

Implementasi Metode *Naïve Bayes* dalam Mengklasifikasi Penerima Program Keluarga Harapan (PKH) Desa Minggiran Kediri

Faisal Kurnia Pratama¹, Danang Wahyu Widodo², Niska Shofia³

Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

E-mail: *¹ salmafaisal428@gmail.com, ²danayudo@yahoo.com, ³niskashofia@gmail.ac.id

Abstrak – Program keluarga harapan atau disingkat PKH merupakan program pemerintah ditujukan untuk masyarakat yang memenuhi kriteria sesuai ketentuan pemerintah. Selama ini penyaluran PKH di desa Minggiran kecamatan Papar kabupaten Kediri kurang tepat karena belum ada sistem pengklasifikasi penerima bantuan. Tujuan penelitian ini adalah untuk mempermudah perangkat desa dalam mengklasifikasi calon penerima bantuan sesuai kriteria yang sudah ditentukan. Proses klasifikasi dilakukan dengan mengambil data penduduk kemudian dihitung menggunakan metode *naïve bayes*. *Naïve bayes* yaitu Metode pengklasifikasian dengan menggunakan metode probabilitas dan statistik, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. Data dikumpulkan dengan teknik observasi, wawancara, dan study literatur. Penelitian ini menghasilkan klasifikasi calon penerima bantuan berbasis website sehingga mempermudah pendamping PKH untuk mengetahui calon penerima baru yang layak menerima bantuan tersebut.

Kata Kunci : bantuan program keluarga harapan, klasifikasi, *naïve bayes*

1. PENDAHULUAN

Program Keluarga Harapan (PKH) adalah program pemberian bantuan sosial bersyarat kepada Keluarga Miskin (KM) yang ditetapkan sebagai keluarga penerima manfaat PKH. Sebagai upaya percepatan penanggulangan kemiskinan, sejak tahun 2007 Pemerintah Indonesia telah melaksanakan PKH. Bantuan seperti ini banyak ditunggu oleh warga yang memiliki ekonomi lemah karna dapat membantu mengurangi angka kemiskinan di desa tersebut [1].

Berdasarkan hasil wawancara dengan pendamping PKH di Desa Minggiran Kecamatan Papar Kabupaten Kediri terdapat 15 penerima namun penduduk yang benar-benar menerima sebanyak 9 sedangkan yang lain tidak menerima. Pendamping PKH desa Minggiran juga mengalami kesulitan untuk mengklasifikasi penduduk yang berhak menerima bantuan PKH.

Naïve bayes merupakan konsep klasifikasi dengan menggunakan metode probabilitas dan statistik serta dikemukakan oleh ilmuwan Inggris Thomas Bayes dengan memprediksi peluang di masa

depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Menurut Olson Delen (2008) menjelaskan *Naïve Bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dari "master" [2].

Pada penelitian sebelumnya yang dilakukan Annur (2018) membahas klasifikasi masyarakat miskin menggunakan metode *Naïve Bayes* mendapat nilai akurasi sebesar 73% [3]. Sedangkan penelitian lainnya menunjukkan bahwa penerapan metode *Naïve Bayes* dalam klasifikasi kelayakan penerima beras rastra memperoleh akurasi sebesar 78,33 % [4].

Berikutnya penelitian lain juga membahas tentang implementasi metode *naïve bayes* dalam menentukan dan memprediksi tingkat kesejahteraan keluarga dengan tingkat akurasi mencapai 87% [5].

Penelitian terdahulu tentang sistem pakar untuk mendiagnosa penyakit lambung dengan 5 jenis

penyakit serta menggunakan metode *Naïve Bayes* untuk mendapatkan nilai tertinggi pada hasil diagnosa penyakitnya menunjukkan bahwa sistem tersebut dapat digunakan membantu user [6].

Penelitian ini bertujuan membuat sistem klasifikasi penerima bantuan program keluarga harapan (PKH) berbasis website menggunakan metode *naïve bayes*. Adanya sistem ini dapat membantu dan memudahkan pendamping desa Minggiran untuk melakukan klasifikasi calon penerima di desa Minggiran tersebut.

2. METODE PENELITIAN

Dalam penelitian ini, metode pengembangan sistem berdasarkan model pengembangan *waterfall*. Model *waterfall* merupakan model pengembangan perangkat lunak yang dilakukan secara berurutan. Adapun tahapan penelitian ini antara lain analisis kebutuhan, desain, implementasi dan pengujian sistem [7].

