

## Analisis Penerapan Metode *Naïve Bayes* Pada Karakteristik Penerima PKH Untuk Lansia Dan Anak Sekolah

Putri Fatikhatul Jannah<sup>1</sup>, Budanis Dwi Meilani<sup>2\*</sup>, Shah Khadafi<sup>3</sup>

<sup>1,2,3</sup>Institut Teknologi Adhi Tama Surabaya

\*Penulis Korespondensi : [budanis@itats.ac.id](mailto:budanis@itats.ac.id)

### ABSTRACT

The distribution of social assistance through Indonesia's Family Hope Program (PKH) often faces challenges in accurately identifying eligible recipients, particularly among the elderly and school-aged children. This study aims to analyse the characteristics of PKH beneficiaries using the Naïve Bayes classification method in data mining. The research utilized resident data from the Sukolilo Baru Sub-district, Surabaya, comprising 10,676 records. These records underwent data cleaning and discretization to produce a relevant and usable dataset. The attributes included age, name initials, number of dependents, number of elderly individuals, number of school-aged children, head of household income, BPJS membership status, type of vehicle, water source, toilet type, building ownership status, floor and wall materials, building size category, and aid eligibility label. A classification system was then developed as a web-based application using the Naïve Bayes method to assess eligibility for assistance. The results demonstrated that this method effectively supports the accurate identification of aid recipients. Evaluation showed a high level of classification accuracy. Specifically, testing revealed that a combination of three datasets (Mix Data, PKH, and Resident data) achieved the highest and most consistent performance by an average accuracy of 97.83%. Moreover, 8 out of 10 classification patterns reached 100% accuracy. The three-dataset combination proved to be the most stable and suitable model for further development. This study can serve as a reference for related parties to enhance the efficiency and effectiveness of targeted PKH social assistance distribution.

### Article History

Received : 11-07-2026  
Revised : 27-07-2026  
Accepted : 30-07-2026

### Key words

*Naive Bayes,*  
*PKH, Elderly,*  
*School Age*  
*Children,*  
*Data mining.*

### ABSTRAK

Distribusi bantuan sosial Program Keluarga Harapan (PKH) di Indonesia sering kali menemui kendala dalam ketepatan sasaran penerima, khususnya pada kelompok lansia dan anak sekolah. Penelitian ini bertujuan untuk menganalisis karakteristik penerima bantuan PKH dengan menggunakan *data mining* metode klasifikasi *Naive Bayes*. Penelitian dilakukan pada data warga di Kelurahan Sukolilo Baru Surabaya, dengan total 10.676 data, yang kemudian diproses melalui tahap *data cleaning* dan diskritisasi data untuk menghasilkan data yang relevan. Atribut yang dianalisis meliputi usia, inisial nama, tanggungan kepala keluarga, jumlah lansia, jumlah anak wajib sekolah, penghasilan kepala keluarga, status bpjs, tipe kendaraan, sumber air, tipe jamban, status kepemilikan bangunan, bahan lantai, bahan dinding, kategori luas bangunan, dan label. Sistem klasifikasi kemudian dibangun dalam bentuk aplikasi berbasis *website* menggunakan metode *Naive Bayes* untuk menentukan kelayakan penerima bantuan. Hasil pengujian menunjukkan bahwa metode ini mampu membantu dalam memetakan penerima bantuan secara lebih akurat, dengan evaluasi yang menunjukkan tingkat keakuratan tinggi. Dari pengujian, pola kombinasi 3 *file* data (*Mix Data*, PKH, Warga) memberikan hasil akurasi tertinggi dan paling stabil dari pola lainnya dengan rata-rata 97,83% dan 8 dari 10 pola menghasilkan akurasi 100%. Kombinasi pola 3 *file* terbukti lebih stabil dan lebih baik untuk menjadi model pada pengujian selanjutnya. Penelitian ini diharapkan dapat menjadi acuan bagi pihak terkait dalam meningkatkan efisiensi dan efektivitas distribusi bantuan sosial PKH secara tepat sasaran.

