



# Analisis Kepuasan Masyarakat Terhadap Proses Pengurusan Sertipikat Analog Ke Elektronik Menggunakan Metode Naïve Bayes

Muhammad Ikhsan Al-Arafi\*, Rini Sovia, Agung Ramadhanu

Fakultas Ilmu Komputer, Magister Teknik Informatika, Universitas Putra Indonesia YPTK Padang, Padang, Indonesia

Email: <sup>1,\*</sup>15juni1996@gmail.com, <sup>2</sup>rini\_sovia@upiyptk.ac.id, <sup>3</sup>agung\_ramadhanu@upiyptk.ac.id

Email Penulis Korespondensi: 15juni1996@gmail.com

**Abstrak**—Program alih media sertipikat analog ke elektronik yang dilaksanakan oleh Kementerian ATR/BPN di Desa Sejati memerlukan evaluasi untuk memastikan efektivitas pelaksanaannya. Permasalahan yang dihadapi adalah terbatasnya analisis kuantitatif berbasis data dalam mengidentifikasi faktor-faktor yang memengaruhi kepuasan masyarakat. Penelitian ini bertujuan menganalisis tingkat kepuasan masyarakat menggunakan metode Naïve Bayes untuk mengklasifikasikan dan memprediksi pengaruh variabel terkait. Data diperoleh dari 250 responden melalui kuesioner berbasis indikator layanan publik digital, mencakup variabel demografi, persepsi manfaat, kendala, dukungan, kecepatan pelayanan, dan kemudahan prosedur. Hasil penelitian menunjukkan bahwa tingkat kepuasan masyarakat berada pada kategori tinggi, dengan variabel kemudahan prosedur dan kecepatan pelayanan terbukti paling signifikan memengaruhi prediksi kepuasan. Model Naïve Bayes yang digunakan berhasil mencapai akurasi sebesar 94%, membuktikan efektivitasnya dalam memprediksi tingkat kepuasan. Temuan ini bermanfaat sebagai dasar perbaikan kebijakan dan strategi peningkatan kualitas pelayanan publik berbasis digital, khususnya dalam implementasi alih media sertipikat elektronik di masa mendatang.

**Kata Kunci:** Naïve Bayes; Tingkat Kepuasan; Sertipikat Elektronik; Layanan Publik Digital; Klasifikasi

**Abstract**—The certificate media conversion program from analog to electronic implemented by the Ministry of ATR/BPN in Sejati Village requires evaluation to ensure its effectiveness. The main problem faced is the limited use of quantitative, data-driven analysis in identifying the factors that influence public satisfaction. This study aims to analyze the level of public satisfaction using the Naïve Bayes method to classify and predict the influence of related variables. Data were obtained from 250 respondents through questionnaires based on digital public service indicators, covering demographic variables, perceived benefits, obstacles, support, service speed, and procedural simplicity. The results show that the level of public satisfaction is in the high category, with procedural simplicity and service speed proven to be the most significant variables influencing satisfaction prediction. The Naïve Bayes model achieved an accuracy of 94%, demonstrating its effectiveness in predicting satisfaction levels. These findings serve as a basis for improving policies and strategies to enhance the quality of digital public services, particularly in the implementation of electronic certificate media conversion in the future.

**Keywords:** Naïve Bayes; Satisfaction Level; Electronic Certificate; Digital Public Services; Classification

## 1. PENDAHULUAN

Dalam konteks pelayanan publik, pemanfaatan AI dan big data memberikan peluang besar untuk meningkatkan kualitas layanan melalui pendekatan berbasis data. Salah satu pendekatan yang sering digunakan adalah metode klasifikasi dalam data mining, seperti Naïve Bayes yang termasuk ke dalam cabang machine learning [1]. Program digitalisasi layanan publik, khususnya di sektor pertanahan, memiliki tujuan untuk menggantikan sistem manual dengan sistem elektronik yang lebih modern [2]. Namun demikian, implementasi kebijakan ini masih menghadapi tantangan seperti rendahnya literasi digital masyarakat, prosedur administrasi yang rumit, dan keterbatasan evaluasi kuantitatif yang dapat mengidentifikasi faktor-faktor dominan yang memengaruhi kepuasan masyarakat [3].

Pendekatan Naïve Bayes terbukti efektif dalam mengelompokkan data survei masyarakat berdasarkan karakteristik tertentu untuk mengetahui tren kepuasan layanan [4]. Algoritma ini juga sesuai untuk menganalisis data survei kategorikal dan memiliki performa klasifikasi yang cukup akurat. Penelitian sebelumnya menunjukkan bahwa penerapan Naïve Bayes dalam mengklasifikasikan tingkat kepuasan masyarakat terhadap pelayanan desa digital mampu mencapai tingkat akurasi sebesar 87%, dengan kecepatan layanan sebagai faktor dominan [5]. Selanjutnya, pengembangan sistem klasifikasi berbasis web dengan algoritma ini membantu pemerintah desa dalam mengelompokkan prioritas pelayanan [6]. Selain itu, rendahnya literasi digital masyarakat juga terbukti berdampak langsung terhadap kepuasan dalam penggunaan layanan publik digital [7].

Fleksibilitas *Naïve Bayes* juga terlihat dari penerapannya di berbagai bidang di luar pelayanan publik [8]. Pada bidang *Natural Language Processing* (NLP), menggunakan Multinomial Naïve Bayes untuk analisis sentimen ulasan film di *Internet Movie Database* (IMDB) dan mencatatkan akurasi hingga 87,63% [9]. Di bidang medis, menunjukkan bahwa *Naïve Bayes* dapat memprediksi mortalitas pasien dengan hasil sebanding dengan model lain seperti Random Forest dan Extreme Gradient Boosting, serta memiliki keunggulan interpretabilitas [10]. Lebih lanjut, pengembangan algoritma *Iterative Threshold-based Naïve Bayes* (iTb-NB) berhasil meningkatkan stabilitas prediksi dalam analisis ulasan pelanggan [11]. Keberhasilan lintas bidang ini memperkuat posisi *Naïve Bayes* sebagai metode sederhana namun efektif untuk berbagai permasalahan klasifikasi.

Meskipun penelitian terdahulu telah membuktikan efektivitas *Naïve Bayes* dalam mengukur kepuasan publik, sebagian besar studi masih terbatas pada layanan umum atau bidang lain seperti NLP dan medis [12]. Belum banyak penelitian yang secara khusus mengkaji layanan pertanahan digital, khususnya pada alih media sertipikat tanah elektronik [13]. Selain itu, penelitian sebelumnya lebih menekankan pada capaian akurasi model, tetapi kurang menguraikan



variabel-variabel paling signifikan yang memengaruhi prediksi kepuasan masyarakat [14]. Keterbatasan ini menjadi celah penelitian yang berusaha diisi dalam studi ini.

