHP (*Hypertext Preprocessor*) merupakan Bahasa *script* yang dapat ditanamkan atau disisipkan ke dalam HTML, PHP sering kali dipakai untuk membuat program situs *web* dinamis. Salah satu kelebihan PHP yang paling signifikan adalah kemampuannya dalam melakukan koneksi dari berbagai macam *database*. PHP dirancang untuk membentuk satu tampilan berdasarkan permintaan terkini, seperti menampilkan isi dari basis data ke halaman web.

### 2.6.3 HTML

HTML diciptakan oleh Tim Berners Lee dan dipopulerkan pertama kali oleh Browser Mosaic. Menurut Suyanto (2007) yang dikutip pada jurnal (**Lavarino & Yustanti, 2016**) HTML (*Hypertext Markup Language*) adalah bahasa markup yang biasanya digunakan dalam membuat halaman *web*. Sebenarnya HTML bukanlah sebuah bahasa pemrograman, apabila di tinjau dari namanya HTML merupakan bahasa markup atau penandaan yang terhadap pada sebuah dokumen teks. Tugas HTML dalam membangun suatu *website*, diantaranya:

1. Menentukan *layout* dari *website*.
2. Memformat *text* dasar seperti pengaturan paragraph dan format *font*.
3. Membuat list dan formular.

4. Membuat tabel, gambar, video, audio dan *link*.

### 2.6.4 MySQL

MySQL (*My Structure Query Language*) merupakan salah satu Basis Data Management System (DBMS) dari sekian banyaknya DBMS seperti Oracle, MS SQL, Postagre SQL, dan lain-lain. MySQL bersifat *open source* sehingga dapat digunakan secara gratis. MySQL berfungsi untuk mengolah Basis Data menggunakan bahasa SQL (**Lavarino & Yustanti, 2016**). MySQL adalah sistem manajemen database digunakan untuk membuat dan mengelola database, menambah, menghapus, mengubah data yang berada di dalam database. MySQL merupakan database yang sering digunakan dikalangan programmer *web*, dengan alasan program ini merupakan database yang sangat kuat dan cukup stabil untuk digunakan sebagai media penyimpanan data.

## 2.7 UML (Unified Modelling Language)

### 2.7.1 Use Case Diagram

*Use Case Diagram* merupakan pemodelan untuk melakukan sistem informasi yang akan dibuat. *Use Case* digunakan untuk mengetahui fungsi apa saja yang ada di dalam sistem informasi dan siapa saja yang berhak menggunakan fungsi-fungsi tersebut (**Hendini, 2016**).

### 2.7.2 Activity Diagram

*Activity diagram* merupakan gambaran *workflow* (aliran kerja) atau suatu aktivitas dari sebuah sistem atau proses dari bisnis/program yang dibuat (**Hendini, 2016**).

### 2.7.3 Class Diagram

*Class Diagram* merupakan hubungan antar masing-masing kelas dan penjelasan secara detail dari setiap kelas di dalam model desain dari suatu sistem memperlihatkan aturan dan tanggung jawab dari sebuah entitas yang mene-

ntukan perilaku dari suatu sistem (Hendini, 2016).

#### 2.7.4 Collaboration Diagram

Pada *Collaboration Diagram* setiap event mengirimkan informasi dari suatu objek ke objek yang lainnya. *Collaboration Diagram* menggambarkan interaksi antar objek (Sesnika, Andreswari, & Efendi, 2016).

### PEMBAHASAN

#### 3.1 Desain Penelitian

Desain dalam penelitian ini meliputi:

1. Tahapan Awal Penelitian, diantaranya:
  - a. Menentukan beberapa data yang akan dibutuhkan.
  - b. Mengumpulkan semua data-data yang akan dibutuhkan.
  - c. Mempersiapkan metode penelitian.
2. Studi Literatur, meliputi:
  - a. Mempelajari konsep data mining.
  - b. Mempelajari algoritma *Naïve Bayes*.
  - c. Mempelajari *confusion matrix*.
3. Data Penelitian  
Pada tahap ini peneliti mengumpulkan semua data yang dibutuhkan, seperti jenis kelamin, umur, SKS, IPK, status mahasiswa dan kelulusan mahasiswa yang akan digunakan pada data *training*.
4. Penerapan Algoritma *Naïve Bayes*, berupa:
  - a. Menentukan atribut apa saja yang akan digunakan dalam memprediksi kelulusan mahasiswa Sistem Informasi di Universitas XXX.
  - b. Melakukan klasifikasi pada atribut yang telah ditetapkan, agar mendapatkan data yang diinginkan.
  - c. Menghitung probabilitas pada algoritma dan memasukan data

*testing* untuk dilakukan pengujian.

- d. Menentukan prediksi kelulusan setelah dilakukan penghitungan pada data yang diuji.
5. Pengembangan Perangkat Lunak  
Tahapan pengembangan sistem dilakukan dengan menggunakan metode Pengembangan *Waterfall* yang lebih menekankan waktu pengembangan dengan singkat pada pengembangan perangkat lunak yang terdiri dari tahapan-tahapan analisis, perancangan, implementasi dan pengujian.

