

KLASIFIKASI PENENTUAN TINGKAT KESEJAHTERAAN KELUARGA MENGUNAKAN METODE NAIVE BAYES PADA KELURAHAN PEMATANG KANDIS

Hawari Alhaq¹, Widja Yanto², Muhammad Akbar Dwiyantra³

Program Studi Teknologi Informasi Universitas Merangin^{1,3}
Program Studi Sistem Informasi Universitas Merangin²
Studi Teknologi Informasi Universitas Merangin

¹Corresponding Author: arialhaq@gmail.com

ABSTRAK

Family welfare is a crucial factor in determining the social and economic development of a region. This study aims to classify the level of family welfare in Pematang Kandis Subdistrict using machine learning model of the Naïve Bayes method, a probability-based algorithm that is effective in data classification. The data used in this study were obtained through surveys covering various economic, social, and demographic factors, such as income, education level, type of occupation, housing conditions, and the number of dependents. The research process includes data collection, preprocessing, training the Naïve Bayes model, and evaluating the model's performance using accuracy, precision, recall, and F1-score metrics. The results indicate that the Naïve Bayes method achieves a high level of accuracy in classifying family welfare levels. Therefore, this method can serve as a supporting tool in formulating social policies by local governments. Through this approach, the government is expected to identify families in need of assistance and design more targeted welfare programs.

Keywords: Family Welfare, Classification, Naïve Bayes, Machine Learning, Pematang Kandis.

ABSTRAK

Kesejahteraan keluarga merupakan faktor krusial dalam menentukan kemajuan sosial dan ekonomi suatu daerah. Penelitian ini bertujuan untuk mengelompokkan tingkat kesejahteraan keluarga di Kelurahan Pematang Kandis dengan menerapkan pemodelan pembelajaran mesin dengan metode Naïve Bayes, yaitu algoritma berbasis probabilitas yang efektif dalam proses klasifikasi data. Data dalam penelitian ini dikumpulkan melalui survei yang mencakup berbagai aspek ekonomi, sosial, dan demografi, seperti pendapatan, jenjang pendidikan, jenis pekerjaan, kondisi tempat tinggal, serta jumlah tanggungan keluarga. Tahapan penelitian mencakup pengumpulan data, tahap preprocessing, pelatihan model Naïve Bayes, serta evaluasi kinerja model menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil analisis menunjukkan bahwa metode Naïve Bayes mampu mencapai tingkat akurasi yang tinggi dalam menentukan kategori kesejahteraan keluarga. Oleh karena itu, metode ini dapat digunakan sebagai alat pendukung dalam penyusunan kebijakan sosial oleh pemerintah daerah. Dengan penerapan pendekatan ini, pemerintah diharapkan dapat mengidentifikasi keluarga yang membutuhkan bantuan serta merancang program kesejahteraan yang lebih tepat sasaran.

Kata Kunci: Kesejahteraan Keluarga, Klasifikasi, Naïve Bayes, Pembelajaran Mesin, Pematang Kandis.

PENDAHULUAN

Kesejahteraan keluarga merupakan faktor utama dalam mendukung pembangunan sosial dan ekonomi suatu daerah. Tingkat kesejahteraan yang tinggi mencerminkan stabilitas ekonomi, akses pendidikan yang memadai, serta kualitas hidup yang layak bagi setiap anggota keluarga. Sebaliknya, kesejahteraan yang rendah dapat menjadi indikasi berbagai permasalahan sosial, seperti kemiskinan, keterbatasan layanan kesehatan, serta rendahnya tingkat pendidikan dan kesempatan kerja. Oleh karena itu, diperlukan metode yang sistematis untuk mengidentifikasi tingkat kesejahteraan keluarga agar kebijakan yang diterapkan pemerintah lebih efektif dan tepat sasaran.

Dalam beberapa tahun terakhir, penggunaan metode berbasis data mining semakin berkembang dalam analisis sosial dan ekonomi, termasuk untuk klasifikasi tingkat kesejahteraan keluarga. Salah satu teknik yang efektif dalam klasifikasi adalah Naïve Bayes, sebuah algoritma berbasis probabilitas yang mampu mengolah dataset berukuran besar dengan efisien. Metode ini memanfaatkan berbagai variabel, seperti pendapatan, tingkat pendidikan, jenis pekerjaan, kondisi tempat tinggal, serta jumlah tanggungan, untuk menentukan kategori kesejahteraan keluarga secara akurat.

Penelitian ini bertujuan untuk menerapkan metode Naïve Bayes dalam mengklasifikasikan tingkat kesejahteraan keluarga di Kelurahan Pematang Kandis. Dengan menggunakan data hasil survei masyarakat, model ini diharapkan dapat membantu pemerintah dalam mengidentifikasi keluarga yang membutuhkan bantuan serta merancang program kesejahteraan yang lebih optimal. Selain itu, penelitian ini juga mengevaluasi kinerja model berdasarkan metrik akurasi, presisi, recall, dan F1-score untuk memastikan keandalan metode yang digunakan.

