

PENERAPAN ALGORITMA NAÏVE BAYES PADA DATA MINING UNTUK PERBANDINGAN TINGKAT AKURASI KEMISKINAN DI INDONESIA

Aryo Putro Adi Baskorojati¹, Rizky Tahara Shita²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia.

Email: ¹aryo.p.a.bj@email.com, ²rizky.taharashita@email.com
(* : corresponding author)

Abstrak- Badan Pusat Statistik (BPS) memiliki peran penting dalam menyediakan data statistik yang akurat dan terkini untuk Indonesia, namun pengolahan dan analisis data dalam jumlah besar menjadi tantangan tersendiri. Penerapan sistem prediksi berbasis web menggunakan metode Naïve Bayes dapat menjadi solusi yang tepat untuk mengatasi tantangan ini, memungkinkan BPS memproses dan menganalisis data secara lebih efektif dan efisien. Data mining, yang merupakan teknik penggalian informasi dari data, dapat digunakan untuk menganalisis data statistik dari berbagai sektor, seperti ekonomi, kesehatan, dan pendidikan. Permasalahan yang dihadapi yaitu Badan Pusat Statistik (BPS) membutuhkan aplikasi untuk menghadapi tantangan tingkat akurasi data kemiskinan, aplikasi tersebut perlu dirancang untuk menangani masalah yang mempengaruhi akurasi. Metode Naïve Bayes, yang menggunakan algoritma klasifikasi berbasis probabilitas, dapat membantu BPS memprediksi kategori data baru dengan cukup akurat. Sistem prediksi ini terdiri dari beberapa tahap, yaitu pengumpulan data, training model, prediksi, dan pengambilan keputusan. Implementasi sistem ini diharapkan dapat memberikan beberapa keuntungan, seperti meningkatkan akurasi dan efisiensi dalam pengolahan dan analisis data, memudahkan BPS dalam memprediksi tren dan pola data, serta mendukung pengambilan keputusan yang lebih tepat berdasarkan data. Secara keseluruhan, penerapan sistem prediksi berbasis web menggunakan metode Naïve Bayes dapat menjadi solusi yang efektif dan efisien untuk membantu BPS dalam mengelola data statistik dan mendukung pengambilan keputusan. Setelah dilakukan pengolahan data menggunakan metode preprocessing, proses algoritma naïve bayes dijalankan menggunakan data training 10% dan data testing 90% dari data kemiskinan di Indonesia periode tahun 2023 sampai dengan tahun 2024 dengan data keseluruhan sebanyak 210 data. Berdasarkan hasil pengujiannya diperoleh hasil accuracy 94,2%. Tujuan dari penelitian ini adalah untuk menilai efektivitas dan keakuratan algoritma dalam mengklasifikasikan data kemiskinan.

Kata Kunci: Perbandingan Tingkat Akurasi, Algoritma Naïve Bayes, Data Mining, Kemiskinan.

IMPLEMENTATION OF THE NAÏVE BAYES ALGORITHM IN DATA MINING FOR COMPARING POVERTY ACCURACY LEVELS IN INDONESIA

Abstract- The Central Bureau of Statistics (BPS) plays a crucial role in providing accurate and up-to-date statistical data for Indonesia, but processing and analyzing large amounts of data poses its own challenges. The implementation of a web-based prediction system using the Naïve Bayes method can be an effective solution to address these challenges, enabling BPS to process and analyze data more effectively and efficiently. Data mining, which is a technique for extracting information from data, can be used to analyze statistical data from various sectors such as economics, health, and education. The issue faced by the Central Bureau of Statistics (BPS) is the need for an application to address the challenge of data accuracy in poverty statistics, and this application needs to be designed to handle issues affecting accuracy. The Naïve Bayes method, which uses a probability-based classification algorithm, can help BPS predict new data categories with reasonable accuracy. This prediction system consists of several stages: data collection, model training, prediction, and decision-making. The implementation of this system is expected to offer several advantages, such as improving accuracy and efficiency in data processing and analysis, facilitating BPS in predicting data trends and patterns, and supporting more accurate decision-making based on data. Overall, the implementation of a web-based prediction system using the Naïve Bayes method can be an effective and efficient solution to assist BPS in managing statistical data and supporting decision-making. After data processing using the preprocessing method, the Naïve Bayes algorithm was executed using 10% training data and 90% testing data from poverty data in Indonesia for the period from 2023 to 2024, with a total of 210 data points. The test results showed an accuracy of 94.2%. The purpose of this study is to assess the effectiveness and accuracy of the algorithm in classifying poverty data.

