



Application of Supervised Learning Algorithm for Classification of Family Hope Program

Penerapan Algoritma Supervised Learning untuk Klasifikasi Program Keluarga Harapan

**Muhammad Syarif Hartawan¹, Moh. Erkamim², Sitti Rachmawati Yahya³,
Nirma Ceisa Santi⁴, Legito⁵, Sepriano⁶**

¹Program Studi Sistem Informasi, Fakultas Teknik, Universitas Krisnadwipayana, Indonesia

²Program Studi Sistem Informasi Kota Cerdas, Universitas Tunas Pembangunan Surakarta, Indonesia

³Program Studi Sistem Informasi, Universitas Siber Asia, Jakarta, Indonesia

⁴Program Studi Sistem Informasi, Fakultas Sains dan Teknologi,
Universitas Nahdlatul Ulama Sunan Giri, Indonesia

⁵Jurusan Teknik Informatika, Sekolah Tinggi Teknologi Sinar Husni Deliserdang Sumatera Utara, Indonesia

⁶Program Studi Sistem Informasi, Universitas Islam Negeri Sulthan Thaha Saifuddin Jambi, Indonesia

E-Mail: ¹muhammadsyarif@unkris.ac.id, ²erkamim@lecture.utp.ac.id, ³sitti.rachma@gmail.com,
⁴nirmaceisa@unugiri.ac.id, ⁵legitostt@gmail.com, ⁶sepriano@uinjambi.ac.id

*Received Mar 28th 2023; Revised Jul 3rd 2023; Accepted Aug 10th 2023
Corresponding Author: Muhammad Syarif Hartawan*

Abstract

This research is intended to classify the beneficiaries of the Family Hope Program (PKH) Pekanbaru City by comparing three methods at once, namely K-Nearest Neighbor (KNN), Probabilistic Neural Network (PNN) and Nave Bayes (NBC). (The attributes used in the classification process are the number of elementary school children, the number of junior high school children, the number of pregnant women, and the number of children under five years (toddlers), these attributes are based on the guidelines of the National Team for Poverty Reduction and Handling (TNP2K) Ministry of Social Affairs of the Republic of Indonesia. The classification process is carried out on very poor households (RTSM) as training data with a total of 450 data as testing data with a total of 10 data, so that by making comparisons the accuracy results vary between the three methods. The Nave Bayes (NBC) method has the highest accuracy results with 80% accuracy results, which are compared to the KNN 20% and PNN 10% methods, so the Nave Bayes (NBC) method is determined to be the best method for the case of classifying the family hope program (PKH) Pekanbaru City.

Keyword: K-Nearest Neighbor, Naive Bayes Classifier, Probabilistic Neural Network, Family Hope Program, Households Extremely Poor Households

Abstrak

Penelitian ini dimaksudkan untuk melakukan klasifikasi terhadap penerima bantuan Program Keluarga Harapan (PKH) Kota Pekanbaru dengan membandingkan tiga metode sekaligus yaitu K-Nearest Neighbor (KNN), Probabilistic Neural Network (PNN) dan Naive Bayes Classifier (NBC). Atribut yang digunakan dalam proses klasifikasi adalah Jumlah Anak SD, Jumlah Anak SMP, Jumlah Ibu Hamil, dan Jumlah Anak dibawah lima tahun (Balita), atribut ini berdasarkan panduan Tim Nasional Penanggulangan dan Penanganan Kemiskinan (TNP2K) Kementerian Sosial Republik Indonesia. Proses klasifikasi dilakukan terhadap Rumah Tangga Sangat Miskin (RTSM) sebagai data training dengan jumlah 450 data sebagai data testing dengan jumlah 10 data, sehingga dengan melakukan perbandingan didapatkan hasil akurasi yang bervariasi diantara ke-tiga metode. Metode Nave Bayes (NBC) memiliki hasil akurasi yang paling tinggi dengan hasil akurasi 80%, yang dibandingkan dengan metode KNN 20% dan PNN 10% maka Metode Nave Bayes(NBC) ditetapkan menjadi metode terbaik untuk kasus pengklasifikasian program keluarga harapan (PKH) Kota Pekanbaru.

Kata Kunci: K-Nearest Neighbor, Naive Bayes Classifier, Probabilistic Neural Network, Program Keluarga Harapan, Rumah Tangga Sangat Miskin

1. PENDAHULUAN

Kemiskinan merupakan masalah global yang dihadapi negara di dunia termasuk Indonesia. Kemiskinan digambarkan dengan kondisi seseorang yang tidak dapat memenuhi kebutuhan pokoknya meliputi, sandang, pangan, dan papan [1][2]. Kemiskinan telah membatasi hak rakyat untuk mendapatkan pendidikan yang layak, mendapatkan pekerjaan yang memadai, mengakses kesehatan yang terjangkau, dan kemiskinan menjadi alasan rendahnya Indeks Pembangunan Manusia di Indonesia [3]. Salah satu program pemerintah dalam penanganan masalah kemiskinan adalah melalui program berbasis rumah tangga yang disebut Program Keluarga Harapan (PKH) [4].