Tahap-tahap penelitian yang akan dilakukan, yaitu:

1. Menentukan kebutuhan dari data yang akan digunakan nantinya, seperti data kelulusan mahasiswa Sistem Informasi di Universitas XXX dan metode klasifikasi dari algoritma *Naïve Bayes*.
2. Mengumpulkan data-data yang telah didapat untuk selanjutnya diolah.
3. Menyiapkan alat-alat dan bahan yang akan digunakan untuk penelitian. Alat yang digunakan yaitu berupa laptop sebagai perangkat keras dan beberapa *software* untuk menunjang penelitian, sedangkan bahan-bahan yang digunakan yaitu sekumpulan data-data yang telah dikumpulkan untuk selanjutnya menentukan atribut pada sistem yang akan dibuat.

Tahapan diatas bisa disebut sebagai studi literatur dan observasi, yaitu data-data yang telah diproses selanjutnya akan diolah sehingga terciptanya sebuah aplikasi. Tahapan tersebut akan diimplementasi untuk memprediksi tingkat kelulusan mahasiswa yang akan diterapkan pada Program Studi Sistem Informasi di Universitas XXX.

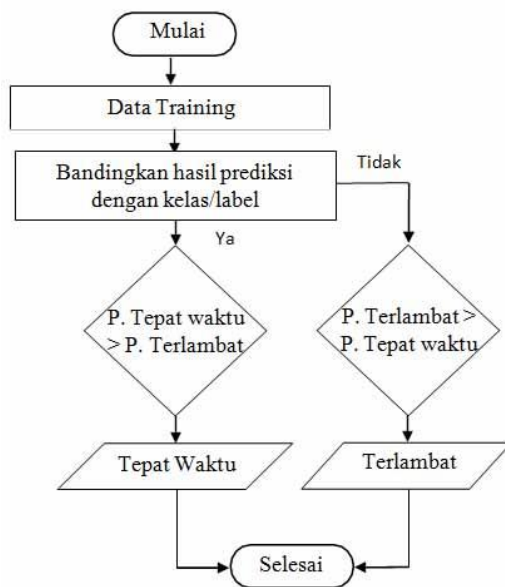
#### 3.2 Pengolahan Data Awal

Pada proses pengolahan data awal terdapat beberapa tahapan, yaitu:

### 1. Data selection

Dalam penelitian ini data yang akan digunakan adalah data kelulusan seluruh mahasiswa pada Program Studi Sistem Informasi Universitas XXX berjumlah 100 mahasiswa yang di peroleh dari PUSKOMINFO, BAAK, dan Kaprodi SI Universitas XXX. Data terdiri dari NIM, Nama, Nama, jenis kelamin, tempat tanggal lahir, SKS lulus, angkatan, tanggal yudisium, tahun lulus, nama program studi, dan IPK. Dilakukan seleksi dan menghilangkan atribut yang tidak dibutuhkan.

### 3.2.1 Flowchart Algoritma Naïve Bayes



### Flowchart Algoritma Naïve Bayes pada Sistem

Berdasarkan Gambar 3.3 diatas algoritma *Naïve Bayes* adalah memasukan data yang akan diolah (data *training*), yaitu data yang sudah diseleksi dan dipilih atributnya untuk dihitung jumlah dan probabilitas kondisi dari data yang telah di *input*, setelah mendapatkan hasil probabilitas kondisi dari data

*training* maka data *testing* dapat di *input* untuk dihitung jumlah dan probabilitas prediksi sehingga dapat di bandingkan hasil prediksi dengan kelas/label yang ada.

Apabila probabilitas tepat waktu lebih besar daripada probabilitas terlambat maka data masuk ke kelas “Tepat Waktu”, apabila probabilitas terlambat lebih besar daripada probabilitas tepat waktu maka data masuk ke kelas “Terlambat”.

### Eksperimen dan Pengujian Metode

#### 1. Data Training

Data *training* yang akan dihitung menggunakan algoritma *Naïve Bayes* menampilkan atribut yang terdapat pada tabel 3.4 berdasarkan data mahasiswa yang di dapat, sejumlah 100 data kelulusan seluruh mahasiswa Sistem Informasi Univeritas XXX yang akan dilakukan pemilihan data *training* seperti pada tabel 3.4 dibawah ini.

Tabel 3.4 Data Training

NO	JENIS KELAMIN	UMUR	SKS	IPK	STATUS MHS	KELULUSAN
1	L	24-29	rerata sks >=19	ipk >=3.00	bekerja	terlambat
2	L	24-29	rerata sks >=19	ipk <3.00	bekerja	terlambat
3	L	lebih dari 29	rerata sks >=19	ipk <3.00	bekerja	terlambat
4	L	lebih dari 29	rerata sks >=19	ipk >=3.00	bekerja	terlambat
5	L	18-23	rerata sks >=19	ipk >=3.00	mahasiswa	terlambat
6	P	18-23	rerata sks >=19	ipk >=3.00	mahasiswa	terlambat
7	L	24-29	rerata sks >=19	ipk >=3.00	bekerja	terlambat
8	L	24-29	rerata sks >=19	ipk <3.00	bekerja	terlambat
9	L	24-29	rerata sks >=19	ipk <3.00	bekerja	terlambat
10	L	24-29	rerata sks >=19	ipk <3.00	bekerja	terlambat
11	L	lebih dari 29	rerata sks >=19	ipk >=3.00	bekerja	terlambat

Dikarenakan keterbatasannya halaman, maka penulis hanya akan menampilkan 41 data untuk dijadikan tampilan awalnya saja, tampilan lengkap data *training* akan dilampirkan pada lampiran I.

