

Prediksi Kelulusan Mahasiswa Tepat Waktu atau Tidak Menggunakan Metode *Naïve Bayes Classifier* (Studi kasus : Mahasiswa Fakultas Teknik)

Ayyub, Iddrus

Informatika, Universitas Wiraraja, Indonesia

ayyubalfredo@gmail.com , iddrus@wiraraja.ac.id

Diterima : 17 Oktober. Disetujui : 05 November 2024. Dipublikasikan : 05 November 2024.

ABSTRACT - Higher education institutions have the obligation to produce competent graduates. This can be assessed by the graduation rate of their students. Timely graduation is one of the evaluation points in the accreditation process of higher education institutions. However, student graduation cannot always be detected quickly. There are many factors that influence delayed graduation, such as the students' understanding of the course material, which can be reflected in their GPA. Given these issues, there is a need for a system to predict student graduation rates based on existing variables. With such a system, higher education institutions are expected to develop policies that allow students to graduate on time. This study uses data from 358 students from the Faculty of Engineering, graduating from 2017 to 2022, with the Naive Bayes method, consisting of 286 training data and 72 testing data. The attributes used are name, school (Public/Private), type of school (High School/Vocational School/Islamic School), and GPA (Semester 1 - Semester 4). The stages include problem identification, data collection, data cleaning, data integration, data transformation (divided into training and testing data),

classification with Naive Bayes Classifier, validation, evaluation, and results. The research findings show an accuracy of 91.7%, which falls into the good classification category.

Keywords : Prediction, Naïve Bayes Classifier

ABSTRAK - Perguruan tinggi memiliki kewajiban untuk menghasilkan lulusan yang kompeten. Hal tersebut dapat dinilai dari tingkat kelulusan mahasiswanya. Kelulusan mahasiswa tepat waktu merupakan salah satu point penilaian dalam proses akreditasi perguruan tinggi. Namun kelulusan mahasiswa tidak selalu dapat dideteksi secara cepat, Banyak faktor yang menjadi pengaruh kelulusan mahasiswa terlambat, seperti tingkat pemahaman mahasiswa terhadap materi kuliah yang dapat dilihat dari IPK mahasiswa. Dari permasalahan yang ada, perlu adanya sistem untuk memprediksi tingkat kelulusan mahasiswa berdasarkan variabel-variabel yang ada. Dengan sistem yang dibuat diharapkan perguruan tinggi bisa membuat kebijakan sehingga mahasiswa dapat lulus tepat waktu. Penelitian ini menggunakan 358 data mahasiswa fakultas teknik lulusan angkatan 2017 -

2022, dengan metode Naive bayes, dengan rincian data training 286 data dan data testing 72 data. Atribut yang digunakan nama, sekolah (Negeri/Swasta), jenis sekolah (SMA/SMK /MA), dan IPS-1 – IPS4. Dengan tahapan identifikasi masalah, pengumpulan data, data cleaning, data integration data transformation (dibagi menjadi data training dan data tesing), klasifikasi dengan Naive bayes Classifier, validasi, evaluasi dan hasil. Hasil penelitian yang diperoleh yaitu akurasi = 91,7%, termasuk dalam kategori good classification.

Kata kunci : Prediksi, Naive Bayes Classifier

I. PENDAHULUAN

Perguruan tinggi yang memiliki pendidikan berkualitas, dapat menghasilkan mahasiswa yang kompeten, dan kreatif. Oleh karena itu setiap perguruan tinggi dituntut untuk melakukan hal itu, untuk melihat baik atau kurangnya kualitas Pendidikan di Indonesia bisa dilihat dari Akreditasi pada Badan Akreditasi Nasional Perguruan Tinggi atau BAN-PT. Salah satu yang bisa menjadi acuan untuk penilaian akreditasi, seperti yang tertera pada lampiran Pengaturan Badan Akreditasi Nasional Perguruan Tinggi No. 23 tahun 2022. Universitas Wiraraja (UNIJA) yang berdiri pada tahun 1986 bertempat dikawasan Sumenep Madura ini, termasuk salah satu perguruan tinggi yang memiliki Pendidikan berkualitas. Mahasiswa di UNIJA setiap tahunnya mengalami kenaikan dan penurunan yang tidak seimbang antara mahasiswa yang masuk dan yang keluar setelah menyelesaikan studinya. Ini bisa berpengaruh terhadap kualitas dan akreditasi dari UNIJA. UNIJA selama ini melakukan evaluasi tingkat kelulusan mahasiswa hanya melihat pada jumlah data pendaftaran wisuda, tanpa memperhatikan faktor-faktor permasalahan akademis maupun administratif yang dialami mahasiswa. Sedangkan mahasiswa yang tidak lulus tepat waktu, UNIJA hanya menggrahkan dan membimbing mahasiswa untuk segera menyelesaikan studinya.