METODE PENELITIAN

Penelitian ini memanfaatkan data primer yang dikumpulkan melalui survei langsung terhadap keluarga di Kelurahan Pematang Kandis. Pengumpulan data dilakukan menggunakan kuesioner yang mencakup berbagai aspek, termasuk sosial, ekonomi, dan demografi. Beberapa variabel utama yang menjadi fokus dalam penelitian ini meliputi:

1. Aspek Ekonomi: Tingkat pendapatan keluarga, sumber utama penghasilan, serta kepemilikan aset.
2. Aspek Sosial: Pendidikan kepala keluarga, akses terhadap layanan kesehatan, dan

jumlah tanggungan.

3. Aspek Demografi: Jenis pekerjaan, kondisi tempat tinggal, serta jumlah anggota keluarga.

Selain data primer, penelitian ini juga mengandalkan data sekunder yang diperoleh dari instansi pemerintah setempat, seperti data statistik kependudukan dan laporan kesejahteraan masyarakat.

Penelitian ini menggunakan metode Naïve Bayes untuk mengklasifikasikan tingkat kesejahteraan keluarga di Kelurahan Pematang Kandis. Adapun tahapan dalam penelitian ini meliputi pengumpulan data, preprocessing data, pelatihan model, serta evaluasi model.

1. Pengumpulan Data

Data penelitian diperoleh melalui survei yang mencakup faktor-faktor ekonomi, sosial, dan demografi. Beberapa variabel utama yang dikumpulkan antara lain pendapatan keluarga, tingkat pendidikan kepala keluarga, jenis pekerjaan, kondisi tempat tinggal, serta jumlah tanggungan. Data ini dikategorikan berdasarkan tingkat kesejahteraan keluarga untuk digunakan dalam proses klasifikasi.

2. Preprocessing Data

Sebelum dilakukan pemodelan, data yang telah dikumpulkan mengalami proses preprocessing yang mencakup pembersihan data (cleaning), transformasi data, serta pemilihan fitur yang relevan. Tahapan ini bertujuan untuk meningkatkan kualitas data agar lebih siap untuk dianalisis dan dimodelkan menggunakan algoritma Naïve Bayes.

3. Pelatihan Model Naïve Bayes

Model Naïve Bayes dilatih menggunakan dataset yang telah diproses sebelumnya. Algoritma ini bekerja dengan menghitung probabilitas setiap kelas berdasarkan variabel yang tersedia, sehingga dapat memprediksi kategori kesejahteraan keluarga berdasarkan input data baru.

4. Evaluasi Model

Performa model dievaluasi menggunakan beberapa metrik seperti akurasi, presisi, recall, dan F1-score. Evaluasi ini bertujuan untuk mengukur sejauh mana model dapat mengklasifikasikan tingkat kesejahteraan keluarga dengan baik.

Melalui pendekatan ini, diharapkan hasil klasifikasi yang diperoleh dapat digunakan sebagai dasar dalam perumusan kebijakan sosial dan program kesejahteraan

yang lebih tepat sasaran.

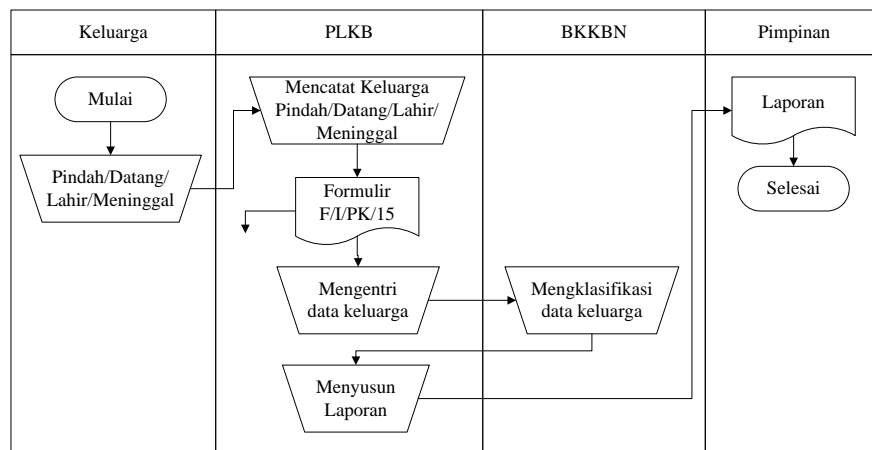
HASIL DAN PEMBAHASAN

Analisa Sistem Yang Sedang Berjalan

Untuk mengetahui dan mempelajari sistem yang ada diperlukan suatu penggambaran aliran-aliran informasi dari bagian-bagian yang terkait baik dari dalam maupun luar sistem. Proses yang berjalan selama ini, yaitu :

1. PLKB (Petugas Lapangan Keluarga Berencana) Kecamatan Pasar Jambi bekerja sama dengan kader KB yang ada disetiap Kelurahan Sekecamatan Pasar Jambi melakukan pendataan keluarga dilakukan setiap ada mutasi (pindah/lahir/meninggal) penduduk di Kecamatan Pasar Jambi.
2. Pencatatan dilakukan pada formulir yang sudah tersedia, yaitu F/I/PK.15.
3. Hasil pendataan selanjutnya di entry oleh operator di Balai KB.
4. Selanjutnya PLKB membuat laporan pemutahiran basis data keluarga pada Kecamatan Pasar Jambi. Pelaporan dibuat berkala setiap bulannya.