PKH merupakan bantuan langsung tunai dengan sasaran Rumah Tangga sangat Miskin (RTSM) bertujuan untuk meningkatkan partisipasi pendidikan dan kesehatan bagi para RTSM [5][6]. PKH merupakan program yang berada dibawah naungan kementerian sosial republik indonesia melalui Tim nasional penanganan dan penanggulangan kemiskinan (TNP2K) [7] dan akan dilaksanakan oleh unit pelaksana program keluarga harapan (UPPKH)[8].

Banyaknya data penduduk yang bervariasi maka digunakanlah klasifikasi. Adapun teknik yang sering digunakan pada klasifikasi adalah algoritma K-Nearest Neighbor (KNN). Algoritma KNN adalah suatu metode yang menggunakan algoritma supervised [9]. Untuk menentukan jumlah bantuan PKH yaitu berdasarkan atribut jumlah anak SD, jumlah anak SMP, jumlah ibu hamil, dan jumlah balita [10]. Penerimaan jumlah bantuan PKH terdiri dari beberapa macam yaitu Rp. 800.000, Rp. 1.300.000, Rp. 1.800.000, Rp. 2.300.000, dan Rp. 2.800.000. dari perbedaan jumlah bantuan yang diterima setiap RTSM dapat artikan bahwa ada 5 (lima) tingkatan klasifikasi dalam pemberian bantuan PKH.

Khusus Kota Pekanbaru, PKH berada dibawah naungan dinas sosial kota pekanbaru, PKH Pekanbaru di koordinasi oleh UPPKH Kota Pekanbaru dan telah dilaksanakan sejak tahun 2013 sampai sekarang [11]. Dari tahun 2013 sebanyak 2546 data, 2014 sebanyak 2467 data, dan 2015 sebanyak 2382 data. Dengan adanya perbedaan penerimaan bantuan PKH, maka akan dilakukan klasifikasi berdasarkan komponen kesehatan, pendidikan, dan kesehatan pendidikan dengan 4 atribut yaitu jumlah anak SD, jumlah anak SMP, jumlah ibu hamil, dan jumlah anak balita.

Pada penelitian sebelumnya telah dilakukan pengklasifikasi dengan metode yang berbeda yaitu metode Modified K-Nearest Neighbor (MKNN) maka menghasilkan akurasi dengan nilai 99,51%. Pengujian algoritma ini dilakukan menggunakan perhitungan confusion matriks yaitu membandingkan kelas label faktual dengan kelas label klasifikasi. Namun [12] pada penelitiannya membandingkan MKNN dengan K-NN didapatkan hasil performance KNN lebih tinggi dibandingkan dengan MKNN. Kemudian pada penelitian [13] yaitu melakukan pengklasifikasian dengan membandingkan Naïve Bayes Classifier (NBC) dan Decision Tree, didapatkan hasil NBC lebih akurat dibandingkan dengan Decision tree dengan tingkat akurasi sebesar 93.96%. Selanjutnya pada penelian lainnya dilakukan pengklasifikasian dengan menggunakan Algoritma Probabilistic Neural Network (PNN) didapatkan hasil bahwa kinerja dari PNN sangat baik dibandingkan dengan algoritma lain [14].

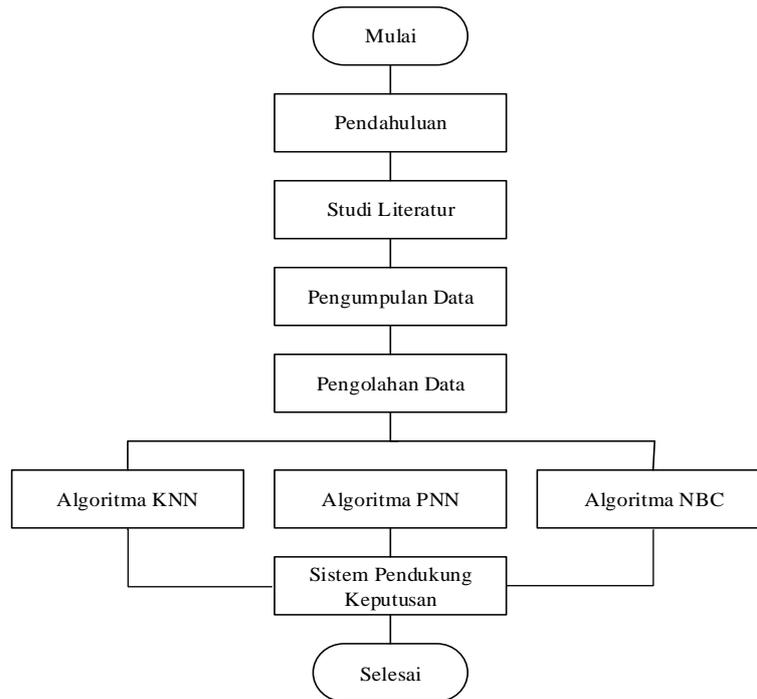
Oleh karena itu pada penelitian ini dilakukanlah perbandingan dengan menggunakan tiga metode yang berbeda yaitu KNN, NBC dan PNN. Sehingga dapat menemukan tingkat akurasi yang paling tinggi diantara ketiga metode tersebut. Tujuan dari penelitian ini yaitu mencari algoritma yang paling tepat untuk dijadikan sebagai sistem prediksi pada kasus penentu tingkatan jumlah bantuan yang akan diterima penduduk. Algoritma yang memiliki nilai akurasi yang paling tinggi, maka ditetapkan sebagai metode yang paling efektif dan efisien untuk dikembangkan kedepannya dan menjadi rekomendasi dinas terkait.