### PENDAHULUAN

Kesejahteraan sosial adalah salah satu prioritas utama pemerintah Indonesia dalam upaya mengurangi kemiskinan dan meningkatkan kualitas hidup masyarakat. Pada Maret 2024, jumlah penduduk miskin di Indonesia mencapai 25,22 juta orang [1].



Program-program seperti Bantuan Pangan Non Tunai (BPNT), Bantuan Sosial Tunai (BST), Program Keluarga Harapan (PKH), Penerima Bantuan Iuran - Jaminan Kesehatan (PBI-JK), Bantuan Langsung Tunai – Bahan Bakar Minyak (BLT-BBM), Bantuan Yatim Piatu, Rumah Sejahtera Terpadu (RST), Permakanan, dan Sembako Adaptif menjadi andalan pemerintah.

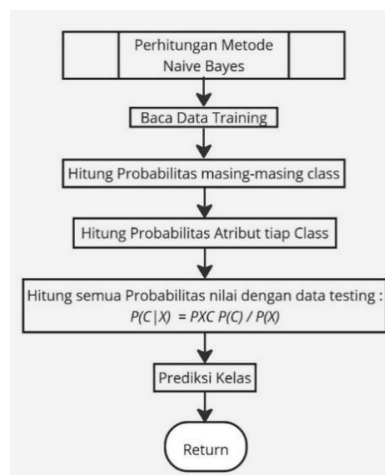
Namun, distribusi bansos di Indonesia masih menghadapi sejumlah masalah, terutama ketidaktepatan sasaran penerima manfaat. Salah satu elemen penting yang mendukung strategi penanggulangan kemiskinan adalah ketersediaan statistik kemiskinan yang tepat dan terfokus. Untuk membantu orang-orang yang membutuhkan, pemerintah biasanya mengadakan program bantuan sembako. Keluarga Penerima Manfaat (KPM) menerima bantuan pangan nontunai (BPNT) dan bantuan sosial (Program Sembako) setiap bulan melalui sistem perbankan. Namun, dalam praktiknya, pemberian dorongan sembako tidak sesuai dengan sarannya. Akibatnya, pendataan yang lebih akurat perlu dilakukan terhadap keluarga yang layak maupun tidak layak menerima dorongan sembako. Sebuah pendekatan *data mining* dapat digunakan untuk menentukan kelayakan penerima sembako [2].

Pelayanan di Desa Pesanggrahan yang menerima bantuan Surat Keterangan Tidak Mampu (SKTM). Program bantuan seperti Surat Keterangan Tidak Mampu (SKTM) adalah surat yang dikeluarkan oleh kelurahan atau desa dan kemudian diketahui oleh camat bagi keluarga miskin untuk mendapatkan bantuan dalam hal kesehatan, keuangan, dan pendidikan. Namun proses pemilihan penerima bantuan sering kali menghadapi tantangan, seperti ketidak akurat dalam penentuan kriteria penerima dan potensi salah sasaran. Oleh karena itu aparat desa akan lebih sistematis dan berbasis data untuk meningkatkan efisiensi dalam klasifikasi penerima bantuan [3]. Penelitian ini bertujuan untuk meningkatkan akurasi penentuan penerima bantuan sosial menggunakan metode klasifikasi *Naive Bayes*. Dengan keakuratan yang lebih baik, program PKH diharapkan dapat semakin efektif dalam membantu keluarga miskin dan memutus rantai kemiskinan [4].

## METODE

### Klasifikasi

Klasifikasi adalah teknik *data mining* yang sering digunakan untuk membuat prediksi atau keputusan, seperti menentukan kelayakan penerima bantuan sosial. Tujuan klasifikasi adalah untuk mengelompokkan data ke dalam kategori tertentu berdasarkan karakteristik yang telah ditentukan [6].