2.1 Analisis Kebutuhan

Kegiatan analisis kebutuhan meliputi identifikasi data, mendefinisikan serta mengumpulkan data yang diperlukan dalam penelitian yaitu berkaitan dengan data penduduk desa Minggiran calon penerima program PKH. Penelitian ini menggunakan data training penerima PKH tahun 2020. Selain itu juga mengumpulkan referensi baik dari buku maupun artikel ilmiah berkaitan dengan metode yang digunakan yaitu *naïve bayes* serta konsep sistem pakar.

2.2 Desain Sistem

Kegiatan perancangan sistem meliputi perancangan data, perancangan proses dan perancangan interface.

2.3 Implementasi

Kegiatan pada tahap ini yaitu membangun aplikasi menggunakan bahasa pemrograman untuk mengembangkan sistem berbasis web.

2.4 Pengujian Sistem

Pengujian sistem dilakukan dengan melihat fungsionalitas sistem. Yaitu menguji kesesuaian sistem yang berjalan dengan tujuan tampilan tersebut dibuat. Selain itu juga dilakukan evaluasi terhadap hasil klasifikasi.

2.5 Sistem Pakar

Sistem pakar adalah sistem berbasis komputer yang menggunakan pengetahuan, fakta, dan teknik penalaran dalam memecahkan masalah yang biasanya hanya dapat dipecahkan oleh seorang pakar dalam bidang tersebut [8].

Pada dasarnya sistem pakar diterapkan untuk mendukung aktivitas pemecahan masalah. Beberapa aktivitas pemecahan yang dimaksud antara lain : pembuatan keputusan, pemanduan pengetahuan, pembuatan desain, perencanaan, prakiraan, pengaturan, pengendalian, diagnosis, perumusan, penjelasan, pemberian nasehat, dan pelatihan. Selain itu sistem pakar juga dapat berfungsi sebagai asisten yang pandai dari seorang pakar [8].

2.6 Naïve Bayes

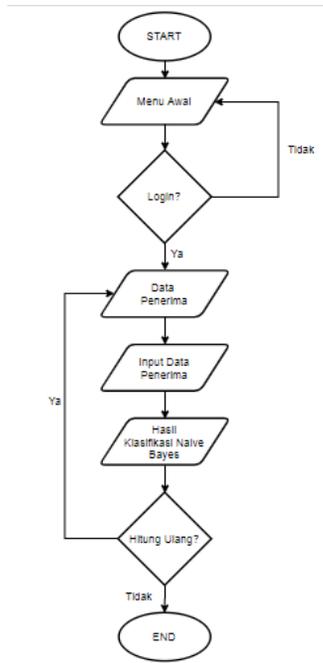
Klasifikasi Bayesian merupakan proses klasifikasi statistik yang digunakan dalam memprediksi probabilitas keanggotaan suatu *class* berdasarkan teorema teorema *bayes*. [9].

Menurut Olson Delen (2008) menjelaskan *Naïve Bayes* untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa ini mengasumsikan bahwa atribut obyek adalah independen. Probabilitas yang terlibat dalam memproduksi perkiraan akhir dihitung sebagai jumlah frekuensi dr " master " tabel keputusan [2].

3. HASIL DAN PEMBAHASAN

3.1 Alur Proses Sistem

Alur proses klasifikasi penerima PKH Desa Minggiran Kecamatan Papar Kabupaten Kediri antara lain ditunjukkan gambar 1. Alur proses diawali dengan login kemudian setelah berhasil login maka dilakukan input data penerima. Setelah itu dilakukan perhitungan dengan menggunakan metode *naïve bayes*. Setelah perhitungan selesai maka data yang diinputkan akan ditampilkan dan pengguna dapat melakukan input data penerima dan di lakukan hitungan ulang jika tidak maka selesai.



Gambar 1. Flowchart Sistem Penerima PKH

3.2 Tampilan *Login*

Tampilan halaman *login* ditunjukkan gambar 2 digunakan ketika admin atau pembina PKH mengakses sistem yaitu dengan menggunakan *username* dan *password* yang sudah terdaftar.



Gambar 2. Tampilan Halaman Login

3.3 Halaman *Home*

Halaman *home* merupakan menu awal ketika berhasil login sehingga dapat melanjutkan ke tahap selanjutnya.