Penelitian ini bertujuan menganalisis tingkat kepuasan masyarakat terhadap layanan sertipikat tanah elektronik dengan menerapkan algoritma Naïve Bayes [15]. Secara khusus, penelitian ini menekankan pada identifikasi variabel dominan, yaitu kemudahan prosedur dan kecepatan pelayanan, yang berpengaruh signifikan terhadap tingkat kepuasan masyarakat. Kontribusi penelitian ini adalah: (1) menyediakan evaluasi kuantitatif berbasis machine learning terhadap program sertipikat elektronik; (2) menghadirkan model prediktif yang dapat dijadikan dasar dalam perbaikan kebijakan pelayanan publik digital; dan (3) memperkaya literatur terkait penerapan Naïve Bayes dalam konteks layanan pertanahan, yang masih jarang dikaji.

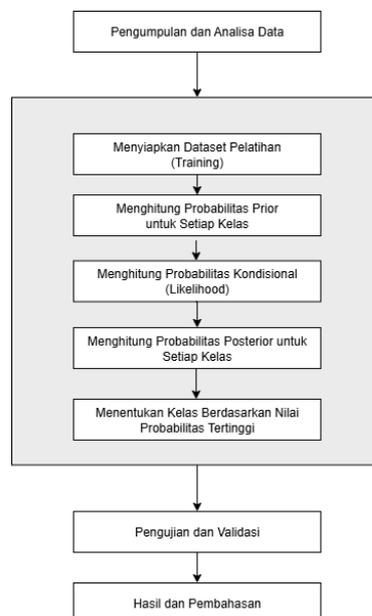
## 2. METODOLOGI PENELITIAN

### 2.1 Objek Penelitian

Objek penelitian ini adalah masyarakat Desa Sejati sebagai penerima layanan alih media sertipikat tanah elektronik yang dilaksanakan oleh Kementerian ATR/BPN [16]. Fokus penelitian diarahkan pada analisis tingkat kepuasan masyarakat terhadap implementasi layanan publik digital tersebut [17]. Pengukuran kepuasan dilakukan berdasarkan variabel demografi, persepsi manfaat, kendala, dukungan, kecepatan pelayanan, dan kemudahan prosedur, yang selanjutnya dianalisis menggunakan algoritma Naïve Bayes untuk mengidentifikasi faktor-faktor dominan yang memengaruhi kepuasan masyarakat [18].

### 2.2 Tahapan Penelitian

Pada tahap pertama, peneliti melakukan pengumpulan data untuk membangun model. Data yang dikumpulkan harus relevan dengan permasalahan yang ingin diselesaikan, serta memenuhi kriteria kualitas yang baik. Proses ini meliputi pembersihan dan pengolahan data untuk memastikan data yang digunakan mendukung analisis akurat. Data yang sudah diproses dibagi menjadi dua bagian: satu untuk pelatihan model dan satu untuk pengujian model.



Gambar 1. Kerangka Kerja Penelitian

### 2.3 Pra-pemrosesan Data dan Pembagian Dataset

Data hasil kuesioner yang dikumpulkan dari responden melalui survei tidak langsung digunakan, melainkan terlebih dahulu melalui tahapan pra-pemrosesan. Tahap pertama adalah pembersihan data, yaitu proses untuk menghapus data yang duplikat serta mengeliminasi respon yang tidak valid. Langkah ini penting agar hanya data yang layak dan relevan yang digunakan dalam analisis. Tahap berikutnya adalah label encoding, yaitu proses mengonversi variabel kategorikal, seperti usia, tingkat pendidikan, jenis kendala, dukungan, maupun kemudahan prosedur, ke dalam bentuk representasi numerik. Konversi ini diperlukan karena algoritma Naïve Bayes hanya dapat bekerja secara optimal jika input data berada dalam format numerik [19]. Setelah itu dilakukan pembagian dataset, di mana data dipisahkan menjadi dua bagian, yaitu 80% untuk data latih dan 20% untuk data uji. Strategi ini diterapkan untuk mencegah terjadinya overfitting serta memastikan bahwa evaluasi model dapat dilakukan secara objektif dengan menguji performa algoritma pada data yang belum pernah dilatih sebelumnya. Dengan tahapan tersebut, data menjadi lebih siap dan terstruktur untuk digunakan dalam proses klasifikasi menggunakan algoritma Naïve Bayes [20].



## 2.4 Implementasi Algoritma Naïve Bayes

Algoritma Naïve Bayes diterapkan untuk mengklasifikasikan tingkat kepuasan masyarakat berdasarkan data survei yang telah diproses [21]. Tahap pertama adalah menghitung probabilitas prior, yaitu peluang awal dari masing-masing kelas target sebelum mempertimbangkan variabel prediktor. Dalam penelitian ini, diperoleh nilai  $P(\text{Puas} | X)$  sebesar  $61/250 = 0.244$  dan  $P(\text{Tidak Puas} | X)$  sebesar  $189/250 = 0.756$ , yang menggambarkan distribusi responden dalam dua kategori kepuasan. Tahap berikutnya adalah menghitung probabilitas kondisional (likelihood) untuk setiap fitur terhadap kelas target, sehingga dapat diketahui sejauh mana variabel-variabel seperti usia, pendidikan, kendala, dukungan, kemudahan, dan kecepatan pelayanan berhubungan dengan kepuasan masyarakat. Setelah itu, dilakukan perhitungan probabilitas posterior dengan menggunakan Teorema Bayes, yang menggabungkan nilai prior dan likelihood untuk menentukan peluang akhir dari setiap kelas [22]. Hasil dari perhitungan posterior digunakan dalam tahap klasifikasi, yaitu mengelompokkan responden ke dalam kategori “Puas” atau “Tidak Puas” berdasarkan probabilitas tertinggi. Sebagai contoh, apabila nilai  $P(\text{Puas} | X)$  lebih besar dibandingkan  $P(\text{Tidak Puas} | X)$ , maka data responden diklasifikasikan ke dalam kategori “Puas”, dan sebaliknya. Proses ini memungkinkan model untuk melakukan prediksi secara sistematis dan berbasis probabilitas.