Keywords: Comparison of Accuracy Levels, Naïve Bayes Algorithm, Data Mining, Poverty.

1. PENDAHULUAN

Badan Pusat Statistik (BPS) memiliki peran penting dalam menyediakan data statistik yang akurat dan terkini untuk Indonesia, namun pengolahan dan analisis data dalam jumlah besar menjadi tantangan tersendiri. Penerapan sistem prediksi berbasis web menggunakan metode Naïve Bayes dapat menjadi solusi yang tepat untuk mengatasi tantangan ini, memungkinkan BPS memproses dan menganalisis data secara lebih efektif dan efisien. Data mining, yang merupakan teknik penggalian informasi dari data, dapat digunakan untuk menganalisis data statistik dari berbagai sektor, seperti ekonomi, kesehatan, dan pendidikan. Metode Naïve Bayes, yang menggunakan algoritma klasifikasi berbasis probabilitas, dapat membantu BPS memprediksi kategori data baru dengan cukup akurat. Sistem prediksi ini terdiri dari beberapa tahap, yaitu pengumpulan data, training model, prediksi, dan pengambilan keputusan. Implementasi sistem ini diharapkan dapat memberikan beberapa keuntungan, seperti meningkatkan akurasi dan efisiensi dalam pengolahan dan analisis data, memudahkan BPS dalam memprediksi tren dan pola data, serta mendukung pengambilan keputusan yang lebih tepat berdasarkan data. Secara keseluruhan, penerapan sistem prediksi berbasis web menggunakan metode Naïve Bayes dapat menjadi solusi yang efektif dan efisien untuk membantu BPS dalam mengelola data statistik dan mendukung pengambilan keputusan.

Penelitian sebelumnya telah menggunakan algoritma Naïve Bayes untuk membantu memprediksi klasifikasi penerima bantuan sembako, di mana algoritma ini mampu menemukan pola kesamaan karakteristik dalam kelompok atau kelas tertentu, sehingga dapat membantu dalam menentukan apakah seseorang layak menerima bantuan sembako atau tidak. Dalam penelitian ini, data sampel dari kabupaten/kota XYZ digunakan untuk melatih dan menguji model prediksi, dan algoritma Naïve Bayes diimplementasikan serta dianalisis menggunakan aplikasi berbasis web yang khusus dikembangkan untuk tujuan ini. Secara singkat, penelitian ini bertujuan untuk memprediksi klasifikasi penerima bantuan sembako menggunakan algoritma Naïve Bayes, mengembangkan aplikasi berbasis web untuk implementasi dan analisis algoritma Naïve Bayes, serta menguji akurasi model prediksi pada data sampel dari kabupaten/kota XYZ. Penelitian ini diharapkan dapat membantu pemerintah dalam mendistribusikan bantuan sembako secara lebih tepat dan efisien [1].

Agar data mining menghasilkan informasi yang bermanfaat, penting untuk melakukan serangkaian langkah-langkah mulai dari pengumpulan data yang cermat, preprocessing untuk memastikan data dalam kondisi optimal, penerapan algoritma Naïve Bayes untuk mengklasifikasikan data, hingga evaluasi model untuk memastikan akurasi dan keandalannya [2]. Proses ini memastikan bahwa data mining menghasilkan informasi yang akurat dan bermanfaat untuk pengambilan keputusan [3].