2. METODOLOGI

Metode penelitian yang diterapkan dalam penelitian ini dapat dilihat pada Gambar 1. Secara umum penelitian utama dibagi menjadi 4 bagian utama, yaitu studi literatur dan pengumpulan data, pra proses dan pengolahan data, perbandingan algoritma dan pengambilan keputusan. Adapun algoritma yang digunakan dalam penelitian ini adalah membandingkan performa dari algoritma KNN, PNN dan NBC. Percobaan akan dilakukan dengan menggunakan simulasi 10 percobaan untuk masing-masing algoritma, 10 percobaan ini menjadi dasar utama dalam proses klasifikasi. Jika dalam 10 percobaan tidak ada respon efektif atau perubahan pada grafik, maka akan diteruskan menjadi 20 percobaan, hingga seterusnya. Namun jika pada 10 percobaan terdapat nilai pembeda pada hasil klasifikasi maka percobaan dihentikan.

2.1 Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [15][16]. Data mining merupakan metode untuk mengekstraksi dan mengidentifikasi informasi pengetahuan potensial dari suatu informasi tersembunyi yang tersimpan di dalam database besar. Tujuan utama data mining adalah untuk menemukan, menggali dan atau menambang suatu informasi dan pengetahuan dari data yang dimiliki [17].



Gambar 1. Metodologi Penelitian

2.2 Algoritma K-Nearest Neighbor (KNN)

KNN adalah salah satu algoritma machine learning dengan pendekatan *supervised* yang bekerja dengan membuat kelas data baru menggunakan kemiripan antar data baru dengan sejumlah data (k) pada lokasi terdekat [18]. KNN termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. KNN dilakukan dengan mencari kelompok k objek dalam data *training* yang paling dekat (mirip) dengan objek pada data baru atau data *testing*. Dekat atau jauhnya jarak suatu data dapat dihitung berdasarkan jarak Euclidian menggunakan Persamaan 1 [19][20].

$$d(a, b) = \sum_{i=0}^n (X_i - Y_i)^2 \quad (1)$$

2.3 Algoritma Probabilistic Neural Network (PNN)

PNN dikembangkan oleh Donald F. Specht pada tahun 1988 yang merupakan salah satu metode klasifikasi jaringan syaraf tiruan [21] dan termasuk kedalam struktur metode *feedforward* [22]. PNN didasarkan pada jaringan syaraf radial. PNN menggunakan fungsi aktivasi yang berasal dari statistic sebagai pengganti fungsi aktivasi sigmoid [23]. Hasil dari algoritma ini akan menghasilkan klasifikasi berdasarkan nilai yang terdapat pada summation layer dengan nilai terbesar yang dinyatakan pada Persamaan 2 [24].

$$d(x) = C_k \text{ jika } p(C_k) > p(X|C_i) \text{ Pr}(c_j) \quad (2)$$

Model jaringan syaraf tiruan probabilistik yang dibuat oleh Cain memperbolehkan setiap kelas memiliki parameter penghalus, σ_k , yang berbeda satu dengan yang lain dan menerapkan algoritma belajar yang baru untuk memperoleh σ_k secara otomatis. Apabila tiap kelas memiliki parameter yang memiliki fungsi peluang dapat ditulis pada persamaan 3.

$$p(x|C_k) = \frac{1}{(2\pi)^{\frac{m}{2}} \sigma_k^m} \sum_{p_i \in C_k} \exp[-||x - w_i||^2 (2\sigma_k^2)] \quad (3)$$

2.4 Algoritma Naïve Bayes Classification (NBC)

NBC adalah pengklasifikasian statistic yang dapat digunakan untuk memprediksi probabilistic keanggotaan suatu *class* [25]. NBC didasarkan pada teorema *Bayes* yang memiliki kemampuan klasifikasi serupa dengan *Decision Tree* dan *Neural Network*. NBC terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam *Database* dengan data yang besar. NBC merupakan metode pengklasifikasian probabilistic sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan variasi nilai dari dataset yang diberikan [26]. Teorema bayes dapat menggunakan Persamaan 4 [27].

$$P(X|C_i) = \frac{P(X|C_i)P(C_i)}{P(X)} \quad (4)$$

$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i) \quad (5)$$

3. HASIL DAN PEMBAHASAN

3.1. Mendefinisikan Kriteria

Sesuai dengan metodologi penelitian pada pembahasan sebelumnya, beberapa hal penting yang akan dilakukan dalam menyelesaikan penelitian ini terdiri atas pengumpulan data, pra proses dan pengolahan data, perbandingan algoritma dan merancang Sistem Pendukung Keputusan (SPK). Dataset yang digunakan bersumber dari UPPKH, berdasarkan penetapan kelas UPPKH dapat dilihat pada Tabel 1.

