

Penerapan Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Rekayasa Perangkat Lunak Politeknik Negeri Bengkalis

¹May Sinta Samosir, ²Lidya Wati
^{1,2}Politeknik Negeri Bengkalis
Bengkalis, Riau, Indonesia

¹maysintasamosir@polbeng.ac.id, ²lidyawati@polbeng.ac.id

*Penulis Korespondensi

Diajukan : 01/08/2024

Diterima : 05/08/2024

Dipublikasi : 05/08/2024

ABSTRAK

Penelitian ini bertujuan untuk menerapkan algoritma Naive Bayes sebagai metode prediksi kelulusan mahasiswa Rekayasa Perangkat Lunak Politeknik Negeri Bengkalis. Masalah utama di Politeknik Negeri Bengkalis adalah sulitnya memprediksi kelulusan mahasiswa secara akurat, yang berdampak pada perencanaan akademik dan pembinaan mahasiswa. Populasi penelitian ini adalah mahasiswa Rekayasa Perangkat Lunak yang lulus antara tahun 2020 hingga 2023. Teknik pengambilan sampel yang digunakan adalah sampling dengan 163 data alumni sebagai sampel. Naive Bayes digunakan untuk mengelompokkan data mahasiswa berdasarkan atribut tertentu guna memprediksi kemungkinan kelulusan. Dataset meliputi IPS, IPK, SKS, organisasi, kompetisi skill programming, keaktifan di kelas, dan tahun kelulusan. Metode pengembangan yang digunakan adalah Rapid Application Development (RAD) yang memungkinkan pengembangan aplikasi secara cepat dan iteratif. Hasil prediksi kelulusan dari algoritma Naive Bayes menunjukkan precision sebesar 92,31%, recall 80%, dan akurasi 76,47% untuk perbandingan data latih dan uji (90% dan 10%). Hasil pengelompokan mahasiswa terbagi menjadi dua kategori: berpotensi lulus tepat waktu dan lulus terlambat. Dengan menerapkan algoritma Naive Bayes, diharapkan Politeknik Negeri Bengkalis dapat memberikan dukungan tambahan kepada mahasiswa yang memerlukan dan mengoptimalkan tingkat kelulusan. Penelitian ini diharapkan dapat memberikan rekomendasi yang berguna bagi pengambil keputusan di Polbeng untuk meningkatkan kualitas lulusan sehingga dapat bersaing secara nasional maupun internasional. Dengan demikian, penelitian ini dapat memberikan kontribusi pada pengembangan sistem prediksi kelulusan yang lebih akurat dan efektif.

Kata Kunci: Kelulusan, Klasifikasi, Naive Bayes, Web

I. PENDAHULUAN

Pendidikan tinggi yang semakin berkembang dan maju di Indonesia, khususnya di Politeknik Negeri Bengkalis, menuntut mahasiswa untuk memiliki kemampuan yang mumpuni dalam menghadapi tantangan dunia kerja. Selain itu setiap kampus juga pasti menginginkan mahasiswanya untuk lulus tepat waktu, karena hal ini bermanfaat bagi pihak kampus untuk meningkatkan reputasi dan akreditasinya (Apridiansyah et al., 2021). Kelulusan merupakan langkah awal yang menentukan arah karir mereka di masa depan atau mempersiapkan mereka untuk jenjang pendidikan yang lebih tinggi (Sukarna et al., 2022). Kemampuan seorang mahasiswa menyelesaikan studi sesuai dengan jangka waktu yang ditetapkan juga menjadi tolak ukur keberhasilan akademik (Apridiansyah et al., 2021).

Permasalahan yang terjadi pada sebuah lembaga pendidikan terutama pada Perguruan Tinggi baik itu negeri maupun swasta ialah mencari prediksi tingkat kelulusan mahasiswa tepat

waktu(David Imanuel et al., 2024), karena dengan lulus tepat waktu merupakan salah satu penilaian dalam proses akreditasi Perguruan Tinggi, sehingga semakin banyak mahasiswa yang tepat waktu kelulusannya maka semakin baik penilaian yang didapat pada saat akreditasi(Apridiansyah et al., 2021). Penelitian yang dilakukan oleh Chandra Babu dan Shantharajah (2019) menggunakan teknik klasifikasi algoritma seperti Decision Tree, Naive Bayes, dan Support Vector Machine (SVM) pada KPI (Key Performance Indicator). Hasil penelitian menunjukkan bahwa penggunaan algoritma ini dapat memprediksi performa siswa dan meningkatkan metode pembelajaran mereka(Studi et al., n.d.). Oleh karena itu, diperlukan suatu sistem yang dapat memprediksi kelulusan mahasiswa dengan akurat sehingga dapat dilakukan tindakan preventif terhadap mahasiswa yang berpotensi tidak lulus(Khasanah et al., 2022)(Case et al., 2022).

Data mining adalah teknik yang digunakan untuk membangun model pembelajaran mesin. Pembelajaran mesin (machine learning) adalah teknik kecerdasan buatan modern yang belajar membangun model dengan menggunakan data empiris(Liliana et al., 2021). Tujuan data mining adalah menemukan informasi tersembunyi dalam suatu dataset yang tidak dapat terlihat melalui analisis statistik manual(Amelia et al., 2022). Kelebihan data mining dapat mempercepat proses pengklasifikasian data dengan tingkat akurasi yang tinggi, sehingga sangat efektif(Attamami et al., 2023).