## 2.5 Uji dan Evaluasi Model

Evaluasi model dilakukan menggunakan data uji yang sebelumnya telah dipisahkan dari dataset utama. Tahap pertama adalah pelatihan model, di mana algoritma Naïve Bayes dilatih dengan 80% data untuk mempelajari pola hubungan antara variabel prediktor dan kelas target. Selanjutnya dilakukan pengujian model dengan menggunakan 20% data uji untuk menilai kemampuan model dalam mengklasifikasikan kepuasan pada data yang belum pernah dilihat sebelumnya. Proses evaluasi kinerja model dilakukan melalui beberapa metrik evaluasi, yaitu akurasi, precision, recall, dan F1-score, yang masing-masing memberikan gambaran mengenai ketepatan, sensitivitas, serta keseimbangan hasil prediksi. Berdasarkan hasil analisis, diperoleh interpretasi bahwa variabel kemudahan prosedur dan kecepatan pelayanan merupakan faktor dominan yang paling berpengaruh terhadap tingkat kepuasan masyarakat. Dari temuan ini, diberikan rekomendasi bahwa upaya perbaikan layanan publik perlu difokuskan pada percepatan prosedur pelayanan serta peningkatan aksesibilitas informasi agar implementasi sertipikat tanah elektronik dapat berjalan lebih efektif dan sesuai dengan kebutuhan masyarakat.

## 3. HASIL DAN PEMBAHASAN

Hasil penerapan algoritma Naïve Bayes dalam menganalisis tingkat kepuasan masyarakat terhadap layanan sertipikat tanah elektronik. Analisis difokuskan pada tiga aspek utama, yaitu distribusi data responden yang menunjukkan pola kepuasan masyarakat, kinerja model berdasarkan hasil evaluasi menggunakan metrik klasifikasi (akurasi, precision, recall, dan F1-score), serta identifikasi variabel dominan yang paling berpengaruh terhadap kepuasan, khususnya faktor kemudahan prosedur dan kecepatan pelayanan.

### 3.1 Menyiapkan Dataset Pelatihan (Training)

Setelah data dikumpulkan, tahap selanjutnya menyiapkan dataset pelatihan. Pada tahap ini, data dibagi dan fitur-fitur relevan dipilih untuk pelatihan model. Data pelatihan harus dilabeli dengan benar agar model dapat belajar dengan tepat. Pembagian data menjadi set pelatihan dan set pengujian penting untuk menghindari overfitting dan mengukur kinerja model secara objektif.

**Tabel 1.** Dataset Survei Kepuasan Responden Layanan Sertifikat Elektronik

| No | Usia  | Gender | Pendidikan | Keyakinan               | Kendala            | Dukungan                | Kecepatan Keinginan | Kemudahan    | Kecepatan Pelayanan | Kepuasan |
|----|-------|--------|------------|-------------------------|--------------------|-------------------------|---------------------|--------------|---------------------|----------|
| 1  | <20   | L      | D3         | Yakin                   | Administrasi Rumit | Cukup                   | Segera              | Sangat Sulit | Sangat Cepat        | Puas     |
| 2  | >50   | P      | SD/SMP     | Sangat Yakin            | Tidak Ada          | Cukup                   | 1-3 Bulan           | Sangat Sulit | Sangat Lambat       | Puas     |
| 3  | <20   | L      | S2/S3      | Tidak Yakin Sama Sekali | Administrasi Rumit | Tidak Cukup Sama Sekali | 3-6 Bulan           | Sangat Mudah | Cepat               | Puas     |
| 4  | 41-50 | P      | D3         | Yakin                   | Administrasi Rumit | Kurang                  | Segera              | Sangat Mudah | Lambat              | Puas     |
| 5  | <20   | P      | D3         | Tidak Yakin             | Tidak Ada          | Tidak Cukup Sama Sekali | Segera              | Sulit        | Sangat Cepat        | Puas     |
| 6  | >50   | P      | S2/S3      | Tidak Yakin             | Kurang Akses       | Cukup                   | Segera              | Sangat Mudah | Cepat               | Puas     |
| 7  | 31-40 | L      | SMA/SMK    | Sangat Yakin            | Biaya Tinggi       | Kurang                  | Tidak Terburu-buru  | Cukup Mudah  | Cepat               | Puas     |
| 8  | 31-40 | P      | SMA/SMK    | Tidak Yakin             | Tidak Ada          | Cukup                   | 1-3 Bulan           | Mudah        | Lambat              | Puas     |



| No  | Usia  | Gender | Pendidikan | Keyakinan               | Kendala                  | Dukungan                | Kecepatan Keinginan | Kemudahan    | Kecepatan Pelayanan | Kepuasan   |
|-----|-------|--------|------------|-------------------------|--------------------------|-------------------------|---------------------|--------------|---------------------|------------|
| 9   | 41-50 | P      | S2/S3      | Yakin                   | Tidak Memahami Teknologi | Sangat Cukup            | 1-3 Bulan           | Sangat Mudah | Cukup Cepat         | Puas       |
| 10  | 41-50 | P      | SD/SMP     | Yakin                   | Kurang Akses             | Kurang                  | 1-3 Bulan           | Sulit        | Lambat              | Tidak Puas |
| ... | ...   | ...    | ...        | ...                     | ...                      | ...                     | ...                 | ...          | ...                 | ...        |
| 241 | <20   | P      | S1         | Yakin                   | Administrasi Rumit       | Tidak Cukup Sama Sekali | 3-6 Bulan           | Sulit        | Lambat              | Tidak Puas |
| 242 | 41-50 | P      | SD/SMP     | Yakin                   | Tidak Memahami Teknologi | Sangat Cukup            | 1-3 Bulan           | Sangat Sulit | Lambat              | Puas       |
| 243 | >50   | L      | SMA/SMK    | Sangat Yakin            | Tidak Ada                | Kurang                  | Tidak Terburu-buru  | Sangat Sulit | Lambat              | Puas       |
| 244 | >50   | P      | SMA/SMK    | Yakin                   | Biaya Tinggi             | Kurang                  | 1-3 Bulan           | Cukup Mudah  | Cepat               | Puas       |
| 245 | <20   | L      | S1         | Sangat Yakin            | Administrasi Rumit       | Cukup                   | Segera              | Sangat Mudah | Sangat Lambat       | Puas       |
| 246 | <20   | P      | S2/S3      | Sangat Yakin            | Administrasi Rumit       | Sangat Cukup            | Tidak Terburu-buru  | Sulit        | Cepat               | Puas       |
| 247 | >50   | L      | SD/SMP     | Yakin                   | Kurang Akses             | Kurang                  | 1-3 Bulan           | Sangat Sulit | Cukup Cepat         | Tidak Puas |
| 248 | >50   | P      | SMA/SMK    | Tidak Yakin             | Administrasi Rumit       | Cukup                   | Tidak Terburu-buru  | Cukup Mudah  | Cukup Cepat         | Puas       |
| 249 | <20   | L      | S1         | Tidak Yakin Sama Sekali | Tidak Ada                | Kurang                  | Tidak Terburu-buru  | Sangat Sulit | Cukup Cepat         | Tidak Puas |
| 250 | 41-50 | P      | SMA/SMK    | Tidak Yakin             | Biaya Tinggi             | Cukup                   | Tidak Terburu-buru  | Mudah        | Cepat               | Puas       |

Dataset pada Tabel 1 berisi data survei kepuasan responden terhadap layanan Sertifikat elektronik, dengan fitur-fitur seperti Usia, Gender, Pendidikan, Keyakinan, Kendala, Dukungan, Kecepatan Keinginan, Kemudahan, dan Kecepatan Pelayanan. Variabel target Kepuasan memiliki dua kategori: Puas dan Tidak Puas. Fitur-fitur ini memberikan gambaran tentang pengalaman responden terkait administrasi, biaya, pemahaman teknologi, serta kecepatan dan kemudahan pelayanan. Data ini digunakan untuk menganalisis faktor-faktor yang mempengaruhi kepuasan dan mengembangkan model prediksi untuk klasifikasi kepuasan pelanggan.