Penghitungan dibawah ini menggunakan aplikasi *Microsoft Excel* untuk mempermudah perhitungan algoritma *Naïve Bayes*. Setelah mengetahui data

*training*, maka dapat dihitung probabilitas kondisi dari data tersebut sehingga menghasilkan data seperti pada tabel dibawah ini.

**Tabel 3.5 Probabilitas Kondisi Data Training**

P(TEPAT/TERLAMBAT)	0,500	0,500
Total	1,000	
P(JENIS KELAMIN)	TEPAT	TERLAMBAT
L	0,600	0,808
P	0,400	0,192
Total	1,000	1,000
P(UMUR)	TEPAT	TERLAMBAT
18-23	0,400	0,240
24-29	0,520	0,600
lebih dari 29	0,080	0,160
Total	1,000	1,000
P(SKS)	TEPAT	TERLAMBAT
rerata sks <19	0,260	0,100
rerata sks >=19	0,740	0,900
Total	1,000	1,000
P(IPK)	TEPAT	TERLAMBAT
ipk >=3.00	0,900	0,720
ipk <3.00	0,100	0,280
Total	1,000	1,000
P(STATUS MHS)	TEPAT	TERLAMBAT
mahasiswa	0,400	0,280
bekerja	0,600	0,720
Total	1,000	1,000

## 2. Data Testing

Data *testing* yang digunakan dipilih berdasarkan data kelulusan mahasiswa Program Studi Sistem Informasi yang terdapat pada data *training*. Berdasarkan referensi dari peneliti lain, maka peneliti akan mengambil 60 % dari data latih yang ada untuk dijadikan data uji, maka terdapat data *testing* sejumlah 60 data. Data akan dipilih secara acak dan dinilai dari atribut yang berbagai macam pilihannya sehingga data yang diuji akan menghasilkan akurasi yang baik.

## 3. Pengujian Algoritma Naïve Bayes

Setelah mengetahui probabilitas kondisi dan data *testing*, maka akan dilakukan pengujian menggunakan algoritma *Naïve Bayes* untuk memprediksi kelulusan mahasiswa. Setelah dihitung menggunakan *Microsoft Excel*, maka hasil yang didapat seperti pada tabel 3.7.

**Tabel 3.7 Probabilitas Prediksi Data Testing**

NO	TEPAT	TERLAMBAT	CONFUSION MATRIX	PREDIKSI
1	0.073	0.087	TN	TERLAMBAT
2	0.008	0.034	TN	TERLAMBAT
3	0.001	0.009	TN	TERLAMBAT
4	0.011	0.023	TN	TERLAMBAT
5	0.037	0.013	TP	TEPAT
6	0.016	0.008	TP	TEPAT
7	0.073	0.087	TN	TERLAMBAT
8	0.008	0.034	TN	TERLAMBAT
9	0.008	0.034	TN	TERLAMBAT
10	0.008	0.034	TN	TERLAMBAT
11	0.011	0.023	TN	TERLAMBAT
12	0.011	0.023	TN	TERLAMBAT
13	0.073	0.087	TN	TERLAMBAT
14	0.003	0.021	TN	TERLAMBAT
15	0.011	0.023	TN	TERLAMBAT
16	0.004	0.003	TN	TEPAT
17	0.008	0.002	TP	TEPAT
18	0.016	0.008	TP	TEPAT
19	0.037	0.013	TP	TEPAT
20	0.056	0.035	TP	TEPAT
21	0.037	0.013	TP	TEPAT
22	0.024	0.021	TP	TEPAT
23	0.073	0.087	FN	TERLAMBAT
24	0.024	0.021	TP	TEPAT
25	0.056	0.035	TP	TEPAT
60	.....	.....	.....	.....

Dikarenakan keterbatasannya halaman, maka penulis hanya akan menampilkan 60 data untuk dijadikan tampilan awalnya saja.

## 3.3 Evaluasi dan Validasi Hasil

Pada tahap ini data yang telah diuji menggunakan metode klasifikasi algoritma *Naïve Bayes* akan dilihat evaluasi dan validasi dari hasil pengujian, yaitu dengan menggunakan metode *confusion matrix* untuk mengetahui *accuracy*, *recall*, dan *precision* dari data yang telah diuji. Peneliti menggunakan *Microsoft Excel* sebagai penghitungan manual dan hasilnya dapat dilihat pada tabel dibawah ini.

