

**IMPLEMENTASI METODE *NAÏVE BAYES* UNTUK PERANCANGAN SISTEM
KLASIFIKASI REKOMENDASI CALON PENERIMA BANTUAN SOSIAL**

Wahyu lazzuardy¹, Umi chotijah², Henny dwi bakti³

Program Studi Teknik Informatika, Fakultas Teknik Universitas Muhammadiyah Gresik

Email: wahyulazzuardy@gmail.com, umi.chotijah@umg.ac.id, hennydwi@umg.ac.id

Abstract

Received: 02-09-2022
Accepted: 09-09-2022
Published: 22-09-2022
Keywords: Social Assistance,
Classification,
Naïve Bayes
Classifier

The Doudo Village Government in determining prospective recipients of social assistance to be recommended to the central government is still not computerized so that the work takes a long time. The use of a computerized system in an agency is a very important tool to support smoothness and accuracy in achieving goals. In this research, the writer uses the Naïve Bayes Classifier method to solve the existing problems. The Naïve Bayes method has the advantage that it is easy to understand and faster in calculations. The criteria used to support this decision are the type of work, the amount of income, the number of family members, the amount of expenditure per day and the status of residence. The training data used in the testing process greatly influences the classification results. If the training data used is more and more varied, the accuracy results obtained will be better.

Abstrak

Kata kunci: Bantuan Sosial,
Klasifikasi, *Naïve Bayes Classifier*

Pemerintahan Desa Doudo dalam menentukan calon penerima bantuan sosial untuk direkomendasikan ke pemerintahan pusat masih belum terkomputerisasi sehingga pekerjaan terbilang memakan waktu yang lama. Penggunaan sistem yang terkomputerisasi disuatu instansi merupakan alat yang sangat penting untuk pendukung kelancaran dan keakuratan dalam pencapaian tujuan. Penelitian ini penulis menggunakan metode *Naïve Bayes Classifier* untuk menyelesaikan permasalahan yang ada. Metode *Naïve Bayes* memiliki keunggulan yaitu mudah dipahami dan lebih cepat dalam perhitungan. Kriteria yang digunakan untuk mendukung pengambilan keputusan ini adalah jenis pekerjaan, jumlah penghasilan, jumlah anggota keluarga, jumlah pengeluaran perhari dan status tempat tinggal. Data *training* yang dipakai pada proses pengujian sangat mempengaruhi hasil klasifikasi. Jika data *training* yang digunakan semakin banyak dan bervariasi, hasil akurasi yang didapat akan lebih baik.

Corresponding Author: Wahyu lazzuardy
E-mail: wahyulazzuardy@gmail.com



PENDAHULUAN

Kemiskinan dianggap sebagai suatu kondisi seseorang atau sekelompok orang, laki-laki atau perempuan yang hak-hak dasarnya tidak terpenuhi secara layak untuk hidup dan mengembangkan kehidupan yang bermartabat (Pulungan, Harahap, & Manurung, 2017). Pada saat ini masyarakat Indonesia masih banyak yang tergolong masyarakat miskin atau kurang mampu (Angga & Arifin, 2019). Tidak meratanya pendapatan menimbulkan terjadinya

ketidakseimbangan pendapatan yang merupakan awal dari terjadinya masalah kemiskinan (Indra Suhendra, Ginanjar, & Navik Istikomah, 2021).

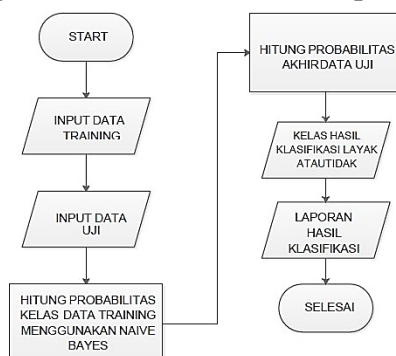
Dalam mengatasi permasalahan kemiskinan di masyarakat Indonesia, pemerintahan pusat mengadakan berbagai program seperti yang terdapat di Desa Doudo, salah satunya dengan membagikan bantuan sosial untuk masyarakat kurang mampu (Nawastuti, 2018). Sebelum melakukan pembagian bantuan, seksi kesejahteraan melakukan proses penyeleksian untuk merekomendasikan masyarakat yang layak mendapatkan bantuan sosial tersebut (Wahidah, Widayani, Wardani, Rachmawati, & Latifah, 2022). Selama ini yang dilakukan seksi kesejahteraan pemerintahan Desa Doudo dalam menentukan calon penerima bantuan sosial untuk direkomendasikan ke pemerintahan pusat masih belum terkomputerisasi sehingga pekerjaan terbilang memakan waktu yang lama (Chalik, 2015).