UNIJA belum memiliki sistem proses prediksi kelulusan mahasiswa tepat waktu, ini menjadi permasalahan yang selama ini terjadi di UNIJA. Sehingga Bagian Biro Administrasi

Akademik dan Kemahasiswaan (BAAK) tidak bisa mengetahui satu angkatan tersebut lulus tepat waktu atau tidak, dan mahasiswa tidak mendapat solusi atas keterlambatan kelulusan yang terjadi sBerdasarkan permasalahan di atas, maka peneliti melakukan penelitian terhadap mahasiswa di Universitas Wiraraja Fakultas Teknik Prodi Teknik Sipil untuk membuat aplikasi prediksi kelulusan mahasiswa tepat waktu dengan menggunakan teknik data mining yaitu menggunakan Metode Algoritma Naive Bayes Classifier. Sitem ini nantinya bebrbasis Website yang memilki fitur untuk memprediksi kelulusan mahasiswa tepat waktu atau tidak berdasarkan probabilitas dari faktor faktor yang mempengaruhi kelulusan mahasiswa, sistem ini menggunakan konsep probabilitas yang belajar dari data sebelumnya sehingga bisa memprediksi peluang dimasa depan.

II. TEORI DASAR

2.1 Penelitian Terdahulu

Terkait kasus prediksi kelulusan mahasiswa tepat waktu atau tidak, telah dilakukan penelitian dari berbagai universitas dengan algoritma yang beragam. Penelitian sebelumnya dilakukan oleh Yovi (dkk) “Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode *Naive Bayes*”[1], dengan permasalahan yang digunakan berdasarkan rendahnya jumlah lulusan yang bisa mempengaruhi akreditasi dari sebuah kampus. Algoritma yang digunakan pada penelitian prediksi kelulusan mahasiswa tepat waktu ini adalah *naive bayes classifier*, pada penelitian ini menggunakan 20 data mahasiswa sebagai data uji, hasil akurasi yang diperoleh mencapai 90% , precision 90%, dan recal sebesar 100%. dapat disimpulkan bahwa dengan menerapkan metode Naive Bayes dalam prediksi kelulusan mahasiswa dikategorikan berhasil dengan nilai persentase akurasi mencapai 90%.

2.2 Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining bervariasi. Pemilihan metode atau algoritma

yang tepat sangat bergantung pada tujuan proses *Knowledge Discovery in Database (KDD)* secara keseluruhan.

Data mining merupakan proses iterative dan interaktif untuk menemukan pola atau model baru yang sempurna, bermanfaat dan dapat dimengerti dalam suatu database yang sangat besar (massive database). Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambil keputusan diwaktu yang akan datang, pola-pola ini dikenali perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lain[2].

2.3 Klasifikasi

Klasifikasi adalah proses menemukan model (atau fungsi) yang menggambarkan dan membedakan kelas atau konsep data. Model diturunkan berdasarkan analisis sekumpulan data pelatihan (yaitu, objek data yang label kelasnya diketahui). Model yang digunakan untuk memprediksi label kelas objek yang label kelasnya tidak diketahui. Tahapan klasifikasi data dibagi dalam 2 langkah proses. Langkah pertama adalah tahapan learning yang disebut sebagai fase training. Fase ini merupakan tahapan algoritma klasifikasi untuk membentuk aturan klasifikasi data yang diambil melalui analisi data training. Tahap kedua adalah klasifikasi, dimana data tes digunakan untuk memperkirakan akurasi dari aturan klasifikasi[3].

2.4 Prediksi:

Prediksi adalah kegiatan yang dapat dijalankan guna memperkirakan hal yang dimungkinkan terjadi pada masa mendatang, melalui pemanfaatan berbagai data lama dengan berdasarkan pada suatu indikator tertentu. Atau belajar dari data lama yang besar lalu diolah menjadi informasi maka didapatkan pengetahuan dari data tersebut sehingga bisa menjadi kebijakan di masa depan. Berbagai permasalahan yang memerlukan prediksi, antara lain seperti dalam hal harga, hasil produksi, prediksi tingkat kelulusan, ataupun lain sebagainya. Prediksi bertujuan untuk meminimalkan kesalahan, yaitu selisih antara hasil perkiraan dan apa yang sebenarnya terjadi. [4].