Penelitian ini berfokus pada penerapan algoritma Naïve Bayes untuk mengklasifikasikan tingkat kemiskinan di Indonesia dengan tujuan membantu pemerintah merancang program pengentasan kemiskinan yang lebih efektif dan efisien. Penerapan algoritma Naïve Bayes ini diharapkan dapat memberikan beberapa manfaat, yaitu membantu pemerintah mengidentifikasi individu dan keluarga yang miskin, menentukan jenis program pengentasan kemiskinan yang tepat untuk setiap individu dan keluarga, serta memantau dan mengevaluasi efektivitas program pengentasan kemiskinan. Penelitian ini juga diharapkan dapat memberikan kontribusi yang berarti dalam dua bidang, yaitu bidang data mining dengan menunjukkan bahwa algoritma Naïve Bayes dapat digunakan secara efektif untuk klasifikasi kemiskinan, membuka peluang baru untuk penelitian data mining di bidang lain, serta implementasi kebijakan publik dengan menunjukkan bahwa data mining dapat digunakan untuk mendukung pengambilan keputusan berbasis data dalam kebijakan publik, membantu pemerintah merancang dan melaksanakan kebijakan yang lebih efektif dan efisien. Secara keseluruhan, penelitian ini diharapkan dapat membantu pemerintah dalam mencapai tujuannya untuk mengentaskan kemiskinan di Indonesia [4].

Pada penelitian yang dilakukan oleh [5] yang berjudul “Naïve Bayes Classification untuk Penentuan Status Penduduk Miskin”, penelitian tersebut mengelompokkan penduduk untuk menilai tingkat kemiskinan dengan menerapkan metode Naïve Bayes dengan menggunakan tujuh atribut untuk memprediksi tingkat kemiskinan di Kecamatan Warudoyong.

Perbedaan penelitian tersebut dengan penelitian ini terletak pada penerapan permasalahannya. Penelitian sebelumnya mengelompokkan 7 atribut yang digunakan yaitu umur, tanggungan, pekerjaan, pendidikan, pendapatan, keikutsertaan program JKN/KIS dan PKH, sedangkan penelitian ini mengakurasi kenaikan kemiskinan penduduk dari tahun 2023 hingga 2024 dengan pengelompokan 4 atribut [5].

Pada penelitian ini, hasil Classification yang didapat dari perhitungan Naïve Bayes dievaluasi menggunakan metode Classification, yang dimana penelitian ini dapat mengevaluasi hasil perhitungan dengan Confusion Matrix, sehingga dapat menentukan akurasi yang lebih baik.

2. METODE PENELITIAN

2.1 Data Penelitian

Sumber data yang digunakan dalam penelitian ini bersumber dari data Badan Pusat Statistik, berdasar pada garis kemiskinan menurut kabupaten/kota yang digunakan selama periode 2023 hingga 2024 yang dilakukan proses penginputan oleh admin. Langkah pertama adalah dengan membaca data awal sebanyak 210 data yang tidak memiliki label atau kelas. analisis metode *naive bayes* dilakukan pada data awal ini, yang disajikan dalam tabel 2.1 berikut:

Tabel 1. Data Penelitian

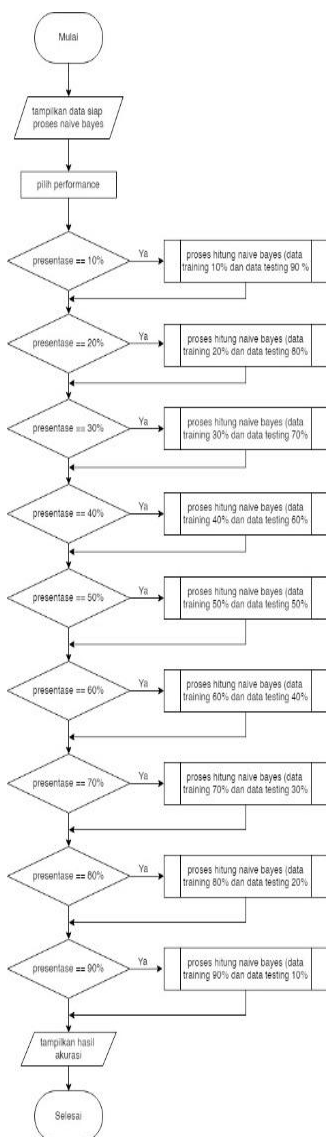
No	Wilayah	Kemiskinan 2023	Penduduk Miskin 2023	Kemiskinan 2024	Penduduk Miskin 2024	Kondisi
1	Simeulue	538693	17.62	576505	17.59	Naik
2	Aceh Singkil	568691	24.62	609322	24.84	Turun
3	Aceh Selatan	494565	30.36	528243	30.53	Turun
4	Aceh Tenggara	471301	27.96	496074	27.28	Naik
5	Aceh Timur	530934	60.63	557943	60.86	Naik
6	Aceh Tengah	584863	31.68	626090	31.85	Naik
7	Aceh Barat	616091	38.84	644009	38.79	Naik
8	Aceh Besar	564431	58.94	586860	58.98	Naik
9	Pidie	579450	86.79	607117	86.89	Turun
10	Bireuen	486667	59.21	494866	59.93	Turun