Metode *naïve bayes classifier* merupakan metode klasifikasi yang berakar pada teorema bayes. Metode pengklasifikasian dengan menggunakan probabilitas dan statistik yang dikemukakan oleh *Thomas Bayes*, yaitu memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya (Misdram, Syarifuddin, & Widodo, 2020). Dalam penelitian ini penulis menggunakan metode *naïve bayes classifier* untuk menyelesaikan permasalahan yang ada, dikarenakan metode tersebut menurut penulis sesuai dengan data yang akan digunakan untuk penelitian, juga memiliki keunggulan yaitu mudah dipahami, lebih cepat dalam perhitungan dan dalam lima tahun terakhir metode tersebut masih diterapkan dalam penelitian yang serupa. Seperti penelitian dengan judul klasifikasi masyarakat miskin dengan metode *Naïve Bayes* oleh Haditsah Annur pada tahun 2018 (Annur, 2018). Pada penelitian ini menjelaskan beberapa acuan yang bisa digunakan untuk mengkategorikan masyarakat miskin atau tidak miskin yaitu Pekerjaan, Penghasilan, Tanggungan. Hasil pengujian *confussion matrix* dengan teknik split validasi, penggunaan metode klasifikasi *naïve bayes* diperoleh tingkat akurasi sebesar 73% (Annur, 2018). Kedua, penelitian dengan judul Implementasi Klasifikasi *Naïve Bayes* Untuk Prediksi Kelayakan Pemberian Pinjaman Pada Koperasi Anugerah Bintang Cemerlang oleh Siti Lestari, Akmaludin dan Mohammad Badrul pada tahun 2020. Dalam penelitian ini diperoleh tingkat akurasi sebesar 78,08 % (Lestari, Akmaludin, & Badrul, 2020).

METODE PENELITIAN

Deskripsi Sitem

Sistem yang akan dibangun adalah sistem klasifikasi rekomendasi calon penerima bantuan sosial menggunakan metode *Naïve Bayes*. Berikut adalah rancangan *flowcart* sistem klasifikasi rekomendasi calon penerima bantuan sosial, seperti pada gambar 1.



Gambar 1. *Flowcart* sistem klasifikasi rekomendasi calon penerima bantuan sosial.

Berdasarkan gambar 1 tahapan sistem klasifikasi rekomendasi calon penerima bantuan dimulai dengan memasukkan data *training*. Kemudian memasukkan data *testing*.

Setelah itu sistem melakukan perhitungan probabilitas tiap kelas pada data *training* menggunakan *Naïve Bayes Classifier*. Setelah itu menghitung probabilitas data uji / *testing*. Setelah menghitung probabilitas data uji / *testing* selesai, maka sistem akan melakukan *output*, apakah warga itu masuk kategori layak atau tidak layak untuk menerima bantuan sosial.

Algoritma *Naïve Bayes*

Naïve bayes merupakan salah satu algoritma data mining klasifikasi yang dikembangkan oleh ilmuwan inggris *Thomas Bayes*. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistic, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya (Syarli & Muin, 2016).

Persamaan Teorema *Bayes* sebagai berikut :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$: Probabilitas hipotesis H (prior probability)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Adapun proses klasifikasi dengan menggunakan data numeric menggunakan rumus *Densitas Gauss*, seperti persamaan dibawah :

$$P(X_i = x_i | C = c_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \quad (2)$$

Keterangan :

P : Peluang

X_i : Atribut ke- i

C : Kelas yang dicari

μ : Mean, yaitu rata-rata dari seluruh atribut

σ : Standar Deviasi, yaitu akar kuadrat dari varian

π : 3,14

Alur metode *Naïve Bayes* sebagai berikut :

1. Baca data training.
2. Menghitung jumlah dan probabilitas, tapi jika data numeric maka :
 - a. Mencari nilai *mean* dan standar deviasi dari masing –masing parameter yang merupakan data numerik. Persamaan untuk mencari nilai *mean* sebagai berikut :

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (3)$$

Penjelasan :

μ = Nilia rata-rata (*mean*)

x_i = Nilai x ke- i

n = Jumlah Sampel

Persamaan unutup menghitung standar deviasi sebagai berikut :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}} \quad (4)$$

Penjelasan :

σ = Standar Deviasi

x_i = Nilai x ke-i

μ = Nilia rata-rata (*mean*)

n = Jumlah Sampel

- b. Mencari nilai probabilistik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
3. Mendapatkan nilai dalam tabel *mean*, standar deviasi dan probabilitas
4. Solusi yang dihasilkan.