Gambar 3. Tampilan halaman *home*

3.4 Tampilan Data Klasifikasi

Tampilan ini berisi data yang dimasukkan oleh admin yaitu data calon penerima PKH yang berhak menerima atau tidaknya bantuan tersebut. Data calon penerima PKH ditunjukkan gambar 4.

No	Nama	No Kk	NIK	Status	Tanggungan	Umur	Penghasilan	Menerima	Aksi
1	Suyatno	350618022000005	350614470564001	Kawin	3	Muda	rendah	Ya	
2	Suwarno	350618022000001	350614410756004	Cera	2	Tua	sedang	Tidak	
3	Dwi Utami	350618022000008	350614500880002	Cera	2	Tua	sedang	Tidak	
4	Rati Agustina	350618022000008	35061449008790002	Cera	3	Tua	tinggi	Ya	
5	Marni	350618022000022	35061431107720002	Kawin	2	Tua	sedang	Tidak	
6	Suwarno	350618022000028	3506143307810000	Kawin	3	Tua	sedang	Ya	
7	Mardi Hidayat	350618022000040	350614000500001	Cera	1	Muda	rendah	Ya	

Gambar 4. Halaman Hasil Data Klasifikasi

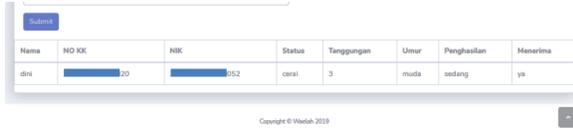
3.5 Tampilan Data Input Calon Penerima Bantuan

Tampilan ini berisi kolom untuk mengisi data diri calon penerima bantuan terdiri dari nama, nomor kartu keluarga (KK), nomor induk kependudukan (NIK), status, umur, tanggungan dan penghasilan dari calon penerima sehingga dapat diklasifikasi menerima atau tidaknya data dari calon penerima tersebut.

Gambar 3. 1 Tampilan Halaman Input

3.6 Tampilan Hasil Klasifikasi

Tampilan ini berisi data hasil klasifikasi calon penerima PKH tersebut berhak atau tidak menerima bantuan.



Gambar 3. 2 Tampilan Hasil Input dan Klasifikasi

3.7 Kebutuhan Data

Penelitian ini menggunakan data training yang diperoleh dari penerima PKH tahun 2020. Data training ditunjukkan pada tabel 1 berjumlah 16 penerima terdiri dari 6 penduduk tidak menerima dan 9 penduduk menerima.

Tabel 1. Data Training

No	Nama	Status	Tang- gungan	Umur	Penghasil- an	Mene- rima
1	Supiatun	Kawin	3	Muda	Rendah	Ya
2	Sunarsih	Cerai	2	Tua	Sedang	Tidak
3	Reti Agustina	Cerai	3	Tua	Tinggi	Ya
4	Marmi	Cerai	2	Tua	Sedang	Tidak
5	Supriatin	Kawin	3	Tua	Sedang	Ya
6	Nurul Hidayat	Cerai	1	Muda	Rendah	Ya
7	Kusrini	Kawin	2	Muda	Rendah	Ya
8	Ripuan	Cerai	1	Muda	Sedang	Tidak
9	Ismiati	Cerai	3	Tua	Sedang	Tidak
10	Sumiatun	Kawin	3	Tua	Rendah	Ya
11	Nawiasih	Kawin	3	Tua	Tinggi	Ya
12	Retnowati	Kawin	2	Tua	Tinggi	Tidak
13	Summinah	Kawin	3	Muda	Rendah	Tidak
14	Musri	Cerai	1	Muda	Sedang	Ya
15	Romelah	Kawin	3	Tua	Sedang	Ya

Keterangan:

- 1) Umur Muda = ≤ 30 tahun
Tua = > 30 tahun
- 2) Penghasilan Rendah = Kurang dari 1 juta
Sedang = 1 juta – 2 juta
Tinggi = lebih dari 2 juta

3.7 Perhitungan Sederhana

Perhitungan sederhana dilakukan dengan menggunakan rumus teorema bayes.

$$P(C|X) = \frac{P(x|c)P(c)}{P(x)} \dots\dots\dots(1)$$

Keterangan :