Metode Naïve Bayes merupakan metoda machine learning, yang berada dalam kelompok klasifikasi dalam Supervised Learning (Saputra et al., 2023)(Sistem et al., 2021). Machine learning adalah bidang kecerdasan buatan atau artificial intelligence yang berkaitan dengan pengembangan teknik-teknik yang dapat diprogram dan dipelajari dari data masa lalu(Ericha Apriliyani & Salim, 2022). Metode ini merupakan salah satu metode dalam data mining yang melakukan perhitungan probabilitas (Susana & Suarna, 2022)(Poerwandono & Perwitosari, 2023), biasanya digunakan untuk mengelompokkan data-data yang memiliki karakteristik yang mirip(Setiyani et al., 2020). Naive Bayes Gaussian adalah metode klasifikasi yang memanfaatkan distribusi Gaussian untuk merepresentasikan variabel-variabel yang digunakan dalam proses klasifikasi(Rosyid Mubarak, 2024). Teorema Bayes, yang juga dikenal sebagai aturan Bayes, adalah alat yang berguna untuk menghitung probabilitas bersyarat. Peluang bersyarat dari A ketika B dilambangkan dengan $P(A | B)$. Distribusi gaussian adalah salah satu metode paling umum dan penting dalam menghitung probabilitas dan statistik(Ericha Apriliyani & Salim, 2022). Kelebihan Naive Bayes adalah mudah diimplementasikan dan sering memberikan hasil yang baik dalam banyak kasus. Namun, kekurangannya adalah asumsi independensi antar fitur(Chen et al., 2021), padahal dalam kenyataannya, keterkaitan antar fitur sering kali ada dan tidak dapat dimodelkan oleh Naive Bayesian Classifier(Teknologi et al., 2021). Naive Bayes Classifier adalah salah satu pengklasifikasi statistik yang dapat memprediksi probabilitas keanggotaan suatu data dalam kelas tertentu berdasarkan perhitungan probabilitas(Simanjuntak et al., 2022).

Untuk memprediksi kelulusan mahasiswa menggunakan metoda Naïve Bayes juga dilakukan oleh [6](Riyadsyah et al., 2024) dengan hasil akurasi 90%, penelitian lain dilakukan oleh [4] dengan atribut yang digunakan adalah nama, status mahasiswa, status perkawinan, IPS, IPK, dan status kelulusan. Hasil yang didapatkan yaitu akurasi = 88,16%, precision = 93,62% dan recall = 88%, termasuk dalam kategori good classification. Penelitian lain dilakukan oleh [3] hasil yang didapatkan yaitu nilai precision sebesar 90%, recall 100% dan akurasi sebesar 90%. Penelitian lain juga dilakukan oleh [2] hasil yang didapatkan yaitu nilai precision sebesar 83,56%, recall 88,41% dan akurasi sebesar 81% . Selain untuk memprediksi kelulusan metoda Naïve Bayes juga digunakan untuk memprediksikan penggunaan internet(Susana & Suarna, 2022), prediksi hargaonseluntuk mentukan kelayakan kredit rumah bersubsidi(Kelayakan & Rumah, 2020).

Penelitian ini menggunakan metode Naïve Bayes(Afriansyah et al., 2024), untuk memprediksi kelulusan mahasiswa Rekayasa Perangkat Lunak Politeknik Negeri Bengkalis. Pada algoritma ini, pembelajaran lebih difokuskan pada estimasi probabilitas(Mardiani et al., 2023)(Romano et al., 2024). Parameter yang digunakan dalam dataset mengenai data kelulusan mahasiswa(Eka et al., 2021), seperti data mahasiswa dari nama mahasiswa, ips semester 1-8, ipk, sks, tahun kelulusan mahasiswa, organisasi, kompetisi, skill programming, keaktifan di kelas. Sampel data diambil dari alumni mahasiswa rekayasa perangkat lunak sebagai penentu dalam hal prediksi tingkat

kelulusan mahasiswa tepat waktu, normalnya mahasiswa yang dapat lulus tepat waktu adalah mahasiswa yang masa pendidikannya mampu ditempu selama empat (4) tahun masa studi, jika lewat dari itu maka mahasiswa tersebut digolongkan mahasiswa tidak tepat waktu.

Hasil pengelompokan mahasiswa terbagi menjadi dua kategori yaitu kelompok yang berpotensi lulus tepat waktu dan kelompok yang akan lulus terlambat. Dengan menerapkan Algoritma Naive Bayes dalam prediksi kelulusan mahasiswa, diharapkan Politeknik Negeri Bengkalis (Polbeng) dapat mengambil langkah-langkah yang lebih tepat dalam memberikan dukungan tambahan kepada mahasiswa yang memerlukan, serta mengoptimalkan tingkat kelulusan. Dengan demikian, penelitian ini diharapkan dapat memberikan rekomendasi yang berguna bagi pengambil keputusan di Polbeng untuk meningkatkan kualitas lulusan Rekayasa Perangkat Lunak sehingga dapat bersaing dengan lulusan dari perguruan-perguruan tinggi lainnya secara nasional maupun internasional.