Gambar 1. Flowchart algoritma *Naive Bayes*



## Naïve Bayes

Metode *Naïve Bayes* adalah algoritma klasifikasi berbasis probabilitas yang menggunakan *Teorema Bayes* dan menganggap bahwa setiap fitur atau atribut berbeda dari yang lain. Metode ini bekerja dengan menghitung kemungkinan kondisi untuk setiap kelas dan menemukan kelas dengan kemungkinan tertinggi. Karena sederhana dan cepat dalam komputasi, algoritma ini sangat cocok untuk menangani data berskala besar [7].

Proses implementasi algoritma *Naïve Bayes* terdapat beberapa Langkah yang harus dilakukan, yang ditunjukkan pada gambar 1.

Saat implementasi algoritma *Naïve Bayes*, ada beberapa langkah yang harus dilakukan. Langkah pertama yang dilakukan yaitu membaca data yang akan digunakan dalam proses algoritma *Naïve Bayes*. Data yang digunakan dapat dilihat pada gambar 2.

| NO | Usia | Initial | Tanggungan Kepala Keluarga | Tanggungan Lansia | Tanggungan Anak Sekolah | Penghasilan Kepala Keluarga | Status Asuransi    | Tipe Kendaraan      | Sumber Air Bersih | Tipe Jamban     | Status Bangunan | Tipe Lantai | Tipe Dinding | Luas Bangunan | Label          |
|----|------|---------|----------------------------|-------------------|-------------------------|-----------------------------|--------------------|---------------------|-------------------|-----------------|-----------------|-------------|--------------|---------------|----------------|
| 1  | 5    | RV      | Sedang                     | Tidak Ada         | Ada                     | Rendah                      | Tidak Ada Asuransi | Motor               | Sumur             | Jamban Keluarga | Milik Sendiri   | Keramik     | Batu Bata    | Besar         | Bukan Penerima |
| 2  | 10   | EV      | Sedang                     | Tidak Ada         | Ada                     | Tinggi                      | Punya Asuransi     | Sepeda              | Sumur             | Jamban Bersama  | Milik Sendiri   | Ubin        | Tembok       | Besar         | Bukan Penerima |
| 3  | 44   | PK      | Sedikit                    | Tidak Ada         | Ada                     | Tinggi                      | Tidak Ada Asuransi | Sepeda              | Pompa Air         | Tidak Ada       | Milik Sendiri   | Semen       | Tembok       | Kecil         | Penerima       |
| 4  | 40   | DF      | Sedang                     | Tidak Ada         | Ada                     | Tinggi                      | Tidak Ada Asuransi | Tidak Ada Kendaraan | Air Ledeng        | Jamban Bersama  | Milik Sendiri   | Tanah       | Kayu         | Besar         | Penerima       |
| 5  | 77   | GB      | Tidak Ada                  | Ada               | Tidak Ada               | Sedang                      | Tidak Ada Asuransi | Sepeda              | Air Ledeng        | Tidak Ada       | Milik Sendiri   | Semen       | Tembok       | Sedang        | Penerima       |
| 6  | 34   | JN      | Sedikit                    | Tidak Ada         | Ada                     | Sedang                      | Tidak Ada Asuransi | Motor               | Pompa Air         | Jamban Bersama  | Menumpang       | Ubin        | Tembok       | Besar         | Penerima       |
| 7  | 48   | ZY      | Banyak                     | Ada               | Ada                     | Sedang                      | Tidak Ada Asuransi | Mobil               | Air Isi Ulang     | Jamban Keluarga | Sewa            | Ubin        | Tembok       | Besar         | Penerima       |
| 8  | 68   | RP      | Tidak                      | Ada               | Tidak                   | Rendah                      | Tidak Ada Asuransi | Sepeda              | Sumur             | Jamban Bersama  | Milik Sendiri   | Ubin        | Kayu         | Kecil         | Penerima       |
| 9  | 62   | IV      | Tidak                      | Ada               | Tidak                   | Tinggi                      | Tidak Ada Asuransi | Sepeda              | Pompa Air         | Jamban Bersama  | Menumpang       | Keramik     | Tembok       | Besar         | Penerima       |
| 10 | 18   | EH      | Sedang                     | Tidak             | Ada                     | Sedang                      | Punya Asuransi     | Mobil Pribadi       | Sumur             | Jamban Bersama  | Sewa            | Ubin        | Tembok       | Besar         | Bukan Penerima |

Gambar 2. Data training untuk proses *Naïve Bayes*

Selanjutnya, data dapat diproses dan dilakukan perhitungan dengan menggunakan algoritma *Naïve Bayes*.

