

KLASIFIKASI KELAYAKAN PENERIMA BANTUAN PROGRAM KELUARGA HARAPAN (PKH) MENGGUNAKAN METODE *WEIGHTED NAÏVE BAYES* DENGAN *LAPLACE SMOOTHING*

Diyah Utami¹⁾, Putri Aisyiyah Rakhma Devi²⁾

^{1,2)}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Gresik
Jl. Sumatra 101 Gresik Kota Baru (GKB), Randuagung, Kebomas, Gresik
e-mail: diyahutami3010@gmail.com¹⁾, deviaisyiyah@umg.ac.id²⁾

ABSTRAK

Program Keluarga Harapan (PKH) yaitu keluarga yang telah ditetapkan menjadi penerima PKH dapat menerima bantuan sosial bersyarat dalam program PKH. Pemerintah Indonesia mencanangkan PKH sebagai inisiatif pemerintah pada tahun 2007. PKH berupaya untuk meningkatkan kualitas sumber daya manusia (SDM) di bidang kesehatan, pendidikan, dan kesejahteraan dengan meningkatkan kualitas hidup. Program tersebut juga dapat membantu keluarga miskin dalam mengurangi pengeluaran. Proses penerima bantuan PKH masih dilakukan secara manual dan penyebaran penerima bantuan PKH tidak tepat sasaran bagi warga yang menerima bantuan PKH. Oleh karena itu, dalam penelitian ini dapat memberikan prediksi klasifikasi kelayakan penerima bantuan PKH sehingga dapat membantu perangkat desa dalam menentukan warga desa yang layak menerima bantuan tersebut dan dapat mendapatkan hasil yang lebih cepat, akurat dan tepat sasaran bagi yang menerima. Dalam penelitian ini akan dilakukan klasifikasi kelayakan penerima bantuan PKH menggunakan metode Weighted Naïve Bayes dengan Laplace Smoothing. Dengan melakukan penambahan bobot terhadap atribut kelas pada algoritma naïve bayes, maka akurasi klasifikasi algoritma naïve Bayes berbobot tidak hanya didasarkan pada probabilitas tetapi juga pada pembobotan fitur ke kelas, yang ditambahkan ke atribut kelas dalam algoritma naïve bayes dan dengan menggunakan laplace smoothing dapat menghindari nilai probabilitas 0. Dari hasil perhitungan klasifikasi menggunakan metode naïve bayes dengan laplace smoothing pada 56 data training dan 24 data uji didapatkan hasil pengujian evaluasi performa dengan menggunakan confusion matrix dengan nilai Accurasi 95,83%, error 4,17%, Sensitivitas 100,00%, dan Spesifitas 94,12%.

Kata Kunci: *Klasifikasi, PKH, Weighted Naïve Bayes, Laplace Smoothing.*

ABSTRACT

The Family Hope Program (PKH) means that families who have been designated as PKH recipients can receive conditional social assistance in the PKH program. The Indonesian government launched PKH as a government initiative in 2007. PKH seeks to improve the quality of human resources (HR) in the fields of health, education, and welfare by improving the quality of life. The program can also help poor families reduce expenses. The process of receiving PKH assistance is still done manually and the distribution of PKH recipients is not well targeted for residents who receive PKH assistance. eligible to receive the assistance and can get faster, accurate and targeted results for those who receive it. In this study, the eligibility classification of PKH beneficiaries will be carried out using the Weighted Naïve Bayes method with Laplace Smoothing. By adding weights to class attributes in the naïve Bayes algorithm, the classification accuracy of the weighted naïve Bayes algorithm is not only based on probability but also on the weighting of features to class, which are added to class attributes in the naïve Bayes algorithm and by using laplace smoothing, can avoid the probability value of 0. From the results of classification calculations using the naïve Bayes method with laplace smoothing on 56 training data and 24 test data, the results of performance evaluation tests using confusion matrix with Accuracy value 95,83%, error 4,17%, Sensitivity is 100,00%, and Sensitivity is 94.12%.

Keywords: *Classification, PKH, Weighted Naïve Bayes, Laplace Smoothing*

I. PENDAHULUAN

PROGRAM Keluarga Harapan (PKH) yaitu keluarga yang telah ditetapkan menjadi penerima PKH dapat menerima bantuan sosial bersyarat dalam program PKH. Pemerintah Indonesia mencanangkan PKH sebagai inisiatif pemerintah pada tahun 2007. PKH berupaya untuk meningkatkan kualitas sumber daya manusia (SDM) di bidang kesehatan, pendidikan, dan kesejahteraan dengan meningkatkan kualitas hidup. Program tersebut juga dapat membantu keluarga miskin dalam mengurangi pengeluaran[1].

Proses penerimaan bantuan PKH di Dusun Mando, Desa Pandu, Kecamatan Cerme, Kabupaten Gresik masih dilakukan secara manual yaitu dengan cara menyeleksi daftar calon dari penerima bantuan PKH yang telah diusulkan oleh ketua Rumah Tangga (RT) dan Kepala Dusun. Kemudian data tersebut diserahkan kepada sekretaris

desa dan perangkat desa setempat untuk dilakukan penilaian. Adapun pada praktisnya, terdapat beberapa penerima bantuan PKH yang tidak tepat sasaran, seperti contoh kasus di dusun Mando, terdapat penerima manfaat yang termasuk dalam kategori layak mendapatkan bantuan PKH justru tidak mendapatkan yang seharusnya layak mendapatkan bantuan PKH justru tidak mendapatkan bantuan PKH dan sebaliknya.

Salah satu solusi untuk menyelesaikan ketidaksesuaian penerima bantuan PKH adalah dengan melakukan klasifikasi. Klasifikasi kelayakan penerima bantuan PKH di Dusun Mando Desa Pandu agar penerima bantuan bisa sesuai dan tepat sasaran. Klasifikasi kelayakan penerima bantuan dapat dilakukan dengan menggunakan metode data mining. Data mining yaitu proses pengumpulan data-data dengan cara yang baru, mudah dipahami, dan bermanfaat bagi pemilik data[2]. Klasifikasi adalah salah satu teknik data mining. Menemukan pola yang menjelaskan atau mendefinisikan suatu kelas data adalah proses klasifikasi, yang bertujuan untuk meramalkan kelas suatu item yang labelnya tidak diketahui. Model yang dapat membedakan antara berbagai data dan kelas berdasarkan aturan fungsi adalah apa yang ingin dihasilkan oleh proses klasifikasi.

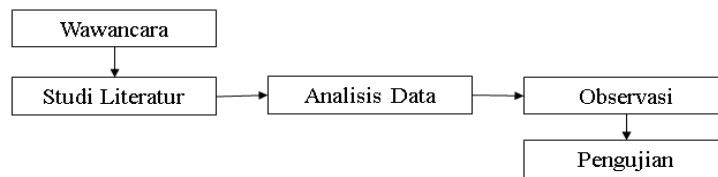
Adapun salah satu metode populer dalam klasifikasi adalah *naïve bayes*. *Naïve bayes* adalah metode klasifikasi probabilistik dasar yang menghitung nilai probabilitas dengan menjumlahkan frekuensi dan jumlah nilai dari kumpulan data yang diberikan. Untuk menghasilkan nilai variabel, metode ini menggunakan teorema Bayes dan mengasumsikan bahwa kualitas-kualitas tersebut bersifat independen atau tidak dependen. Berdasarkan kejadian sebelumnya, *Naïve Bayes* meramalkan peluang masa depan[3]. Pada praktisnya, penerapan metode *naïve bayes* pada data dapat mengakibatkan kesalahan dalam klasifikasi jika data training jumlahnya hanya sedikit dan menyebabkan kesalahan dalam proses klasifikasi. Oleh karena itu, untuk mengatasi hal tersebut dapat diperbaiki dengan menggunakan teknik *smoothing*. Selain itu beberapa dari penelitian sebelumnya menjelaskan bahwa teknik *smoothing* dapat meningkatkan performansi dari klasifikasi *naïve bayes*[4]. Teknik *smoothing* yang digunakan dalam penelitian ini yaitu *laplace smoothing*. Teknik *Laplace Smoothing* digunakan untuk mengatasi nilai probabilitas pada klasifikasi *Naïve Bayes* yang bernilai 0 dengan cara menambahkan angka 1 pada perhitungan[5].