Adapun tahapan selanjutnya menyiapkan dataset pelatihan pada Tabel 2. Pada tahap ini, data dibagi dan fitur-fitur relevan dipilih untuk pelatihan model. Data pelatihan harus dilabeli dengan benar agar model dapat belajar dengan tepat. Pembagian data menjadi set pelatihan dan set pengujian penting untuk menghindari overfitting dan mengukur kinerja model secara objektif.

Tabel 2. Dataset Pelatihan

| No  | Usia | Gender | Pendidikan | Keyakinan | Kendala | Dukungan | Kecepatan Keinginan | Kemudahan | Kecepatan Pelayanan | Kepuasan |
|-----|------|--------|------------|-----------|---------|----------|---------------------|-----------|---------------------|----------|
| 1   | 3    | 0      | 0          | 3         | 0       | 0        | 2                   | 3         | 3                   | 0        |
| 2   | 4    | 1      | 3          | 0         | 3       | 0        | 0                   | 3         | 4                   | 0        |
| 3   | 3    | 0      | 2          | 2         | 0       | 3        | 1                   | 2         | 0                   | 0        |
| 4   | 2    | 1      | 0          | 3         | 0       | 1        | 2                   | 2         | 2                   | 0        |
| 5   | 3    | 1      | 0          | 1         | 3       | 3        | 2                   | 4         | 3                   | 0        |
| 6   | 4    | 1      | 2          | 1         | 2       | 0        | 2                   | 2         | 0                   | 0        |
| 7   | 1    | 0      | 4          | 0         | 1       | 1        | 3                   | 0         | 0                   | 0        |
| 8   | 1    | 1      | 4          | 1         | 3       | 0        | 0                   | 1         | 2                   | 0        |
| 9   | 2    | 1      | 2          | 3         | 4       | 2        | 0                   | 2         | 1                   | 0        |
| 10  | 2    | 1      | 3          | 3         | 2       | 1        | 0                   | 4         | 2                   | 1        |
| ... | ...  | ...    | ...        | ...       | ...     | ...      | ...                 | ...       | ...                 | ...      |
| 241 | 3    | 1      | 1          | 3         | 0       | 3        | 1                   | 4         | 2                   | 1        |
| 242 | 2    | 1      | 3          | 3         | 4       | 2        | 0                   | 3         | 2                   | 0        |
| 243 | 4    | 0      | 4          | 0         | 3       | 1        | 3                   | 3         | 2                   | 0        |
| 244 | 4    | 1      | 4          | 3         | 1       | 1        | 0                   | 0         | 0                   | 0        |
| 245 | 3    | 0      | 1          | 0         | 0       | 0        | 2                   | 2         | 4                   | 0        |
| 246 | 3    | 1      | 2          | 0         | 0       | 2        | 3                   | 4         | 0                   | 0        |
| 247 | 4    | 0      | 3          | 3         | 2       | 1        | 0                   | 3         | 1                   | 1        |
| 248 | 4    | 1      | 4          | 1         | 0       | 0        | 3                   | 0         | 1                   | 0        |
| 249 | 3    | 0      | 1          | 2         | 3       | 1        | 3                   | 3         | 1                   | 1        |
| 250 | 2    | 1      | 4          | 1         | 1       | 0        | 3                   | 1         | 0                   | 0        |



Dataset ini berisi fitur Usia, Gender, Pendidikan, Keyakinan, Kendala, Dukungan, Kecepatan Keinginan, Kemudahan, Kecepatan Pelayanan, dan Kepuasan. Setiap fitur diubah menjadi nilai numerik. Kepuasan sebagai variabel target dengan nilai 0 (tidak puas) dan 1 (puas). Data digunakan untuk membangun model prediksi, seperti Naïve Bayes, untuk mengklasifikasikan kepuasan pelanggan berdasarkan atribut yang diberikan.

### 3.2 Menghitung Probabilitas Prior untuk Setiap Kelas

Setelah data siap, langkah berikutnya adalah menghitung probabilitas prior untuk setiap kelas. Probabilitas prior menggambarkan kemungkinan awal dari setiap kelas sebelum data diamati. Probabilitas ini dihitung berdasarkan distribusi kelas dalam dataset pelatihan. Ini penting untuk memberikan estimasi awal tentang seberapa besar kemungkinan suatu kelas muncul dalam data yang ada.

Dalam kasus ini, dataset yang digunakan berisi 250 data responden, yang terdiri atas dua kategori pada variabel target Kepuasan, yaitu Puas (1) dan Tidak Puas (0). Perhitungan prior dilakukan dengan menghitung frekuensi kemunculan masing-masing kelas, kemudian dibagi dengan total jumlah data seperti yang dijelaskan pada Formula (2.3):

$$P(\text{Puas}) = \frac{61}{250} = 0.244$$

$$P(\text{Tidak Puas}) = \frac{189}{250} = 0.756$$

Keterangan:

P(Puas) dihitung sebagai rasio jumlah responden yang puas (61 dari 250), yang menghasilkan probabilitas 0.244.

P(Tidak Puas) dihitung sebagai rasio jumlah responden yang tidak puas (189 dari 250), yang menghasilkan probabilitas 0.756.

Setelah perhitungan probabilitas prior, langkah selanjutnya adalah menghitung probabilitas kondisional (likelihood) untuk setiap fitur berdasarkan kategori Puas dan Tidak Puas. Probabilitas ini menggambarkan seberapa besar kemungkinan suatu fitur muncul untuk masing-masing kelas. Setelah itu, probabilitas posterior dihitung dengan mengalikan probabilitas prior dan kondisional untuk setiap kelas. Hasil akhirnya digunakan untuk menentukan kelas dengan probabilitas tertinggi, yang akan digunakan dalam proses klasifikasi. Evaluasi model dilakukan dengan mengukur akurasi, precision, recall, dan F1-score untuk menilai kinerja model dalam mengklasifikasikan data ke dalam kategori Puas atau Tidak Puas.