2.5 Naïve Bayes Classifier

Naive Bayes adalah salah satu metode pengklasifikasian sederhana yang sering digunakan karena mudah diterapkan dan memiliki hasil yang baik saat diterapkan pada banyak kasus. Sedangkan kelemahan dari metode ini sendiri yaitu adanya asumsi atau dengan kata lain kondisi kelas saling bebas, sehingga kurang akurat. adapun pada prakteknya, kebergantungan ada diantara variable. Misalnya rumah sakit: pasien, umur, keluarga. Kebergantungan diantara variable ini tidak dapat dimodelkan dengan menggunakan *Naive Bayes Classifier*[5]. Dalam klasifikasi Bayesian, hal yang paling penting menemukan probabilitas posterior yaitu probabilitas label yang diberikan beberapa fitur yang diamati, (L). Dengan bantuan teorema Bayes, kita dapat menyatakannya dalam bentuk formula berikut ini :

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

Dimana,

X, merupakan data dengan kelas yang belum diketahui

H, merupakan hipotesis data yang merupakan suatu kelas spesifik

$P(H|X)$, merupakan probabilitas hipotesis H berdasarkan pada kondisi X (posterior probabilitas)

$P(H)$, merupakan probabilitas hipotesis H (prior probabilitas)

$P(X|H)$, merupakan probabilitas X berdasarkan pada kondisi pada hipotesis H, dan

$P(X)$, merupakan probabilitas X

2.6 PHP (Hypertext Preprocessor)

PHP(Hypertext Preprocessor) merupakan bahasa pemrograman yang digunakan untuk membuat aplikasi berbasis website. Sebagai sebuah aplikasi, website tersebut hendaknya memiliki sifat dinamis dan interaktif. Memiliki sifat dinamis artinya, website tersebut bisa berupa tampilan kontennya sesuai, kondisi tertentu (misalnya menampilkan produk yang berbeda-beda untuk setiap pengunjung). Interaktif artinya, website tersebut dapat memberi feedback bagi user (misalnya, menampilkan hasil pencarian produk). PHP merupakan bahasa pemrograman berjenis server-

side. Dengan demikian, PHP akan diproses oleh server yang hasil olahannya akan dikirim kembali ke browser. Oleh karena itu, salah-satu tool yang harus tersedia sebelum memulai pemrograman PHP adalah server[6].

2.7 Sublime Text

Menurut Saputra (dkk 2021), Sublime Text adalah texteditor berbasis Python, sebuah text editor yang elegan, kaya fitur, cross platform, mudah Jurnal Ekonomi, Bisnis dan Pendidikan, 1(9), 2021, 914-921 916 dan simple yang cukup terkenal dikalangan developer (pengembang) dan desainer. Sublime Text digunakan sebagai editor dari bahasa pemrograman PHP dalam melakukan pengelolaan konten di dalam aplikasi server [7].

2.8 Unified Modeling Language (UML)

Unified Modeling Language (UML) menurut Syarif (dkk 2020), merupakan bahasa visual untuk memperoleh pemodelan dan komunikasi mengenai sebuah sistem dengan menggunakan diagram dan teks-teks pendukung. UML menggunakan notasi grafis untuk mengekspresikan desain proyek perangkat lunak. UML membantu tim proyek berkomunikasi, mengeksplorasi desain potensial, dan memvalidasi desain arsitektur perangkat lunak Beberapa pemodelan termasuk kedalam pemodelan UML seperti *Use Case Diagram*, *Class Diagram*, *Activity Diagram*, dan *Squence Diagram* [8].