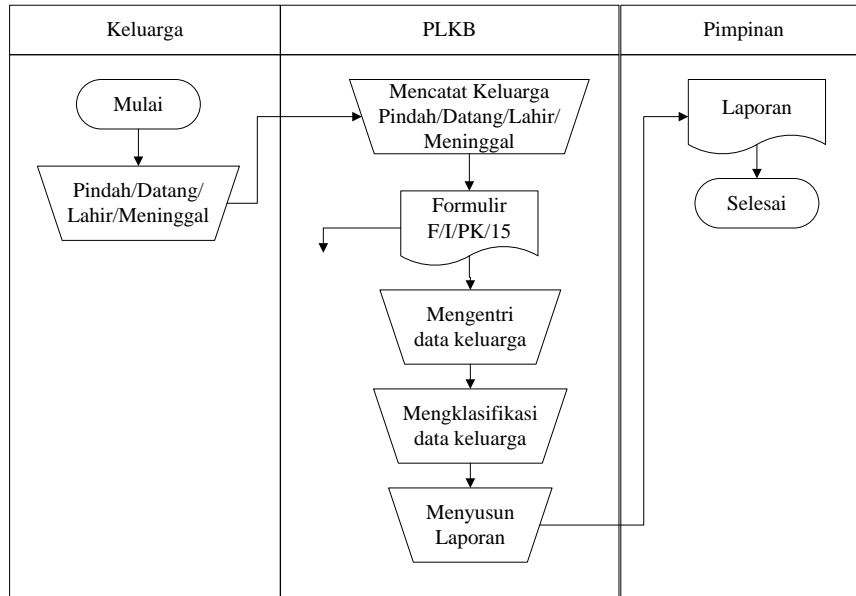
Berikut flowchart dokumen dari sistem yang berjalan :



Gambar 1. Flowchart Dokumen Sistem Berjalan

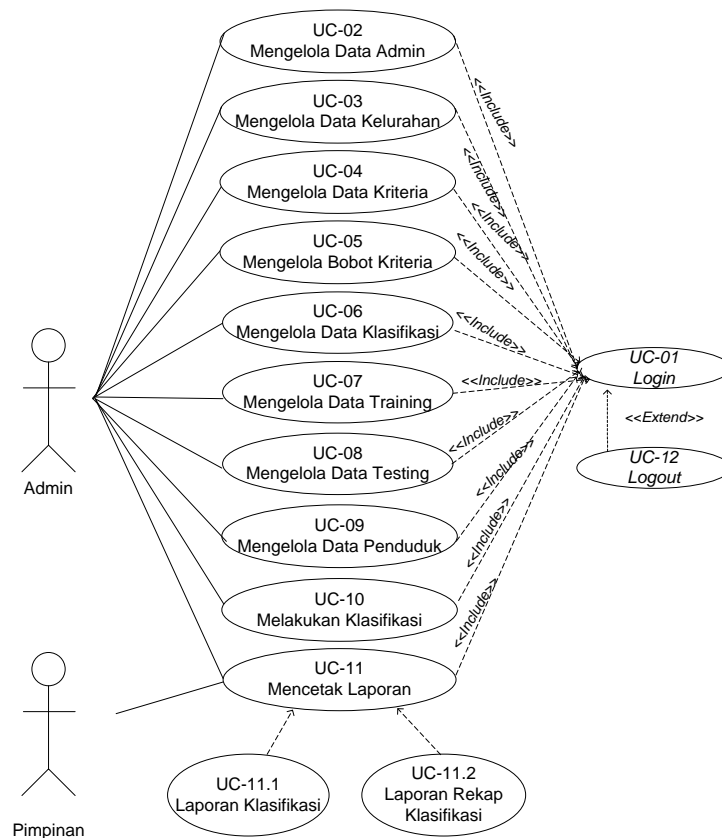
Solusi Sistem Yang Ditawarkan

Solusi sistem yang ditawarkan untuk mengatasi kelemahan pada sistem yang berjalan dapat digambarkan melalui flowchart dibawah ini.