Tabel 1. Kelas Berdasarkan UPPKH 2013

Nominal Bantuan	Bantuan Tetap	Komponen			Keterangan
		Bumil/ Nifas/ Balita	Anak SD	Anak SMP	
800.000	300.000	0	500.000	0	bila 1 anak SD
	300.000	1.000.000	0	0	bila ada bumil/nifas/balita
1.300.000	300.000	0	1.000.000	0	bila ada 2 anak SD
	300.000	0	0	1.000.000	bila ada 1 anak SMP
	300.000	1.000.000	500.000	0	bila ada bumil/nifas/balita dan 1 anak SD
1.800.000	300.000	0	500.000	1.000.000	bila 1 anak SD dan 1 Anak SMP
	300.000	0	1.500.000	0	bila 3 anak SD
	300.000	1.000.000	0	1.000.000	bila ada bumil/nifas/balita dan 1 anak SMP
	300.000	0	1.000.000	1.000.000	bila 2 anak SD dan 1 anak SMP
2.300.000	300.000	1.000.000	1.000.000	0	bila ada bumil/nifas/balita dan 2 anak SD
	300.000	0	0	2.000.000	bila ada 2 anak SMP
	300.000	1.000.000	500.000	1.000.000	bila ada bumil/nifas/balita dan 1 anak SD dan 1 anak SMP
	300.000	1.000.000	1.500.000	0	bila ada bumil/nifas/balita dan 3 anak SD
2.800.000	300.000	0	1.500.000	1.000.000	bila ada 3 anak SD dan 1 anak SMP
	300.000	0	500.000	2.000.000	bila 1 anak SD dan 2 anak SMP

Pembersihan data dilakukan untuk mengurangi efek *noise* pada saat proses perhitungan dan menghilangkan atribut yang tidak digunakan. Dan selanjutnya proses normalisasi data. Hasil detail Normalisasi data dapat dilihat pada tabel 2. Untuk data training dan data testing untuk percobaan pada ketiga algoritma yang digunakan ditunjukkan pada tabel 3 dan tabel 4.

Tabel 2. Normalisasi Data

JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA
0,2	0	0	0
0,2	0,5	0	0,25
0,2	0	0	0,25
0,6	0	0	0,25
0,4	0	0	0,5
0,8	0	0	0,25
0,6	0	0	0,25
0,4	1	0	0,25

Tabel 3. Data Training

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	TAHUN	KELAS
IG-1	1	0	0	0	2013	1
IG-2	1	1	0	1	2013	5
IG-3	1	0	0	1	2013	3
IG-4	3	0	0	1	2013	5
IG-5	2	0	0	2	2013	4
IG-6	4	0	0	1	2013	5
IG-7	2	2	0	1	2013	5

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	TAHUN	KELAS
IG-8	3	3	1	2	2013	4
IG-9	2	1	1	1	2013	3
...
IG-1294	1	1	1	1	2013	5

10 dari 1304 data awal digunakan sebagai data testing dinormalisasi menggunakan *min-max training* pada simulasi metode KNN *normalization*.

Tabel 4. Data Testing

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	TAHUN
IG-1295	1	0	0	1	2013
IG-1296	1	0	0	1	2013
IG-1297	2	0	0	0	2013
IG-1298	2	1	1	1	2013
IG-1299	1	0	0	1	2013
IG-1300	3	1	0	1	2013
IG-1301	1	1	0	1	2013
IG-1302	0	0	0	1	2013
IG-1303	1	1	0	1	2013
IG-1304	2	0	0	2	2013

Tabel 3 dan tabel 4 merupakan simulasi preprocessing data sebelum dilakukan proses mining menggunakan algoritma KNN, PNN dan NBC. Preprocessing dilakukan dengan menggunakan bahasa pemrograman python.

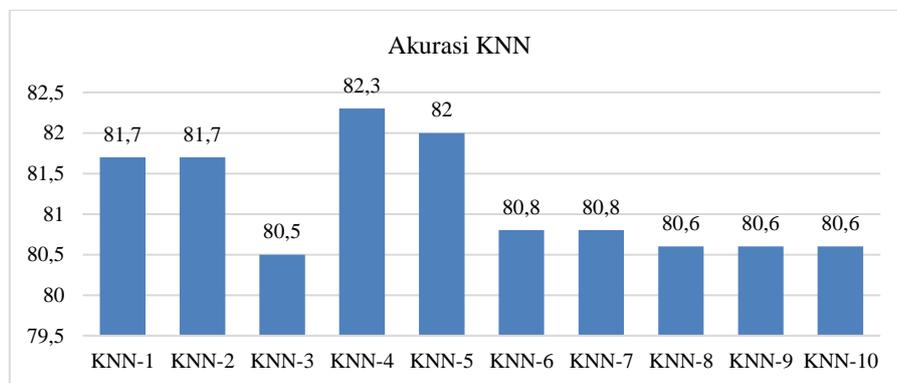
3.2. Percobaan dengan Metode KNN

Percobaan dengan menggunakan algoritma KNN digunakan persamaan 1 sebagai penentuan nilai jarak antar satu data dengan data yang lain. Dari hasil proses klasifikasi diperoleh hasil klasifikasi yang ditunjukkan pada table 5.