2.9 Flowchart

Flowchart atau disebut juga diagram alir merupakan notasi yang menggambarkan alur atau tahapan suatu proses. Dimana, suatu alur proses disimbolkan (dinotasikan) dengan bentuk-bentuk simbol standar untuk menunjukkan aktivitas, kondisi, dan alur logika dari proses yang digambarkan. Biasanya yang dimaksud dengan notasi ini berupa tahapan-tahapan seperti memulai proses, memproses sesuatu yang bersifat berulang, proses pengambilan keputusan, masukan, keluaran dan lain sebagainya Flowchart juga disebut penggambaran secara grafik dari langkah-langkah dan urutan prosedur suatu program,. Biasanya mempengaruhi penyelesaian masalah yang khususnya perlu dipelajari dan dievaluasi lebih

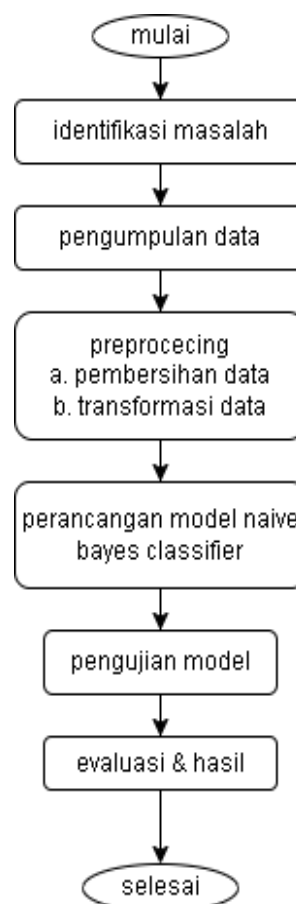
lanjut. Flowchart dapat digunakan untuk menyajikan kegiatan manual, kegiatan pemrosesan ataupun keduanya [9].

2.10 Use Case Diagram

Use case merupakan deskripsi fungsi dari sebuah sistem dari perspektif atau sudut pandang para pengguna sistem. Use case mendefinisikan apa yang akan diproses oleh sistem dan komponen – komponennya. Use case bekerja dengan menggunakan scenario yang merupakan deskripsi dari urutan atau langkah –langkah yang menjelaskan apa yang dilakukan oleh user terhadap sistem maupun sebaliknya. Use case mengidentifikasi fungsionalitas yang dipunya sistem, interaksi user dengan sistem dan keterhubungan antara user dengan fungsionalitas sistem[10].

III. PERANCANGAN

3.1 Rancangan Penelitian



Gambar 3.1 Rancangan Penelitian

1. Identifikasi Masalah

Pada tahap ini peneliti mengidentifikasi hal-hal yang harus diperhatikan dalam pemilihan atribut yang signifikan dalam dataset, pengumpulan dataset, analisis outlier dari dataset dan analisis kebutuhan untuk implementasi algoritma. Peneliti melakukan identifikasi masalah di universitas wiraraja tepatnya di fakultas teknik prodi teknik sipil.

2. Pengumpulan data

Teknik yang digunakan dalam mengumpulkan data adalah dataset statistik dimana biasanya data yang digunakan sudah dikumpulkan pihak ke-3 dalam hal ini instansi yang terkait. Peneliti menggunakan data mahasiswa FT prodi teknik sipil lulusan angkatan 2017-2022.

3. Preprocecing data

Tahap preprocecing, setelah peneliti mengumpulkan data berupa dataset yang didapatkan dari BAAK selanjutnya peneliti melakukan preprocecing seperti berikut:

a. Pembersihan data

Pada tahap ini peneliti membersihkan data. artinya, data mentah yang didapatkan dari pihak ke-tiga (BAAK) akan diseleksi kembali, seperti menghapus atau menghilangkan data yang tidak lengkap, tidak relevan, tidak akurat. Selanjutnya peneliti memilih atribut atau variable (X) yang menjadi faktor dalam kelulusan mahasiswa atau variable (Y).

b. Transformasi data

Proses *preprocecing* berikutnya yang harus dilakukan adalah transformasi data. Seperti yang telah dijelaskan di atas, data akan diambil dari berbagai sumber yang kemungkinan memiliki perbedaan format. Peneliti harus menyamakan seluruh data yang terkumpul supaya dapat mempermudah proses analisis data. Pada proses transformasi data ini peneliti melakukan transformasi data untuk value dari variabel yang tadinya numerik dirubah menjadi text, serta mengubah atribut tersebut sesuai kebutuhan penelitian.