Gambar 2 Flowchart Dokumen Solusi Sistem Yang Ditawarkan

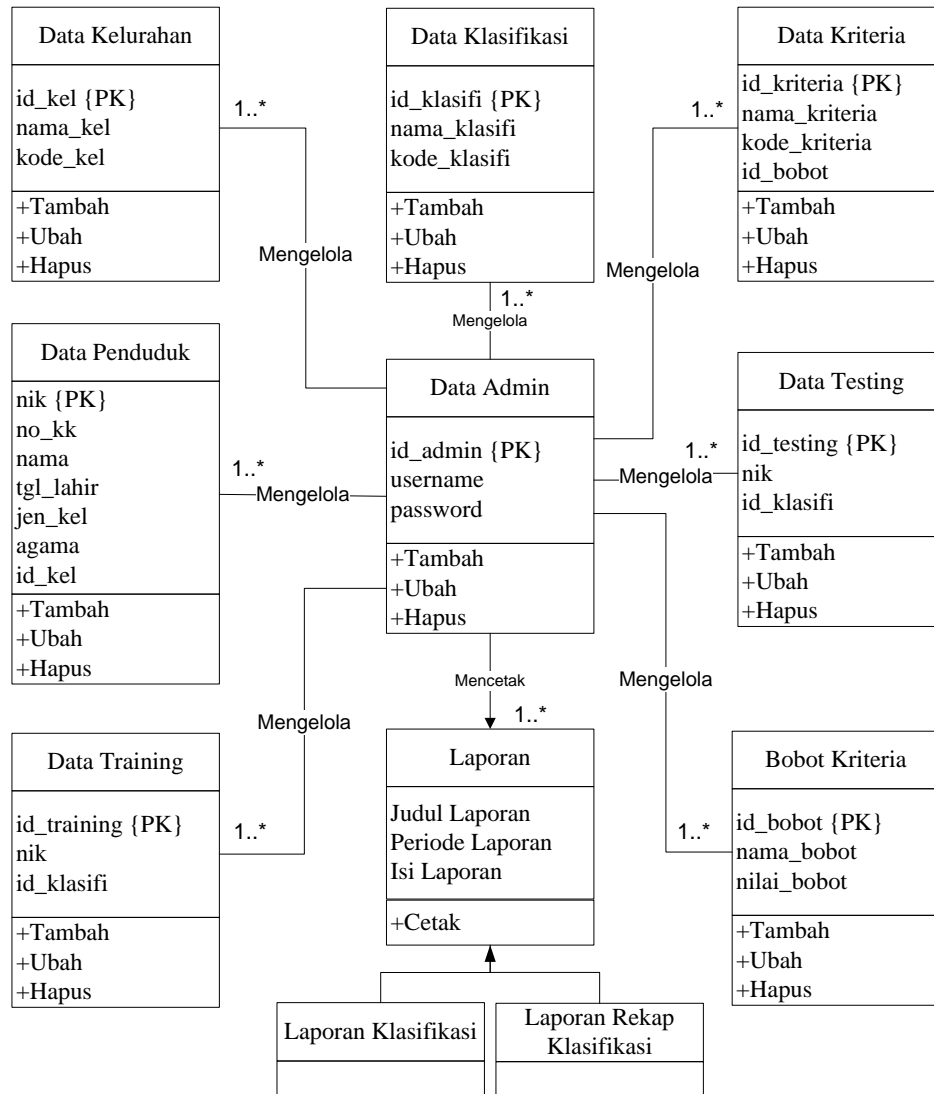
Use Case Diagram (Diagram Use Case)



Gambar 3 Use Case Diagram

Class Diagram

Kebutuhan data untuk perangkat lunak yang akan dibuat dapat digambarkan dengan class diagram berikut ini :



Gambar 4 Class Diagram

Pengujian Metode Klasifikasi Menggunakan Data Training

Jumlah data yang peneliti dapatkan berjumlah 3.567 data. Kemudian data tersebut dibagi 2 (dua) menjadi data latih (training) dan data uji (testing). Data latih diambil dari 3 kelurahan berjumlah 3.407 data, sementara data uji diambil dari 1 kelurahan berjumlah 160 data.

Berikut hasil prediksi klasifikasi dari data latih menggunakan tools weka.