Tahap 1, menghitung jumlah *class*/label :

$$P(C) = \left( \frac{\text{jumlah kasus dalam kelas C}}{\text{total jumlah kasus}} \right)$$

$$P(Y=\text{Penerima}) = 7/10 \mid 0.70$$

$$P(N=\text{Bukan Penerima}) = 3/10 \mid 0.30$$

Tahap 2, menghitung jumlah kasus yang sama dan memiliki *class* yang sama.

$$P(X_i|C) = \left( \frac{\text{jumlah kasus dengan atribut } X_i \text{ dalam kelas C}}{\text{total kasus dalam kelas C}} \right)$$

Perhitungan semua variabel/kategori penelitian dapat dilihat pada tabel 1.

Table 1 Perhitungan kasus dengan class yang sama

| Kategori                      |                            | Perhitungan |                | Hasil    |                |
|-------------------------------|----------------------------|-------------|----------------|----------|----------------|
|                               |                            | Penerima    | Bukan Penerima | Penerima | Bukan Penerima |
| Tanggungan Kepala Keluarga    | Tidak Ada                  | 3/7         | 0/3            | 0.428571 | 0              |
|                               | Sedikit                    | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Sedang                     | 1/7         | 3/3            | 0.142857 | 1              |
|                               | Banyak                     | 1/7         | 0/3            | 0.142857 | 0              |
| Tanggungan Lansia             | Tidak Ada                  | 3/7         | 3/3            | 0.428571 | 1              |
|                               | Ada                        | 4/7         | 0/3            | 0.571429 | 0              |
| Tanggungan Anak Wajib Sekolah | Tidak Ada                  | 3/7         | 0/3            | 0.428571 | 0              |
|                               | Ada                        | 4/7         | 3/3            | 0.571429 | 1              |
| Penghasilan Kepala Keluarga   | Rendah                     | 1/7         | 1/3            | 0.142857 | 0.333333333    |
|                               | Sedang                     | 3/7         | 1/3            | 0.428571 | 0.333333333    |
|                               | Tinggi                     | 3/7         | 1/3            | 0.428571 | 0.333333333    |
| Status Asuransi               | Tidak Ada Asuransi         | 7/7         | 1/3            | 1        | 0.333333333    |
|                               | Punya Asuransi             | 0/7         | 2/3            | 0        | 0.666666667    |
| Tipe Kendaraan                | Tidak Ada Kendaraan        | 1/7         | 0/3            | 0.142857 | 0              |
|                               | Sepeda                     | 4/7         | 1/3            | 0.571429 | 0.333333333    |
|                               | Motor                      | 1/7         | 1/3            | 0.142857 | 0.333333333    |
|                               | Mobil Pribadi              | 1/7         | 1/3            | 0.142857 | 0.333333333    |
| Sumber Air Bersih             | Sumur dan Pompa Air        | 4/7         | 3/3            | 0.571429 | 1              |
|                               | Air ledeng                 | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Air Isi Ulang              | 1/7         | 0/3            | 0.142857 | 0              |
| Tipe Jamban                   | Tidak Ada                  | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Jamban Bersama             | 4/7         | 2/3            | 0.571429 | 0.666666667    |
|                               | Jamban Keluarga            | 1/7         | 1/3            | 0.142857 | 0.333333333    |
| Status Kepemilikan Bangunan   | Menumpang                  | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Sewa, Kontrak, dan lainnya | 1/7         | 1/3            | 0.142857 | 0.333333333    |
|                               | Milik Sendiri              | 4/7         | 2/3            | 0.571429 | 0.666666667    |
| Bahan Lantai                  | Ubin dan Keramik           | 1/7         | 0/3            | 0.142857 | 0              |
|                               | Semen                      | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Tanah                      | 4/7         | 3/3            | 0.571429 | 1              |
|                               | Tembok                     | 2/7         | 0/3            | 0.285714 | 0              |
| Bahan Dinding                 | Batu Bata                  | 0/7         | 1/3            | 0        | 0.333333333    |
|                               | Kayu                       | 5/7         | 2/3            | 0.714286 | 0.666666667    |
| Luas Bangunan                 | Kecil                      | 2/7         | 0/3            | 0.285714 | 0              |
|                               | Sedang                     | 1/7         | 0/3            | 0.142857 | 0              |
|                               | Besar                      | 4/7         | 3/3            | 0.571429 | 1              |