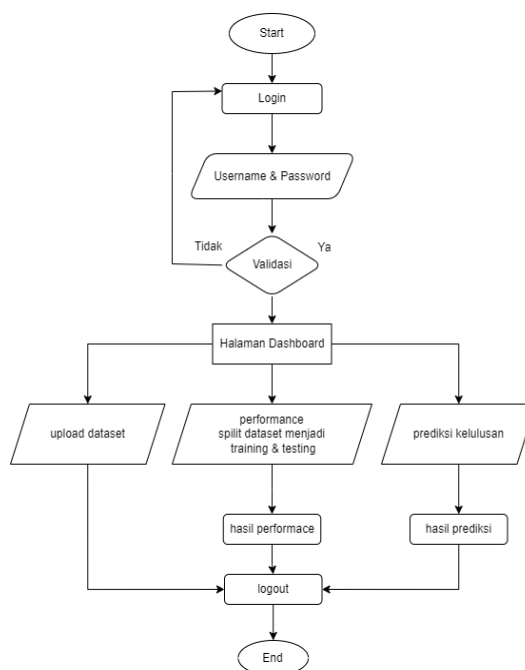
4. Implementasi Naïve Bayes

Pada tahap ini peneliti melakukan pengujian model algoritma naive bayes menggunakan tools Rapidminer dengan menggunakan dataset yang sudah dikumpulkan, dan nanti hasilnya dijadikan rujukan oleh peneliti pada penghitungan manual yang dimasukkan ke dalam koding program. Koding program berbasis website dengan menggabungkan PHP, JavaScript, Html, Css, Editor Sublime Text, dll.

3.2 Rancangan Sistem

Rancangan sistem pada aplikasi prediksi kelulusan mahasiswa tepat waktu melibatkan beberapa aspek yang penting untuk memastikan aplikasi dapat berfungsi dengan baik, memberikan prediksi yang akurat, dan mudah digunakan oleh pengguna. Perancangan sistem sangat penting dalam pembuatan aplikasi karena berperan sebagai panduan atau rencana yang terstruktur untuk mengembangkan, mengimplementasikan, dan memelihara aplikasi dengan efektif dan efisien. Berikut Perancangan sistem prediksi kelulusan mahasiswa tepat waktu menggunakan algoritma *Naive bayes* bisa meliputi langkah-langkah berikut:

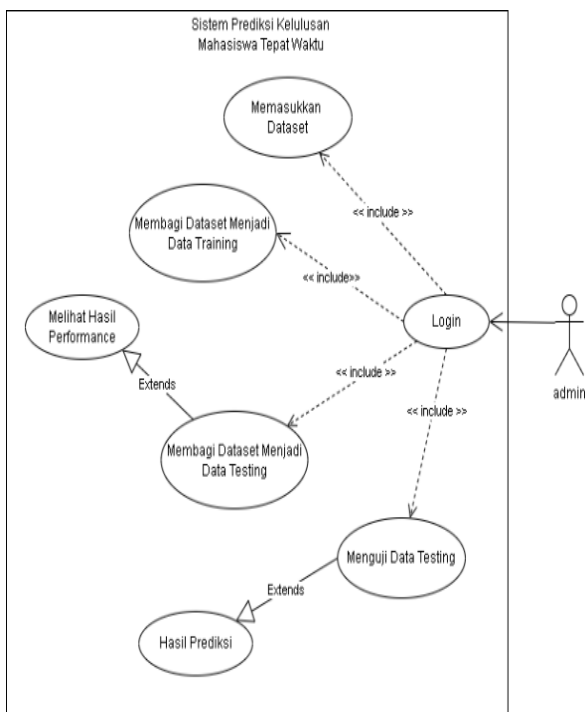
1. Flowchart Rancangan Sistem



Gambar 3.2 Flowchart Rancangan Sistem

Pada Gambar 3.2 merupakan gambar *flowchart* dari rancangan sistem dan aplikasi prediksi kelulusan mahasiswa tepat waktu. Secara umum alur dari *flowchart* ini adalah user akan login terlebih dahulu diawali dengan menu login jika user berhasil login maka akan ditampilkan ke menu dashboard, dan jika tidak berhasil login akan kembali ke menu form login. kemudian jika berhasil login akan ditampilkan menu dashboard dan juga menu-menu dalam prediksi kelulusan mahasiswa tepat waktu seperti input upload dataset, initial proses atribut pendukung dan atribut kelas, membagi dataset menjadi data *training* dan data *testing*, melihat performance atau akurasi pada sistem. Dan menguji data *testing*. Didalam menu-menu tersebut terdapat form-form yang berfungsi sebagaimana nama dari sistem tersebut kemudian langkah selanjutnya jika user atau pengguna sudah melakukan aktivitas dapat logout dan *flowchart* ini selesai. Berikut adalah *flowchart* rancangan sistem prediksi kelulusan mahasiswa tepat waktu / tidak.