Pada penelitian sebelumnya dilakukan oleh M. riski Qishiano, dkk tahun 2021 yang berjudul “Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode *Naïve Bayes*” mendapatkan kesimpulan, bahwa dengan mengimplementasikan metode *naïve bayes* mendapatkan nilai akurasi 82%[6]. Penelitian yang dilakukan oleh Indah Listiowarni dan Nilam Ramadhani tahun 2019 yang berjudul “Implementasi *Naïve Bayesian* dengan *Laplacian Smoothing* untuk Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan” mendapatkan kesimpulan, bahwa akurasi perhitungan klasifikasi menggunakan *naïve bayes* dengan *laplace smoothing* adalah 92,11%, dan tingkat eror adalah 7,02%. Sehingga hasil dari klasifikasi bisa dijadikan pertimbangan dalam menentukan kelas peminatan dan lintas minat siswa SMAN 5 Pamekasan[7]. Penelitian yang dilakukan Qisthiano, M Riski, Tri Basuki Kurniawan, Edi Surya Negara, dan Muhammad Akbar tahun 2021 dengan judul penelitian “Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode *Naïve Bayes*” mendapatkan kesimpulan, bahwa dengan mengimplementasikan metode *naïve bayes* mendapatkan nilai akurasi 81% [6]. Namun, metode *naïve bayes* memiliki kekurangan karena probabilitas tidak dapat mengukur seberapa besar tingkat akurat atau kebenaran dari sebuah prediksi, sehingga diperlukan pembobotan *weighted* untuk mengoptimalkan metode *naïve bayes* dengan memberikan bobot pada setiap fitur. Penelitian yang dilakukan oleh Azminuddin I. S. Azis, Budy Santoso, dan Serwin tahun 2020 dengan judul penelitian “LL-KNNACW-NB: Local Learning K-Nearest Neighbor in Absolute Correlation Weighted *Naïve Bayes* untuk Klasifikasi Data Numerik” mendapatkan kesimpulan, dapat meningkatkan kinerja *Naïve Bayes* dengan akurasi rata-rata 91,48%, 1,92% lebih baik dari *naïve bayes* dan 2,86% lebih baik daripada *Kernel Naïve Bayes*[8].

Berdasarkan studi literatur pada penelitian sebelumnya, terbukti bahwa penerapan metode *Weighted Naïve Bayes* menggunakan teknik *Laplace Smoothing* berhasil memberikan hasil klasifikasi yang baik. Dalam penelitian ini akan dilakukan proses klasifikasi kelayakan penerima bantuan PKH di Dusun Mando Desa Pandu menggunakan *weighted naïve bayes* dengan *laplace smoothing*. Penelitian ini akan menghasilkan *output* klasifikasi kelayakan penerima bantuan PKH di Dusun Mando Desa Pandu. Sehingga terdapat 2 kelas yang akan digunakan dalam proses klasifikasi ini yaitu kelas layak dan kelas tidak layak. Adapun 8 kriteria yang digunakan dalam penelitian ini yaitu usia, pekerjaan, penghasilan, jumlah keluarga, jumlah tanggungan anak sekolah, status rumah, jenis lantai dan jenis dinding. Penelitian bertujuan untuk memberikan prediksi klasifikasi kelayakan penerima bantuan Program Keluarga Harapan (PKH) sehingga dapat membantu perangkat desa dalam menentukan warga desa yang layak menerima bantuan tersebut dan dapat mendapatkan hasil yang lebih cepat, akurat dan tepat sasaran bagi yang menerima.

II. METODE PENELITIAN

A. Tahapan Penelitian



Gambar 1. Tahapan Penelitian

Pada Gambar 1 merupakan tahapan penelitian yang dijelaskan sebagai berikut:

1. Tahap pertama yaitu melakukan wawancara dengan pihak yang terkait dalam penelitian guna mendapatkan informasi dan data yang diperlukan.
2. Tahapan kedua, melakukan studi literatur yaitu mengumpulkan bahan referensi yang sesuai dengan topik penelitian melalui buku, jurnal, artikel, makalah, situs internet, maupun skripsi. Referensi tersebut bertujuan untuk menentukan variabel-variabel dan metode penyelesaian yang akan digunakan.
3. Tahapan ketiga, analisis data melakukan perhitungan dengan menggunakan metode *weighted naïve bayes* dengan *laplace smoothing*, sehingga didapatkan hasil klasifikasi kelayakan penerima bantuan program keluarga harapan (PKH) secara cepat dan tepat.
4. Tahapan keempat, berdasarkan hasil observasi dan pengumpulan data, untuk memilih fitur mana yang akan digunakan, dilakukan analisis kebutuhan[9].
5. Pengujian dilakukan pada langkah kelima untuk memastikan bahwa teknik tersebut dapat memberikan informasi secara akurat.

B. Preprocessing

Preprocessing merupakan teknik yang dipergunakan untuk membarui data mentah pada format yang bermanfaat serta efisien. Inisiatif ini diharapkan karena data mentah seringkali tak lengkap serta mempunyai format yang tidak konsisten. *Preprocessing* sendiri melibatkan validasi dan imputasi data[10]. Ada tahapan dalam preprocessing data yaitu mengatasi nilai null atau kosong, normalisasi, dan membagi data latih dan data uji[11]. Proses ini utamanya dilakukan buat memastikan kualitas data baik sebelum digunakan waktu menganalisis data. dalam proses ini dapat memastikan enam hal, yakni akurasi data, kelengkapan, konsistensi, ketepatan, tepercaya, serta dapat di interpretasikan dengan baik[12].

C. Klasifikasi

Untuk memprediksi kelas dari suatu objek yang labelnya tidak diketahui, klasifikasi adalah proses mengidentifikasi model atau fungsi yang membedakan dan menggambarkan kelas data yang digunakan. Konsep klasifikasi dapat digunakan untuk meramalkan bagaimana setiap kumpulan data akan diklasifikasikan. Ada dua cara atau proses klasifikasi untuk sampai pada kesimpulan: yang pertama adalah proses pembelajaran (*training*), dan yang kedua adalah proses klasifikasi, juga dikenal sebagai pengujian (*testing*). Proses tersebut sesuai dengan metode klasifikasi yang digunakan[13].