Tabel 5. Hasil Klasifikasi dengan KNN

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	KELAS
IG-1295	1	0	0	1	4
IG-1296	1	0	0	1	4
IG-1297	2	0	0	0	2
IG-1298	2	1	1	1	1
IG-1299	1	0	0	1	4
IG-1300	3	1	0	1	2
IG-1301	1	1	0	1	2
IG-1302	0	0	0	1	5
IG-1303	1	1	0	1	2
IG-1304	2	0	0	2	4

Dari 4 nilai jarak, data menunjukkan nilai peredikat sangat bervariasi, sehingga untuk nilai prediksi kalsifikasi data baru yang pertama. Dari hasil pengukuran dan evaluasi data dapat ditunjukkan bahwa akurasi yang dihasilkan algoritma KNN adalah 82,3% pada percobaan ke 4 dengan nilai akurasi detail pada 10 percobaan ditunjukkan pada gambar 2.



Gambar 2. Hasil Percobaan Algoritma KNN

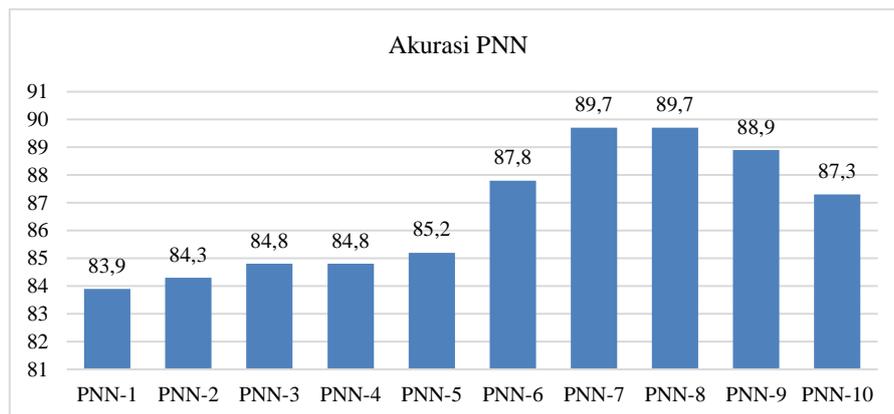
3.3. Percobaan dengan Metode PNN

Percobaan algoritma PNN menggunakan beberapa skema atau prosedur yaitu Misalkan terdapat n kelas, $C_0, C_1, C_2, \dots, C_{n-1}$; diasumsikan pola yang diamati adalah variabel acak x dengan m -dimensi dan fungsi pada peluang bersyarat x , bila diketahui bahwa pola tersebut berasal dari kelas C_k , dinotasikan dengan $P = (x|C_k)$. Dengan menerapkan aturan pertama dari metode jaringan syaraf tiruan, dapat mengacu kepada persamaan 3. Dengan mengacu pada prosedur dan persamaan tersebut dihasilkan klasifikasi yang dapat ditunjukkan pada tabel 6.

Tabel 6. Hasil Klasifikasi dengan PNN

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	KELAS
IG-1295	1	0	0	1	4
IG-1296	1	0	0	1	4
IG-1297	2	0	0	0	2
IG-1298	2	1	1	1	3
IG-1299	1	0	0	1	4
IG-1300	3	1	0	1	5
IG-1301	1	1	0	1	2
IG-1302	0	0	0	1	1
IG-1303	1	1	0	1	2
IG-1304	2	0	0	2	2

Tabel 6 menunjukkan hasil dari klasifikasi menggunakan algoritma PNN didapatkan hasil dari pengklasifikasian yang berbeda dari metode sebelumnya. Hasil dari 10 percobaan dengan PNN diperoleh akurasi terbaik adalah 89,7% pada percobaan ke 7 dan percobaan ke 8 dengan detail hasil percobaan ditunjukkan pada gambar 3.



Gambar 3. Hasil Percobaan Algoritma KNN

3.4. Percobaan dengan Metode NBC

Percobaan dengan algoritma NBC dilakukan dengan mengacu kepada persamaan 4, dengan beberapa prosedur:

1. Setiap data dipresentasikan sebagai vector berdimensi- n yaitu $X = (x_1, x_2, x_3, \dots, x_n)$, dimana n adalah gambaran dari ukuran yang dibuat ditest dari n atribut yaitu $A_1 - A_2, A_3 \dots A_n$.
2. m adalah kumpulan kategori yaitu $C_1, C_2, C_3 \dots, C_m$. Diberikan data test X yang tidak diketahui kategorinya, maka classifier akan memprediksi bahwa X adalah milik kategori dengan posterior probability tertinggi berdasarkan kondisi X . Oleh karena itu, NBC menandai bahwa test X yang tidak diketahui tadi ke kategori C_i jika dan hanya jika:

$$P(C_i|X) > P(C_j|X) \text{ Untuk } 1 \leq j \leq m, j \neq i$$

Kemudian memaksimalkan: $P(C_i|X)$. Class C_i dari $P(C_i|X)$ yang dimaksimalkan biasa disebut dengan *maximum posteriori hypothesis*. Kondisi ini diformulasikan pada persamaan 4.