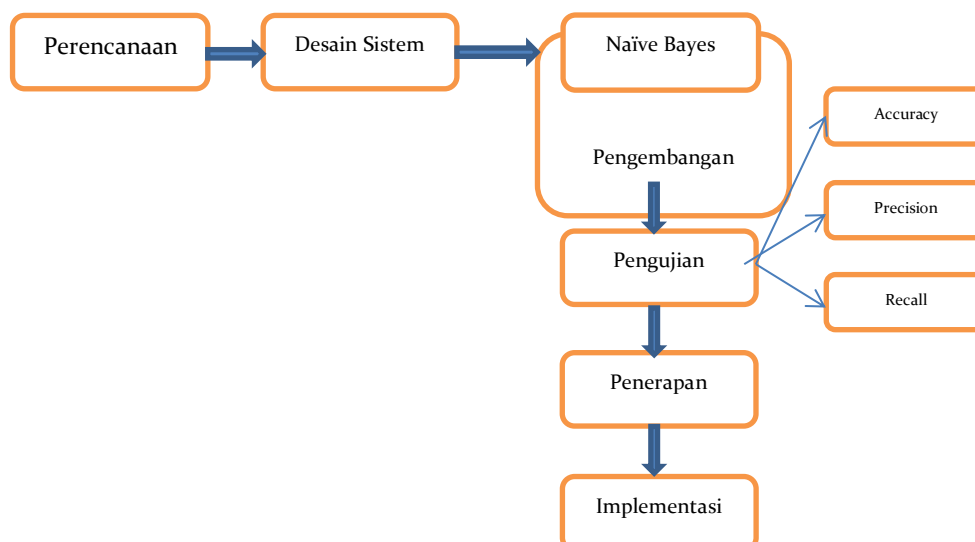
II. STUDI LITERATUR

Penelitian terdahulu untuk memprediksi kelulusan mahasiswa yang dilakukan oleh Apridiansyah dkk menggunakan metoda naïve bayes tahun 2021. Hasil dari penelitian ini adalah sistem dapat memprediksi kelulusan mahasiswa dengan tepat waktu berdasarkan nilai IPK mahasiswa dan lama Studi mahasiswa, menggunakan algoritma naïve bayes menghasilkan nilai precision sebesar 90 %, recall 100 % dan akurasi sebesar 90% (Apridiansyah et al., 2021).

Penelitian lain untuk tentang kelulusan mahasiswa juga dilakukan oleh David Imanuel dkk tahun 2024. Setelah menerapkan algoritma naïve bayes hasil dari penelitian ini adalah Accuracy 93,10%, Precision 95,24%, dan Recall 90%, dengan atribut yang digunakan yaitu Status pekerjaan, IPK, Total SKS lulus, Total mata kuliah lulus, dan juga Status pembayaran (David Imanuel et al., 2024). Penelitian serupa juga dilakukan oleh Setiani dkk tahun 2020. Hasil yang didapat adalah accuracy sebesar 90% (Setiyani et al., 2020).

III. METODE

Metode penelitian ini menggunakan Rapid Application Development (RAD) untuk mengembangkan sistem prediksi kelulusan mahasiswa menggunakan metode Naïve Bayes (Minahasa et al., n.d.). Pendekatan RAD memungkinkan pengembangan sistem yang lebih cepat melalui prototyping dan iterasi cepat (Pratama & Zufria, 2022).



Gambar 1. Tahapan Penelitian

Berikut adalah tahapan-tahapan dalam metodologi RAD yang digunakan dalam penelitian ini,

1. Perencanaan Kebutuhan pada tahap ini, dilakukan identifikasi dan pengumpulan kebutuhan pengguna. Kebutuhan ini mencakup fitur-fitur yang diinginkan dalam sistem prediksi kelulusan mahasiswa. Pengumpulan data dilakukan melalui wawancara dengan dosen dan mahasiswa serta analisis dokumen terkait. Data-data yang diperlukan untuk

yaitu nama, ips semester 1 sampai semester 7, ipk, sks, tahun lulus, organisasi, kompetisi, skill programing data tersebut di ambil dari alumni rpl tahun 2020 sampai tahun 2023 yang telah lulus.

2. Desain Sistem, Setelah kebutuhan dikumpulkan, tahap selanjutnya adalah merancang sistem. Desain sistem meliputi pembuatan model data, desain antarmuka pengguna, dan arsitektur sistem. Dalam tahap ini, juga dibuat prototipe awal sistem untuk mendapatkan umpan balik dari pengguna.
3. Pengembangan Pada tahap ini, pengembangan sistem dilakukan berdasarkan desain yang telah disetujui. Pengembangan mencakup implementasi metode Naïve Bayes untuk prediksi kelulusan mahasiswa, pengolahan data mahasiswa, dan integrasi dengan database. Sistem ini menggunakan bahasa pemrograman PHP dan database MySQL dengan XAMPP sebagai server lokal.

Persamaan dari teorema Bayes sebagai berikut

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

X : Data dengan class yang belum diketahui

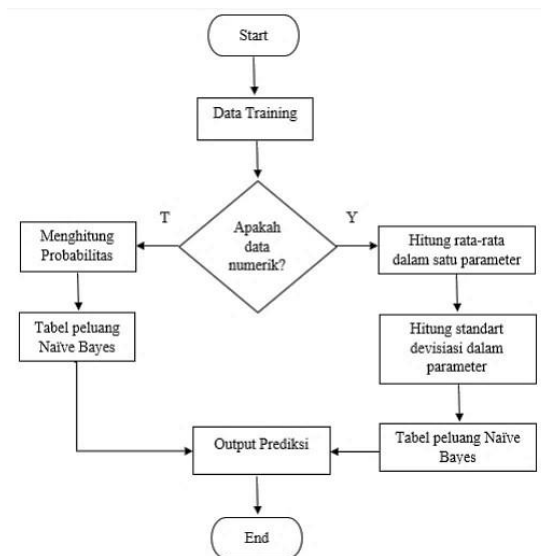
H : Hipotesis data X merupakan suatu class spesifik