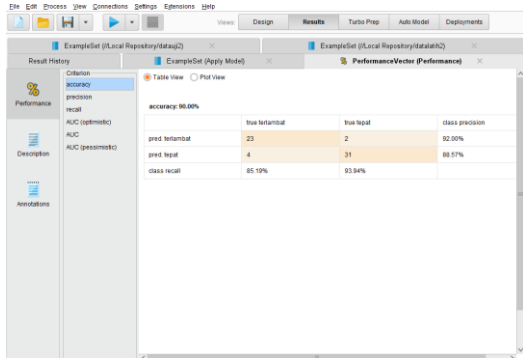
**Tabel 3.8 Hasil Confusion Matrix**

Predicted ↓	CLASS	
	TEPAT	TERLAMBAT
TEPAT	31	4
TERLAMBAT	2	23

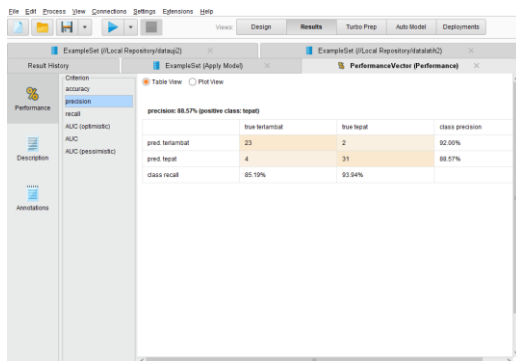
Untuk menghasilkan nilai yang akurat, maka peneliti memakai aplikasi RapidMiner agar mengetahui nilai aku-



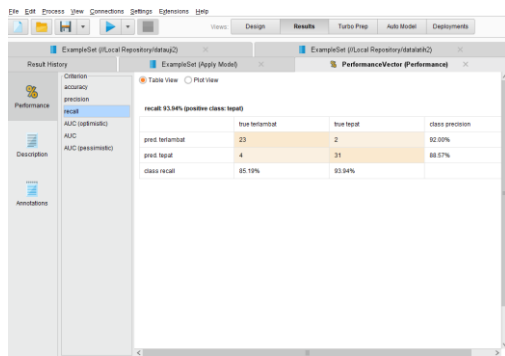
rasi yang di hitung memiliki kesamaan apabila diolah dengan sistem, sehingga dapat menghasilkan nilai yang sama seperti pada gambar dibawah ini.



**Gambar 3.4 Hasil Accuracy dengan RapidMiner**



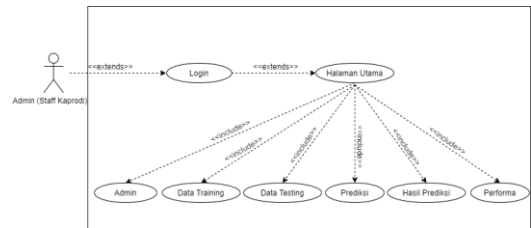
**Hasil Precision dengan RapidMiner**



**Hasil Recall dengan RapidMiner**

Setelah dilakukan penghitungan menggunakan *Microsoft Excel* dan aplikasi *RapidMiner*, dapat dilihat keduanya menghasilkan nilai *accuracy*, *precision*, dan *recall* yang sama, sehingga penghitungan ini berhasil untuk diuji.

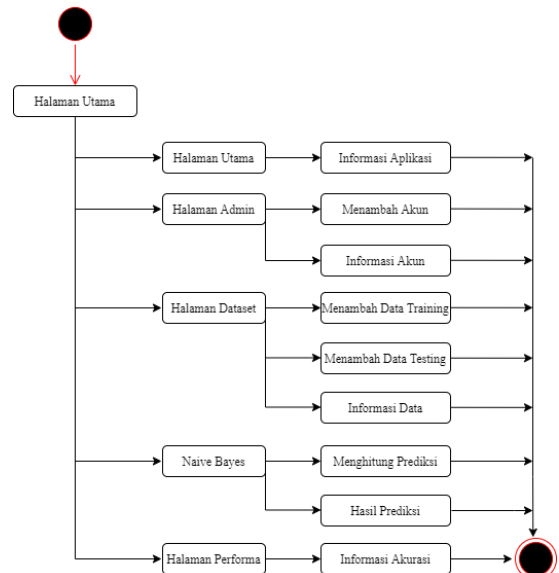
### 3.4 Use Case Diagram



**Gambar 3.7 Use Case Diagram pada Sistem yang Diusulkan**

#### 3.4.1 Activity Diagram

*Activity diagram* dari perancangan *use case* sistem yang diusulkan sebagai berikut:

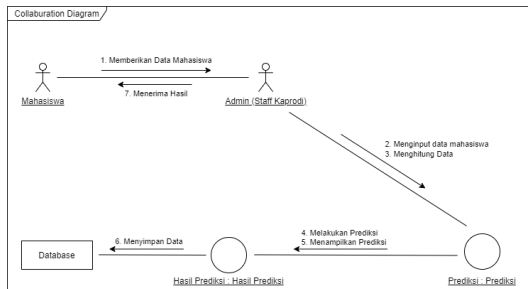


**Gambar 3.8 Activity Diagram Sistem yang Diusulkan**

#### 3.4.2 Class Diagram



**Gambar 3.14 Class Diagram**  
**3.4.3 Collaboration Diagram**



**Gambar 3.15 Collaboration Diagram**

**HASIL DAN PEMBAHASAN**

**4.1 Hasil**

**4.1.1 Hasil Perhitungan Data Training dengan Algoritma Naïve Bayes**

Berikut merupakan hasil perhitungan data *training* menggunakan algoritma *Naïve Bayes*, terdapat pada Tabel 4.1.