2.2 Penerapan Metode

Pada tahap ini dilakukan proses KDD (*Knowledge Discovery Database*) berikut ini tahapan-tahapan yang dilakukan yaitu *Data Seleccion, Pre-processing, Transformation, Data Mining, Evaluation* [6].

2.3 Penerapan *Naïve Bayes*

Proses dataset, kinerja, dan prediksi dijelaskan pada Gambar 1 dari *flowchart* *naïve bayes* ini. Di sini, dataset yang diimport dari file excel, kemudian dihasilkan Accuracy atau akurasi, Dimana setiap presentase yang dipilih akan melakukan proses menghitung. Jika memilih presentase 10% maka proses hitung *Naïve Bayes* akan menggunakan 10% data training dan 90% data testing, apabila memilih presentase 90% maka proses hitung *Naïve Bayes* akan menggunakan 90% data training dan 10% data testing.



Gambar 1. Flowchart Naïve Bayes

2.4 Confussion Matrix

Klasifikasi data seperti proses detektif yang mencari pola dan hubungan dalam data untuk memecahkan kasus. Tujuannya adalah menemukan model atau aturan yang dapat menjelaskan dan memprediksi kelas data atau konsep tertentu. Proses klasifikasi data umumnya terdiri dari beberapa langkah, seperti pengumpulan data, pemilihan fitur, pemilihan algoritma, pelatihan model, evaluasi model, dan penyebaran model [7].

2.5 Data Penelitian

Sumber data yang digunakan dalam penelitian ini bersumber dari data Badan Pusat Statistik, berdasar pada garis kemiskinan menurut kabupaten/kota yang digunakan selama periode 2023 hingga 2024 yang dilakukan proses penginputan oleh admin.

2.6 Pengumpulan Data

Dalam melakukan penelitian, peneliti mengumpulkan data dengan tujuan untuk memperoleh informasi yang relevan terkait dengan topik yang dibahas, yaitu :

Menggunakan website Badan Pusat Statistik untuk mendapatkan dataset data tentang kemiskinan di Indonesia dikumpulkan dari Badan Pusat Statistik, yang telah divalidasi oleh Ketua Program Studi Teknik

Informatika. Dalam dokumentasi, peneliti meminta Badan Pusat Statistik (BPS) untuk menyediakan data yang diperlukan untuk proses pengembangan sistem aplikasi.

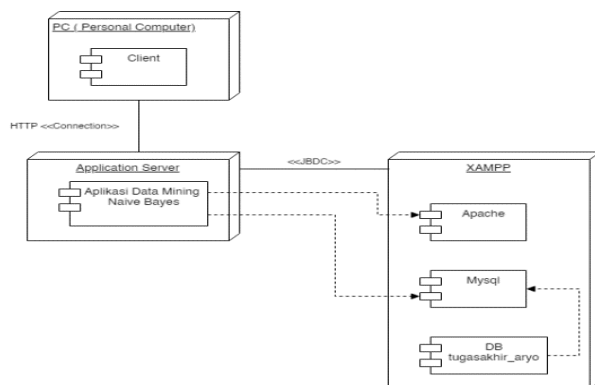
3. HASIL DAN PEMBAHASAN

Sumber Data yang digunakan dalam penelitian saat ini diambil dari sumber resmi seperti Badan Pusat Statistik (BPS) dan beberapa data dari sumber lain yang mendukung dipertimbangkan dalam penelitian ini. Data yang digunakan adalah data kemiskinan yang diambil di website resmi BPS sebanyak 210 data. Penelitian ini hanya untuk menghitung tingkat akurasi kemiskinan di Indonesia. Data yang digunakan adalah data kemiskinan pada tahun 2023 sampai dengan tahun 2024 [8].