Kebutuhan Data

Kriteria yang digunakan untuk mendukung pengambilan keputusan ini adalah jenis pekerjaan, jumlah penghasilan, jumlah anggota keluarga, jumlah pengeluaran perhari, status tempat tinggal. Data yang sudah didapat kemudian dibagi menjadi 5 atribut, seperti yang ada dalam tabel 1.

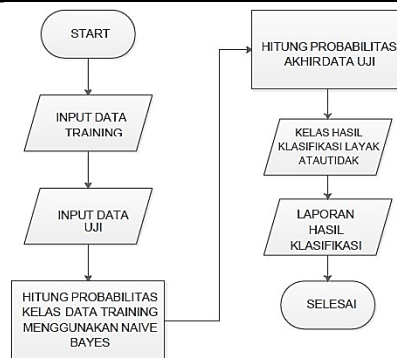
Tabel 1. Tabel Atribut

Atribut	Keterangan	Nilai
Jenis pekejaan	Jenis pekejaan warga	Petani
		Pedagang
		Wiraswasta
		PNS
Jumlah penghasilan	Jumlah penghasilan warga perbulan	Jumlah penghasilan warga
Jumlah Anggota keluarga	Jumlah keseluruhan keluarga	Jumlah keseluruhan keluarga
Jumlah pengeluaran perhari	Pengeluaran warga perhari	< 50.000
		50.000 – 100.000
		> 100.000
Status tempat tinggal	Status tempat tinggal	Numpang
		Sewa
		Milik Sendiri

HASIL DAN PEMBAHASAN

Deskripsi Sitem

Sistem yang akan dibangun adalah sitem klasifikasi rekomendasi calon penerima bantuan sosial menggunakan metode *Naïve Bayes*. Berikut adalah rancangan *flowcart* sistem klasifikasi rekomendasi calon penerima bantuan sosial, seperti pada gambar 1.



Gambar 1. Flowcart sitem klasifikasi rekomendasi calon penerima bantuan sosial.

Berdasarkan gambar 1 tahapan sistem klasifikasi rekomendasi calon penerima bantuan dimulai dengan memasukkan data *training*. Kemudian memasukkan data *testing*. Setelah itu sistem melakukan perhitungan probabilitas tiap kelas pada data *training* menggunakan *Naïve Bayes Classifier*. Setelah itu menghitung probabilitas data uji / *testing*. Setelah menghitung probabilitas data uji / *testing* selesai, maka sistem akan melakukan *output*, apakah warga itu masuk kategori layak atau tidak layak untuk menerima bantuan sosial.

Algoritma *Naïve Bayes*

Naïve bayes merupakan salah satu algoritma data mining klasifikasi yang dikembangkan oleh ilmuan inggris *Thomas Bayes*. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistic, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya (Syarli & Muin, 2016).

Persamaan Teorema *Bayes* sebagai berikut :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (1)$$

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$: Probabilitas hipotesis H (prior probability)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Adapun proses klasifikasi dengan menggunakan data numeric menggunakan rumus *Densitas Gauss*, seperti persamaan dibawah :

$$P(X_i = x_i | C = c_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \quad (2)$$

Keterangan :

P : Peluang

X_i : Atribut ke- i

C : Kelas yang dicari

μ : Mean, yaitu rata-rata dari seluruh atribut

σ : Standar Deviasi, yaitu akar kuadrat dari varian

π : 3,14

Alur metode *Naïve Bayes* sebagai berikut :

5. Baca data training.
6. Menghitung jumlah dan probabilitas, tapi jika data numeric maka :

- c. Mencari nilai *mean* dan standar deviasi dari masing –masing parameter yang merupakan data numerik. Persamaan untuk mencari nilai *mean* sebagai berikut :

$$\mu = \frac{\sum_{i=1}^n x_i}{n} \quad (3)$$

Penjelasan :

μ = Nilia rata-rata (*mean*)

x_i = Nilai x ke-i

n = Jumlah Sampel

Persamaan unutupuk menghitung standar deviasi sebagai berikut :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n - 1}} \quad (4)$$

Penjelasan :

σ = Standar Deviasi

x_i = Nilai x ke-i

μ = Nilia rata-rata (*mean*)

n = Jumlah Sampel

- d. Mencari nilai probablistik dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut.
7. Mendapatkan nilai dalam tabel *mean*, standar deviasi dan probabilitas
 8. Solusi yang dihasilkan.