3. Dari persamaan 4, $P(X)$ adalah konstan untuk semua kategori, hanya $P(X | C_i)$. $P(C_i)$ yang perlu dimaksimalkan. Jika class prior probability tidak diketahui, maka akan diasumsikan sama dengan hasil dari kategori-kategori yang lain seperti $P(C_1) = P(C_2) = \dots = P(C_m)$ dan oleh karena itu kita akan memaksimalkan $P(X|C_i)$, $P(C_i)$. Perlu dicatat bahwa class prior probability mungkin diperkirakan

dengan perhitungan $P(C_i) = s_i$ dimana s_i adalah jumlah dari data training s dari kategori C_i dan s adalah jumlah total data training.

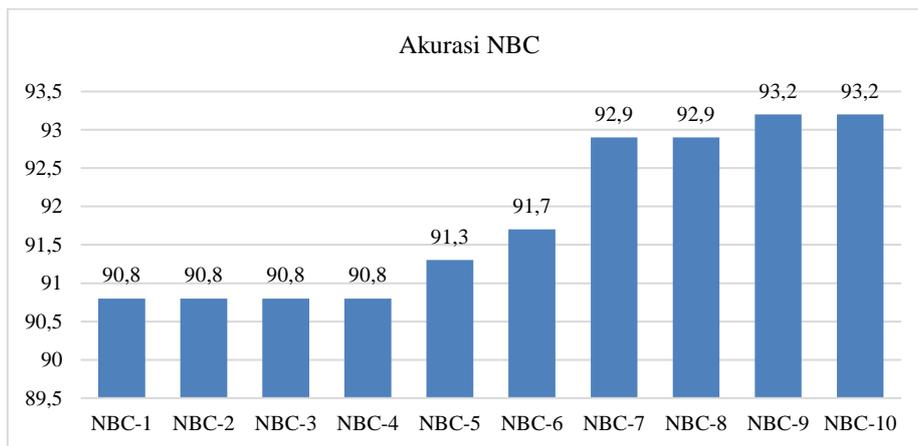
4. Diberikan data dengan banyak atribut, ini akan menjadi komputasi yang kompleks untuk mengkomputasi $P(X|C_i)$. Untuk mengurangi komputasi pada saat mengevaluasi $P(X|C_i)$, maka dapat dihitung menggunakan persamaan 5.

Dari prosedur an proses yang dilakukan, maka didapatkan hasil klasifikasi menggunakan algoritma NBC yang ditunjukkan pada table 7.

Tabel 7. Hasil Klasifikasi dengan PNN

NO_DATA	JML_ANAK_SD	JML_ANAK_SMP	JML BUMIL	JML_BALITA	KELAS
IG-1295	1	0	0	1	3
IG-1296	1	0	0	1	3
IG-1297	2	0	0	0	4
IG-1298	2	1	1	1	4
IG-1299	1	0	0	1	3
IG-1300	3	1	0	1	5
IG-1301	1	1	0	1	5
IG-1302	0	0	0	1	2
IG-1303	1	1	0	1	5
IG-1304	2	0	0	2	4

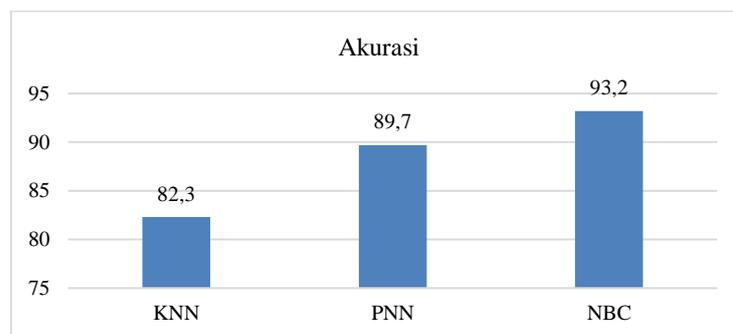
Tabel 7 menunjukkan hasil dari klasifikasi menggunakan algoritma NBC diperoleh hasil dari pengklasifikasian yang berbeda dari algoritma KNN dan PNN. Hasil dari 10 percobaan dengan NBC diperoleh akurasi terbaik adalah 93,2% pada percobaan ke 9 dengan detail hasil percobaan ditunjukkan pada gambar 4.



Gambar 4. Hasil Percobaan Algoritma KNN

3.5. Perbandingan Keakuratan

Perbandingan keakuratan dari evaluasi algoritma KNN, PNN dan NBC mengacu kepada hasil akurasi tertinggi, secara umum dapat dilihat pada gambar 5.