**Tabel 4.1 Hasil Perhitungan Data Training**

Probabilitas (TEPAT/TERLAMBAT)	C (Tepat)= 50/100=0.5	C (Terlambat)= 50/100 =0.5
<b>Probabilitas (GENDER)</b>	<b>TEPAT</b>	<b>TERLAMBAT</b>
L	X   C (Laki-laki   Tepat) = 35/50 = 0.7	X   C (Laki-laki   Terlambat) = 31/50 = 0.62
P	X   C (Perempuan   Tepat) = 15/50 = 0.3	X   C (Perempuan   Terlambat) = 19/50 = 0.38
<b>Probabilitas (UMUR)</b>	<b>TEPAT</b>	<b>TERLAMBAT</b>
18-23	X   C (18-23   Tepat) = 20/50 = 0.4	X   C (18-23   Terlambat) = 12/50 = 0.24
24-29	X   C (24-29   Tepat) = 26/50 = 0.52	X   C (24-29   Terlambat) = 30/50 = 0.6
lebih dari 29	X   C (lebih dari 29   Tepat) = 4/50 = 0.08	X   C (lebih dari 29   Terlambat) = 8/50 = 0.16
<b>Probabilitas (SKS)</b>	<b>TEPAT</b>	<b>TERLAMBAT</b>
rerata sks <19	X   C (rerata sks < 19   Tepat) = 13/50 = 0.26	X   C (rerata sks < 19   Terlambat) = 5/50 = 0.1
rerata sks >=19	X   C (rerata sks >= 19   Tepat) = 37/50 = 0.74	X   C (rerata sks >= 19   Terlambat) = 45/50 = 0.9
<b>Probabilitas (IPK)</b>	<b>TEPAT</b>	<b>TERLAMBAT</b>
>= 3.00	X   C (ipk >= 3.00   Tepat) = 45/50 = 0.9	X   C (ipk >= 3.00   Terlambat) = 36/50 = 0.72
< 3.00	X   C (ipk < 3.00   Tepat) = 5/50 = 0.1	X   C (ipk < 3.00   Terlambat) = 14/50 = 0.28
<b>Probabilitas (STATUS)</b>	<b>TEPAT</b>	<b>TERLAMBAT</b>
MHS)		
mahasiswa	X   C (mahasiswa   Tepat) = 20/50 = 0.4	X   C (mahasiswa   Terlambat) = 14/50 = 0.28
bekerja	X   C (bekerja   Tepat) = 30/50 = 0.6	X   C (bekerja   Terlambat) = 36/50 = 0.72

Hasil ini didapatkan setelah dilakukannya perhitungan terhadap data kelulusan mahasiswa Sistem Informasi sebanyak 100 data mahasiswa yang didapat.

Berdasarkan hasil perhitungan dari data *training* diatas, maka akan menghasilkan probabilitas kondisi dari seluruh kelulusan Tepat sejumlah 0.5 dan Terlambat sejumlah 0.5. Hasil probabilitas kondisi kelulusan Gender Laki-laki yang Tepat sejumlah 0.7 dan Terlambat sejumlah 0.62, untuk kelulusan Perempuan yang Tepat sejumlah 0.3 dan Terlambat sejumlah 0.38. Hasil probabilitas kondisi kelulusan Umur 18-23 yang Tepat sejumlah 0.4 dan Terlambat sejumlah 0.24, untuk kelulusan 24-29 yang Tepat sejumlah 0.52 dan Terlambat sejumlah 0.6, untuk kelulusan Lebih dari 29 yang Tepat sejumlah 0.08 dan Terlambat sejumlah 0.16.

Hasil probabilitas kondisi kelulusan SKS Rerata SKS <19 yang Tepat sejumlah 0.26 dan Terlambat sejumlah 0.1, untuk kelulusan Rerata SKS >=19 yang Tepat sejumlah 0.74 dan Terlambat sejumlah 0.9.

Hasil probabilitas kondisi kelulusan IPK >=3.00 yang Tepat sejumlah 0.9 dan Terlambat sejumlah 0.72, untuk kelulusan <3.00 yang Tepat sejumlah 0.1 dan Terlambat sejumlah 0.28.

Hasil probabilitas kondisi kelulusan Status MHS Mahasiswa yang Tepat sejumlah 0.4 dan Terlambat sejumlah 0.28, untuk kelulusan Bekerja yang Tepat sejumlah 0.6 dan Terlambat sejumlah 0.72.

**4.1.2 Hasil Perhitungan Data Testing dengan Algoritma Naïve Bayes**

Berikut merupakan hasil perhitungan data *testing* menggunakan algoritma *Naïve Bayes*, terdapat pada Tabel 4.2.