D. Naïve Bayes

Naïve bayes adalah salah satu metode klasifikasi yang digolongkan kedalam pembelajaran *supervised learning*, dimana diperlukan pengolahan data awal untuk pelatihan agar dapat menghasilkan penilaian. Tujuan dari pelatihan ini adalah untuk menetapkan bobot probabilitas untuk setiap parameter. Metode *naive bayes* adalah metodologi klasifikasi yang dapat digunakan untuk menyelesaikan masalah klasifikasi atau melakukan penambangan data dengan cepat dan efisien[14].

E. Weighted Naïve Bayes

Dasar akurasi klasifikasi ditentukan dengan menghitung bobot masing-masing karakteristik dalam kelas, tidak hanya probabilitas[8]. Pembobotan *Naïve Bayes* dihitung dengan cara menambahkan bobot pada setiap fitur atau atribut[15].

Perhitungan berikut dilakukan dengan menggunakan metode *Weighted Nave Bayes* dengan *Laplace Smoothing* untuk menentukan klasifikasinya:

1. Perhitungan nilai probabilitas pada setiap kelas, dirumuskan sebagai berikut:

$$P(C_i) = \frac{\sum C_i}{n} \quad (1)$$

Keterangan:

$P(C_i)$: Probabilitas label pada kelas C_i

$\sum C_i$: Jumlah data dengan label pada kelas C_i

n : Jumlah total pada data latih

2. Perhitungan nilai probabilitas pada setiap fitur. Rumus *laplace smoothing* yang dapat dituliskan sebagai berikut, digunakan untuk menentukan jumlah dan probabilitas data dengan nilai $K = 1$ untuk mencegah probabilitas yang bernilai 0.

$$P(x_k|C_i) = \frac{\sum x_k|C_i + K}{\sum C_i + K |x|} \quad (2)$$

Keterangan:

$P(x_k|C_i)$: Probabilitas fitur x_k dengan label pada kelas C_i

$\sum x_k|C_i$: Jumlah data fitur x_k dengan label pada kelas C_i

$\sum C_i$: Jumlah data dengan label pada kelas C_i

K : Parameter *smoothing*, $K=1$

$|x|$: Jumlah kelas pada sampel

3. Perhitungan nilai probabilitas pada setiap kelas dalam tiap data, dirumuskan sebagai berikut:

$$P(C_i|X) = P(C_i) \prod_{k=1}^n P(x_k|C_i)^{w_k} \quad (3)$$

Keterangan:

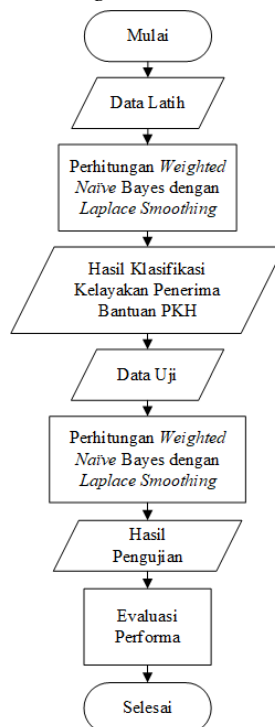
$P(C_i|X)$: Probabilitas pada kelas C_i dalam data X

$P(C_i)$: Probabilitas pada label kelas C_i

$P(x_k|C_i)$: Probabilitas pada fitur x_k dengan label kelas C_i

w_k : Bobot atribut

Pada Gambar 2 merupakan *flowchart* dari sistem klasifikasi kelayakan penerima bantuan PKH menggunakan metode *weighted naïve bayes* dengan *laplace smoothing*.



Gambar 2. *Flowchart* Sistem

F. Rank Order Centroid (ROC)

Rank Order Centroid (ROC) yaitu teknik pembobotan pada tiap kriteria sesuai dengan tingkat prioritas penilaian[16]. Konsep dasar dari metode pembobotan adalah dengan memberikan nilai bobot berdasarkan tingkat prioritas pada kriteria[17]. Biasanya dibentuk dengan pernyataan “kriteria 1 lebih penting dari kriteria 2” dan seterusnya sampai kriteria ke-n. Sehingga dapat dirumuskan sebagai berikut:

$$W_m = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{i}\right) \quad (4)$$

dengan W_m adalah Bobot kriteria, m adalah Jumlah kriteria, dan i adalah Kriteria.

G. Evaluasi Confusion Matrix

Confusion matrix yaitu tabel yang menunjukkan kumpulan data uji yang benar dan kumpulan data uji yang salah[18].

TABEL I
KONDISI *CONFUSION MATRIX*

| Two-Class Prediction | | Predicted Class | |
|----------------------|-----|-----------------|----------------|
| | | Yes | No |
| Actual Class | Yes | True Positive | False Negative |
| | No | False Positive | True Negative |

Pada Tabel 1 merupakan kondisi *confusion matrix*. Metode ini terdiri dari TP yang merupakan banyaknya data pada kelas positif yang dikategorikan dalam kelas positif, TN yaitu banyaknya data pada kelas negatif yang dikategorikan pada kelas negatif, FP yaitu banyaknya data kelas negatif yang dikategorikan dalam kelas positif, dan FN yaitu banyaknya pada data kelas positif yang dikategorikan dalam kelas negatif. Persamaan (5) digunakan untuk menghitung *accuracy*, sedangkan persamaan (6) untuk menghitung laju error, *Sensitivitas* dihitung menggunakan persamaan (7), dan *Spesifisitas* dihitung menggunakan persamaan (8)[19].

$$\begin{aligned} \text{Akurasi} &= \frac{\text{Jumlah data yang diprediksi secara benar}}{\text{Jumlah prediksi yang dilakukan}} \times 100 \\ &= \frac{TP+TN}{TP+FN+FP+TN} \times 100 \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Laju Error} &= \frac{\text{Jumlah data yang diprediksi secara salah}}{\text{Jumlah prediksi yang dilakukan}} \times 100 \\ &= \frac{FN+FP}{TP+FN+FP+TN} \times 100 \end{aligned} \quad (6)$$

$$\text{Sensitivitas} = \frac{TP}{TP+FN} \times 100 \quad (7)$$

$$\text{Spesifisitas} = \frac{TN}{FP+TN} \times 100 \quad (8)$$

III. HASIL DAN PEMBAHASAN

A. Dataset

Data latih dan data uji didapat dari data kepala keluarga Dusun Mando Desa Pandu Kecamatan Cerme Kabupaten Gresik Data yang digunakan tahun 2021. Dalam penelitian ini, atribut yang digunakan berasal dari data kepala keluarga yang sudah diperoleh dari desa. Atribut yang digunakan untuk menentukan kelayakan penerima bantuan PKH yaitu atribut usia, pekerjaan, penghasilan, jumlah keluarga, jumlah tanggungan anak sekolah, status rumah, jenis lantai dan jenis dinding dapat dilihat pada Tabel 2, sedangkan untuk mengklasifikasikan digunakan 2 kelas atau label, yaitu layak dan tidak layak.

TABEL II
ATRIBUT DATA

| No | Atribut | Tipe Atribut | Keterangan |
|----|------------------------|--------------|---|
| 1 | Usia | Numerik | Usia Kepala Keluarga |
| 2 | Pekerjaan | Kategorial | Pekerjaan Kepala Keluarga |
| 3 | Penghasilan | Numerik | Penghasilan Per bulan |
| 4 | Jumlah Keluarga | Numerik | Jumlah keluarga yang ada dalam kartu keluarga |
| 5 | Jumlah Tanggungan Anak | Numerik | Jumlah tanggungan anak yang masih bersekolah |
| 6 | Status Rumah | Kategorial | Status rumah yang di tempati |
| 7 | Jenis Lantai | Kategorial | Jenis lantai rumah |
| 8 | Jenis Dinding | Kategorial | Jenis dinding rumah |

Jumlah data yang digunakan dalam perhitungan adalah 80 data yang terbagi menjadi 56 data latih dengan 28 data sebaran disetiap kelas dan 24 data uji. Data latih berguna untuk menentukan nilai probabilitas dari setiap nilai fitur dan label kelas, dan data uji berguna sebagai pengujian terhadap label kelas prediksi dengan label kelas dari data nyata atau aktual.