Gambar 5. Perbandingan Tingkat Akurasi

Dari hasil pengklasifikasian menggunakan 3 algoritma dapat disimpulkan bahwa algoritma terbaik dari sisi akurasi adalah NBC dibandingkan metode KNN dan PNN untuk kasus klasifikasi PKH Kota Pekanbaru. Untuk selanjutnya pengambilan keputusan diambil berdasarkan hasil rekomendasi algoritma terbaik yaitu NBC dalam penentuan penerima bantuan.

4. KESIMPULAN

Berdasarkan hasil analisa dan pembahasan maka dapat ditarik kesimpulan bahwasannya pemilihan metode terbaik adalah yang mempunyai tingkat akurasi yang paling tinggi. Algoritma NBC memiliki hasil akurasi yang paling tinggi dengan presentase 93,2%, dibandingkan algoritma yang lain yaitu KNN dengan akurasi 82,3% dan algoritma PNN dengan akurasi 89,7%. Dengan demikian untuk kasus pengklasifikasian PKH kota Pekanbaru algoritma NBC merupakan algoritma yang paling tepat untuk dijadikan pemodelan dalam pengambilan keputusan.

REFERENCES

- [1] M. F. Dakhilullah, "Pengaruh Upah Minimum Dan Pengangguran Terhadap Kemiskinan di Pulau Jawa Sebelum dan Selama Pandemi," *J. Ilmu Manajemen, Ekon. dan Kewirausahaan*, vol. 1, no. 1, pp. 60–75, 2023.
- [2] M. Irawati and M. A. I. Pakereng, "Analisis Pengaruh Jumlah Pengangguran Terhadap Jumlah Kemiskinan Menggunakan Metode Regresi Linier (Studi Kasus: Kota Salatiga)," *J. EMT KITA*, vol. 7, no. 2 SE-Articles, pp. 401–408, Apr. 2023, doi: 10.35870/emt.v7i2.1013.
- [3] I. P. F. Rorong, "Jurnal Pembangunan Ekonomi dan Keuangan Daerah Vol 23. No 4 (2022)," vol. 23, no. 4, pp. 398–415, 2022.
- [4] D. I. Kelurahan, P. Kecamatan, and B. Utara, "Journal I La Galigo : Public Administration Journal PEMBERDAYAAN DAN PENINGKATAN PROGRAM KELUARGA HARAPAN Journal I La Galigo : Public Administration Journal," vol. 6, no. April, pp. 40–49, 2023.
- [5] R. A. U. K. Umbu, Yoyok Seby Dwanoko, and Hari Lugis Purwanto, "Sistem Rekomendasi Penerima Bantuan PKH Dengan Metode Smart Guna Mendukung Keputusan Pemerintah Desa," *RAINSTEK J. Terap. Sains Teknol.*, vol. 4, no. 3 SE-Articles, pp. 182–191, Sep. 2022, doi: 10.21067/jtst.v4i3.7818.
- [6] I. A. Sobari and R. A. Zuama, "Pendekatan Machine Learning dalam Memprediksi Keluarga Penerima Program PKH," vol. 9, no. 1, pp. 61–64, 2023, doi: 10.31294/jtk.v4i2.
- [7] S. Latore, R. Harold, S. Bumulo, and A. Ali, "Dampak Program Keluarga Harapan dalam Pengentasan Kemiskinan di Desa Bangga Kecamatan Paguyaman Pantai The Impact of the Family Hope Program in Poverty Alleviation in Bangga Village , Paguyaman Pantai District," vol. 01, no. 01, pp. 10–19, 2023.
- [8] M. M. Mustain and H. Purnomo, "Keefektifan Program Keluarga Harapan Dalam Mewujudkan Kesejahteraan Keluarga Penerima Manfaat," vol. 2021, pp. 85–92, 2022.
- [9] I. Colanus, R. Drajana, and M. H. Botutihe, "Perbandingan Metode Forecasting K-NN , NN dan SVM Untuk Peramalan Jumlah Produksi Coconut Oil sehingga diperlukan prediksi . Produksi minyak kelapa yang terjadi".
- [10] M. M. Alfitri and D. Rusda, "Evaluasi Performa Algoritma Naïve Bayes Dalam Mengklasifikasi Penerima Bantuan Pangan Non Tunai," vol. 7, pp. 1433–1445, 2023, doi: 10.30865/mib.v7i3.6151.
- [11] E. F. Lubis, E. Zubaidah, P. Jumlah, K. Pekanbaru, T. Jumlah, and P. Miskin, "Efektivitas Program Keluarga Harapan Dalam Penanggulangan Kemiskinan Abstrak jumlah penduduk 268 juta jiwa dan terbagi dalam 34 provinsi . Dengan jumlah Sumber : Badan Pusat Statistik Kota Pekanbaru Dari tabel 1 . 1 diatas terlihat bahwa jumlah penduduk miskin terbanyak terletak dilaksanakannya program PKH dan juga dikarenakan bahwa Kecamatan Tenayan Raya akan dibangun sebagai pusat pemerintahan Kota Pekanbaru dengan dibangunnya kantor walikota yang baru . Berdasarkan Peraturan Menteri Sosial Nomor 1 Tahun," vol. 9, no. 2, pp. 88–99, 2020.
- [12] I. Muslim, K. Karo, A. Tsany, R. Dzaky, and M. A. Saputra, "Comparative Analysis of K-Nearest Neighbor and Modified K-Nearest Neighbor Algorithm for Financial Well-Being Data Classification," *J. Comput.*, vol. 6, no. 3, pp. 25–34, 2021, doi: 10.34818/indojc.2021.6.3.593.
- [13] B. H. Agtira, H. H. Handayani, and A. F. N. Masruriyah, "Perbandingan Algoritma NBC dan Decision Tree pada Sentimen Analisis Mengenai Vaksinasi Covid-19 Di Indonesia," *Remik*, vol. 7, no. 1, pp. 704–712, 2023, doi: 10.33395/remik.v7i1.12151.
- [14] Y. Sari, M. Alkaff, and M. Arif Rahman, "Identifikasi Penyakit Tanaman Ubi Kayu Berdasarkan Citra Daun Menggunakan Metode Probabilistic Neural Network (PNN)," *J. Komtika (Komputasi dan Inform.*, vol. 5, no. 1, pp. 1–9, 2021, doi: 10.31603/komtika.v5i1.4605.
- [15] H. Gunawan and V. Purwayoga, "Data Mining Menggunakan Algoritma K-Means Clustering Untuk Mengetahui Potensi Penyebaran Virus Corona Di Kota Cirebon," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 11, no. 1, pp. 1–8, 2022, doi: 10.32736/sisfokom.v11i1.1316.
- [16] S. Amjad, M. Younas, M. Anwar, Q. Shaheen, M. Shiraz, and A. Gani, "Data Mining Techniques to