P(H|X): Probabilitas hipotesis H berdasarkan Kondisi X (posteriori probability)

P(H) : Probabilitas hipotesis H (prior probability)

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

P(X) : Probabilitas X



Gambar 2. Tahapan Algoritma Naive Bayes

4. Pengujian, Setelah sistem dikembangkan, dilakukan pengujian untuk memastikan bahwa sistem berfungsi sesuai dengan kebutuhan. Penulis memilih metode pengujian sistem menggunakan Blackbox untuk menguji sistem secara keseluruhan. Blackbox testing digunakan untuk menguji fungsionalitas atau nonfungsionalitas sistem tanpa mengetahui struktur internal modul atau program yang akan diuji apakah sudah berjalan dengan baik apa belum. Selanjutnya pengujian dengan menggunakan precision and recall untuk mencari tingkat keakurasian dari metode Naive Bayes.

Berikut rumus dari Precision, Recall, dan akurasi:

$$\text{Precision} = \frac{\text{Jumlah prediksi benar (positif)}}{\text{Prediksi benar (positif)} + \text{Prediksi salah (positif)}} \times 100\%$$

$$\text{Recall} = \frac{\text{Jumlah prediksi benar (positif)}}{\text{Prediksi benar (positif)} + \text{Prediksi salah (negatif)}} \times 100\%$$

$$\text{Akurasi} = \frac{\text{Jumlah prediksi benar (positif dan negatif)}}{\text{Jumlah keseluruhan data}} \times 100\%$$

5. Penerapan, Setelah sistem diuji dan dinyatakan siap, tahap selanjutnya adalah penerapan sistem di lingkungan nyata. Penerapan mencakup instalasi sistem, pelatihan pengguna, dan pengaturan lingkungan operasional. Pada tahap ini, juga dilakukan pemantauan untuk memastikan sistem berjalan dengan baik dan memberikan hasil prediksi yang akurat.
6. Evaluasi dan Pemeliharaan, Tahap terakhir adalah evaluasi dan pemeliharaan sistem. Evaluasi dilakukan untuk menilai kinerja sistem dan mengidentifikasi area yang memerlukan perbaikan. Pemeliharaan mencakup perbaikan bug, peningkatan fitur, dan penyesuaian dengan kebutuhan baru. Umpan balik dari pengguna digunakan untuk terus mengoptimalkan sistem.

IV. HASIL DAN PEMBAHASAN

Hasil penelitian ini berupa rancang bangun sistem yang bisa memprediksi kelulusan mahasiswa Rekayasa Perangkat Lunak (RPL) Politeknik Negeri Bengkalis. Sebelum menerapkan kedalam aplikasi dilakukan perhitungan proses prediksi menggunakan metoda naïve bayes, dengan menggunakan 162 data lulusan, yang terdiri dari 132 yang lulus tepat waktu dan 30 yang lulus terlambat. Dalam perhitungan naïve bayes tahap pertama adalah melakukan perhitungan probabilitas untuk jumlah lulusan yang tepat waktu dan terlambat,

$$P(\text{jumlah lulus}) = \frac{P(\text{jumlah lulus})}{P(\text{jumlah data})} = \frac{132}{162} = 0,8148$$

$$P(\text{jumlah lulus terlambat}) = \frac{P(\text{jumlah lulus terlambat})}{P(\text{jumlah data})} = \frac{30}{162} = 0,1852$$

selanjutnya menghitung nilai probabilitas untuk atribut organisasi, atribut kompetensi, atribut keaktifan dan skill programaming,

Tabel 1 Nilai Probabilitas atribut organisasi

| Organisasi Label | Lulus | lulus terlambat |
|---|--------------|-----------------|
| Bem Ukm Olahraga Ukmi Hmti Ippermasik | 0,034 | 0,077 |
| Bem, Hmti | 0,008 | 0,000 |
| Bem, Uimi, Hmti | 0,008 | 0,000 |
| Bem, Ukm Kesenian, Hmti, Ukm Pramuka | 0,008 | 0,000 |
| Bem, Ukm Olahraga, Ukmi Hmti Ippermasik | 0,008 | 0,000 |
| Forum Mahasiswa Bidikmisi | 0,017 | 0,038 |
| Hmti | 0,202 | 0,038 |
| : | : | : |
| Umkm, Olahraga, Hmti | 0,008 | 0,000 |
| Total | 1,000 | 1,000 |

Tabel 2 Nilai Probabilitas Atribut Kompetensi

| Kompetisi Label | Lulus | lulus terlambat |
|--|-------|-----------------|
| Algoritma Pemrograman | 0,034 | 0,000 |
| Atraksi Kimia | 0,008 | 0,000 |
| Bola, Ukm Olahraga Cup Futsal | 0,000 | 0,000 |
| Csf Unri | 0,008 | 0,000 |
| Desain Poster Formadiksi, Video Komtik Hmti | 0,000 | 0,038 |
| Dicoding | 0,218 | 0,154 |
| Dicoding Android Pemula, Bnsp Kominfo Medan Junior Web Development | 0,008 | 0,000 |
| : | : | : |
| Web junior | 0,008 | 0,000 |
| total | 1 | 1 |