2. Use Case Diagram Rancangan Sistem



Gambar 3.3 Use Case Diagram

Pada gambar 3.3 yang merupakan gambar use case diagram sistem administrator. Pada gambar ini menjelaskan bahwa ada 1 aktor pada sistem prediksi kelulusan mahasiswa tepat waktu, yaitu sistem administrator yang berkaitan dengan form yang ada di sistem aplikasi di prediksi kelulusan mahasiswa tepat waktu. Sistem administrator memiliki hak akses ke semua form yang akan dirancang dengan catatan Sistem administrator harus login terlebih dahulu, yang terlihat pada gambar 3.2 ada keterangan include antara form dengan menu atau use case login.

3. Rancangan antarmuka (*Interface*)

Antarmuka pada aplikasi prediksi kelulusan mahasiswa tepat waktu merupakan bagian yang sangat penting karena memungkinkan pengguna untuk berinteraksi dengan aplikasi, memasukkan data, melihat hasil prediksi, dan mengambil keputusan berdasarkan informasi yang disajikan. Berikut adalah penjelasan mengenai antarmuka pada aplikasi prediksi kelulusan mahasiswa tepat waktu:

a. Halaman Login

Gambar 3.4 Halaman Login

Halaman login pada aplikasi prediksi kelulusan mahasiswa tepat waktu seperti gambar 3.4 memiliki beberapa element

inputan diantaranya: element input type username, element input type password dan element button. Berikut tampilan pada sistem

b. Halaman Prediksi

Halaman Prediksi pada aplikasi prediksi kelulusan mahasiswa tepat waktu seperti gambar 3.5 memiliki 4 bagian yaitu navbar, sidebar, content, dan footer, pada navbar terdapat 3 menu seperti Dashboard, Naive Bayes, dan Logout, bagian sidebar memiliki sub menu seperti Dataset, Initial Process, Performance, dan Prediksi. Pada content berisi halaman form inputan mahasiswa untuk melakukan prediksi.



Gambar 3.5 Halaman Prediksi

c. Halaman Hasil Prediksi

Pada gambar 3.6 memiliki 4 bagian yaitu navbar, sidebar, content, dan footer, pada navbar terdapat 3 menu seperti Dashboard, Naive Bayes, dan Logout, sidebar memiliki sub menu seperti Dataset, Initial Process, Performance.



Gambar 3.6 Halaman Hasil Prediksi

IV. IMPLEMENTASI

4.1. Pembersihan Data

Tabel 4. 1 Data Asli Kelulusan

Npm	Nama	Sekolah	Jenis Kelamin		
713520001	Mahasiswa 1	Negeri	Laki-Laki		
Ips-1	Ips-2	Ips-3	Ips-4	Tahun Masuk	Tahun Keluar
2.68	3.64	3.29	3.07	2013	2017

Sumber table : data penelitian

Tahap ini peneliti melakukan pembersihan data dan atribut atau variabel yang tidak berpengaruh dalam perhitungan sehingga data yang dihasilkan sebagai berikut: Tabel 4.2. Tabel sample data penelitian setelah dilakukan pembersihan data dan atribut.

Tabel 4. 2 Hasil Pembersihan Data

Sekolah	JK	Ips-1	Ips-2	Ips-3	Ips-4	TH-K	TH-M
Negeri	L	2.68	3.64	3.29	3.07	2013	2017

Sumber table : data penelitian

4.2. Transformasi Data

Tahap ini peneliti melakukan transformasi data atribut dan value dirubah menjadi text yang sesuai dengan kebutuhan peneliti. Transformasi data ini dilakukan agar mempermudah peneliti untuk memproses data ke dalam aplikasi prediksi kelulusan mahasiswa tepat waktu. Hasil dari proses transformasi data, sebagai berikut:

Tabel 4. 3 Hasil Transformasi Data

Sekolah	Jenis Sekolah	JK	Ips-1 >= 2.74
Negeri	SMA	L	Tidak
Ips-2 >= 2.71	Ips-3 >= 2.92	Ips-4 >= 2.76	Lulus
Ya	Ya	Ya	Tepat

Sumber table : data penelitian

4.3. Analisis Pemodelan Naive Bayes

Analisa menggunakan algoritma Naive Bayes dimaksudkan untuk memperoleh model naive bayes yang terbaik dalam mengklasifikasikan mahasiswa yang lulus tepat waktu atau tidak. Data yang digunakan dalam perhitungan ini sebanyak 358 data. Data training (80%) = 80% * 358 = 286 data, Data testing (20%) = 20% * 358 = 72 data. Data training akan dihitung manual dengan sampel perhitungan yang diambil secara acak sejumlah 286 mahasiswa.