No.	1: C1	2: C2	3: C3	4: C4	5: C5	6: C6	7: C7	8: C8	9: C9	10: C10	11: C11	12: C12	13: C13	14: C14	15: C15	16: C16	17: C17	18: C18	19: C19	20: prediction margin	21: predicted Tingkat KS	22: Tingkat KS
	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Nominal	Numeric	Nominal	Nominal
1	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.961785	Sejahtera 1	Sejahtera 1
2	SLT...	PED...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.954297	Sejahtera 1	Sejahtera 1
3	SLT...	PED...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.96664	Sejahtera 1	Sejahtera 1
4	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Sedikit	0.962826	Sejahtera 1	Sejahtera 1
5	DIP...	MEN...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Tida...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.990402	Sejahtera 1	Sejahtera 1
6	DIP...	KAR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.984865	Sejahtera 1	Sejahtera 1
7	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Tida...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.986863	Sejahtera 1	Sejahtera 1
8	SLT...	WIR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Tida...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.989742	Sejahtera 1	Sejahtera 1
9	SLT...	KAR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.785381	Sejahtera 2	Sejahtera 2
10	SLT...	WIR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.96885	Sejahtera 1	Sejahtera 1
11	TAM...	BUR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.784372	Sejahtera 2	Sejahtera 2
12	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Seda...	0.941922	Sejahtera 1	Sejahtera 1
13	AKA...	KAR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Sedikit	0.978382	Sejahtera 1	Sejahtera 1
14	SLT...	PED...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.954297	Sejahtera 1	Sejahtera 1
15	DIP...	KAR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Tida...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.990343	Sejahtera 1	Sejahtera 1
16	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.80824	Sejahtera 2	Sejahtera 2
17	SLT...	MEN...	JAN...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Sedikit	0.966603	Sejahtera 1	Sejahtera 1
18	SLT...	WIR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.971411	Sejahtera 1	Sejahtera 1
19	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.961785	Sejahtera 1	Sejahtera 1
20	DIP...	PEG...	JAN...	Tida...	Me...	Pen...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Sedikit	0.473905	Prasejahtera	Prasejahtera
21	SLT...	KAR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Sedikit	0.960356	Sejahtera 1	Sejahtera 1
22	SLT...	WIR...	KAW...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.961079	Sejahtera 1	Sejahtera 1
23	SLT...	MEN...	BEL...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Tida...	seng...	Besar	Seda...	0.946336	Sejahtera 1	Sejahtera 1
24	TAM...	MEN...	JAN...	BPJ...	Me...	Me...	Me...	Me...	Me...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.753226	Sejahtera 2	Sejahtera 2
25	SLT...	PED...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.651492	Sejahtera 2	Sejahtera 2
26	TAM...	BUR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Seda...	0.95864	Sejahtera 1	Sejahtera 1
27	TAM...	PED...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	lainn...	Besar	Seda...	0.96161	Sejahtera 1	Sejahtera 1
28	SLT...	WIR...	KAW...	Tida...	Me...	Me...	Me...	Me...	Me...	Mem...	Peng...	Mem...	Mem...	Mem...	Mem...	Mem...	seng...	Besar	Sedikit	0.974411	Sejahtera 1	Sejahtera 1

Gambar 5 Hasil Prediksi Data Klasifikasi Menggunakan Data Training

Dari hasil klasifikasi yang disajikan diatas, dapat diperoleh tabel confusion matrix sebagai berikut :

Tabel 1 Tabel Confusion Matrix Hasil Klasifikasi Data Training

record	Sejahtera 1	Sejahtera 2	Prasejahtera	Jumlah
Sejahtera 1	2331	77	1	2409
Sejahtera 2	66	747	21	834
Prasejahtera	19	77	68	164
Jumlah	2416	901	90	3407

Keterangan Tabel :

- Dari 3.407 data hasil klasifikasi PLKB menyatakan jumlah kategori Prasejahtera sebanyak 164 keluarga, Sejahtera 1 sebanyak 2.409 keluarga, dan Sejahtera 2 sebanyak 834 keluarga.
- Sementara, hasil prediksi klasifikasi Naive bayes menggunakan tools weka didapatkan jumlah kategori Prasejahtera sebanyak 90 keluarga, Sejahtera 1 sebanyak 2.416 keluarga, dan Sejahtera 2 sebanyak 901 keluarga.
- Prediksi berdasarkan tabel matrix konfusi diatas, dapat dilihat data yang diagonal merupakan data yang terklasifikasi benar sebanyak 3.146 sedangkan data yang belum tepat terklasifikasi sebanyak 261.
- Penjabaran prediksi data yang belum tepat klasifikasinya :

1. Kelas Sejahtera 1 yang masuk ke kelas Sejahtera 2 sebanyak 77
 2. Kelas Sejahtera 1 yang masuk ke kelas Prasejahtera sebanyak 1
 3. Kelas Sejahtera 2 yang masuk ke kelas Sejahtera 1 sebanyak 66
 4. Kelas Sejahtera 2 yang masuk ke kelas Prasejahtera sebanyak 21
 5. Kelas Prasejahtera yang masuk ke kelas Sejahtera 1 sebanyak 19
 6. Kelas Prasejahtera yang masuk ke kelas Sejahtera 2 sebanyak 77
5. Nilai Akurasi = $(TP+TN) / \text{TOTAL DATA} = (2.331 + 747 + 68) / 3.407 = 3.147 / 3.407 = 0,92$. Ini berarti nilai akurasi dari klasifikasi cukup tinggi karena mendekati 1.
6. Nilai Error Rate (ER) = $FP+FN / \text{TOTAL DATA} = (77 + 1 + 66 + 21 + 19 + 77) / 3.407 = 261/3.407 = 0,08$. Ini berarti nilai eror dari hasil klasifikasi cukup rendah karena mendekati 0.

Perhitungan Manual Metode Klasifikasi

Untuk pengujian manual peneliti mengambil sampel data dari data latih, sementara data uji digunakan peneliti untuk memprediksi data baru yang belum terklasifikasi. Berikut langkah perhitungan manual dengan rumus Naive bayes menggunakan sampel data baru yang belum terklasifikasi dengan kriteria seperti tabel dibawah ini.