Penentuan kategori layak pada penerimaan bantuan PKH yaitu masyarakat yang memenuhi kriteria-kriteria yang telah ditentukan oleh perangkat desa dengan jumlah atribut atau kriteria yang memenuhi berjumlah lebih dari 4. Kriteria dalam kategori layak dapat dilihat pada Tabel 3.

TABEL III
KRITERIA KATEGORI LAYAK PADA PENERIMA BANTUAN PKH

| No | Atribut | Keterangan |
|----|------------------------|------------------|
| 1 | Usia | ≥ 60 |
| 2 | Penghasilan | $\leq 1.000.000$ |
| 3 | Jumlah Keluarga | ≥ 4 |
| 4 | Jumlah Tanggungan Anak | ≥ 2 |
| 5 | Status Rumah | Sewa |
| 6 | Jenis Lantai | Tanah, Semen |
| 7 | Jenis Dinding | Kayu, Batu-Bata |

Data latih penerimaan bantuan PKH dapat dilihat pada Tabel 4 dan data uji penerimaan bantuan PKH dapat dilihat pada Tabel 5.

TABEL IV
DATA LATIH KLASIFIKASI KELAYAKAN PENERIMA BANTUAN PKH

| NO. | NAMA | USIA | PEKERJAAN | PENGHASILAN | JUMLAH KELUARGA | JUMLAH TANGGUGAN ANAK | STATUS RUMAH | JENIS LANTAI | JENIS DINDING | KELAS ASLI |
|-----|------|------|---------------|-------------|-----------------|-----------------------|--------------|--------------|---------------|-------------|
| 1 | L1 | 66 | Tidak Bekerja | 550.000 | 3 | 2 | Sewa | Kramik | Tembok | Layak |
| 2 | L2 | 42 | Pedagang | 950.000 | 3 | 2 | Sendiri | Semen | Batu Bata | Layak |
| 3 | L3 | 54 | Buruh Harian | 800.000 | 4 | 2 | Sendiri | Semen | Tembok | Layak |
| 4 | L4 | 68 | Tidak Bekerja | 650.000 | 4 | 0 | Sendiri | Semen | Tembok | Layak |
| 5 | L5 | 62 | Tidak Bekerja | 1.600.000 | 5 | 2 | Sendiri | Kramik | Batu Bata | Layak |
| 6 | L6 | 68 | Pedagang | 1.000.000 | 5 | 3 | Sendiri | Kramik | Tembok | Layak |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 56 | L56 | 35 | Wiraswasta | 950.000 | 3 | 1 | Sewa | Kramik | Tembok | Tidak Layak |

TABEL V
DATA UJI KLASIFIKASI KELAYAKAN PENERIMA BANTUAN PKH

| NO. | NAMA | USIA | PEKERJAAN | PENGHASILAN | JUMLAH KELUARGA | JUMLAH TANGGUGAN ANAK | STATUS RUMAH | JENIS LANTAI | JENIS DINDING |
|-----|------|------|-----------------|-------------|-----------------|-----------------------|--------------|--------------|---------------|
| 1 | U1 | 78 | Tidak Bekerja | 800.000 | 4 | 0 | Sewa | Semen | Batu Bata |
| 2 | U2 | 65 | Buruh Harian | 1.000.000 | 4 | 2 | Sewa | Semen | Batu Bata |
| 3 | U3 | 41 | Pedagang | 900.000 | 3 | 1 | Sewa | Semen | Batu Bata |
| 4 | U4 | 44 | Buruh Harian | 950.000 | 4 | 2 | Sendiri | Kramik | Tembok |
| 5 | U5 | 64 | Buruh Harian | 1.500.000 | 2 | 1 | Sendiri | Semen | Batu Bata |
| 6 | U6 | 82 | Buruh Harian | 700.000 | 2 | 1 | Sendiri | Semen | Batu Bata |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 24 | U24 | 48 | Karyawan Swasta | 3.000.000 | 3 | 2 | Sendiri | Kramik | Tembok |

B. Data Processing

Proses *preprocessing* dilakukan pada data latih dan data uji terlebih dahulu untuk mengubah tipe fitur atau atribut agar sesuai dengan tipe kategorial atribut, dapat dilihat pada Tabel 6.

TABEL VI
TIPE ATRIBUT

| No | Atribut | Tipe Atribut | Tipe Atribut Setelah Preprocessing |
|----|------------------------|--------------|--|
| 1 | Usia | Numerik | Kategorial (Rendah, Sedang, Tinggi) |
| 2 | Pekerjaan | Kategorial | Kategorial (Tidak Bekerja, Pedagang, Karyawan Swasta, Wiraswasta, PNS, Buruh Harian, Petani) |
| 3 | Penghasilan | Numerik | Kategorial (Rendah, Sedang, Tinggi) |
| 4 | Jumlah Keluarga | Numerik | Kategorial (Rendah, Sedang, Tinggi) |
| 5 | Jumlah Tanggungan Anak | Numerik | Kategorial (Rendah, Sedang, Tinggi) |

| | | | |
|---|---------------|------------|--------------------------------------|
| 6 | Status Rumah | Kategorial | Kategorial (Sewa, Sendiri) |
| 7 | Jenis Lantai | Kategorial | Kategorial (Tanah, Semen, Kramik) |
| 8 | Jenis Dinding | Kategorial | Kategorial (Kayu, Batu-bata, Tembok) |

Mengubah tipe atribut numerik menjadi kategorial

1. Menentukan Nilai Kategorial Usia, dapat dilihat pada Tabel 7.

TABEL VII
KATEGORIAL USIA

| No | Nilai Usia | Kategorial Usia |
|----|------------|-----------------|
| 1 | ≤ 40 | Rendah |
| 2 | 41- 64 | Sedang |
| 3 | ≥ 65 | Tinggi |

2. Menentukan Nilai Kategorial Penghasilan, dapat dilihat pada Tabel 8.

TABEL VIII
KATEGORIAL PENGHASILAN

| No | Nilai Penghasilan | Kategorial Penghasilan |
|----|-----------------------|------------------------|
| 1 | ≤ 1.000.000 | Rendah |
| 2 | 1.100.000 - 3.000.000 | Sedang |
| 3 | ≥ 3.100.000 | Tinggi |

3. Menentukan Nilai Kategorial Jumlah Keluarga, dapat dilihat pada Tabel 9.

TABEL IX
KATEGORIAL JUMLAH KELUARGA

| No | Nilai Jml Keluarga | Kategorial Jml Keluarga |
|----|--------------------|-------------------------|
| 1 | ≤ 1 | Rendah |
| 2 | 1- 3 | Sedang |
| 3 | ≥ 4 | Tinggi |

4. Menentukan Nilai Kategorial Jumlah Tanggungan Anak, dapat dilihat pada Tabel 10.

TABEL X
KATEGORIAL JUMLAH TANGGUNGAN ANAK

| No | Nilai Jml Tanggungan anak | Kategorial Jml Tanggungan Anak |
|----|---------------------------|--------------------------------|
| 1 | 0 | Rendah |
| 2 | 1- 2 | Sedang |
| 3 | ≥ 3 | Tinggi |

Pada Tabel 11 merupakan hasil *preprocessing* data latih dan pada Tabel 12 merupakan hasil *preprocessing* data uji.