3.1 Lingkungan Percobaan

Pengimplementasian untuk aplikasi ini membutuhkan kelengkapan baik dari segi *software* maupun *hardware*. Adapun kelengkapan yang diperlukan adalah minimum 1 buah set komputer/laptop dan software pendukung. Penjelasan meliputi, antara lain: Processor AMD Ryzen 3 3200U CPU @ 2.60GHz, RAM / Memory 12 GB, SSD Harddisk 512 GB, Sistem Operasi Microsoft Windows 11 Pro, Text Editor Visual Studio Code, PHP 7.4.3, XAMPP, Browser (Google Chrome).

Deployment diagram menunjukkan konfigurasi komponen selama proses eksekusi aplikasi. Nilai bukti dalam satu sampel tetap konstan untuk setiap kelas. Nilai bukti ini atau dibandingkan dengan nilai posterior dari kelas lain untuk menentukan kelas mana sampel tersebut harus diklasifikasikan. *Deployment diagram* bisa dilihat pada Gambar 2 berikut.



Gambar 2. *Deployment diagram*

3.1.1 Pengumpulan Data

Langkah pertama adalah dengan membaca data awal sebanyak 210 data yang tidak memiliki label atau kelas. analisis metode *naive bayes* dilakukan pada data awal ini [9], yang disajikan dalam tabel 2 berikut:

Tabel 2. Pengumpulan Data

No	Wilayah	Kemiskinan 2023	Penduduk Miskin 2023	Kemiskinan 2024	Penduduk Miskin 2024	Kondisi
1	Simeulue	538693	17.62	576505	17.59	Naik
2	Aceh Singkil	568691	24.62	609322	24.84	Turun
3	Aceh Selatan	494565	30.36	528243	30.53	Turun
4	Aceh Tenggara	471301	27.96	496074	27.28	Naik
5	Aceh Timur	530934	60.63	557943	60.86	Naik
6	Aceh Tengah	584863	31.68	626090	31.85	Naik
7	Aceh Barat	616091	38.84	644009	38.79	Naik
8	Aceh Besar	564431	58.94	586860	58.98	Naik
9	Pidie	579450	86.79	607117	86.89	Turun
10	Bireuen	486667	59.21	494866	59.93	Turun

3.1.2 Probabilitas

Pada table-table berikut ini adalah hasil dari perhitungan tota dataset kemiskinan di Indonesia yang sudah dipreprocessing sebanyak 210, dengan hasil probabiitas dari setiap atribut tahun 2023, tahun 2024, kategori:

Tabel 3. Jumlah Kategori Kemiskinan di Indonesia

Atribut	Kategori
Tahun_2023	BR AR
Jumlah Penduduk	Rendah Sedang Tinggi
Tahun_2024	BR AR
Jumlah Penduduk	Rendah Sedang Tinggi
Kondisi_Naik	Naik Turun

3.1.3 Data Preprocessing

Data *preprocessing* pada tabel berikut ini adalah data yang sudah ada label kelasnya [7], dihitung dari kategori sesuai dengan datasetnya yaitu: Wilayah, Tahun 2023, Tahun 2024, dan Kondisi Naik (apakah Naik atau Turun).