Tabel 3 Nilai Probabilitas Atribut Keaktifan

| Keaktifan di kelas Label | Lulus | lulus terlambat |
|----------------------------|----------|-----------------|
| Aktif | 0,815 | 0,731 |
| Pasif | 0,185 | 0,269 |
| total | 1 | 1 |

Tabel 4 Nilai Probabilitas Atribut Skill Programing

| Skill Dalam Programming Label | Lulus | lulus terlambat |
|---------------------------------|----------|-----------------|
| Iya | 0,924 | 0,885 |
| Tidak | 0,067 | 0,115 |
| total | 1 | 1 |

Selanjutnya melakukan perhitungan mean, standar deviasi dan gaussian,

Tabel 5 Nilai Perhitungan Mean

| MEAN | | | | | | | | | | | |
|-----------------|------|------|------|------|------|------|------|------|------|--------|---------|
| Label Kelas | IPS1 | IPS2 | IPS3 | IPS4 | IPS5 | IPS6 | IPS7 | IPS8 | IPK | SKS | Tahun |
| LULUS | 3,02 | 3,06 | 3,11 | 3,18 | 3,20 | 3,42 | 3,45 | 3,69 | 3,23 | 135,31 | 2022,02 |
| LULUS TERLAMBAT | 2,80 | 2,91 | 2,96 | 3,11 | 3,13 | 3,26 | 2,85 | 0,91 | 2,86 | 134,92 | 2022,00 |

Tabel 6 Nilai Perhitungan standar deviasi

| STANDAR DEVIASI | | | | | | | | | | | |
|-----------------|------|------|------|------|------|------|------|------|------|------|-------|
| Label Kelas | IPS1 | IPS2 | IPS3 | IPS4 | IPS5 | IPS6 | IPS7 | IPS8 | IPK | SKS | Tahun |
| LULUS | 0,33 | 0,30 | 0,34 | 0,25 | 0,25 | 0,18 | 0,30 | 0,22 | 0,22 | 3,76 | 0,85 |
| LULUS TERLAMBAT | 0,31 | 0,26 | 0,36 | 0,23 | 0,26 | 0,20 | 0,66 | 0,49 | 0,23 | 3,26 | 0,75 |

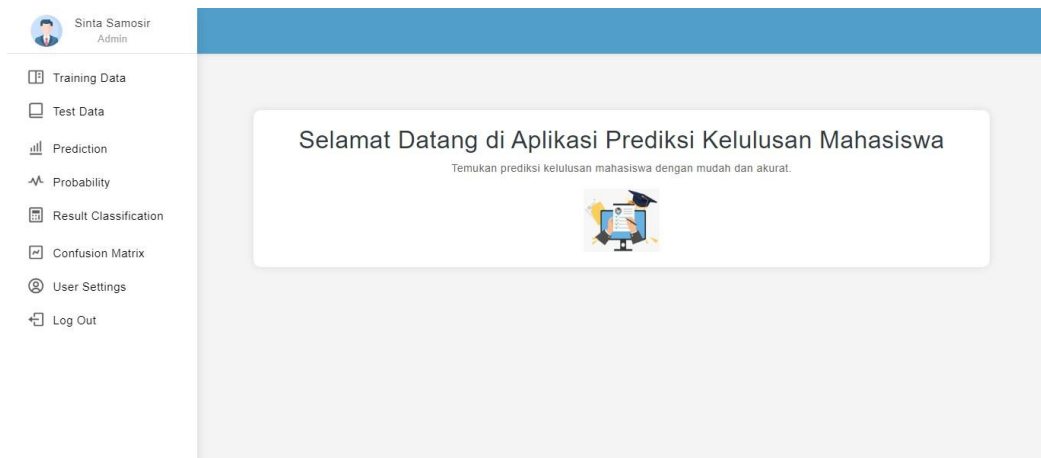
Untuk melakukan perhitungan gaussian digunakan salah satu sampel data mahasiswa yaitu,

| | Ips1 | Ips2 | Ips3 | Ips4 | Ips5 | Ips6 | Ips7 | Ips8 | Ipk | Sks | Tahun Lulus | Organisasi | Kompetensi | Kekurangan di kelas | Skil Dalam Programming | Prediksi |
|-----------------|-------------|-------------|------|-------------|-----------|-------------|----------|----------|----------|----------|-------------|------------|------------|---------------------|------------------------|-------------|
| Data Uji ke-n | 2,76 | 2,89 | 2,95 | 3 | 2,9 | 3,21 | 2,89 | 3,67 | 3,01 | 134 | 2023 | Ukmt, Hmti | Tidak Ada | Pasif | Iya | Lulus |
| LULUS | 0,50659228 | 0,619571528 | 0,61 | 0,623405041 | 0,4008933 | 0,495414347 | 0,126852 | 0,850825 | 0,509809 | 0,193672 | 0,222475 | 0,025 | 0,387 | 0,185 | 0,924 | 7,68561E-08 |
| LULUS TERLAMBAT | 0,712278286 | 0,781698156 | 0,66 | 0,743150987 | 0,5327771 | 0,868211266 | 0,492307 | 9,82E-08 | 0,674361 | 0,212299 | 0,188891 | 0,038 | 0,500 | 0,269 | 0,885 | 1,36595E-13 |