TABEL XI
HASIL *PREPROCESSING* DATA LATIH

| NO. | NAMA | USIA | PEKERJAAN | PENGHASILAN | JUMLAH KELUARGA | JUMLAH TANGGUNGAN ANAK | STATUS RUMAH | JENIS LANTAI | JENIS DINDING | KELAS ASLI |
|-----|------|--------|---------------|-------------|-----------------|------------------------|--------------|--------------|---------------|-------------|
| 1 | L1 | Tinggi | Tidak Bekerja | Rendah | Sedang | Sedang | Sewa | Kramik | Tembok | Layak |
| 2 | L2 | Sedang | Pedagang | Rendah | Sedang | Sedang | Sendiri | Semen | Batu Bata | Layak |
| 3 | L3 | Sedang | Buruh Harian | Rendah | Tinggi | Sedang | Sendiri | Semen | Tembok | Layak |
| 4 | L4 | Tinggi | Tidak Bekerja | Rendah | Tinggi | Rendah | Sendiri | Semen | Tembok | Layak |
| 5 | L5 | Sedang | Tidak Bekerja | Sedang | Tinggi | Sedang | Sendiri | Kramik | Batu Bata | Layak |
| 6 | L6 | Tinggi | Pedagang | Rendah | Tinggi | Tinggi | Sendiri | Kramik | Tembok | Layak |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 56 | L56 | Rendah | Wiraswasta | Rendah | Sedang | Sedang | Sewa | Kramik | Tembok | Tidak Layak |

TABEL XII
HASIL *PREPROCESSING* DATA UJI

| NO. | NAMA | USIA | PEKERJAAN | PENGHASILAN | JUMLAH KELUARGA | JUMLAH TANGGUNGAN ANAK | STATUS RUMAH | JENIS LANTAI | JENIS DINDING |
|-----|------|--------|-----------------|-------------|-----------------|------------------------|--------------|--------------|---------------|
| 1 | U1 | Tinggi | Tidak Bekerja | Rendah | Tinggi | Rendah | Sewa | Semen | Batu Bata |
| 2 | U2 | Tinggi | Buruh Harian | Rendah | Tinggi | Sedang | Sewa | Semen | Batu Bata |
| 3 | U3 | Sedang | Pedagang | Rendah | Sedang | Sedang | Sewa | Semen | Batu Bata |
| 4 | U4 | Sedang | Buruh Harian | Rendah | Tinggi | Sedang | Sendiri | Kramik | Tembok |
| 5 | U5 | Sedang | Buruh Harian | Sedang | Sedang | Sedang | Sendiri | Semen | Batu Bata |
| 6 | U6 | Tinggi | Buruh Harian | Rendah | Sedang | Sedang | Sendiri | Semen | Batu Bata |
| .. | .. | .. | .. | .. | .. | .. | .. | .. | .. |
| 24 | U24 | Sedang | Karyawan Swasta | Sedang | Sedang | Sedang | Sendiri | Kramik | Tembok |

C. Perhitungan Algoritma Weighted Naïve Bayes dengan Laplace Smoothing

Langkah-langkah dalam perhitungan untuk mendapatkan nilai probabilitas pada label kelas Kelayakan Penyerimaan bantuan PKH pada Hasil *Preprocessing* Data Latih sebagai berikut:

1. Perhitungan nilai probabilitas pada setiap kelas dengan menggunakan persamaan (1)

Label Kelas layak:

$$P(1S) = \frac{\sum 1}{n} = \frac{28}{56} = 0,5$$

Label Kelas tidak layak:

$$P(1S) = \frac{\sum 1}{n} = \frac{28}{56} = 0,5$$

2. Perhitungan nilai probabilitas pada setiap fitur. Rumus *Laplace smoothing* yang dapat dituliskan sebagai berikut, digunakan untuk menentukan jumlah dan probabilitas data dengan nilai $K = 1$ untuk mencegah probabilitas yang bernilai 0 dengan menggunakan persamaan (2).

Contoh perhitungan probabilitas pada fitur Usia:

Kelas Layak

$$P_{(R|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{2 + 1}{28 + 1(2)} = \frac{3}{30} = 0,1$$

Kelas Tidak Layak

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{7 + 1}{28 + 1(2)} = \frac{8}{30} = 0,27$$

$$P_{(S|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{17 + 1}{28 + 1(2)} = \frac{18}{30} = 0,6$$

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{20 + 1}{28 + 1(2)} = \frac{21}{30} = 0,7$$

$$P_{(T|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{9 + 1}{28 + 1(2)} = \frac{10}{30} = 0,33$$

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{1 + 1}{28 + 1(2)} = \frac{2}{30} = 0,07$$

3. Perhitungan Sebaran Klasifikasi Data Latih menggunakan metode *Weighted Naïve Bayes* dengan *Laplace Smoothing* menggunakan persamaan (3).

Metode *Rank Order Centroid* digunakan untuk menghitung bobot preferensi atribut data menggunakan penentuan prioritas atribut oleh sekretaris Desa Pandu dengan rumus perhitungan menggunakan persamaan (4). Bobot preferensi atribut atribut dapat dilihat pada Tabel 13.

TABEL XIII
 BOBOT PREFERENSI ATRIBUT

| No | Kode | Keterangan | Bobot |
|----|------|------------------------|-------|
| 1 | C1 | Penghasilan | 0,34 |
| 2 | C2 | Jumlah Tanggungan Anak | 0,32 |
| 3 | C3 | Usia | 0,31 |
| 4 | C4 | Jumlah Keluarga | 0,29 |
| 5 | C5 | Pekerjaan | 0,26 |
| 6 | C6 | Status Rumah | 0,23 |
| 7 | C7 | Jenis Lantai | 0,19 |

Contoh perhitungan pada data latih ke-1

Kelas Layak

$$P(\text{layak}|X) = 0,5 \times 0,33^{0,31} \times 0,37^{0,26} \times 0,60^{0,34} \times 0,30^{0,29} \times 0,57^{0,32} \times 0,27^{0,23} \times 0,30^{0,19} \times 0,47^{0,13}$$

$$P(\text{layak}|X) = 0,07$$

Kelas Tidak Layak

$$P(\text{tidak layak}|X) = 0,5 \times 0,27^{0,31} \times 0,03^{0,27} \times 0,07^{0,34} \times 0,60^{0,29} \times 0,50^{0,32} \times 0,13^{0,23} \times 0,83^{0,19} \times 0,93^{0,13}$$

$$P(\text{tidak layak}|X) = 0,01$$

Dari perhitungan tersebut, hasil terbesar diperoleh pada $P(\text{layak}|X)$. Maka untuk data latih ke-1 diklasifikasikan pada kelas layak. Hasil klasifikasi pada data latih kelayakan penerimaan bantuan PKH dapat dilihat pada tabel 14.