Tabel 4. *Preprocessing*

No	Wilayah	Kemiskinan 2023	Penduduk Miskin 2023	Kemiskinan 2024	Penduduk Miskin 2024	Kondisi
1	Simeulue	BR	Rendah	AR	Rendah	Naik
2	Aceh Singkil	AR	Rendah	BR	Rendah	Turun
3	Aceh Selatan	BR	Rendah	BR	Rendah	Turun
4	Aceh Tenggara	BR	Rendah	BR	Rendah	Naik
5	Aceh Timur	BR	Sedang	AR	Sedang	Naik
6	Aceh Tengah	AR	Rendah	AR	Rendah	Naik
7	Aceh Barat	AR	Rendah	AR	Rendah	Naik
8	Aceh Besar	AR	Sedang	AR	Sedang	Naik
9	Pidie	AR	Sedang	BR	Sedang	Turun
10	Bireuen	BR	Sedang	BR	Sedang	Turun

3.1.4 Proses Pengujian Verifikasi Ke Data Baru

- 1) Jika diberi input untuk data uji
 - a) Wilayah: Simeulue
Naik: 1, Turun: 0,
Simeulue
* Naik: $1 = 1/193=0.0051813471502591$
* Turun: $0 = 0/17=0$
 - b) Kemiskinan_Tahun_2023 :: BR
Naik : 113, Turun : 16,
BR
* Naik: $113 = 113/193=0.58549222797927$
* Turun: $16 = 16/17=0.94117647058824$
AR
* Naik: $80 = 80/193=0.41450777202073$
* Turun: $1 = 1/17=0.058823529411765$

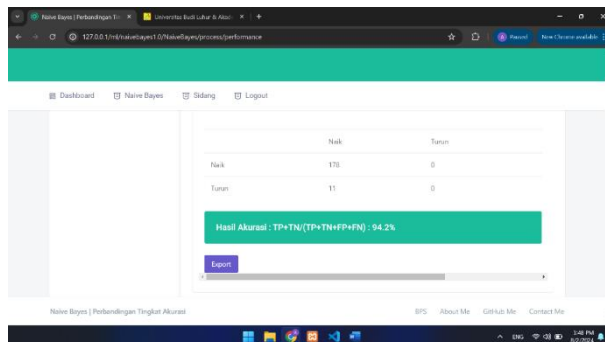
- c) Penduduk_Miskin_2023 :: Rendah
Naik : 115, Turun : 12,
Rendah
* Naik: $115 = 115/193=0.59585492227979$
* Turun: $12 = 12/17=0.70588235294118$
Sedang
* Naik: $37 = 37/193=0.19170984455959$
* Turun: $4 = 4/17=0.23529411764706$
Tinggi
* Naik: $41 = 41/193=0.21243523316062$
* Turun: $1 = 1/17=0.058823529411765$
- d) Kemiskinan_Tahun_2024 :: BR
Naik : 106, Turun : 16,
BR
* Naik: $106 = 106/193=0.54922279792746$
* Turun: $16 = 16/17=0.94117647058824$
AR
* Naik: $87 = 87/193=0.45077720207254$
* Turun: $1 = 1/17=0.058823529411765$
- e) Penduduk_Miskin_2024 :: Rendah
Naik : 117, Turun : 12,
Rendah
* Naik: $117 = 117/193=0.60621761658031$
* Turun: $12 = 12/17=0.70588235294118$
Sedang
* Naik: $36 = 36/193=0.18652849740933$
* Turun: $4 = 4/17=0.23529411764706$
Tinggi
* Naik: $40 = 40/193=0.20725388601036$
* Turun: $1 = 1/17=0.058823529411765$
- f) Total Label
Naik : 193 (0.00055312015013856)
Turun : 17 (0)
Total : 210
- g) Hasil Prediksi
Turun : 0
Naik : 0.00055312015013856
Naik
Hasil array ([Turun] => 0 [Naik] => 0.0034632034632035) Turun: 0
Maka, dari hasil prediksi diatas untuk status kemiskinan wilayah tersebut adalah naik.