**Tabel 4.2 Hasil Perhitungan Data Testing**

NO	TEPAT	TERLAMBAT	CONFUSION MATRIX	PREDIKSI
1	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.9 \times 0.6 = 0.073$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.72 \times 0.72 = 0.087$	TN	TERLAMBAT
2	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.1 \times 0.6 = 0.008$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.28 \times 0.72 = 0.034$	TN	TERLAMBAT
3	$=0.5 \times 0.7 \times 0.08 \times 0.74 \times 0.1 \times 0.6 = 0.001$	$=0.5 \times 0.62 \times 0.16 \times 0.9 \times 0.28 \times 0.72 = 0.009$	TN	TERLAMBAT
4	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.9 \times 0.6 = 0.011$	$=0.5 \times 0.62 \times 0.6 \times 0.99 \times 0.72 \times 0.72 = 0.023$	TN	TERLAMBAT
5	$=0.5 \times 0.7 \times 0.4 \times 0.74 \times 0.9 \times 0.4 = 0.037$	$=0.5 \times 0.62 \times 0.24 \times 0.9 \times 0.72 \times 0.28 = 0.013$	TP	TEPAT
6	$=0.5 \times 0.3 \times 0.4 \times 0.74 \times 0.9 \times 0.4 = 0.016$	$=0.5 \times 0.38 \times 0.24 \times 0.9 \times 0.72 \times 0.28 = 0.008$	TP	TEPAT
7	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.9 \times 0.6 = 0.073$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.72 \times 0.72 = 0.087$	TN	TERLAMBAT
8	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.1 \times 0.6 = 0.008$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.28 \times 0.72 = 0.034$	TN	TERLAMBAT
9	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.1 \times 0.6 = 0.008$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.28 \times 0.72 = 0.034$	TN	TERLAMBAT
10	$=0.5 \times 0.7 \times 0.52 \times 0.74 \times 0.1 \times 0.6 = 0.008$	$=0.5 \times 0.62 \times 0.6 \times 0.9 \times 0.28 \times 0.72 = 0.034$	TN	TERLAMBAT
60	.....	.....	.....	.....

#### 4.1.3 Hasil Perhitungan dengan Confusion Matrix

Berdasarkan hasil perhitungan data training dan data testing maka didapatkan hasil perhitungan menggunakan confusion matrix untuk mengetahui accuracy, recall, dan precision dari data yang telah dihitung, dapat dilihat pada tabel 4.3.

**Tabel 4.3 Confusion Matrix**

Predicted ↓	CLASS	
	TEPAT	TERLAMBAT
TEPAT	31	4
TERLAMBAT	2	23

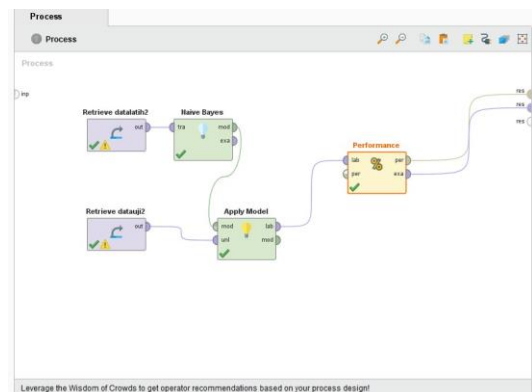
Jumlah yang telah diketahui lalu dihitung accuracy, recall dan precision pada perhitungan dibawah ini:

**Tabel 4.4 Hasil Perhitungan Confusion Matrix**

	Perhitungan	Hasil
Accuracy	$\frac{31+14}{31+2+4+23} \times 100\%$	90%
Recall	$\frac{31}{31+4} \times 100\%$	88.57%
Precision	$\frac{31}{31+4} \times 100\%$	93.94%

Penelitian ini dilakukan dengan

menggunakan software data mining RapidMiner. Penilaian akan dilakukan dengan melihat akurasi yang dihasilkan, semakin tinggi presentase dari hasil klasifikasi maka semakin tinggi pula akurasi dari metode yang digunakan.



#### Pengaturan Proses Utama RapidMiner

Setelah dilakukan pengujian nilai accuracy, recall, dan precision dari algoritma Naive Bayes menggunakan software RapidMiner maka menghasilkan nilai accuracy 90%, recall 88.57%, dan precision 93.93%. Hasil ini sama nilainya dengan perhitungan yang dilakukan manual. Hal ini membuktikan bahwa algoritma Naive Bayes dapat diterapkan untuk memprediksi tingkat kelulusan mahasiswa Sistem Informasi di Universitas XXX.

#### 4.2 Rancangan Antarmuka

Rancangan antarmuka bertujuan untuk memberikan gambaran bentuk antarmuka dari perangkat lunak yang akan digunakan oleh pengguna agar dapat berinteraksi dengan perangkat lunak.

##### 4.2.1 Rancangan Antarmuka Menu Login

Sebelum masuk pada aplikasi, pengguna akan terlebih dahulu disajikan menu halaman Login untuk masuk ke menu utama, seperti pada gambar 4.2.

Halaman Login

Enter Username ...

Password

Login

**Rancangan Antarmuka Menu Login**

#### 4.2.2 Rancangan Antarmuka Menu Utama

Setelah berhasil masuk pengguna akan disajikan menu utama dari *website* Prediksi Kelulusan, didalamnya terdapat nama admin, logo admin, dan sekilas tentang *website* Prediksi Kelulusan, seperti pada gambar 4.3.