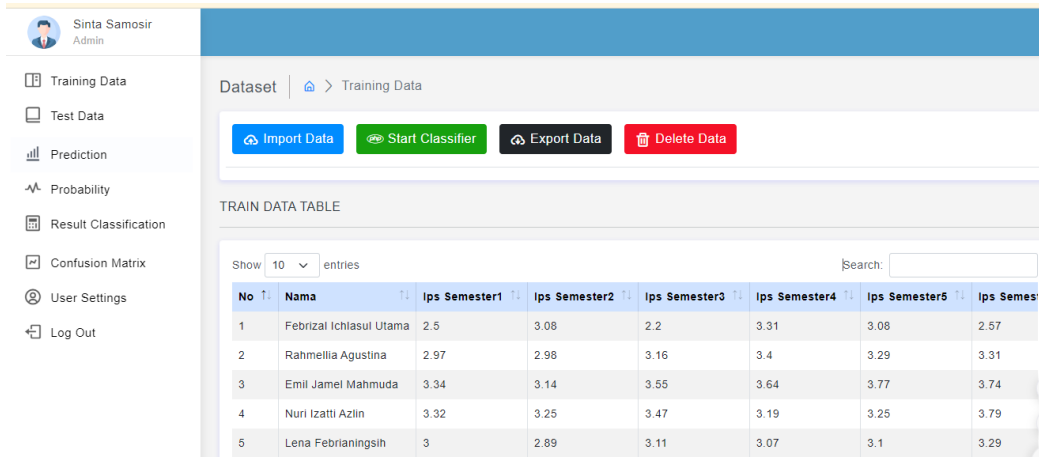
Gambar 3. Nilai Perhitungan Gaussian

Hasil dari nilai probabilitas Lulus 7,68561E-08 dan nilai probabilitas Lulus Terlambat 1,36595E-13. Dari kedua nilai tersebut, nilai probabilitas Lulus lebih besar dari pada nilai probabilitas Lulus Terlambat, maka hasil perhitungan diatas diprediksi bahwa siswa yang memiliki nilai atribut tersebut memiliki hasil prediksi ”Lulus”.

Setelah dilakukan analisis menggunakan metoda naive bayes maka selanjutnya diterapkan di sistem. Pada bagian hasil menampilkan dari tampilan website yang telah diselesaikan sebelumnya, Adapun beberapa tampilan tersebut adalah



Gambar 4. Halaman Beranda



Gambar 5. Halaman Training Data

| Ipsi | Ipsi | Ipsi | Ipsi | Ipsi | Ipk | Sks | Tahun Lulus | Organisasi | Kompetisi | Keaktifan | Skill | Kelas | Prediksi |
|------|------|------|------|------|------|-----|-------------|--|-------------------------------|-----------|-------|-----------------|-----------------|
| 3.48 | 3.25 | 3.48 | 3.79 | 3.94 | 3.33 | 134 | 2021 | Tidak Ada | Php, Javascript | Aktif | Iya | Lulus | Lulus |
| 3.21 | 3.45 | 3.21 | 3.33 | 0.58 | 2.75 | 134 | 2022 | Mpm | Tidak Ada | Aktif | Iya | Lulus Terlambat | Lulus Terlambat |
| 3.38 | 3.38 | 3.31 | 3.39 | 4 | 3.33 | 134 | 2022 | Ukm Olahraga | Bola, Ukm Olahraga Cup Futsal | Aktif | Iya | Lulus | Lulus |
| 3.1 | 3.24 | 3.5 | 3.22 | 3.07 | 3.18 | 134 | 2023 | Hmti Dan Ukmi | Tidak Ada | Pasif | Iya | Lulus | Lulus |
| 2.81 | 2.98 | 3.71 | 3.17 | 0.67 | 2.82 | 134 | 2023 | Ukmi | Junior Web Developer-Dicoding | Aktif | Iya | Lulus Terlambat | Lulus |
| 3.31 | 3.17 | 3.24 | 3.82 | 4 | 3.28 | 134 | 2021 | Mpm, Ukmi | Tidak Ada | Aktif | Iya | Lulus | Lulus |
| 3.88 | 3.98 | 4 | 4 | 4 | 3.94 | 134 | 2023 | Ukm Olahraga, Bem Polbeng, Pkmb Polbeng, Imkd Bengkulu | Digitalent | Aktif | Iya | Lulus | Lulus |
| 3 | 2.9 | 3.21 | 2.89 | 3.67 | 3.01 | 134 | 2023 | Ukmi, Hmti | Tidak Ada | Pasif | Iya | Lulus | Lulus |
| 3.33 | 3.29 | 3.33 | 3.5 | 4 | 3.15 | 134 | 2022 | Hmti, Pramuka | Tidak Ada | Aktif | Iya | Lulus | Lulus |
| 2.88 | 2.85 | 3 | 1.82 | 1 | 2.88 | 134 | 2021 | Hmti, Bem | Tidak Ada | Aktif | Iya | Lulus Terlambat | Lulus Terlambat |

Gambar 6. Hasil Prediksi Naïve Bayes Pada Sistem

Sebelum sistem ini digunakan atau diimplementasikan maka perlu dilakukan pengujian. Pengujian ini bermaksud untuk meminimalkan jumlah kesalahan atau kegagalan sistem pada saat implementasi sistem. Metode pengujian sistem yang digunakan adalah metode blackbox testing. Dalam pengujian sistem ini digunakan perbandingan (90% data latih dan 10% data uji) sehingga didapatkan 146 data latih, 17 data uji mendapatkan akurasi 76,47%, presisi 92,31% dan recall 80%. Sedangkan jika perbandingan data latih dan data uji sebesar (80% dan 20%) maka didapatkan data latih = 130 dan data uji = 32, hasil akurasi 84,38%, presisi = 96.3%, recall = 87%.