Setelah menghitung pada data training, muncul data testing. Berikut data testing serta perhitungan penentuan data testing memiliki label penerima atau bukan penerima pada tabel 2.

Dari perhitungan data testing, menunjukkan bahwa data *dummy* tersebut memiliki label bukan penerima, dikarenakan data tersebut memiliki beberapa kategori yang menunjukkan kesejahteraan hidup seperti bangunan milik sendiri, penghasilan tinggi, luas bangunan besar, dan sebagainya.

**Table 2 Perhitungan Data Testing**

| Perhitungan                   |                    | Label      |                | Hasil    |                |
|-------------------------------|--------------------|------------|----------------|----------|----------------|
|                               |                    | Penerima   | Bukan Penerima | Penerima | Bukan Penerima |
| Tanggungan Kepala Keluarga    | Sedang             | 0.14285714 | 1              | 0.00026  | 0.003292181    |
| Tanggungan Lansia             | Tidak Ada          | 0.42857142 | 1              |          |                |
| Tanggungan Anak Wajib Sekolah | Ada                | 0.57142857 | 1              |          |                |
| Penghasilan Kepala Keluarga   | Tinggi             | 0.42857142 | 0.33333333     |          |                |
| Status Asuransi               | Tidak Ada Asuransi | 1          | 0.33333333     |          |                |
| Tipe Kendaraan                | Sepeda             | 0.57142857 | 0.33333333     |          |                |
| Sumber Air Bersih             | Sumur              | 0.57142857 | 1              |          |                |
| Tipe Jamban                   | Jamban Bersama     | 0.57142857 | 0.66666667     |          |                |
| Status Kepemilikan Bangunan   | Milik Sendiri      | 0.57142857 | 0.66666667     |          |                |
| Bahan Lantai                  | Ubin               | 0.57142857 | 1              |          |                |
| Bahan Dinding                 | Tembok             | 0.71428571 | 0.66666667     |          |                |
| Luas Bangunan                 | Besar              | 0.57142857 | 1              |          |                |
| Label                         |                    | 0.7        | 0.3            |          |                |

## HASIL DAN PEMBAHASAN

### Pengujian Pertama

Pada pengujian pertama dilakukan setiap *file* untuk mengetahui akurasi masing-masing.