TABEL XIV
 HASIL KLASIFIKASI DATA LATIH

| No | Layak | Tidak Layak | Status |
|----|-------|-------------|-------------|
| 1 | 0,07 | 0,01 | Layak |
| 2 | 0,09 | 0,03 | Layak |
| 3 | 0,16 | 0,05 | Layak |
| 4 | 0,11 | 0,01 | Layak |
| 5 | 0,12 | 0,05 | Layak |
| 6 | 0,05 | 0,01 | Layak |
| .. | .. | .. | .. |
| 56 | 0,03 | 0,04 | Tidak Layak |

4. Perhitungan nilai probabilitas akhir pada setiap kelas dengan menggunakan persamaan (1)

Label Kelas layak:

$$P(\text{layak}) = \frac{\sum 1S}{n} = \frac{30}{56} = 0,54$$

Label Kelas tidak layak:

$$P(\text{tidak layak}) = \frac{\sum 2S}{n} = \frac{26}{56} = 0,46$$

5. Perhitungan nilai probabilitas akhir tiap fitur dengan menggunakan persamaan (2)
Perhitungan probabilitas akhir pada fitur Usia:

$$P_{(R|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{2 + 1}{30 + 1(2)} = \frac{3}{32} = 0,09$$

$$P_{(R|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{19 + 1}{30 + 1(2)} = \frac{20}{32} = 0,63$$

$$P_{(R|L)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{9 + 1}{30 + 1(2)} = \frac{10}{32} = 0,31$$

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{7 + 1}{26 + 1(2)} = \frac{8}{28} = 0,29$$

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{18 + 1}{26 + 1(2)} = \frac{19}{28} = 0,68$$

$$P_{(R|TL)} = \frac{\sum x_k | C_i + K}{\sum C_i + K | x|} = \frac{1 + 1}{26 + 1(2)} = \frac{2}{28} = 0,07$$

Setelah dilakukan perhitungan pada data latih, nilai probabilitas akhir yang didapatkan pada data latih digunakan dalam pengujian data uji dengan menggunakan metode *weighted naïve bayes* dengan *laplace smoothing*. Hasil pengujian klasifikasi pada data uji dapat dilihat pada Tabel 14.

TABEL XIV
HASIL PENGUJIAN KLASIFIKASI DATA UJI

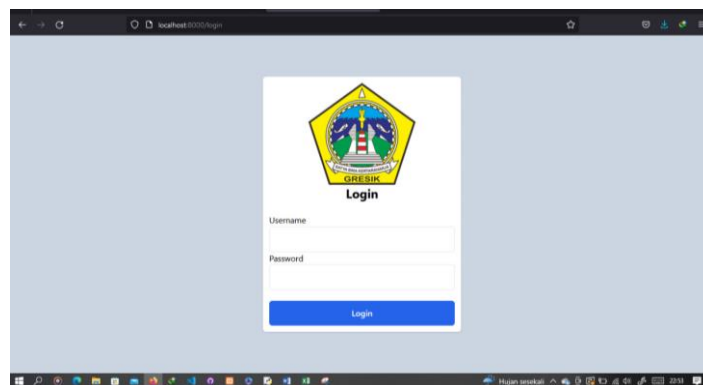
| No. | Kelas | | Status | | Akurasi |
|-----|-------|-------------|-------------|-------------|---------|
| | Layak | Tidak Layak | Prediksi | Asli | |
| 1 | 0,09 | 0,00 | Layak | Layak | Benar |
| 2 | 0,12 | 0,01 | Layak | Layak | Benar |
| 3 | 0,07 | 0,02 | Layak | Layak | Benar |
| 4 | 0,16 | 0,05 | Layak | Layak | Benar |
| 5 | 0,13 | 0,04 | Layak | Layak | Benar |
| 6 | 0,11 | 0,01 | Layak | Layak | Benar |
| .. | .. | .. | .. | .. | .. |
| 24 | 0,06 | 0,15 | Tidak Layak | Tidak Layak | Benar |

D. Implementasi Sistem

Pengembangan sistem klasifikasi kelayakan penerima bantuan PKH menggunakan software web server “XAMPP” dan menggunakan teks editor “Visual Studio Code (VsCode)”. Pengembangan sistem berbasis website dengan bahasa pemrograman HTML, CSS, PHP, dan JavaScript.

1. Halaman Login

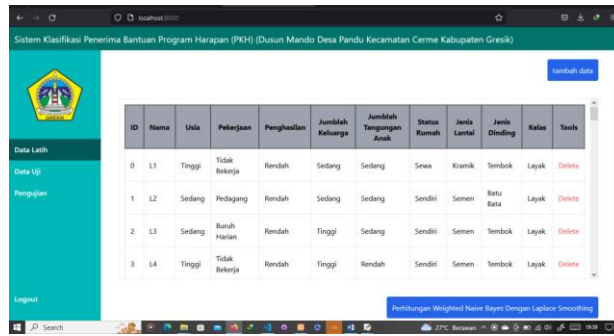
Pada halaman *login*, *user* memasukkan *email* dan *password* yang sesuai dengan database yang telah ditentukan oleh sistem. Jika *user* tidak dapat masuk pada sistem, maka *email* dan *password* yang dimasukkan tidak sesuai dengan *database*. Halaman *login* dapat dilihat pada Gambar 3.



Gambar 3. Halaman Login

2. Halaman Data latih

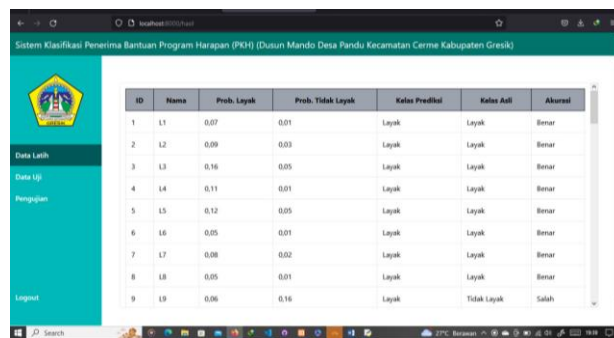
Halaman data latih berisi data warga dusun berupa nama dan kriteria penilaian. Halaman ini juga dapat digunakan untuk menghapus atau menambahkan data latih baru kedalam *database* sistem. Pada halaman data latih terdapat fungsi perhitungan *weighted naïve bayes* dengan *laplace smoothing*. Disajikan pada Gambar 4.



| ID | Nama | Ura | Pekerjaan | Penghasilan | Jumlah Keluarga | Jumlah Tanggungan Anak | Status Rumah | Jenis Lantai | Jenis Dinding | Kelas | Tools |
|----|------|--------|---------------|-------------|-----------------|------------------------|--------------|--------------|---------------|-------|--------|
| 0 | L1 | Tinggi | Tidak Bekerja | Rendah | Sedang | Sedang | Sewa | Kramik | Tembok | Layak | Delete |
| 1 | L2 | Sedang | Pedagang | Rendah | Sedang | Sedang | Sendiri | Semen | Batu Batu | Layak | Delete |
| 2 | L3 | Sedang | Buruh Harian | Rendah | Tinggi | Sedang | Sendiri | Semen | Tembok | Layak | Delete |
| 3 | L4 | Tinggi | Tidak Bekerja | Rendah | Tinggi | Rendah | Sendiri | Semen | Tembok | Layak | Delete |

Gambar 4. Halaman Data Latih

Pada halaman hasil klasifikasi data latih menampilkan hasil perhitungan *weighted naïve bayes* dengan *laplace smoothing* berupa nilai probabilitas yang akan digunakan pada data uji. Disajikan pada Gambar 5.