3.2 Langkah Pengujian

Tampilan layar yang ada pada sistem Prediksi Tingkat Akurasi dataset kemiskinan ini dibuat semudah mungkin untuk dipahami sehingga pengguna mudah dalam menggunakannya. Penjelasan dan gambar dari tampilan layar sistem dapat ditemukan berikut ini:

3.2.1. Tampilan Hasil Uji Akurasi

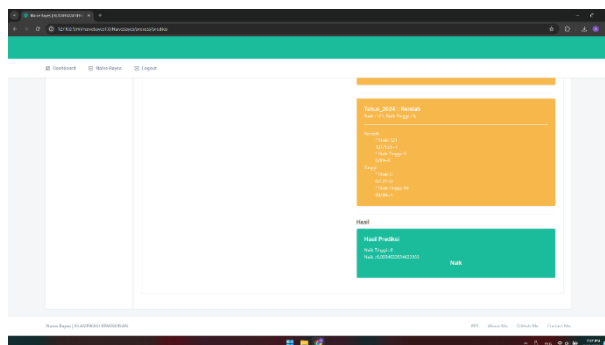
Pada tampilan hasil uji akurasi metode yang ada pada menu Uji Akurasi merupakan data training dan data testing yang telah di pilih presentase data training 80% dan data testing 20% [10]. Berdasarkan hasil pengujiannya diperoleh hasil *Accuracy* 94,2%, dibagian bawah juga terdapat export untuk mengunduh data yang sudah di presentase menjadi file berformat PDF. Tampilan layar pengujian dapat dilihat pada Gambar 3 dan seterusnya berikut ini:



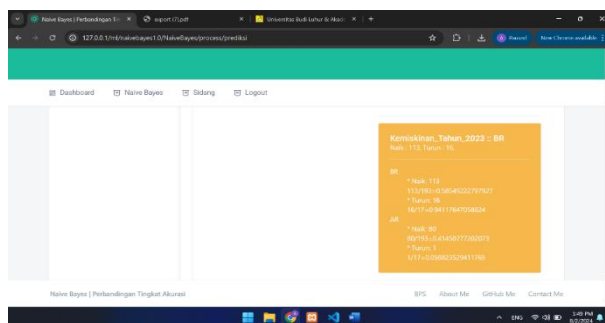
Gambar 3. Tampilan Layar Hasil Akurasi

3.2.2. Tampilan Layar Prediksi

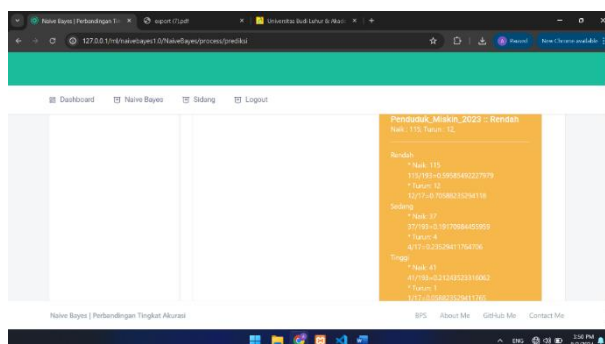
Tampilan layar Prediksi ini adalah untuk menampilkan tabel [11], perhitungan prediksi lalu hasil prediksi apakah kemiskinan tersebut naik atau turun. Pada gambar 3.2 hingga 3.7 menunjukkan tampilan layar:



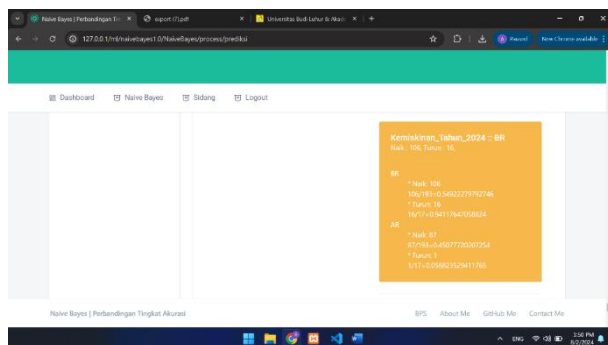
Gambar 4. Tampilan Layar Prediksi



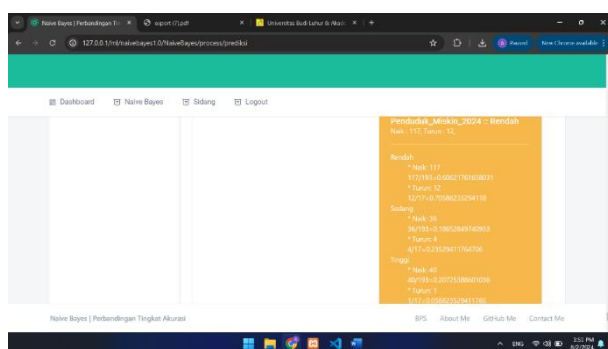
Gambar 5. Tampilan Layar Prediksi Kemiskinan Tahun 2023