**Table 3 Pengujian Pertama**

| WARGA (Bukan Penerima)                 |                    |    |          |                      |                     |               |
|----------------------------------------|--------------------|----|----------|----------------------|---------------------|---------------|
| NO                                     | Kategori           | RW | Jml Data | Data <i>Training</i> | Data <i>Testing</i> | Hasil Akurasi |
| 1                                      | Bayi Balita        | 5  | 87       | 78                   | 9                   | 100%          |
| 2                                      | Anak Wajib Sekolah | 1  | 290      | 261                  | 29                  | 100%          |
| 3                                      | Usia Produktif     | 1  | 489      | 440                  | 49                  | 100%          |
| 4                                      | Lansia             | 1  | 140      | 126                  | 14                  | 100%          |
| PKH (Penerima)                         |                    |    |          |                      |                     |               |
| NO                                     | Kategori           | RW | Jml Data | Data <i>Training</i> | Data <i>Testing</i> | Hasil Akurasi |
| 1                                      | Usia Produktif     | 1  | 47       | 42                   | 5                   | 100%          |
| 2                                      | Lansia             | 1  | 27       | 24                   | 3                   | 100%          |
| MIX DATA (Bukan Penerima dan Penerima) |                    |    |          |                      |                     |               |
| NO                                     | Kategori           | RW | Jml Data | Data <i>Training</i> | Data <i>Testing</i> | Hasil Akurasi |
| 1                                      | Usia Produktif     | 1  | 536      | 482                  | 54                  | 12.96%        |
| 2                                      | Lansia             | 1  | 167      | 150                  | 17                  | 76.47%        |

Dari hasil pengujian pertama, dapat dilihat bahwa apabila dilakukan pengujian pada *file* yang hanya memiliki satu label, maka hasil akurasi pasti akan sempurna. Namun apabila saat pengujian pada *file* terdapat lebih dari satu label, hasil akurasi akan bervariasi bergantung pada persentase data training dan data testing yang sedang digunakan saat pengujian berlangsung.

### Pengujian Kedua

Pengujian kedua dilakukan untuk mengamati pengaruh penggabungan beberapa *file*. Penggabungan *file* berguna untuk menguji performa sistem Algoritma *Naive Bayes*, pengujian ini dibagi menjadi 3 pola, yakni, 2 kombinasi, 3 kombinasi, dan 4 kombinasi *file*.

**Table 4 Pengujian Kedua**

| N<br>O | Kelompok<br>Data | Kategori           | RW | Jml<br>Data | Total All<br>Data | Data<br><i>Training</i> | Data<br><i>Testing</i> | Hasil<br>akurasi |
|--------|------------------|--------------------|----|-------------|-------------------|-------------------------|------------------------|------------------|
| 1      | Warga            | Usia Produktif     | 1  | 489         | 525               | 473                     | 52                     | 30.77%           |
|        | PKH              | Lansia             | 2  | 36          |                   |                         |                        |                  |
| 2      | Warga            | Anak Wajib Sekolah | 2  | 610         | 657               | 591                     | 66                     | 28.79%           |
|        | PKH              | Usia Produktif     | 1  | 47          |                   |                         |                        |                  |
| 3      | Mix Data         | Lansia             | 4  | 129         | 486               | 437                     | 49                     | 4.08%            |
|        | Warga            | Anak Wajib Sekolah | 5  | 357         |                   |                         |                        |                  |
| 4      | PKH              | Usia Produktif     | 2  | 103         | 238               | 214                     | 24                     | 100%             |
|        | Warga            | Usia Produktif     | 6  | 135         |                   |                         |                        |                  |
| N<br>O | Kelompok<br>Data | Kategori           | RW | Jml<br>Data | Total All<br>Data | Data<br><i>Training</i> | Data<br><i>Testing</i> | Hasil<br>akurasi |