- x : Data dengan class yang belum diketahui
- c : Hipotesis data merupakan suatu class spesifik
- $P(c/x)$: Probabilitas hipotesis berdasar kondisi (posteriori probability)
- $P(c)$: Probabilitas class yang di cari (prior probability) dari keseluruhan data
- $P(x/c)$: Probabilitas berdasarkan kondisi pada class
- $P(x)$: Probabilitas c

Persamaan pertama menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (*Posterior*) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut *prior*), dikali dengan peluang kemunculan karakteristik sampel pada kelas C (disebut juga *likelihood*), dibagi dengan peluang kemunculan karakteristik sampel secara global (disebut juga *evidence*). Karena itu, rumus diatas dapat pula ditulis sebagai berikut :

$$\text{Posterior} = \frac{\text{Prior} \times \text{likelihood}}{\text{Evidence}} \dots\dots\dots(2)$$

Nilai *evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut selanjutnya akan dibandingkan dengan nilai posterior kelas lainnya untuk menentukan ke kelas apa suatu sampel akan diklasifikasikan. Penjabaran lebih lanjut rumus Bayes tersebut dilakukan dengan rincian sebagai berikut:

Penerima ‘ya’ = $9 / 15 = 0,6$
Penerima ‘tidak’ = $6/15 = 0,4$

Perhitungan atribut yang lain sama seperti rumus di atas.

Tabel 2. Perhitungan setiap atribut

Atribut	Nilai	
	Ya	Tidak
Probabilitas Penerima		
Menerima	0,6	0,4
Status		
Kawin	0,56	0,50
Cerai	0,44	0,50
Umur		
Tua	0,56	0,67
muda	0,44	0,33
Tanggungan		
1	0,22	0,17
2	0,11	0,5
3	0,67	0,33
Penghasilan		
Rendah	0,44	0,17
Sedang	0,33	0,67
Tinggi	0,22	0,17

3.8 Hasil Uji Klasifikasi

Hasil uji klasifikasi berasal dari perhitungan secara sederhana sesuai dengan data yang diinputkan dengan menggunakan rumus berikut ini.

$$\begin{aligned}
 &P(C|X_1, \dots, X_n) \\
 &= P(C)P(X_1, \dots, X_n|C) = P(C)P(X_1, \dots, X_n|C) \\
 &= P(C)P(X_1|c)P(X_2|C, X_1)P(X_3, \dots, X_n|C, X_1, X_2) \\
 &= P(C)P(X_1, \dots, X_n|C) \\
 &= P(C)P(X_1|c)P(X_2|C, X_1)P(X_3, \dots, X_n|C, X_1, X_2) \\
 &= P(C)P(X_1, \dots, X_n|C) \\
 &= P(C)P(X_1|c)P(X_2|C, X_1)P(X_3, \dots, X_n|C, X_1, X_2) \\
 &\quad (X_n|C, X_1, X_2, \dots, X_{n-1}) \dots\dots(3)
 \end{aligned}$$

Data calon penerima:

$$\begin{aligned}
 \text{Sulami Ya} &= 0,6 \cdot 0,44 \cdot 0,56 \cdot 0,67 \cdot 0,44 \\
 &= 0.044
 \end{aligned}$$

$$\begin{aligned}
 \text{Sulami Tidak} &= 0,4 \cdot 0,50 \cdot 0,67 \cdot 0,33 \cdot 0,17 \cdot \\
 &= 0.007
 \end{aligned}$$

Berdasarkan perkalian dapat diketahui bahwa hasil penjabaran tersebut menyebabkan semakin banyak dan semakin kompleksnya faktor faktor syarat yang mempengaruhi nilai probabilitas, yang hampir mustahil untuk dianalisa satu persatu. Akibatnya,

perhitungan tersebut menjadi sulit untuk dilakukan. Maka digunakan asumsi independensi yang sangat tinggi (naif), bahwa masing masing petunjuk saling bebas (independen) satu sama lain [8].

$$\text{Sulami Ya} = 0,6 \cdot 0,44 \cdot 0,56 \cdot 0,67 \cdot 0,44 = 0.044$$

$$\text{Sulami Tidak} = 0,4 \cdot 0,50 \cdot 0,67 \cdot 0,33 \cdot 0,17 = 0.007$$

Penggunaan persamaan untuk setiap penduduk sesuai atribut beserta masing-masing nilainya ditunjukkan pada tabel 3.