Tabel 2 Data Baru Untuk Di Testing

Kode	Kriteria	Prasejahtera	Sejahtera 1	Sejahtera 2	Prasejahtera	Sejahtera 1	Sejahtera 2
C1	SLTP	11	497	198	0,12	0,21	0,22
C2	Pedagang	22	384	143	0,24	0,16	0,16
C3	Janda	16	288	96	0,18	0,12	0,11
C4	BPJS-PBI	20	276	159	0,22	0,11	0,18
C5	Memenuhi	62	2416	901	0,69	1,00	1,00
C6	Memenuhi	42	2416	901	0,47	1,00	1,00
C7	Memenuhi	62	2416	901	0,69	1,00	1,00
C8	Memenuhi	62	2416	901	0,69	1,00	1,00
C9	Memenuhi	62	2416	901	0,69	1,00	1,00
C10	Memenuhi	62	2416	901	0,69	1,00	1,00

C11	Memenuhi	41	0	901	0,46	0,00	1,00
C12	Tidak Memenuhi	68	58	13	0,76	0,02	0,01
C13	Tidak Memenuhi	68	13	0	0,76	0,01	0,00
C14	Tidak Memenuhi	28	200	84	0,31	0,08	0,09
C15	Memenuhi	20	2404	888	0,22	1,00	0,99
C16	Tidak Memenuhi	90	903	339	1,00	0,37	0,38
C17	seng/tembok/ semen	3	114	32	0,03	0,05	0,04
C18	Sedang	28	953	378	0,31	0,39	0,42
C19	Sedang	20	778	425	0,22	0,32	0,47

Berikut proses perhitungan manual menggunakan formula Naive bayes :

- a. Menghitung jumlah class / label

$$P(Y=\text{Prasejahtera}) = 90/3407 = 0,03$$

“Jumlah data Prasejahtera pada data latih dibagi dengan jumlah keseluruhan data latih”

$$P(Y=\text{Sejahtera 1}) = 2416/3407 = 0,71$$

“Jumlah data Sejahtera 1 pada data latih dibagi dengan jumlah keseluruhan data latih”

$$P(Y=\text{Sejahtera 2}) = 901/3407 = 0,26$$

“Jumlah data Sejahtera 2 pada data latih dibagi dengan jumlah keseluruhan data latih”

- b. Menghitung jumlah kasus yang sama dengan class yang sama

$$P(C1=\text{SLTP} | Y=\text{Prasejahtera}) = 11/90 = 0,12$$

$$P(C1=\text{SLTP} | Y=\text{Sejahtera 1}) = 497/2416 = 0,21$$

$$P(C1=\text{SLTP} | Y=\text{Sejahtera 2}) = 198/901 = 0,22$$

$$\begin{aligned} P(C2=\text{Pedagang} \mid Y=\text{Prasejahtera}) &= 22/90 = 0,24 \\ P(C2=\text{Pedagang} \mid Y=\text{Sejahtera 1}) &= 384/2416 = 0,16 \\ P(C2=\text{Pedagang} \mid Y=\text{Sejahtera 2}) &= 143/901 = 0,16 \\ P(C3=\text{Janda} \mid Y=\text{Prasejahtera}) &= 16/90 = 0,18 \\ P(C3=\text{Janda} \mid Y=\text{Sejahtera 1}) &= 288/2416 = 0,12 \\ P(C3=\text{Janda} \mid Y=\text{Sejahtera 2}) &= 96/901 = 0,11 \\ P(C4=\text{BPJS-PBI} \mid Y=\text{Prasejahtera}) &= 20/90 = 0,22 \\ P(C4=\text{BPJS-PBI} \mid Y=\text{Sejahtera 1}) &= 276/2416 = 0,11 \\ P(C4=\text{BPJS-PBI} \mid Y=\text{Sejahtera 2}) &= 159/901 = 0,18 \\ P(C5=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 62/90 = 0,69 \\ P(C5=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C5=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C6=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 42/90 = 0,47 \\ P(C6=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C6=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C7=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 62/90 = 0,69 \\ P(C7=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C7=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C8=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 62/90 = 0,69 \\ P(C8=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C8=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C9=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 62/90 = 0,69 \\ P(C9=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C9=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C10=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 62/90 = 0,69 \\ P(C10=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 2416/2416 = 1,00 \\ P(C10=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C11=\text{Memenuhi} \mid Y=\text{Prasejahtera}) &= 41/90 = 0,46 \\ P(C11=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) &= 0/2416 = 0 \\ P(C11=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) &= 901/901 = 1,00 \\ P(C12=\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) &= 68/90 = 0,76 \\ P(C12=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 1}) &= 58/2416 = 0,02 \end{aligned}$$

$$P(C12=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 2}) = 13/901 = 0,01$$

$$P(C13=\text{Memenuhi} \mid Y=\text{Prasejahtera}) = 68/90 = 0,76$$

$$P(C13=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) = 13/2416 = 0,01$$

$$P(C13=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) = 0/901 = 0$$

$$P(C14=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 1}) = 28/90 = 0,31$$

$$P(C14=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 2}) = 200/2416 = 0,08$$

$$P(C14=\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) = 84/901 = 0,09$$