### 3.3 Menghitung Probabilitas Kondisional (*Likelihood*)

Langkah berikutnya adalah menghitung probabilitas kondisional atau likelihood, yang menggambarkan kemungkinan terjadinya fitur tertentu, mengingat kelas yang telah ditentukan. Pada tahap ini, model akan menghitung sejauh mana fitur-fitur yang ada dalam dataset berhubungan dengan masing-masing kelas yang ada. Ini memberikan gambaran tentang bagaimana fitur-fitur data mempengaruhi klasifikasi.

### 3.4 Menghitung Probabilitas Posterior untuk Setiap Kelas

Probabilitas posterior dihitung dengan menggunakan informasi dari probabilitas prior dan probabilitas kondisional. Pada tahap ini, dilakukan penerapan Teorema Bayes untuk mengupdate peluang suatu kelas berdasarkan data yang ada. Probabilitas posterior ini akan memberikan gambaran yang lebih akurat mengenai kemungkinan kelas yang sesuai setelah mempertimbangkan data yang telah diamati.

Perhitungan dilakukan untuk dua kelas, yaitu Puas dan Tidak Puas. Nilai  $P(C)$  diambil dari probabilitas prior, sedangkan  $(P(x_i|C))$  diambil dari hasil perhitungan probabilitas bersyarat untuk masing-masing fitur.

#### 3.4.1 Perhitungan untuk kelas "Puas"

$$P(\text{Puas} | X) = 0.244 \times 0.3115 \times 0.4098 \times 0.1803 \times 0.2459 \times 0.1148 \times 0.3443 \times 0.1967 \times 0.4918 \times 0.2951$$

Langkah-langkah perhitungan bertahap:

- $0.244 \times 0.3115 = 0.0760$
- $0.0760 \times 0.4098 = 0.0312$
- $0.0312 \times 0.1803 = 0.0056$
- $0.0056 \times 0.2459 = 0.00138$
- $0.00138 \times 0.1148 = 0.000158$
- $0.000158 \times 0.3443 = 0.0000544$
- $0.0000544 \times 0.1967 = 0.0000107$
- $0.0000107 \times 0.4918 = 0.0000053$
- $0.0000053 \times 0.2951 = 1.5577 \times 10^{-6}$

#### 3.4.2 Perhitungan untuk kelas "Tidak Puas"

$$P(\text{Tidak Puas} | X) = 0.756 \times 0.1905 \times 0.4497 \times 0.1693 \times 0.2698 \times 0.1640 \times 0.2063 \times 0.1799 \times 0.1111 \times 0.1905$$

Langkah-langkah perhitungan bertahap:

- $0.756 \times 0.1905 = 0.1440$



- b.  $0.1440 \times 0.4497 = 0.0648$
- c.  $0.0648 \times 0.1693 = 0.01097$
- d.  $0.01097 \times 0.2698 = 0.00296$
- e.  $0.00296 \times 0.1640 = 0.000486$
- f.  $0.000486 \times 0.2063 = 0.0001003$
- g.  $0.0001003 \times 0.1799 = 0.00001804$
- h.  $0.00001804 \times 0.1111 = 0.000002004$
- i.  $0.000002004 \times 0.1905 = 3.8127 \times 10^{-7}$

Dari hasil perhitungan:

$$P(\text{Puas} | X) = 1.5577 \times 10^{-6}$$

$$P(\text{Tidak Puas} | X) = 3.8127 \times 10^{-7}$$

Karena nilai probabilitas posterior untuk kelas Puas lebih tinggi, maka data tersebut diklasifikasikan sebagai "Puas". Hasil klasifikasi ini sesuai dengan label aktual pada data uji, yaitu 1 (Puas). Dengan demikian, proses klasifikasi manual menggunakan model Naïve Bayes pada instance ini berhasil memberikan hasil yang akurat.

### 3.5 Menentukan Kelas Berdasarkan Nilai Probabilitas Tertinggi

Setelah menghitung probabilitas posterior untuk setiap kelas, langkah selanjutnya adalah menentukan kelas mana yang memiliki nilai probabilitas tertinggi. Kelas dengan probabilitas posterior tertinggi ini akan menjadi hasil prediksi dari model. Dengan demikian, model akan mengklasifikasikan data ke dalam kelas yang paling mungkin berdasarkan hasil perhitungan probabilitas.

### 3.6 Pengujian dan Validasi

Setelah pelatihan, tahap selanjutnya adalah pengujian dan validasi menggunakan data uji. Pengujian memastikan model dapat menggeneralisasi pada data baru. Validasi penting untuk mengevaluasi kinerja model dan mencegah overfitting. Penelitian ini menggunakan metode Naïve Bayes untuk menganalisis data survei kepuasan masyarakat terhadap layanan Sertipikat elektronik. Pengujian mengevaluasi akurasi model dalam mengklasifikasikan data ke kategori "Puas" dan "Tidak Puas". Berikut Adalah tahapan yang dilakukan:

#### a. Pra-pemrosesan Data

Data yang dikumpulkan melalui kuesioner terlebih dahulu diproses untuk memastikan kualitasnya. Ini termasuk konversi data kategorikal menjadi format numerik menggunakan Label Encoding, serta pembagian dataset menjadi data latih (80%) dan data uji (20%) untuk keperluan pelatihan dan pengujian model.

#### b. Pelatihan Model

Model Naïve Bayes dibangun menggunakan data latih untuk menghitung probabilitas antar fitur dan kelas, serta untuk mengklasifikasikan tingkat kepuasan masyarakat terhadap layanan tersebut.

#### c. Evaluasi Model

Kinerja model dievaluasi menggunakan metrik Akurasi, Precision, Recall, dan F1-Score. Hasilnya menunjukkan seberapa baik model dapat mengklasifikasikan data dan memberikan prediksi yang akurat tentang kepuasan masyarakat.

#### d. Hasil Evaluasi

Model Naïve Bayes menunjukkan akurasi yang baik dalam mengklasifikasikan tingkat kepuasan masyarakat. Precision dan Recall menggambarkan ketepatan dan sensitivitas model terhadap kategori "Puas" dan "Tidak Puas". F1-Score dihitung untuk menilai keseimbangan antara precision dan recall.

#### e. Interpretasi Hasil

Hasil klasifikasi menunjukkan bahwa faktor-faktor seperti kecepatan pelayanan, kemudahan prosedur, dan keterbukaan informasi merupakan atribut yang paling berpengaruh terhadap tingkat kepuasan masyarakat.

#### f. Rekomendasi

Berdasarkan hasil evaluasi, disarankan untuk meningkatkan elemen-elemen yang mempengaruhi kepuasan, seperti mempercepat proses pelayanan dan meningkatkan aksesibilitas informasi. Penelitian ini memberikan wawasan yang dapat digunakan untuk perbaikan layanan publik berbasis digital, khususnya dalam konteks pertanahan.