Home

Admin

Admin

Data Training

Data Testing

Prediksi

Hasil Prediksi

Performa

Slide

**PREDIKSI TINGKAT KELULUSAN MAHASISWA PROGRAM STUDI SISTEM INFORMASI**

Website ini merupakan sebuah aplikasi yang didalamnya terdapat penghitungan algoritma Naive Bayes untuk memprediksi kelulusan mahasiswa dari Program Studi Sistem Informasi Universitas Dirgantara Marsekal Suryadarma

**Rancangan Antarmuka Menu Utama**

#### 4.2.3 Rancangan Antarmuka Menu Admin

Pada halaman ini berisi beberapa akun yang telah terdaftar menjadi admin dan dapat menambah akun pada *button* Tambah Data, dengan mengisi nama admin, *username*, dan *password*, seperti pada gambar 4.4.

Admin

Admin

Tambah Data

ID	Nama Admin	Username	Password	Aksi

Prediksi

Hasil Prediksi

Performa

**Rancangan Antarmuka Menu Admin**

#### 4.2.4 Rancangan Antarmuka Menu Data Training

Pada Menu Data Training, terdapat beberapa data *training* yang telah di *input* sebelumnya. Data ini berisi jenis kelamin, umur, SKS, IPK, status mahasiswa, dan kelulusan mahasiswa, seperti pada gambar 4.5.

Terdapat *button* *Import* untuk memasukan file yang terdapat pada 70computer dan *button*

Data Training

Admin

Admin

choose file

Import

Tambah Data

No	Jenis Kelamin	Umur	SKS	IPK	Status Mhs	Kelulusan	Aksi

Prediksi

Hasil Prediksi

Performa

Tambah Data untuk menambahkan data *training* baru yang diinginkan, seperti pada gambar 4.6.

Tambah Data Training

Jenis Kelamin

Umur

SKS

IPK

Status Mahasiswa

Kelulusan

Reset Close Tambah Data

**Rancangan Form Tambah Data Training**

Pada button Edit Data, digunakan untuk mengedit data *training* yang telah di *input* apabila ada data yang salah dan dapat diperbarui dengan data yang benar, seperti pada gambar 4.7.

Edit Data Training

No Train

Jenis Kelamin

Umur

SKS

IPK

Status Mahasiswa

Kelulusan

Reset Close Edit Data

**Rancangan Edit Data Training**

#### 4.2.5 Rancangan Antarmuka Menu Data Testing

Pengguna akan disajikan dengan beberapa data *testing* yang telah di *input* sebelumnya. Data ini berisi jenis kelamin, umur, SKS, IPK, status mahasiswa, kelulusan mahasiswa dan prediksi dari data yang telah dihitung sebelumnya, seperti pada gambar 4.8.

Data Testing Admin

Admin

choose file import

Data Training Tambah Data

No	Jenis Kelamin	Umur	SKS	IPK	Status Mhs	Kelulusan	Prediksi	Aksi

Prediksi

Hasil Prediksi

Performa

Terdapat button *Import* untuk memasukan file yang terdapat pada 720computer dan button Tambah Data untuk menambahkan data *testing* baru yang diinginkan, seperti pada gambar 4.9.

Tambah Data Testing

Jenis Kelamin

Umur

SKS

IPK

Status Mahasiswa

Kelulusan

Reset Close Tambah Data

**Gambar 4.9 Rancangan Form Tambah Data Testing**

Pada button Edit Data, digunakan untuk mengedit data *testing* yang telah di *input* apabila ada data yang salah dan dapat diperbarui dengan data yang benar, seperti pada gambar 4.10.

**Rancangan Edit Data Testing**

#### 4.2.6 Rancangan Antarmuka Menu Prediksi

Bagi pengguna yang ingin melakukan prediksi pada kelulusan mahasiswa dapat melakukan prediksi pada Menu Prediksi dengan memasukkan jenis kelamin, umur, SKS, IPK dan status mahasiswa, seperti pada gambar 4.11.

**Rancangan Antarmuka Menu Prediksi**

Data yang telah diproses pada Menu Prediksi akan menghasilkan nilai prediksi kelulusan mahasiswa, semua data yang telah dihitung prediksinya akan terdata pada Menu Hasil Prediksi dan dapat di *eksport* ke 74omputer pengguna, seperti pada gambar 4.12.

**Rancangan Antarmuka Menu Hasil Prediksi**

#### 4.2.7 Rancangan Antarmuka Menu Hasil Prediksi

Untuk melihat keterangan dari data yang telah dihitung maka pada kolom aksi terdapat button Detail untuk melihat keterangan yang ada pada data, seperti pada gambar 4.13.

**Rancangan Detail Hasil Prediksi**

#### 4.2.8 Rancangan Antarmuka Menu Performa

Pada Menu Performa, pengguna akan disajikan hasil *accuracy*, *recall*, dan *precision* dari data *testing* yang telah diuji sebelumnya, seperti pada gambar 4.14.

### Rancangan Antarmuka Menu Performa

#### 5.1 Kesimpulan

Berdasarkan hasil perhitungan menggunakan Algoritma *Naive Bayes* dan dibuktikan menggunakan aplikasi RapidMiner menggunakan data *testing* yang diambil sebanyak 60% dari data *training* berhasil mendapatkan nilai *accuracy* 90%, *recall* 93.94%, dan *precision* 88.57%. Hasil ini didapatkan dari data *testing* yang diambil sebanyak

60% dari data *training*. Maka dapat disimpulkan bahwa penerapan Algoritma *Naive Bayes* dapat digunakan untuk menentukan prediksi ketepatan waktu tingkat kelulusan mahasiswa Sistem Informasi di Universitas XXX.