- Analyze the Impact of Social Media on Academic Performance of High School Students,” *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/9299115.
- [17] S. P. Dewi, N. Nurwati, and E. Rahayu, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor,” *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [18] D. T. Nugrahadi, T. Mulyani, and ..., “Efek Transformasi Wavelet Diskrit Pada Klasifikasi Aritmia Dari Data Elektrokardiogram Menggunakan Machine Learning,” *J. Media ...*, vol. 7, pp. 13–21, 2023, doi: 10.30865/mib.v7i1.4859.
- [19] S. Deng, L. Wang, S. Guan, M. Li, and L. Wang, “Non-parametric Nearest Neighbor Classification Based on Global Variance Difference,” *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, 2023, doi: 10.1007/s44196-023-00200-1.
- [20] A. P. Silalahi and H. G. Simanullang, “Supervised Learning Metode K-Nearest Neighbor Untuk Prediksi Diabetes Pada Wanita,” *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 7, no. 1, pp. 144–149, 2023, doi: 10.46880/jmika.vol7no1.pp144-149.
- [21] M. Afdal and L. Waroka, “Shopee Application Review Classification Using Probabilistic Neural Network Algorithm And K-Nearest Neighbor Klasifikasi Ulasan Aplikasi Shopee Menggunakan Algoritma Probabilistic Neural Network Dan K-Nearest Neighbor,” *IJRSE (Indonesian J. Inform. Res. Softw. Eng.*, vol. 2, no. 1, pp. 49–58, 2022.
- [22] A. H. Utomo, M. A. Gumilang, and A. Ahmad, “Agricultural Commodity Sales Recommendation System for Farmers Based on Geographic Information Systems and Price Forecasting Using Probabilistic Neural Network Algorithm,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 980, no. 1, 2022, doi: 10.1088/1755-1315/980/1/012061.
- [23] X. Huang *et al.*, “A Gray Wolf Optimization-Based Improved Probabilistic Neural Network Algorithm for Surrounding Rock Squeezing Classification in Tunnel Engineering,” *Front. Earth Sci.*, vol. 10, no. February, pp. 1–9, 2022, doi: 10.3389/feart.2022.857463.
- [24] S. A. Putri and M. Mustakim, “Dimensional Data Unsupervised Learning Using an Analytic Hierarchy Process in Determining Attributes in the Classification Algorithm,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 1, pp. 235–240, 2022, doi: 10.47065/bits.v4i1.1752.
- [25] N. Aini *et al.*, “Implementasi Algoritma Naïve Bayes untuk Memprediksi Penjualan Lampu Pada Toko Satria,” *J. Soc. Sci. Res.*, vol. 3, pp. 9373–9387, 2023.
- [26] M. S. Syarah, M. Wati, and N. Puspitasari, “Klasifikasi Penderita ISPA Menggunakan Metode Naive Bayes Classifier,” *Innov. Res. Informatics*, vol. 4, no. 1, pp. 11–20, 2022, doi: 10.37058/innovatics.v4i1.4427.
- [27] N. D. Rumlaklak, A. Fanggidae, and Y. T. Polly, “Klasifikasi Penentuan Status Zona di Kota Kupang Menggunakan Algoritma Naive Bayes Classifier,” *J. Komput. dan Inform.*, vol. 10, no. 1, pp. 24–30, 2022, doi: 10.35508/jicon.v10i1.6458.