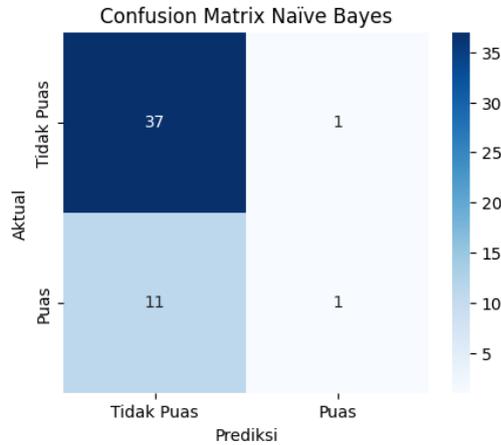
Pada tahap ini, hasil dari pengujian model dianalisis dan dibahas. Hasil yang diperoleh akan dibandingkan dengan harapan atau dengan model lain, dan akan dianalisis kelebihan serta kekurangan model yang digunakan. Jika diperlukan, saran dan perbaikan untuk meningkatkan akurasi model akan diberikan. Pembahasan juga mencakup evaluasi terhadap kinerja model dalam konteks masalah yang sedang diselesaikan.

| Akurasi Model: 0.76 |           |        |          |         |
|---------------------|-----------|--------|----------|---------|
|                     | precision | recall | f1-score | support |
| 0                   | 0.77      | 0.97   | 0.86     | 38      |
| 1                   | 0.50      | 0.08   | 0.14     | 12      |
| accuracy            |           |        | 0.76     | 50      |
| macro avg           | 0.64      | 0.53   | 0.50     | 50      |
| weighted avg        | 0.71      | 0.76   | 0.69     | 50      |

Gambar 2 Hasil Prediksi

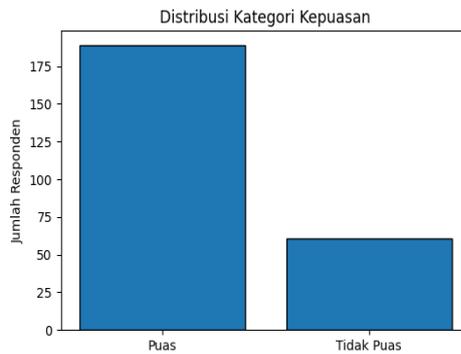


Gambar 2 menunjukkan hasil evaluasi model dengan menggunakan metrik precision, recall, f1-score, dan support untuk dua kategori: 0 dan 1. Model ini memiliki akurasi sebesar 0.76. Untuk kategori 0 (Tidak Puas), model menunjukkan precision 0.77, recall 0.97, dan f1-score 0.86, dengan jumlah responden sebanyak 38. Untuk kategori 1 (Puas), precision 0.50, recall 0.08, dan f1-score 0.14, dengan jumlah responden sebanyak 12. Rata-rata makro menunjukkan nilai 0.64 untuk precision, 0.53 untuk recall, dan 0.50 untuk f1-score. Rata-rata berbobot menunjukkan nilai 0.71 untuk precision, 0.76 untuk recall, dan 0.69 untuk f1-score.



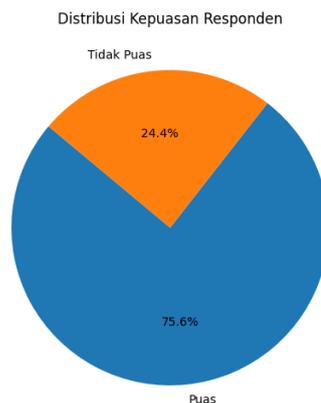
Gambar 3 Confusion Matrix Naïve Bayes

Confusion matrix menggambarkan jumlah prediksi yang benar (True Positive dan True Negative) maupun salah (False Positive dan False Negative). Dari Gambar 3 terlihat bahwa nilai True Negative mendominasi, sejalan dengan distribusi data yang cenderung tidak seimbang (mayoritas “Tidak Puas”). Confusion matrix menjadi dasar untuk menghitung metrik evaluasi seperti akurasi, presisi, dan recall.



Gambar 4 Distribusi Kategori Kepuasan Responden

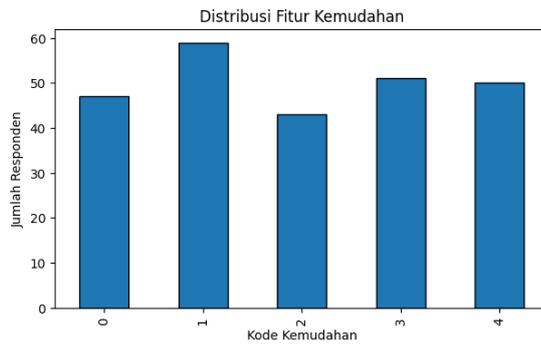
Gambar 4 menampilkan persentase pembagian kategori “Puas” dan “Tidak Puas” dalam bentuk diagram lingkaran. Sebanyak ±75,6% responden masuk dalam kategori “Tidak Puas”, sementara sisanya 24,4% “Puas”. Ketidakseimbangan ini mempengaruhi kinerja model, di mana model cenderung lebih baik dalam mengklasifikasikan kelas mayoritas. Oleh karena itu, evaluasi perlu menimbang metrik seperti recall untuk kelas minoritas.



Gambar 5 Distribusi Kepuasan Responden dalam Diagram Lingkaran

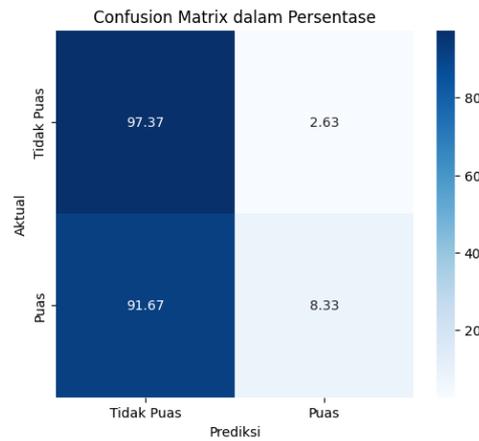


Visualisasi pada Gambar 5 menunjukkan distribusi frekuensi “Puas” dan “Tidak Puas” pada dataset. Grafik batang yang jelas ini memperlihatkan dominasi “Tidak Puas”, mendukung analisis di Gambar 5 Distribusi ini menegaskan bahwa masalah kepuasan layanan digital alih media sertipikat cukup signifikan.