Tabel 7. Perbandingan Skenario Uji

| Data Latih(%) | Data Uji(%) | Data Latih | Data Uji | Akurasi | Presisi | Recall | Error Rate |
|---------------|-------------|------------|----------|---------|---------|--------|------------|
| 90% | 10% | 146 | 17 | 76.47% | 92.31% | 80% | 2% |
| 80% | 20% | 130 | 32 | 84,38% | 96,3% | 87% | 1% |

Tabel 8. Pengujian Black Box

| Skenario Pengujian | Test | Hasil Pengujian | Kesimpulan |
|--------------------|-------------------------|-----------------|------------|
| Login | Input | Bisa diakses | valid |
| Training Data | Import Data | Bisa diakses | valid |
| Save Changes | Save Changes | Bisa diakses | valid |
| Keluar | Keluar | Bisa diakses | valid |
| Start Classifier | Perhitungan Naïve Bayes | Bisa diakses | valid |
| Delete Data | Click | Bisa diakses | valid |
| Test Data | Import Data | Bisa diakses | valid |
| Save Changes | Save Changes | Bisa diakses | valid |
| Keluar | Keluar | Bisa diakses | valid |

Pengujian data dengan Precision, Recall, dan akurasi Dalam pengujian ini data uji yang diinput sebanyak 17 orang nama mahasiswa. Dari ke-17 data uji tersebut dilakukan dengan 17 data testing dan yang mendapatkan 12 orang termasuk kategori True Positif (TP), dan 1 orang

termasuk dalam False Positif (FP), 3 orang termasuk False Negatif (FN) dan 1 orang termasuk True Negatif (TN) yang dapat dilihat pada gambar berikut :

| Data Testing | TP | FP | FN | TN |
|--------------|----|----|----|----|
| 17 | 12 | 1 | 3 | 1 |

Gambar 7. Hasil Prediksi Naïve Bayes Pada Sistem

V. KESIMPULAN

Menghasilkan sistem yang dapat memprediksi dapat memprediksi tingkat kelulusan mahasiswa tepat waktu atau tidak dengan menerapkan algoritma Naïve Bayes. Sistem ini dapat memprediksikan dan memberikan informasi atau masukan bagi pihak perguruan tinggi dalam membuat kebijakan kedepannya, serta bisa melakukan tindakan preventif jika diperlukan. Tingkat keberhasilan prediksi algoritma Naïve Bayes yang ditentukan oleh atribut nama, ips semester 1 sampai semester 7, ipk, sks, tahun lulus, organisasi, kompetisi, skill programing pada lulusan tahun 2020-2023 Program Studi Rekayasa Perangkat Lunak Politeknik Negeri Bengkalis. Dalam pengujian sistem ini digunakan perbandingan data latih dan data uji sebesar (90% dan 10%) sehingga didapatkan 146 data latih, 17 data uji mendapatkan akurasi 76,47%, presisi 92,31% dan recall 80%. Sedangkan jika perbandingan data latih dan data uji sebesar (80% dan 20%) maka didapatkan data latih = 130 dan data uji = 32, hasil akurasi 84,38%, presisi = 96.3%, recall = 87%.

VII. REFERENSI

- Afriansyah, M., Saputra, J., Yoga, V., Ardhana, P., & Sa, Y. (2024). *ALGORITMA NAIVE BAYES YANG EFISIEN UNTUK KLASIFIKASI BUAH*. 1(2), 236–248.
- Amelia, R., Koda, A., Mulyati, I., & Dana, R. D. (2022). *Penerapan Algoritma Frequent Pattern-Growth Dalam Menentukan Pola Penjualan*. 04(02), 65–71.
- Apridiansyah, Y., David, N., Veronika, M., & Putra, E. D. (2021). *Prediksi Kelulusan Mahasiswa Fakultas Teknik Informatika Universitas Muhammadiyah Bengkulu Menggunakan Metode Naive Bayes*. *JSAI : Journal Scientific and Applied Informatics*, 4(2), 236–247.
- Attamami, N., Triayudi, A., & Aldisa, R. T. (2023). *Analisis Performa Algoritma Klasifikasi Naive Bayes dan C4.5 untuk Prediksi Penerima Bantuan Jaminan Kesehatan*. *Jurnal JTik (Jurnal Teknologi Informasi Dan Komunikasi)*, 7(2), 262–269. <https://doi.org/10.35870/jtik.v7i2.756>
- Case, M., Kalimantan, W., Forest, P., Rudiyan, A., Dzulkifli, A. E., & Munazar, K. (2022). *Klasifikasi Kebakaran Hutan Menggunakan Metode K-Nearest Neighbor : Studi Kasus Hutan Provinsi Kalimantan Barat (Forest Fires Classification by Using The K-Nearest Neighbor*. 3(4), 195–202.
- Chen, H., Hu, S., & Hua, R. (2021). *Peningkatan algoritma klasifikasi Naive Bayes untuk manajemen risiko lalu lintas*. 6.
- David Imanuel, A., Nawaningtyas Pusparini, N., & Sani, A. (2024). *Klasifikasi Untuk Memprediksi Tingkat Kelulusan Mahasiswa Stmik Widuri Menggunakan Algoritma Naïve Bayes*. *Jurnal Ilmiah Informatika*, 12(01), 1–7. <https://doi.org/10.33884/jif.v12i01.8201>
- Eka, S., Putri, Y., St, S., Surahman, A., Kom, S., Kom, M., Informatika, S., Teknik, F., Ilmu, D.,