Tabel 3. Hasil Klasifikasi

Nama	Menerima	Ya	Tidak
sulami	ya	0,044	0,007
ana	tidak	0,005	0,022
ani	ya	0,026	0,015
dini	ya	0,026	0,015

3.9 Pengujian Sistem

Pengujian sistem dilakukan dengan melihat fungsionalitas sistem. Sedangkan hasil perhitungan menggunakan metode naïve bayes sesuai data training ditunjukkan seperti gambar 6.

15	Romelah	350618000300033	3506145107090001	Kawin	3	Tua	sedang	Ya
16	ana	350618016014309	3506145007720001	Kawin	2	Muda	sedang	Tidak
17	sulami	350618021000021	3506180210000051	Cerai	3	Tua	rendah	Ya
18	ani	350618021000032	350618021000035	Cerai	3	Muda	sedang	Ya
19	dini	350618021000033	350618021000037	Cerai	3	Muda	sedang	Ya

Gambar 3 Hasil Klasifikasi Data Inputan

Perhitungan *confusion matrix* dilakukan untuk menguji luaran klasifikasi dalam mendapat nilai akurasi. Hasil perhitungan akurasi diperoleh dari tabel 4.

Tabel 4. Confusion Matrix

Aktual	Prediksi	
	Menerima: ya	Menerima: tidak
Menerima: ya	9	0
Menerima: tidak	1	5

Dengan rumus:

$$\text{Akurasi} = (9+5) / (9+0+1+5) / 100\% = 93\%$$

4. KESIMPULAN

Berdasarkan pembahasan tentang Algoritma *Naive Bayes* dalam sistem terdapat akurasi sebesar 93,33%. Sehingga dapat diterapkan dalam memprediksi peluang dimasa depan berdasarkan pengalaman dimasa sebelumnya untuk memudahkan pengurus desa dalam menentukan calon penerima bantuan PKH.

5. SARAN

Dalam rangka meningkatkan performa dari Algoritma *Naive Bayes*, ada beberapa saran yang dapat menjadi masukan pada penelitian ini :

1. Volume data yang digunakan dapat ditambah dengan data yang lebih banyak lagi, di karenak metode *Naive Bayes* semakin banyak data latih yang digunakan maka semakin baik hasil yang diberikan.
2. Penulis mengharapkan penelitian ini dapat diterapkan dengan menggunakan algoritma lain agar dapat mengembangkan penelitian dimasa yang akan datang.

DAFTAR PUSTAKA

- [1] Irwan Susanto. 2019. Program Keluarga Harapan (PKH). <https://kemensos.go.id/program-keluarga-harapan-pkh> diakses pada tanggal 27 November 2020.
- [2] Olson & Delen. 2008. *Advanced Data Mining Techniques*. Springer-Verlag Berlin Heidelberg: USA.
- [3] Annur, H. 2018. Klasifikasi Masyarakat Miskin Menggunakan Metode Naïve Bayes. *ILKOM Jurnal Ilmiah Volume 10 Nomor 2. Agustus 2018, 10*, 161-165.
- [4] Agus Predana, C. S. 2018. Penerapan Metode Naïve Bayes Dalam Klasifikasi Kelayakan Keluarga Penerima Beras Rastra . *JUTIM, Vol 3 No.1*, 1-8.
- [5] Widyadara, M. A., & Irawan, R. H. 2019. Implementasi Metode Naïve Bayes Dalam Penentuan Tingkat Kesejahteraan Keluarga. *Research : Journal of Computer, information system, & technology management Vol. 2, No. 1. April 2019*, 19-24.
- [6] Minarni, & Irawan, P. 2019. Implementasi Metode Naive Bayes Untuk Diagnosa Penyakit Lambung. *Jurnal TEKNOIF Vol. 7 No. 2 Oktober 2019*, 7, 125-123.
- [7] Hidayat, C. 2018. Pengertian Metode Waterfall . <https://ranahresearch.com/metode-waterfall/> diakses tanggal 28 November 2020.
- [8] Kusrini. 2006. *Sistem Pakar Teori dan Aplikasi*. Andi: Yogyakarta.
- [9] Santosa, B. 2007. *Data Mining Teknik Pemanfaata Data untuk Keperluan Bisnis*. Graha Ilmu: Yogyakarta.