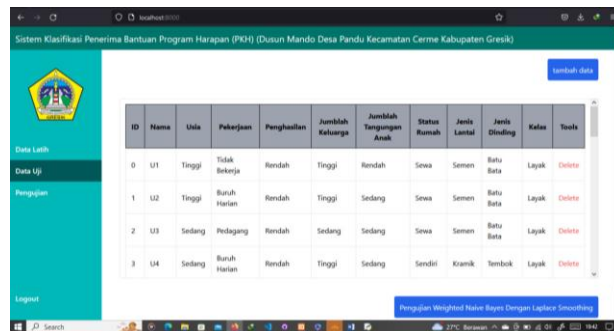


| ID | Nama | Prob. Layak | Prob. Tidak Layak | Kelas Prediksi | Kelas Asli | Akurasi |
|----|------|-------------|-------------------|----------------|-------------|---------|
| 1 | L1 | 0,07 | 0,01 | Layak | Layak | Benar |
| 2 | L2 | 0,09 | 0,03 | Layak | Layak | Benar |
| 3 | L3 | 0,16 | 0,05 | Layak | Layak | Benar |
| 4 | L4 | 0,11 | 0,01 | Layak | Layak | Benar |
| 5 | L5 | 0,12 | 0,05 | Layak | Layak | Benar |
| 6 | L6 | 0,05 | 0,01 | Layak | Layak | Benar |
| 7 | L7 | 0,08 | 0,02 | Layak | Layak | Benar |
| 8 | L8 | 0,05 | 0,01 | Layak | Layak | Benar |
| 9 | L9 | 0,06 | 0,16 | Layak | Tidak Layak | Salah |

Gambar 5. Halaman Hasil Klasifikasi Data Latih

3. Halaman Data Uji


Halaman data uji berisi data warga dusun berupa nama dan kriteria penilaian. Halaman ini juga dapat digunakan untuk menghapus atau menambahkan data latih baru kedalam *database* sistem. Pada halaman data uji terdapat fungsi pengujian metode *weighted naïve bayes* dengan *laplace smoothing*. Disajikan pada Gambar 6.



| ID | Nama | Ura | Pekerjaan | Penghasilan | Jumlah Keluarga | Jumlah Tanggungan Anak | Status Rumah | Jenis Lantai | Jenis Dinding | Kelas | Tools |
|----|------|--------|---------------|-------------|-----------------|------------------------|--------------|--------------|---------------|-------|--------|
| 0 | U1 | Tinggi | Tidak Bekerja | Rendah | Tinggi | Rendah | Sewa | Semen | Batu Batu | Layak | Delete |
| 1 | U2 | Tinggi | Buruh Harian | Rendah | Tinggi | Sedang | Sewa | Semen | Batu Batu | Layak | Delete |
| 2 | U3 | Sedang | Pedagang | Rendah | Sedang | Sedang | Sewa | Semen | Batu Batu | Layak | Delete |
| 3 | U4 | Sedang | Buruh Harian | Rendah | Tinggi | Sedang | Sendiri | Kramik | Tembok | Layak | Delete |

Gambar 6. Halaman Data Uji

Pada halaman hasil klasifikasi data uji menampilkan hasil pengujian *weighted naïve bayes* dengan *laplace smoothing*. Pada halaman ini menampilkan nilai probabilitas dan diperoleh hasil akurasi dari pengujian data uji. Disajikan pada Gambar 7.



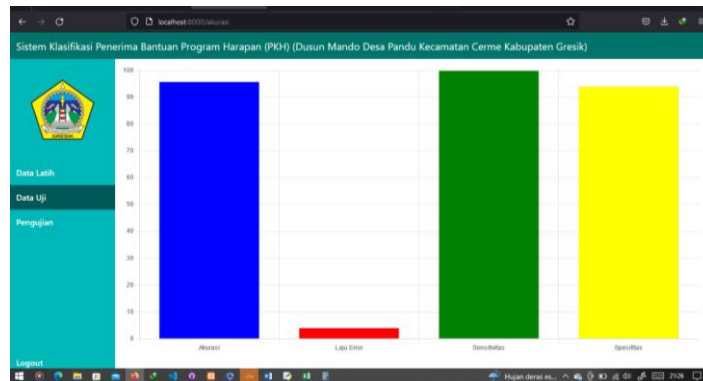
| ID | Nama | Prob. Layak | Prob. Tidak Layak | Kelas Prediksi | Kelas Asli | Akurasi |
|----|------|-------------|-------------------|----------------|------------|---------|
| 1 | U1 | 0,09 | 0,00 | Layak | Layak | Benar |
| 2 | U2 | 0,12 | 0,01 | Layak | Layak | Benar |
| 3 | U3 | 0,07 | 0,02 | Layak | Layak | Benar |
| 4 | U4 | 0,16 | 0,05 | Layak | Layak | Benar |
| 5 | U5 | 0,13 | 0,04 | Layak | Layak | Benar |

Total data: 24 Data Benar: 23 Data Salah: 1 Akurasi: 95.83

Gambar 7. Halaman Hasil Klasifikasi Data Uji

4. Halaman Evaluasi Performa

Halaman evaluasi performa *confussion matrix* menampilkan grafik nilai *accuracy*, *laju eror*, *sensitivitas* dan *spesifisitas* pada data uji. Halaman tersebut dapat dilihat melalui button “Cek Akurasi”. Disajikan pada Gambar 8.



Gambar 8. Halaman Evaluasi Performa *Confussion Matrix*

E. Pengujian Sistem

Pengujian sistem klasifikasi kelayakan penerima bantuan PKH menggunakan metode *confusion matrix*. Pengujian ini akan menghitung nilai dari *accuracy*, *laju eror*, *sensitivitas* dan *spesifisitas* pada data uji. *Accuracy* digunakan untuk menentukan data mana dari kelas asli dan kelas prediksi yang dikategorikan dengan benar sehingga akurasi hasil prediksi dapat ditentukan. *Laju eror* digunakan untuk mengetahui data yang diklasifikasikan atau diprediksi secara salah dari kelas asli dan kelas prediksi data yang di uji. *Sensitivitas* atau di sebut juga *true positive rate* atau *recall* digunakan untuk mengukur proporsi positif asli yang di kenali diprediksi secara benar sebagai positif asli dan *spesifisitas* atau di sebut *true negative rate* digunakan untuk mengukur proporsi negative asli yang dikenal (diprediksi) secara benar sebagai negative asli. Hasil pengujian menggunakan *confusion matrix* dapat dilihat pada Tabel 15.

TABEL XV
HASIL PENGUJIAN *CONFUSION MATRIX*

| | | Kelas Hasil Prediksi | |
|------------|-------------|----------------------|-------------|
| | | Layak | Tidak Layak |
| Kelas Asli | Layak | 7 | 0 |
| | Tidak Layak | 1 | 16 |

$$\text{Akurasi} = \frac{7+16}{7+1+0+16} \times 100 = 95,83$$

$$\text{Laju Error} = \frac{0+1}{7+1+0+16} \times 100 = 4,14$$

$$\text{Sensitivitas} = \frac{7}{7+1} \times 100 = 100,00$$

$$\text{Spesifisitas} = \frac{16}{1+16} \times 100 = 94,12$$

Pada penelitian ini klasifikasi kelayakan penerima bantuan PKH dengan menggunakan *weighted naïve bayes* dengan *laplace smoothing* didapatkan nilai akurasi sebesar 95,83%. Apabila dibandingkan dengan penelitian sebelumnya dengan menggunakan *weighted naïve bayes* pada kasus klasifikasi tanpa menggunakan *laplace smoothing* mendapatkan hasil akurasi sebesar 91,47% [8], dari pembahasan tersebut menjelaskan bahwa dengan menggunakan teknik *laplace smoothing* mendapatkan nilai akurasi yang lebih baik. Pada penelitian kelayakan penerima bantuan PKH mendapatkan nilai laju eror yang bagus yaitu sebesar 4,14%. Nilai *sensitivitas* yang didapatkan sebesar 100% dan nilai dari *spesifisitas* pada penelitian ini sebesar 94,12%.