- Indonesia, U. T., Za, J., Alam, P., Ratu, L., Kedaton, K., & Lampung, K. B. (2021). *PENERAPAN MODEL NAIVE BAYES UNTUK MEMPREDIKSI POTENSI*. 1(1), 82–87.
- Ericha Apriliyani, & Salim, Y. (2022). Analisis performa metode klasifikasi Naïve Bayes Classifier pada Unbalanced Dataset. *Indonesian Journal of Data and Science*, 3(2), 47–54. <https://doi.org/10.56705/ijodas.v3i2.45>
- Kelayakan, M., & Rumah, K. (2020). 1,), 2. 11(2), 101–108.
- Khasanah, N., Salim, A., Afni, N., Komarudin, R., & Maulana, Y. I. (2022). Prediksi Kelulusan Mahasiswa Dengan Metode Naive Bayes. *Technologia : Jurnal Ilmiah*, 13(3), 207. <https://doi.org/10.31602/tji.v13i3.7312>
- Liliana, D. Y., Maulana, H., & Setiawan, A. (2021). *Data Mining untuk Prediksi Status Pasien Covid-19 dengan Pengklasifikasi Naïve Bayes*. 7(1), 48–53.
- Mardiani, E., Rahmansyah, N., Ningsih, S., & ... (2023). Komparasi Metode Knn, Naive Bayes, Decision Tree, Ensemble, Linear Regression Terhadap Analisis Performa Pelajar Sma. *Innovative: Journal Of ...*, 3(2), 13880–13892. <http://j-innovative.org/index.php/Innovative/article/view/1949%0Ahttp://j-innovative.org/index.php/Innovative/article/download/1949/1468>
- Minahasa, W. S., Farmer, R., Service, C., Ringringulu, T. R., Tinno, P., Rompas, D., Kumajas, S. C., Teknik, F., Manado, U. N., & Utara, S. (n.d.). *Aplikasi Layanan Pengaduan Petani Kabupaten Minahasa Selatan Berbasis Web*.
- Poerwandono, E., & Perwitosari, F. J. (2023). *Penerapan Data Mining Untuk Penilaian Kinerja Karya Di PT Riksa Dinar Djaya Menggunakan Metode Naive Bayes Classification*. 5(1), 336–340.
- Pratama, F. D., & Zufria, I. (2022). *Implementasi data mining menggunakan algoritma naïve bayes untuk klasifikasi penerima program indonesia pintar 1) 1,2,3*. 7(1), 77–84.
- Riyadsyah, D., Oktaviana, E., Zahra, A., Anjarwati, P., Gilang, V., Kristiawan, Y. I., Rohman, S. N., Komputer, F. I., Studi, P., Informatika, T., & Bangsa, U. D. (2024). *No Title*. 2, 303–311.
- Romano, M., Zammarchi, G., & Conversano, C. (2024). *1. Perkenalan*.
- Rosyid Mubarak. (2024). Komparasi Performa Naive Bayes Gaussian dan K-NN Untuk Prediksi Kelulusan Mahasiswa dengan CRISP-DM. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 4(6), 2982–2991. <https://doi.org/10.30865/klik.v4i6.1924>
- Saputra, B., Anwar, S., Tohidi, E., Susana, H., Pratama, D., Studi, P., & Informatika, T. (2023). *PENERAPAN ALGORITMA NAÏVE BAYES DALAM KLASIFIKASI HARGA PONSEL*. 7(6), 3587–3594.
- Setiyani, L., Wahidin, M., Awaludin, D., & Purwani, S. (2020). Analisis Prediksi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Data Mining Naïve Bayes : Systematic Review. *Faktor Exacta*, 13(1), 35. <https://doi.org/10.30998/faktorexacta.v13i1.5548>
- Simanjuntak, A. Y., Septian, I., Simatupang, S., & Indonesia, U. P. (2022). *IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER UNTUK DATA KENAIKAN PANGKAT DINAS*. 4307(1), 85–91.

-
- Sistem, R., Lestandy, M., Abdurrahim, A., & Syafa, L. (2021). *Analisis Sentimen Tweet Vaksin COVID-19 Menggunakan Recurrent*. 5(10), 802–808
- Studi, P., Sumber, P., Manusia, D., & Airlangga, P. U. (n.d.). *OPINION MINING FUNGSI KPI (KEY PERFORMANCE INDIKATOR) DENGAN ALGORITMA NAÏVE BAYES CLASIFIER DAN SUPPORT VECTOR MACHINE (SVM)* Elwin Dewi Nurhazizah 1 , Ira Puspitasari 2. 290–302.
- Sukarna, R. H., Ansori, Y., Informatika, T., Komputer, I., & Luhur, U. B. (2022). *IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAIVE BAYES DENGAN FEATURE SELECTION UNTUK*. 6, 50–61.
- Susana, H., & Suarna, N. (2022). *PENERAPAN MODEL KLASIFIKASI METODE NAIVE BAYES*. 4(1), 2–9.
- Teknologi, J., Dan, I., & Vol, K. (2021). *TEMATIK - Jurnal Teknologi Informasi Dan Komunikasi Vol. 8, No. 1 Juni 2021*. 8(1), 96–103.