$$P(C15=\text{Memenuhi} \mid Y=\text{Prasejahtera}) = 20/90 = 0,22$$

$$P(C15=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) = 2404/2416 = 1,00$$

$$P(C15=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) = 888/901 = 0,99$$

$$P(C16=\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) = 90/90 = 1,00$$

$$P(C16=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 1}) = 903/2416 = 0,37$$

$$P(C16=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 2}) = 339/901 = 0,38$$

$$P(C17=\text{Seng/Tembok/Semen} \mid Y=\text{Prasejahtera}) = 3/90 = 0,03$$

$$P(C17=\text{Seng/Tembok/Semen} \mid Y=\text{Sejahtera 1}) = 114/2416 = 0,05$$

$$P(C17=\text{Seng/Tembok/Semen} \mid Y=\text{Sejahtera 2}) = 32/901 = 0,04$$

$$P(C18=\text{Sedang} \mid Y=\text{Prasejahtera}) = 28/90 = 0,31$$

$$P(C18=\text{Sedang} \mid Y=\text{Sejahtera 1}) = 953/2416 = 0,39$$

$$P(C18=\text{Sedang} \mid Y=\text{Sejahtera 2}) = 378/901 = 0,42$$

$$P(C19=\text{Sedang} \mid Y=\text{Prasejahtera}) = 20/90 = 0,22$$

$$P(C19=\text{Sedang} \mid Y=\text{Sejahtera 1}) = 778/2416 = 0,32$$

$$P(C19=\text{Sedang} \mid Y=\text{Sejahtera 2}) = 425/901 = 0,47$$

c. Kalikan semua hasil variabel Prasejahtera, Sejahtera 1 dan Sejahtera 2

$$P(C1=\text{SLTP} \mid Y=\text{Prasejahtera}) * P(C2=\text{Pedagang} \mid Y=\text{Prasejahtera}) *$$

$$P(C3=\text{Janda} \mid Y=\text{Prasejahtera}) * P(C4=\text{BPJS-PBI} \mid Y=\text{Prasejahtera}) *$$

$$P(C5=\text{Memenuhi} \mid Y=\text{Prasejahtera}) * P(C6=\text{Memenuhi} \mid Y=\text{Prasejahtera}) *$$

$$P(C7=\text{Memenuhi} \mid Y=\text{Prasejahtera}) * P(C8=\text{Memenuhi} \mid Y=\text{Prasejahtera}) *$$

$$P(C9=\text{Memenuhi} \mid Y=\text{Prasejahtera}) * P(C10=\text{Memenuhi} \mid Y=\text{Prasejahtera}) *$$

$$P(C11=\text{Memenuhi} \mid Y=\text{Prasejahtera}) * P(C12=\text{Tidak Memenuhi} \mid$$

$$Y=\text{Prasejahtera}) * P(C13=\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) * P(C14=$$

$$\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) * P(C15=\text{Memenuhi} \mid Y=\text{Prasejahtera}) *$$

$$\begin{aligned} & P(C16=\text{Tidak Memenuhi} \mid Y=\text{Prasejahtera}) * P(C17=\text{Seng/Tembok/Semen} \mid \\ & Y=\text{Prasejahtera}) * P(C18=\text{Sedang} \mid Y=\text{Prasejahtera}) * P(C19=\text{Sedang} \mid \\ & Y=\text{Prasejahtera}) \\ & = 0,12 * 0,19 * 0,18 * 0,22 * 0,69 * 0,47 * 0,69 * 0,69 * 0,69 * 0,46 \\ & * 0,76 * 0,76 * 0,31 * 0,22 * 1,00 * 0,03 * 0,31 * 0,22 \\ & = 3,54084E-09 \end{aligned}$$

$$\begin{aligned} & P(C1=\text{SLTP} \mid Y=\text{Sejahtera 1}) * P(C2=\text{Pedagang} \mid Y=\text{Sejahtera 1}) * \\ & P(C3=\text{Janda} \mid Y=\text{Sejahtera 1}) * P(C4=\text{BPJS-PBI} \mid Y=\text{Sejahtera 1}) * \\ & P(C5=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C6=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * \\ & P(C7=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C8=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * \\ & P(C9=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C10=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * \\ & P(C11=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C12=\text{Tidak Memenuhi} \mid Y= \\ & \text{Sejahtera 1}) * P(C13=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C14=\text{Tidak} \\ & \text{Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C15=\text{Memenuhi} \mid Y=\text{Sejahtera 1}) * \\ & P(C16=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 1}) * P(C17=\text{Seng/Tembok/Semen} \mid \\ & Y=\text{Sejahtera 1}) * P(C18=\text{Sedang} \mid Y=\text{Sejahtera 1}) * P(C19=\text{Sedang} \mid Y= \\ & \text{Sejahtera 1}) \\ & = 0,21 * 0,16 * 0,12 * 0,11 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 0 * \\ & 0,02 * 0,01 * 0,08 * 1,00 * 0,37 * 0,05 * 0,39 * 0,32 \\ & = 0 \end{aligned}$$