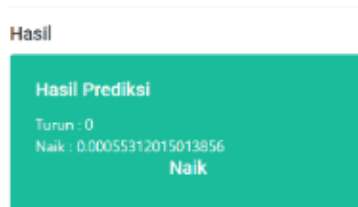
Gambar 6. Tampilan Layar Penduduk Miskin Tahun 2023



Gambar 7. Tampilan Layar Prediksi Kemiskinan Tahun 2024



Gambar 8. Tampilan Layar Penduduk Miskin Tahun 2024



Gambar 9. Tampilan Layar Prediksi Hasil

4. KESIMPULAN

Setelah uji coba dan analisis yang dilakukan dalam penelitian ini, beberapa hal dapat disimpulkan dari penggunaan Algoritma Naïve Bayes untuk aplikasi Perbandingan Tingkat akurasi Kemiskinan di Indonesia yaitu:

- Metode Naïve Bayes dapat memprediksi perbandingan tingkat akurasi pada tingkat Kemiskinan di Indonesia Data pelatihan yang digunakan dalam klasifikasi mempengaruhi hasil uji coba atau pengujian.
- Aplikasi ini juga cocok digunakan dalam berbagai industri, karena sangat mudah dalam penggunaannya.

DAFTAR PUSTAKA

- [1] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, pp. 219-225, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [2] A. Twin, *Data Mining Data mining*, vol. 2, no. 2020. 2023.
- [3] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE: Journal of Business and Audit Information System*, vol. 4, no. 1, pp. 13–21, 2021, doi: 10.30813/jbase.v4i1.2729.
- [4] A. Asri, A. Arifin, W. Handoko, and Z. Efendi, "Implementasi Metode Naive Bayes Untuk Klasifikasi Penerima Program Keluarga Harapan," *J-Com (Journal of Computer)*, vol. 2, no. 1, pp. 21–26, 2022.

- [5] M. Rasyida, “Naïve Bayes Classification untuk Penentuan Status Penduduk Miskin,” *J. Inform. Kaputama*, vol. 4, no. 2, pp. 175–180, 2020, doi: 10.59697/jik.v4i2.329.
- [6] Y. Mardi, “Data Mining: Klasifikasi Menggunakan Algoritma C4.5 Data Mining Merupakan Bagian Dari Tahapan Proses Knowledge Discovery in Database (KDD),” *Jurnal Edik Informatika*, vol. 2, no. 2, pp. 213–219, 2019.
- [7] A. Triayudi and Sumiati, “Implementasi Klasifikasi Data Mining Untuk Penentuan Kelayakan Pemberian Kredit dengan Menggunakan Algoritma Naïve Bayes,” *Jurnal Sistem Komputer dan Informatika*, vol. 4, no. 1, pp. 240–244, 2022, doi: 10.30865/json.v4i1.4653.
- [8] H. Harliana and F. N. Putra, “Klasifikasi Tingkat Rumah Tangga Miskin Saat Pandemi Dengan Naïve Bayes Classifier,” *Jurnal Sains dan Informatika*, vol. 7, no. 2, pp. 165–173, 2021, doi: 10.34128/jsi.v7i2.339.
- [9] W. Yulita, E. D. Nugroho, and M. H. Algifari, “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier,” *Jurnal Data Mining dan Sistem Informasi*, vol. 2, no. 2, pp. 1-9, 2021, doi: 10.33365/jdmsi.v2i2.1344.
- [10] D. Darwis, N. Siskawati, and Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *Jurnal Teknokompak*, vol. 15, no. 1, pp. 131-145, 2021, doi: 10.33365/jtk.v15i1.744.
- [11] H. Susana, “Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet,” *JURSISTEKNI: Jurnal Riset Sistem Informasi dan Teknologi Informasi*, vol. 4, no. 1, pp. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.