1. Data Uji

Untuk data uji yang digunakan, penulis menggunakan 1 sampel data pada Tabel 4.1.

2. Menghitung Probabilitas Kelas

$P(C1) = \text{jumlah tepat/jumlah semua data} = 262/286 * 100\% = 91\%$.

$P(C2) = \text{jumlah terlambat/jumlah semua data} = 24/286 * 100\% = 8\%$.

3. Menghitung Probabilitas Kejadian Atribut X Kelas Tepat Waktu:

Untuk X1 Adalah Jenis Kelamin, probabilitasnya adalah:

$p(x = "l" | y = \text{"tepat waktu"}) = 197/262 = 0,751$

$p(x = "p" | y = \text{"tepat waktu"}) = 65/262 = 0,2$

untuk x2 adalah sekolah negeri/swasta, probabilitasnya adalah:

$p(x = \text{"negeri"} | y = \text{"tepat waktu"}) = 223/262 = 0,85$

$p(x = \text{"swasta"} | y = \text{"tepat waktu"}) = 39/262 = 0,14$

untuk x3 adalah jenis sekolah, probabilitasnya adalah:

$p(x = \text{"sma"} | y = \text{"tepat waktu"}) = 79/262 = 0,30$

$p(x = \text{"smk"} | y = \text{"tepat waktu"}) = 107/262 = 0,40$

$p(x = \text{"ma"} | y = \text{"tepat waktu"}) = 76/262 = 0,29$

untuk x4 adalah index prestasi semester 1, probabilitasnya adalah:

$p(x = \text{">= 2.74"} | y = \text{"tepat waktu"}) = 243/262 = 0.92$

$p(x = \text{"< 2.74"} | y = \text{"tepat waktu"}) = 19/262 = 0.07$

untuk x5 adalah index prestasi semester 2, probabilitasnya adalah:

$p(x = \text{">= 2.71"} | y = \text{"tepat waktu"}) = 245/262 = 0.93$

$p(x = \text{"< 2.71"} | y = \text{"tepat waktu"}) = 17/262 = 0.06$

untuk x6 adalah index prestasi semester 3, probabilitasnya adalah:

$p(x = \text{">= 2.92"} | y = \text{"tepat waktu"}) = 246/262 = 0.93$

$p(x = \text{"< 2.92"} | y = \text{"tepat waktu"}) = 16/262 = 0.06$

untuk x7 adalah index prestasi semester 4, probabilitasnya adalah:

$p(x = \text{">= 2.76"} | y = \text{"tepat waktu"}) = 246/262 = 0.93$

$p(x = \text{"< 2.76"} | y = \text{"tepat waktu"}) = 16/262 = 0.06$

4. probabilitas kejadian atribut x untuk kelas tidak tepat waktu

$p(x = "l" | n = \text{"tidak tepat waktu"}) = 19/24 = 0,79$

$p(x = "p" | n = \text{"tidak tepat waktu"}) = 5/24 = 0,20$

untuk x2 adalah sekolah negeri/swasta, probabilitasnya adalah:

$p(x = \text{"negeri"} | n = \text{"tidak tepat waktu"}) = 19/24 = 0,79$

$p(x = \text{"swasta"} | n = \text{"tidak tepat waktu"}) = 5/24 = 0,20$

untuk x3 adalah jenis sekolah, probabilitasnya adalah:

$p(x = \text{"sma"} | n = \text{"tidak tepat waktu"}) = 6/24 = 0,25$

$p(x = \text{"smk"} | n = \text{"tidak tepat waktu"}) = 7/24 = 0,29$

$p(x = \text{"ma"} | n = \text{"tidak tepat waktu"}) = 11/24 = 0,45$

untuk x4 adalah ip semester 1, probabilitasnya adalah:

$p(x = \text{">= 2.74"} | n = \text{"tidak tepat waktu"}) = 10/24 = 0,41$

$p(x = \text{"< 2.74"} | n = \text{"tidak tepat waktu"}) = 11/24 = 0,58$

untuk x5 adalah ip semester 2, probabilitasnya adalah:

$$p(x = ">= 2.71" | n = \text{"tidak tepat waktu"}) = 10/24 = 0,41$$

$$p(x = "< 2.71" | n = \text{"tidak tepat waktu"}) = 14/24 = 0,58$$

untuk x6 adalah ip semester 3, probabilitasnya adalah:

$$p(x = ">= 2.92" | n = \text{"tidak tepat waktu"}) = 8/24 = 0,33$$

$$p(x = "< 2.92" | n = \text{"tidak tepat waktu"}) = 16/24 = 0,66$$

untuk x7 adalah ip semester 4, probabilitasnya adalah:

$$p(x = ">= 2.76" | n = \text{"tidak tepat waktu"}) = 8/24 = 0,33$$

$$p(x = "< 2.76" | n = \text{"tidak tepat waktu"}) = 16/24 = 0,66$$

5. menghitung probabilitas kelas "tepat waktu"

$$p(y = \text{"tepat waktu"}) = (0,75 * 0,85 * 0,29 * 0,92 * 0,93 * 0,93 * 0,93 * 0,96) = 0.13$$

6. menghitung probabilitas kelas "terlambat"

$$p(y = \text{"tepat waktu"}) = (0,79 * 0,79 * 0,45 * 0,41 * 0,41 * 0,33 * 0,33 * 0,08) = 0.00041$$

7. Membandingkan Probabilitas Kelas

Dengan demikian, didapatkan kesimpulan bahwa dari data pelatihan diatas, mahasiswa tersebut termasuk ke dalam kategori kelas "Tepat Waktu" dengan rinciang probabilitas seperti di atas.

V KESIMPULAN

Penelitian prediksi tingkat kelulusan mahasiswa menggunakan 358 data, dengan rincian data *training* 286 data dan data *testing* 72 data. Atribut yang digunakan antara lain : Sekolah, Jenis Sekolah, Jenis Kelamin, dan Ip Semester 1 – Ip Semester 4. Tahapan yang dilakukan dalam penelitian antara lain: Identifikasi Masalah, Pengumpulan Data, Preprocecing, Klasifikasi dengan menggunakan *Naïve Bayes*, Validasi, dan Hasil. Hasil akurasi yang diperoleh yaitu 91,16%.

REFERENSI

- [1] Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode *Naïve bayes*[Journal] / auth. Apridiansyah yovi , Veronika Nuri David Maria and Putra Erwin Dwika // Journal Scientific and Applied Informatics. - 2021. - pp. 1-12.
- [2] Data Mining : Klasifikasi Menggunakan Algoritma C4.5 [Journal] / auth. Mardi Yuli // Jurnal Edik Informatika. - 2019. - pp. 2-7.
- [3] Prediksi Kelulusan Mahasiswa Dengan Metode *Naïve bayes*[Journal] / auth. Khasanah Nurul [et al.] // Technologia. - 2022. - pp. 2-8.
- [4] Model Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode *Naïve Bayes* [Journal] / auth. Armansyah. - Medan : Medan, 2021. - Vol. 78.
- [5] Implementasi Metode *Naïve bayes*Classifier (NBC) Pada Komentar Warga Sekolah Mengenai Pelaksanaan Pembelajaran Jarak Jauh (PJJ) [Journal] / auth. Siregar dkk // Jurnal Technologia. - 2020. - pp. 3-9.
- [6] Membangun Website Toko Online Pempek Nthree Menggunakan Php Dan Mysql [Journal] / auth. Susilawati Tuti [et al.] // Jurnal Teknik Informatika Mahakarya. - 2020. - pp. 2-10.
- [7] Mobile learning application berbasis sublime text pada mata pelajaran kearsipan [Journal] / auth. Suputra I Nyoman and Itsnania Binti . - Malang : Universitas Negeri Malang,, 2021.
- [8] Pemodelan Diagram Uml Sistem Pembayaran pada Transaksie-Commerce [Journal] / auth. Syarif Muhammad and Nugraha Wahyu. - Pontianak : Universitas Bina Sarana, 2020. - 7 : Vol. IV. - pp. 1-7.
- [9] Analisis Pengendalian Mutu Di Bidang Industri Makanan(Studi Kasus: UMKM Mochi Kaswari Lampion Kota Sukabumi) [Journal] / auth. Ilham Budiman Sopyan Saori. - Suka Bumi : Universitas Muhammadiyah Sukabumi, 2021. - Vol. 1.
- [10] Desain Sistem : Use Case Diagram [Journal] / auth. Setiyani Lila. - Karawang : STMIK RosmaKarawang, Indonesia, 2021.

Halaman ini sengaja dikosongkan