$$\begin{aligned} & P(C1=\text{SLTP} \mid Y=\text{Sejahtera 2}) * P(C2=\text{Pedagang} \mid Y=\text{Sejahtera 2}) * \\ & P(C3=\text{Janda} \mid Y=\text{Sejahtera 2}) * P(C4=\text{BPJS-PBI} \mid Y=\text{Sejahtera 2}) * \\ & P(C5=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C6=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * \\ & P(C7=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C8=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * \\ & P(C9=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C10=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * \\ & P(C11=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C12=\text{Tidak Memenuhi} \mid Y= \\ & \text{Sejahtera 2}) * P(C13=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C14=\text{Tidak} \\ & \text{Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C15=\text{Memenuhi} \mid Y=\text{Sejahtera 2}) * \\ & P(C16=\text{Tidak Memenuhi} \mid Y=\text{Sejahtera 2}) * P(C17=\text{Seng/Tembok/Semen} \mid \\ & Y=\text{Sejahtera 2}) * P(C18=\text{Sedang} \mid Y=\text{Sejahtera 2}) * P(C19=\text{Sedang} \mid Y= \\ & \text{Sejahtera 2}) \end{aligned}$$

$$\begin{aligned} &= 0,22 * 0,16 * 0,11 * 0,18 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * 1,00 * \\ &0,01 * 0 * 0,09 * 0,99 * 0,38 * 0,04 * 0,42 * 0,47 \\ &= 0 \end{aligned}$$

- d. Bandingkan hasil class Prasejahtera, Sejahtera 1 dan Sejahtera 2 dari hasil dari hasil diatas, terlihat bahwa nilai probabilitas tertinggi ada pada kelas (P|Prasejahtera) sehingga dapat disimpulkan bahwa penduduk tersebut masuk dalam klasifikasi “Prasejahtera”.

KESIMPULAN

Berdasarkan pembahasan yang telah diuraikan pada bab-bab sebelumnya, maka peneliti dapat menarik beberapa kesimpulan, sebagai berikut :

1. Penelitian ini menghasilkan klasifikasi tingkat kesejahteraan keluarga yaitu Prasejahtera, Sejahtera 1 dan Sejahtera 2 dengan memanfaatkan Ilmu Data Mining teknik klasifikasi menggunakan algoritma Naive Bayes.
2. Hasil prediksi klasifikasi berdasarkan data latih (training) didapatkan :
 - a. Keluarga Prasejahtera sebanyak 90 keluarga
 - b. Keluarga Sejahtera 1 sebanyak 2.416 keluarga
 - c. Keluarga Sejahtera 2 sebanyak 901 keluarga
3. Hasil akurasi dikatakan baik apabila mendekati 1 atau 100%, berbanding terbalik dengan nilai error dikatakan baik apabila mendekati 0, dan pada penelitian ini didapatkan nilai akurasi pada sistem yaitu 92,33% atau 0,92 dengan nilai error sebesar 7,67% atau 0,08.
4. Teknik Klasifikasi tingkat kesejahteraan keluarga dengan memanfaatkan Data Mining menggunakan algoritma Naive Bayes dapat mempermudah PLKB dalam menentukan tingkat kesejahteraan keluarga.

DAFTAR PUSTAKA

- Alpaydin, E. (2014). *Introduction to Machine Learning* (3rd ed.). MIT Press.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3rd ed.). Elsevier.
- Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2(3), 18-22.

- Mamase, S. (2022). Prediksi Tingkat Kesejahteraan Rakyat Suatu Kecamatan Menggunakan Generalized Regression Neural Network. *Jurnal Teknologi Informasi Indonesia (JTHI)*, 7(1), 62-65.
- Permana, Y., & Lelah, L. (2020). Pengklasifikasian Tingkat Kesejahteraan Keluarga Di Desa Citamiang Dengan Penerapan Logika Fuzzy Model Tahani. *Rabit J. Teknol. dan Sist. Inf. Univrab*, 5(2), 97-107.
- Qusyairi, M. (2024). Analisa Prediksi Tingkat Kesejahteraan Masyarakat Nelayan Lombok Timur Dengan Algoritma Naïve Bayes. *Infotek: Jurnal Informatika dan Teknologi*, 7(2), 563-574.
- Supriana, I. W., & Astuti, L. G. (2019). Implementasi K-Nearest Neighbor Pada Penentuan Keluarga Miskin Bagi Dinas Sosial Kabupaten Tabanan. *Jurnal Teknologi Informasi dan Komputer*, 5(1).
- Tarigan, A., Mustakim, M., Wahyudi, E., & Adhiva, J. (2019). Klasifikasi Status Kesejahteraan Rumah Tangga di Kabupaten Siak Menggunakan Algoritma Naive Bayes Classifier. In *Seminar Nasional Teknologi Informasi Komunikasi dan Industri* (pp. 187-196).
- Wanto, A., & Gunawan, I. (2022). Penerapan Algoritma Decision Tree C4. 5 untuk Klasifikasi Tingkat Kesejahteraan Keluarga pada Desa Tiga Dolok. *Krisnadana Journal*, 1(2), 21-32.