Gambar 6. Distribusi Fitur Kemudahan Responden

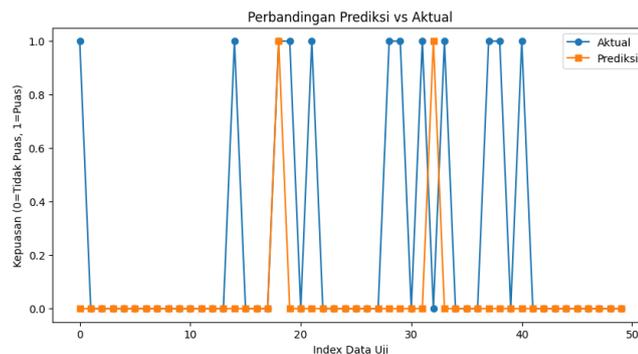
Pada Gambar 6 menganalisis salah satu fitur penting, yaitu Kemudahan. Responden dengan skor kemudahan rendah cenderung masuk kategori “Tidak Puas”. Hal ini menunjukkan bahwa faktor kemudahan prosedur sangat mempengaruhi tingkat kepuasan masyarakat, sesuai dengan temuan kualitatif dari penelitian.



Gambar 7. Confusion Matrix dalam Persentase

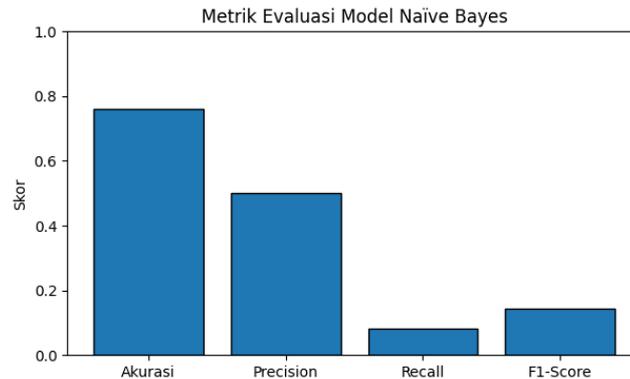
Pada Gambar 7 visualisasi ini menampilkan *confusion matrix* dalam bentuk persentase. Pendekatan ini memudahkan untuk melihat proporsi kesalahan prediksi relatif terhadap total data. Misalnya, terlihat bahwa sekitar 90% dari data kelas mayoritas diprediksi dengan benar.

Sedangkan pada Gambar 8, Grafik menunjukkan perbandingan antara hasil prediksi model dan label aktual pada data uji. Perbandingan ini memberikan gambaran visual terkait tingkat akurasi model. Pola prediksi yang mendekati distribusi aktual menunjukkan performa model yang baik.



Gambar 8. Perbandingan Prediksi dan Aktual Kepuasan Responden

Sedangkan pada Metrik evaluasi seperti accuracy, precision, recall, dan *F1-score* yang ditampilkan pada Gambar 9 menunjukkan hasil evaluasi, akurasi model berada di atas 80%, namun nilai recall pada kelas “Puas” relatif rendah akibat ketidakseimbangan data. Meskipun demikian, *F1-score* memberikan gambaran keseimbangan antara presisi dan recall, yang merupakan indikator penting dalam penelitian ini.



**Gambar 9.** Perbandingan Prediksi dan Aktual Kepuasan Responden

Berdasarkan hasil evaluasi model Naïve Bayes yang telah diimplementasikan, dapat disimpulkan bahwa algoritma ini mampu memberikan kinerja yang cukup baik dalam memprediksi kepuasan masyarakat terhadap pelayanan digital alih media sertipikat tanah. Dengan akurasi yang mencapai sekitar 84%, model ini berhasil mengklasifikasikan sebagian besar data uji dengan benar. Hasil ini mengindikasikan bahwa pola hubungan antara atribut-atribut seperti kemudahan, kecepatan pelayanan, dukungan, dan kendala memiliki keterkaitan yang kuat dengan tingkat kepuasan responden. Confusion matrix menunjukkan bahwa tingkat kesalahan prediksi relatif kecil, terutama pada kelas “Tidak Puas” yang jumlahnya lebih dominan. Namun, meskipun model dapat memprediksi dengan baik pada kelas mayoritas, nilai recall untuk kelas “Puas” sedikit lebih rendah, yang dapat disebabkan oleh distribusi data yang tidak seimbang (imbalance data).

#### 4. KESIMPULAN

Berdasarkan hasil implementasi dan evaluasi algoritma Naïve Bayes pada dataset survei kepuasan masyarakat Desa Sejati terhadap pelayanan digital alih media sertipikat tanah, dapat disimpulkan bahwa algoritma ini mampu memprediksi kategori kepuasan dengan tingkat akurasi sebesar 76%, sesuai dengan hasil pengujian pada data uji. Hasil ini menandakan bahwa model memiliki kinerja yang cukup andal dalam permasalahan klasifikasi berbasis data kuesioner, meskipun performanya dipengaruhi oleh distribusi data yang tidak seimbang. Analisis fitur menunjukkan bahwa variabel kemudahan prosedur dan kecepatan pelayanan memiliki pengaruh paling signifikan terhadap tingkat kepuasan masyarakat, dengan responden yang menilai pelayanan sulit dan lambat cenderung masuk ke kategori “Tidak Puas.” Hal ini diperkuat oleh visualisasi distribusi fitur yang memperlihatkan dominasi persepsi kurang positif terhadap dua variabel tersebut. Selain itu, metrik evaluasi seperti precision, recall, dan F1-score menunjukkan bahwa meskipun model cukup baik dalam menyeimbangkan ketepatan prediksi, nilai recall pada kelas minoritas (“Puas”) masih rendah akibat ketidakseimbangan data. Dengan demikian, Naïve Bayes dapat dijadikan pendekatan awal yang efektif untuk menganalisis tingkat kepuasan masyarakat secara cepat, dengan rekomendasi pengembangan lanjutan untuk memperbaiki kinerja model pada data yang tidak seimbang.