IV. KESIMPULAN

Hasil dari penelitian dengan judul klasifikasi kelayakan penerima bantuan Program Keluarga Harapan (PKH) menggunakan metode *weighted naïve bayes* dengan *laplace smoothing* dapat membantu perangkat desa dalam menentukan warga desa yang layak menerima bantuan tersebut. Metode *laplace smoothing* berhasil meminimalkan kesalahan klasifikasi, karena dapat mencegah probabilitas dengan nilai 0. Dari hasil perhitungan klasifikasi menggunakan metode *naïve bayes* dengan *laplace smoothing* pada 56 data training dan 24 data uji didapatkan hasil pengujian evaluasi performa dengan menggunakan *confusion matrix* dengan nilai *Accurasi* 95,83%, *eror* 4,14%, *Sensitivitas* 100,00%, dan *Spesifisitas* 94,12%. Pada penelitian selanjutnya agar hasil klasifikasi semakin bagus

maka bisa ditambahkan lagi jumlah atribut kriteria yang digunakan dalam klasifikasi ini, selain itu bisa dicoba untuk di implementasikan metode modifikasi lain dari *weighted naive bayes* untuk mengetahui perbandingan hasil performa.

DAFTAR PUSTAKA

- [1] I. P. Pertiwi, F. Fedinandus, dan A. D. Limantara, "Sistem Pendukung Keputusan Penerima Program Keluarga Harapan (PKH) Menggunakan Metode Simple Additive Weighting," *CAHAYATECH*, 2019, doi: 10.47047/ct.v8i2.46.
- [2] M. S. Mustafa, M. R. Ramadhan, dan A. P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier," *Creat. Inf. Technol. J.*, 2018, doi: 10.24076/citec.2017v4i2.106.
- [3] A. C. Fauzan dan K. Hikmah, "IMPLEMENTASI ALGORITMA NAIVE BAYES DALAM ANALISIS POLARISASI OPINI MASYARAKAT TERKAIT VAKSIN COVID-19," *Rabit J. Teknol. dan Sist. Inf. Univrab*, vol. 7, no. 2, hal. 122–128, 2022.
- [4] N. Ramadhani dan N. Fajarianto, "Sistem Informasi Evaluasi Perkuliahan dengan Sentimen Analisis Menggunakan Naive Bayes dan Smoothing Laplace," *JSINBIS (Jurnal Sist. Inf. Bisnis)*, vol. 10, no. 2, hal. 228–234, 2020.
- [5] I. Listiowarni dan E. R. Setyaningsih, "Analisis Kinerja Smoothing pada Naive Bayes untuk Pengkategorian Soal Ujian," *J. Teknol. dan Manaj. Inform.*, vol. 4, no. 2, 2018.
- [6] M. R. Qisthiano, T. B. Kurniawan, E. S. Negara, dan M. Akbar, "Pengembangan Model Untuk Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu dengan Metode Naive Bayes," *J. MEDIA Inform. BUDIDARMA*, 2021, doi: 10.30865/mib.v5i3.3030.
- [7] I. Listiowarni, "Implementasi Naive Bayessian dengan Laplacian Smoothing untuk Peminatan dan Lintas Minat Siswa SMAN 5 Pamekasan," *J. Sisfokom (Sistem Inf. dan Komputer)*, 2019, doi: 10.32736/sisfokom.v8i2.652.
- [8] A. I. S. Azis dan B. Santoso, "LL-KNN ACW-NB: Local Learning K-Nearest Neighbor in Absolute Correlation Weighted Naive Bayes for Numerical Data Classification," *J. RESTI (Rekayasa Sist. Dan Teknol. Informatika)*, vol. 4, no. 1, hal. 28–36, 2020.
- [9] A. W. N. Ulfy dan P. A. R. Devi, "Penentuan Kenaikan Jabatan Menggunakan Pembobotan Metode AHP dan Didukung Metode Complex Proportional Assessment," *J. Sist. Komput. dan Inform.*, vol. 3, no. 3, hal. 232–240, 2022.
- [10] A. Oliver, "Apa Itu Data Preprocessing?," *glints.com*, 2021.
- [11] I. K. Yoga, S. S. Prasetyowati, dan Y. Sibaroni, "PREDICTION AND MAPPING RAINFALL CLASSIFICATION USING NAIVE BAYES AND SIMPLE KRIGING," *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 7, no. 4, hal. 1244–1253, 2022.
- [12] A. D. S. Academy, "FUNGSI PREPROCESSING PADA DATA MINING," *algorit.ma*, 2022.
- [13] I. Mubarog, A. Setyanto, dan H. Sismoro, "Sistem Klasifikasi Pada Penyakit Breast Cancer Dengan Menggunakan Metode Naive Bayes," *Creat. Inf. Technol. J.*, 2021, doi: 10.24076/citec.2019v6i2.246.
- [14] C. P. Lubis, R. Rosnelly, R. Roslina, Z. Situmorang, dan W. Wanayumini, "PENERAPAN METODE NAIVE BAYES DAN C4.5 PADA PENERIMAAN PEGAWAI DI UNIVERSITAS POTENSI UTAMA," *CSRID (Computer Sci. Res. Its Dev. Journal)*, 2021, doi: 10.22303/csrid.12.1.2020.51-62.
- [15] K. Umam, D. Arifianto, dan M. Kom, "Metode Optimasi Pembobotan Gain Ratio Terhadap Metode Klasifikasi Weighted Naive Bayes Studi Kasus Ulasan Produk White Perfect Clinical Day Cream".
- [16] M. Al Farosa, P. Kasih, dan R. H. Irawan, "Pemodelan Algoritma ROC Dalam Pembobotan Kriteria Seleksi Penerima Bantuan Sosial Pendidikan Menggunakan Algoritma CPI," in *Prosiding SEMNAS INOTEK (Seminar Nasional Inovasi Teknologi)*, 2022, vol. 6, no. 1, hal. 332–337.
- [17] R. K. Ndruru, "Penerapan Metode Additive Ratio Assessment (ARAS) dan Rank Order Centroid (ROC) Dalam Pemilihan Jaksa Terbaik Pada Kejaksaan Negeri Medan," *Semin. Nas. Teknol. Komput. Sains*, 2020.
- [18] D. Normawati dan S. A. Prayogi, "Implementasi Naive Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J-SAKTI (Jurnal Sains Komput. dan Inform.)*, vol. 5, no. 2, hal. 697–711, 2021.
- [19] N. M. A. J. Astari, Dewa Gede Hendra Divayana, dan Gede Indrawan, "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier," *J. Sist. dan Inform.*, 2020, doi: 10.30864/jsi.v15i1.332.