|   |          |                    |   |     |      |      |     |        |
|---|----------|--------------------|---|-----|------|------|-----|--------|
| 1 | PKH      | Lansia             | 3 | 20  | 80   | 72   | 8   | 100%   |
|   | PKH      | Usia Produktif     | 3 | 55  |      |      |     |        |
|   | PKH      | Usia Produktif     | 7 | 5   |      |      |     |        |
| 2 | Mix Data | Lansia             | 2 | 308 | 320  | 288  | 32  | 81.25% |
|   | PKH      | Lansia             | 5 | 7   |      |      |     |        |
|   | PKH      | Usia Produktif     | 7 | 5   |      |      |     |        |
| 3 | PKH      | Usia Produktif     | 1 | 47  | 190  | 171  | 19  | 100%   |
|   | PKH      | Usia Produktif     | 4 | 3   |      |      |     |        |
|   | Warga    | Lansia             | 1 | 140 |      |      |     |        |
| 4 | Mix Data | Usia Produktif     | 4 | 296 | 676  | 608  | 68  | 97.06% |
|   | PKH      | Usia Produktif     | 7 | 5   |      |      |     |        |
|   | Warga    | Lansia             | 5 | 375 |      |      |     |        |
| 1 | Mix Data | Lansia             | 3 | 222 | 419  | 377  | 42  | 100%   |
|   | PKH      | Lansia             | 7 | 2   |      |      |     |        |
|   | PKH      | Usia Produktif     | 3 | 55  |      |      |     |        |
|   | Warga    | Lansia             | 1 | 140 |      |      |     |        |
| 2 | PKH      | Usia Produktif     | 4 | 3   | 1247 | 1122 | 125 | 59.20% |
|   | Warga    | Anak Wajib Sekolah | 1 | 290 |      |      |     |        |
|   | Warga    | Anak Wajib Sekolah | 2 | 610 |      |      |     |        |
|   | Warga    | Anak Wajib Sekolah | 3 | 344 |      |      |     |        |
| 3 | Mix Data | Usia Produktif     | 1 | 536 | 1371 | 1234 | 137 | 7.30%  |
|   | Mix Data | Usia Produktif     | 3 | 691 |      |      |     |        |
|   | PKH      | Usia Produktif     | 5 | 17  |      |      |     |        |
|   | Warga    | Lansia             | 4 | 127 |      |      |     |        |
| 4 | PKH      | Lansia             | 3 | 20  | 814  | 733  | 81  | 100%   |
|   | PKH      | Usia Produktif     | 2 | 103 |      |      |     |        |
|   | Warga    | Lansia             | 3 | 202 |      |      |     |        |
|   | Warga    | Usia Produktif     | 1 | 489 |      |      |     |        |

Dapat dilihat pada pengujian kedua, bahwa pola kombinasi 3 *file* terbukti stabil dari segi hasil akurasi dan siap untuk dilakukan pengujian selanjutnya. Hasil akurasi yang sempurna dan tidak kurang dari 75%.

### Pengujian Ketiga

Pengujian ketiga dilakukan untuk menentukan pada persentase berapa data akan mengalami penurunan hasil akurasi.

**Table 5 Pengujian Ketiga**

| Kelompok Data | Kategori       | RW     |
|---------------|----------------|--------|
| Mix Data      | Usia Produktif | 4      |
| PKH           | Usia Produktif | 7      |
| Warga         | Lansia         | 5      |
| Persentase    |                |        |
| 90/10         |                | 97,06% |
| 80/20         |                | 91,11% |
| 70/30         |                | 91,63% |
| 60/40         |                | 78,52% |
| 50/50         |                | 27,22% |
| 40/60         |                | 1,23%  |
| 30/70         |                | 1,48%  |
| 20/80         |                | 1,29%  |
| 10/90         |                | 1,15%  |

Pada skenario pengujian yang dilakukan, pembagian data 90:10, 80:20, 70:30, dan 60:40 menghasilkan akurasi di atas 75%, sedangkan skenario lainnya menghasilkan akurasi lebih rendah. Nilai 75% dalam penelitian ini digunakan sebagai batas deskriptif untuk membandingkan hasil pada skenario pengujian yang dilakukan dan bukan sebagai ambang universal untuk menentukan bahwa suatu model telah memiliki performa yang layak digunakan.

### Keterbatasan Penelitian

Penelitian ini menggunakan rancangan evaluasi sesuai dengan penelitian awal yang dilakukan dalam skripsi, yaitu pembagian data training dan testing dengan beberapa proporsi serta evaluasi menggunakan akurasi. Evaluasi menggunakan stratified sampling, k-fold cross-validation, dan metrik tambahan seperti precision, recall, serta F1-score belum dilakukan dalam penelitian ini dan menjadi peluang pengembangan penelitian selanjutnya.

### KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dengan menggunakan algoritma *Naive Bayes* untuk mengetahui apakah seseorang dengan kriteria tertentu layak mendapatkan bantuan sosial PKH. Dari pengujian pertama hingga ketiga, dapat disimpulkan bahwa kombinasi file dan persentase yang dipakai saat pengujian, hasil akurasi akan semakin rendah dan kurang dari 75%. Sebaliknya, menggunakan persentase tinggi saat pengujian maka hasil akurasi yang didapat juga tinggi dan bisa melampaui 75%. Hasil implementasi algoritma *Naive Bayes* diharapkan bisa membantu pihak terkait untuk lebih adil saat menentukan layak tidaknya menerima bantuan dari segi kesejahteraan sosial yang bisa di data dari kriteria pengujian.

### DAFTAR PUSTAKA

- [1] N. Salsabila, N. Muna, V. H. Pradana, and W. F. Nurcahya, "Analisis efektivitas bantuan sosial (bansos) dalam mengatasi kemiskinan di Indonesia," *Journal of Macroeconomics and Social Development*, vol. 1, no. 4, pp. 1–13, 2024, doi: 10.47134/jmsd.v1i4.317.
- [2] D. A. Setiawan, R. Helilintar, and L. S. Wahyuniar, "Penerapan metode Naive Bayes untuk klasifikasi penentuan penerima bantuan PKH," in *Prosiding Seminar Nasional Inovasi Teknologi*, Jul. 2021, pp. 249–254.
- [3] A. Surahman and U. Hayati, "Implementasi algoritma Naïve Bayes untuk prediksi penerima bantuan sosial," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 7, no. 1, pp. 347–352, 2023, doi: 10.36040/jati.v7i1.6302.

- [4] S. Lee, R. Koffer, and J. Drewelies, "Adults older than age 55 engage in less diverse activities than those 18 years ago," *The Journals of Gerontology: Series B*, vol. 78, no. 9, pp. 1511–1520, 2023, doi: 10.1093/geronb/gbad047.
- [5] A. Gunawan and Z. Fatah, "Klasifikasi penerima bantuan SKTM menggunakan algoritma Naive Bayes: Studi kasus Desa Pesanggrahan," *Jurnal Riset Sistem Informasi*, vol. 2, no. 1, pp. 45–52, 2025, doi: 10.69714/6w34wq73.
- [6] A. W. S. Sriwibowo, A. Alwi, and Sugianti, "Penerapan algoritma Naïve Bayes dalam prediksi penerimaan bantuan Program Keluarga Harapan (PKH)," *JICTE (Journal of Information and Computer Technology Education)*, vol. 4, no. 1, 2020, doi: 10.21070/jicte.v4i1.916.
- [7] Badan Pusat Statistik, "Profil kemiskinan di Indonesia Maret 2024," Jul. 1, 2024. [Online]. Available: <https://www.bps.go.id/id>
- [8] Z. Bami, A. Behnampour, and H. Doosti, "A new flexible train-test split algorithm: An approach for choosing among the hold-out, K-fold cross-validation, and hold-out iteration," *arXiv preprint arXiv:2501.06492*, 2025.
- [9] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. Cambridge, MA, USA: Morgan Kaufmann, 2022.
- [10] D. T. R. Setyawardani, C. J. Paat, and L. Lesawengen, "Dampak bantuan PKH terhadap masyarakat miskin di Kelurahan Bumi Nyiur Kecamatan Wanea Kota Manado," *HOLISTIK, Journal of Social and Culture*, vol. 13, no. 2, 2020.
- [11] E. A. Djamhari et al., *Laporan Riset Kesejahteraan Lansia 2020*. Jakarta, Indonesia: The PRAKARSA, 2020.
- [12] Kementerian Sosial Republik Indonesia, *Pedoman Program Keluarga Harapan (PKH)*. Jakarta, Indonesia: Kementerian Sosial Republik Indonesia, 2021.
- [13] Kementerian Pendidikan Nasional Republik Indonesia, *Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional*. Jakarta, Indonesia, 2003.