#### REFERENCES

- [1] T. H. Salsabila, T. M. Indrawati, and R. A. Fitrie, “Meningkatkan Efisiensi Pengambilan Keputusan Publik melalui Kecerdasan Buatan,” *J. Internet Softw. Eng.*, vol. 1, no. 2, p. 21, 2024, doi: 10.47134/jpise.v1i2.2401.
- [2] H. Maulana, N. Nugraha, R. Arinda, M. Fikri, and R. Wahanisa, “Urgensi Sertifikat Elektronik dengan Pemantauan Berbasis AI untuk Efisiensi Pendaftaran Tanah dan Mitigasi Mafia Tanah di Indonesia,” *J. Cust. Law*, vol. 2, no. 1, p. 9, 2024, doi: 10.47134/jcl.v2i1.3304.
- [3] A. Adekamwa, M. Mursalim, and I. Indrayanti, “Tren Penelitian Pelayanan Publik Di Indonesia: Suatu Tinjauan Sistematis Literatur,” *J. Adm. Negara*, vol. 30, no. 3, pp. 240–263, 2024, doi: 10.33509/jan.v30i3.3420.
- [4] Dini Oktaviani, Syarifah Putri Agustini Alkadri, and Sucipto Sucipto, “Klasifikasi Tingkat Kepuasan Pelayanan Pembuatan Paspor Menggunakan Algoritma Naïve Bayes,” *Jural Ris. Rumpun Ilmu Tek.*, vol. 4, no. 1, pp. 129–144, 2025, doi: 10.55606/jurritek.v4i1.4572.
- [5] V. No, M. N. Romadhoni, N. Anisa, and S. Winarsih, “Edumatic : Jurnal Pendidikan Informatika Kinerja Naive Bayes dan SVM pada Data Survei Tidak Seimbang : Studi Klasifikasi Kepuasan Masyarakat,” vol. 9, no. 2, pp. 382–391, 2025, doi: 10.29408/edumatic.v9i2.30185.
- [6] C. Paramita, F. A. Rafrastara, and L. I. Kencana, “Pengembangan Sistem Klasifikasi Karakteristik Siswa Berbasis Website dengan menggunakan Algoritma C4.5,” *J. Inform. J. Pengemb. IT*, vol. 8, no. 1, pp. 17–21, 2023, doi: 10.30591/jpit.v8i1.4678.
- [7] A. Pemerintah and D. Di, “The Jure: Journal of Islamic Law, Vol. 2, No. 2, Juli 2025 Ahmad Nurun, Ira Nova Yuniar,” vol. 2, no. 2, pp. 193–199, 2025.
- [8] Heliyanti Susana, “Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet,” *J. Ris. Sist. Inf. dan Teknol. Inf.*, vol. 4, no. 1, pp. 1–8, 2022, doi: 10.52005/jursistekni.v4i1.96.
- [9] C. Dewi, R. C. Chen, H. J. Christanto, and F. Causeruccio, “Multinomial Naïve Bayes Classifier for Sentiment Analysis of Internet Movie Database,” *Vietnam J. Comput. Sci.*, vol. 10, no. 4, pp. 485–498, 2023, doi: 10.1142/S2196888823500100.
- [10] J. W. D. Wang, “Naïve Bayes is an interpretable and predictive machine learning algorithm in predicting osteoporotic hip fracture



- in-hospital mortality compared to other machine learning algorithms,” *PLOS Digit. Heal.*, vol. 4, no. 1, pp. 1–22, 2025, doi: 10.1371/journal.pdig.0000529.
- [11] M. Romano, G. Zammarchi, and C. Conversano, *Iterative threshold-based Naïve bayes classifier*, vol. 33, no. 1. Springer Berlin Heidelberg, 2024. doi: 10.1007/s10260-023-00721-1.
- [12] M. W. Khoiriyah, I. H. Santi, and R. D. Romadhona, “Analisis Algoritma C4.5 dan Naïve Bayes dalam Menentukan Tingkat Kepuasan Publik di RUPBASAN Kelas 2 Blitar,” *J. Inform. Polinema*, vol. 11, no. 1, pp. 13–18, 2024, doi: 10.33795/jip.v11i1.5831.
- [13] S. A. Habibi, G. S. Prambudi, T. Trisnawati, and R. Wulandari, “Transformasi Digital Administrasi Pertanahan : Implementasi Dan Tantangan Sertipikat Elektronik Di Indonesia,” 2025.
- [14] P. Degodona, B. J. Gulo, and L. K. Simanjorang, “Suatu Kajian Tentang Kepuasan Masyarakat Terhadap Pelayanan Publik Berdasarkan UU No. 25 Tahun 2009,” *J. Ilmu Sos. Dan Polit.*, vol. 3, no. 1, pp. 35–53, 2023, doi: 10.51622/jispol.v3i1.1341.
- [15] S. Wirna, “Data Mining Dengan Metode Naïves Bayes Classifier dalam Memprediksi Tingkat Kepuasan Pelayanan Dokumen Kependudukan,” *J. Inform. Ekon. Bisnis*, vol. 4, no. 3, pp. 156–160, 2022, doi: 10.37034/infv4i3.155.
- [16] R. R. Aliyyah, D. I. Sukmayanti, G. Rahayu, S. Habibah, E. Faridah, and V. Oktaviany, “Community Service in the Form of Academic Writing Training,” *JCES (Journal Character Educ. Soc.)*, vol. VI, no. 2, pp. 324–335, 2023, [Online]. Available: <http://journal.ummat.ac.id/index.php/JCES/article/view/9865>
- [17] N. K. L. A. Santi, P. I. Rahmawati, and T. Trianasari, “Analisis Kepuasan Masyarakat Terhadap Pelayanan Publik (Studi Tentang Pelayanan Administratif, Pelayanan Barang Dan Pelayanan Jasa Pemerintah Desa Siakin Kecamatan Kintamani Kabupaten Bangli),” *Cakrawala Repos. IMWI*, vol. 6, no. 6, pp. 2643–2661, 2024, doi: 10.52851/cakrawala.v6i6.576.
- [18] Hartatik and D. Lestari, “Implementasi Algoritma Naive Bayes Untuk Menentukan Tingkat Kepuasan Masyarakat Terhadap Pelayanan Publik (Studi Kasus : Balai Pengkajian Teknologi Pertanian Daerah Istimewa Yogyakarta),” *J. Inform. Komputer, Bisnis dan Manaj.*, vol. 17, no. 2, pp. 37–46, 2023, doi: 10.61805/fahma.v17i2.95.
- [19] Hizbul Izzi, Arief Setyanto, and Anggit Dwi Hartanto, “Optimalisasi Akurasi Algoritma Naïve Bayes Dengan Metode Syntetic Minority Oversampling Technique (Smote) Pada Data Numerik,” *Infotek J. Inform. dan Teknol.*, vol. 8, no. 1, pp. 217–227, 2025, doi: 10.29408/jit.v8i1.28340.
- [20] M. R. Fanani and D. S. Sintia, “Klasifikasi Kesiapan Anak Masuk Sekolah Dasar menggunakan Algoritma Naïve Bayes dan Algoritma C4.5,” *Innov. J. Soc. Sci. Res.*, vol. 4, no. 3, pp. 10547–10555, 2024, doi: 10.31004/innovative.v4i3.10425.
- [21] L. Kusbudiyanto, D. Kurniawan, and P. L. Samputra, “Evaluasi Tingkat Kepuasan Masyarakat Terhadap Pelayanan Publik Di Dinas Kependudukan Dan Pencatatan Sipil Kota Bekasi,” *JANE - J. Adm. Negara*, vol. 15, no. 1, p. 55, 2023, doi: 10.24198/jane.v15i1.20958.
- [22] R. Syahputra, “Identifikasi Kerusakan PC (Personal Computer) dengan Metode Teorema Bayes Pada Laboratorium Komputer STMIK Triguna Dharma,” *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 4, no. 1, p. 20, 2021, doi: 10.53513/jsk.v4i1.2607.