Kebutuhan Data

Kriteria yang digunakan untuk mendukung pengambilan keputusan ini adalah jenis pekerjaan, jumlah penghasilan, jumlah anggota keluarga, jumlah pengeluaran perhari, status tempat tinggal. Data yang sudah didapat kemudian dibagi menjadi 5 atribut, seperti yang ada dalam tabel 1.

Tabel 1. Tabel Atribut

Atribut	Keterangan	Nilai
Jenis pekejaan	Jenis pekejaan warga	Petani
		Pedagang
		Wiraswasta
		PNS
Jumlah penghasilan	Jumlah penghasilan warga perbulan	Jumlah penghasilan warga
Jumlah Anggota keluarga	Jumlah keseluruhan keluarga	Jumlah keseluruhan keluarga
Jumlah pengeluaran perhari	Pengeluaran warga perhari	< 50.000
		50.000 – 100.000
		> 100.000
Status tempat tinggal	Status tempat tinggal	Numpang
		Sewa
		Milik Sendiri

KESIMPULAN

Setelah melakukan implementasi dan pengujian sistem, maka menghasilkan kesimpulan, diantaranya:

1. Proses klasifikasi pada sistem yang telah berhasil dibangun terdiri dari dua kelas yaitu layak dan tidak layak. Akurasi terbaik yang didapatkan dari tiga kali percobaan pengujian adalah 83,3% dengan menggunakan data *training* sebanyak 90 data dan data uji sebanyak 30 data.
2. Data *training* yang dipakai pada proses pengujian sangat mempengaruhi hasil klasifikasi. Jika data *training* yang digunakan semakin banyak dan bervariasi, hasil akurasi yang didapat akan lebih baik.

BIBLIOGRAFI

- Angga, Angga, & Arifin, Ridwan. (2019). Penerapan Bantuan Hukum Bagi Masyarakat Kurang Mampu di Indonesia. *DIVERSI: Jurnal Hukum*, 4(2), 218–236.
- Annur, Haditsah. (2018). Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes. *ILKOM Jurnal Ilmiah*, 10(2), 160–165.
- Chalik, Abdul. (2015). *Pelayanan publik tingkat desa*. Interpena.
- Indra Suhendra, S. E., Ginanjar, Rah Adi Fahmi, & Navik Istikomah, S. E. (2021). *Distribusi Pendapatan: Konteks Provinsi di Indonesia*. Media Sains Indonesia.
- Lestari, Siti, Akmaludin, Akmaludin, & Badrul, Mohammad. (2020). Implementasi Klasifikasi Naive Bayes Untuk Prediksi Kelayakan Pemberian Pinjaman Pada Koperasi Anugerah Bintang Cemerlang. *PROSISKO: Jurnal Pengembangan Riset Dan Observasi Sistem Komputer*, 7(1).
- Misdram, Muhammad, Syarifuddin, Fahmi, & Widodo, Anang Aris. (2020). Klasifikasi Data Set Virus Corona Menggunakan Metode Naïve Bayes Classifier. *SPIRIT*, 12(2).
- Nawastuti, Dati. (2018). Kearifan Lokal Masyarakat Pesisir dalam Memahami Teknologi Hasil Perikanan. *Jurnal Maksipreneur: Manajemen, Koperasi, Dan Entrepreneurship*, 8(1), 32–44.
- Pulungan, Khairul Anwar, Harahap, Riva Ubar, & Manurung, Saprial. (2017). Analisis Kebijakan Program Penanggulangan Kemiskinan: Implementasi Manajemen Pemerintahan Kota Tanjung Balai. *Jurnal Riset Sains Manajemen Jurnal Riset Sains Manajemen Jurnal Riset Sains Manajemen*, 1(12), 115–128. <https://doi.org/10.5281/zenodo.1069779>
- Syarli, S., & Muin, A. (2016). Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi). *Jurnal Ilmiah Ilmu Komputer*, 2(1), 22–26.
- Wahidah, Anna Zahrotul, Widayani, Anna, Wardani, Shanti Ike, Rachmawati, Ika, & Latifah, Nunuk. (2022). PROSEDUR PENYALURAN BANTUAN SOSIAL DI ERA PANDEMI COVID-19 PADA DINAS SOSIAL KABUPATEN BLITAR. *Competence: Journal of Management Studies*, 16(1), 51–63.