#### 5.2 Saran

Agar sistem yang diterapkan menggunakan algoritma *Naive Bayes* untuk memprediksi ketepatan waktu kelulusan mahasiswa Sistem Informasi di Universitas XXX berbasis *website* menjadi lebih baik, maka diharapkan pada penelitian berikutnya disarankan sebagai berikut:

1. Menambahkan atribut/indikator baru sehingga menghasilkan lebih banyak informasi dan penghitungan baru.
2. Menggunakan algoritma lainnya sehingga menjadi perbandingan dengan algoritma *Naive Bayes*.

#### DAFTAR PUSTAKA

- Adekitan, A. I., & Salau, O. (2019). The impact of engineering students' performance in the first three years on their graduation result using educational data mining. *Heliyon*, 5(2), e01250. <https://doi.org/10.1016/j.heliyon.2019.e01250>
- Azhar, Z., Hutahaean, J., Siagian, Y., & Syah, A. Z. (2019). Pelatihan Microsoft Excel 2010 Pada Siswa Lembaga Kursus Pendidikan (LKP) Mandiri. *Anadara Pengabdian Kepada Masyarakat*, 1(2), 185.
- Awaludin, M., & Mantik, H. (2023). PENERAPAN METODE SERVQUAL PADA SKALA LIKERT UNTUK MENDAPATKAN KUALITAS PELAYANAN KEPUASAN PELANGGAN Muryan Awaludin 1, Hari Mantik 2, Fadli Fadillah 3 1. *Jurnal Sistem Informasi Univesitas Suryadarma*, 10(1).
- Broto Legowo, M., & Indiarto, B. (2017). Model Sistem Penjaminan Mutu Berbasis Integrasi Standar Akreditasi BAN-PT dan ISO 9001:2008. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(2), 90. <https://doi.org/10.29207/resti.v1i2.51>
- Destiningrum, M., & Adrian, Q. J. (2017). Sistem Informasi Penjadwalan Dokter Berbassis Web Dengan Menggunakan Framework Codeigniter (Studi Kasus: Rumah Sakit Yukum Medical Centre). *Jurnal Teknoinfo*, 11(2), 30. <https://doi.org/10.33365/jti.v11i2.24>

- Dewi, S. (2016). Komparasi 5 Metode Algoritma Klasifikasi Data Mining Pada Prediksi Keberhasilan Pemasaran Produk Layanan Perbankan. *None*, 13(1), 60–66.
- Diqy Fakhrun Shiddieq, S.T., M. K., & Patricia. (2020). *Implementasi Algoritma Data Mining Naive Bayes untuk Prediksi Kelulusan Mahasiswa*. (456), 8–13.
- Etriyanti, E., Syamsuar, D., & Kunang, Y. N. (2020). Implementasi Data Mining Menggunakan Algoritme Naive Bayes Classifier dan C4.5 untuk Memprediksi Kelulusan Mahasiswa. *Telematika*, 13(1), 56–67. <https://doi.org/10.35671/telematika.v13i1.881>
- Febie Elfaladonna, A. R. (2019). Analisa Metode Classification-Decission Tree Dan Algoritma C.45 Untuk Memprediksi Penyakit Diabetes dengan menggunakan Aplikasi Rapid Miner. *Science And Information Technology*, 2(1), 10–17. Retrieved from <http://jurnal.stiki-indonesia.ac.id/index.php/sintechjournal>
- Firman Azhar Riyadi. (2020). *Implementasi Metode naive Bayes Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu Prodi Informatika (Studi Kasus : Universitas Teknologi Yogyakarta)*. 1–9.
- Fitri Ayu and Nia Permatasari. (2018). perancangan sistem informasi pengolahan data PKL pada divisi humas PT pegadaian. *Jurnal Infra Tech*, 2(2), 12–26. Retrieved from <http://journal.amikmahaputra.ac.id/index.php/JIT/article/download/33/25>
- Han, J., & Kamber, M. (2006). Data mining: Data mining concepts and techniques. In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. <https://doi.org/10.1109/ICMIRA.2013.45>
- Han, J., Kamber, M., & Pei, J. (2012). Data mining: Data mining concepts and techniques (Third Edition). In *Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013*. <https://doi.org/10.1109/ICMIRA.2013.45>
- Handayani, S. (2018). Perancangan Sistem Informasi Penjualan Berbasis E- Commerce Studi Kasus Toko Kun Jakarta. *ILKOM Jurnal Ilmiah*, 10(2), 182– 189. <https://doi.org/10.33096/ilkom.v10i2.310.182-189>
- Harmianty, P. (2017). *Aplikasi Prediksi Kelulusan Tepat Waktu Mahasiswa Menggunakan Algoritma C4.5*. 1–9.
- Hendini, A. (2016). Pemodelan UML Sistem Informasi Monitoring Penjualan Dan Stok Barang (Studi Kasus: Distro Zhezha Pontianak). *Crop Science*, 23(2), 201–205. <https://doi.org/10.2135/cropsci1983.0011183x002300020002x>