

PERBANDINGAN ALGORITMA C4.5 DAN NAIVE BAYES UNTUK MENENTUKAN PENERIMA BANTUAN BEDAH RUMAH DI KABUPATEN TANGGAMUS

Arbi Gunawan ^{1*}, M.Said Hasibuan²

Program Pascasarjana Magister IIB Darmajaya¹, IIB Darmajaya²

¹arbigunawa.mti@gmail.com, ²msaid@darmajaya.ac.id

Abstract

A livable house is a basic human need, a livable house is said to be if the home owner has a sense of peace and comfort living in that house. A livable house is interpreted as a house that has facilities to meet human needs in carrying out daily activities. The government assistance has not been distributed evenly to the homes of the residents of Tiga Dolok Village, because the program that has been running so far is very complicated, the system that runs is still manual in processing community data. Therefore, the authors provide a solution to classify the eligibility level of the house using the C4.5 algorithms and the Naïve Bayes algorithm. This algorithm was chosen because it is one of the methods in the Decision Tree that is widely used to make predictions about a case. Thus, government subsidies for renovating uninhabitable houses can be channeled in an appropriate manner. The research objective is to use a decision tree based on the C4.5 algorithms, which is expected to increase the accuracy of receiving housing feasibility renovation assistance. The research method used is Data Mining. Analysis and calculations using this algorithm really help the decision-making process to determine people who are eligible to receive housing renovation assistance and people who are not eligible to receive housing renovation assistance.

Keywords: *Livable House; C4.5 Algorithm; Decision tree, Naïve Bayes*

Abstrak

Rumah yang layak huni merupakan kebutuhan dasar manusia, rumah yang layak huni dikatakan jika pemilik rumah memiliki rasa tenang dan nyaman tinggal di rumah tersebut. Rumah yang layak huni dimaknai sebagai rumah yang memiliki fasilitas untuk memenuhi kebutuhan manusia dalam melaksanakan aktivitas sehari-hari. Belum tersalurkannya bantuan pemerintah secara merata terhadap rumah penduduk Desa Tiga Dolok, disebabkan oleh program yang telah berjalan selama ini sangat rumit, sistem yang berjalan masih secara manual dalam pengolahan data masyarakatnya. Oleh sebab itu penulis memberikan solusi untuk mengklasifikasi tingkat kelayakan rumah dengan menggunakan algoritma C4.5 dan algoritma naïve bayes. Algoritma ini dipilih karena merupakan salah satu metode pada Decision tree/Pohon Keputusan yang banyak dimanfaatkan untuk melakukan prediksi terhadap suatu kasus. Dengan demikian subsidi pemerintah untuk merenovasi rumah yang tidak layak huni dapat disalurkan secara tepat sasaran. Tujuan penelitian yaitu menggunakan pohon keputusan berbasis algoritma C4.5 diharapkan dapat meningkatkan keakuratan penerimaan bantuan renovasi kelayakan rumah. Metode penelitian yang digunakan yaitu Data Mining. Analisa dan perhitungan menggunakan algoritma ini sangat membantu proses penentuan keputusan untuk menentukan masyarakat yang layak menerima bantuan bedah rumah dan masyarakat yang tidak layak menerima bantuan bedah rumah.

Kata Kunci: *Rumah Layak Huni; Algoritma C4.5; Decision tree, Naïve Bayes*

1. PENDAHULUAN

Pekon Sukapadang Kecamatan Cukuh Balak Kabupaten Tanggamus Lampung memang memiliki banyak rumah yang layak huni, dan juga beberapa rumah yang tidak layak huni. Dalam rangka membantu masyarakat yang masih tinggal di Rumah Tidak Layak Huni (RTLH), (Andrianti & Firmansyah, 2020) pemerintah memberikan anggaran untuk mengurangi kemiskinan kepada masyarakatnya yang dianggap memenuhi kriteria tertentu. Pemerintah memberikan dana sebesar satu milyar untuk tiap-tiap kecamatan, dan kemudian dana tersebut dibagikan kepada 200 warga yang

berhak mendapatkan bantuan masing-masing sebesar sepuluh juta rupiah untuk renovasi rumah. Kementerian Pekerjaan Umum dan Perumahan Rakyat (PUPR)(Fitriani, 2020)(Krisdiyanto, 2021) melalui Direktorat Jenderal Penyediaan Perumahan menjalankan program Bantuan Stimulan Perumahan Swadaya (BSPS) atau yang lebih dikenal dengan program bedah rumah untuk membantu meningkatkan kualitas tempat tinggal masyarakat kurang mampu di seluruh Indonesia. (Andrianti & Firmansyah, 2020)

Permasalahan yang sering terjadi adalah bahwa masyarakat yang seharusnya tidak berhak untuk mendapatkan bantuan bedah rumah, justru mendapatkan bantuan tersebut, begitupun sebaliknya masyarakat yang berhak mendapatkan bantuan bedah rumah tidak mendapatkan bantuan tersebut. Selain itu, permasalahan lain yang terjadi adalah bahwa bantuan bedah rumah tersebut sering tidak didukung dengan data yang akurat. Hal tersebut jelas membutuhkan efisiensi waktu pengelolaan data penerima bantuan bedah rumah.(Ibrahim Irawan, Gani Hilmansyah, 2022) (Ibrahim Irawan, Gani Hilmansyah, 2022)

Pengolahan data untuk menentukan bantuan bedah rumah di Pekon Sukapadang Kecamatan Cukuh Balak Kabupaten Tanggamus Lampung yang kemudian dijadikan data referensi dan akan dicocokkan dengan kriteria-kriteria seperti Status Penguasaan bangunan, Tempat tinggal, Jenis atap, Jenis lantai, Jenis dinding, Sumber air, dan Kualitas atap. Nilai dari kriteria akan menjadi patokan untuk penyeleksian penduduk yang menjadi prioritas utama untuk mendapatkan bantuan bedah rumah. Untuk mengetahui informasi tentang warga yang mendapatkan bantuan bedah rumah biasanya diberitahu oleh Ketua RT, dan untuk kejelasan dana akan disalurkan dan kapan rumah mulai dikerjakan renovasinya warga harus menunggu keputusan dari DINSOS. (Paul V.M.; et al., 2021)(Fitriani, 2020)

Berdasarkan latar belakang di atas, penulis akan melakukan sebuah penelitian yang berjudul **“PERBANDINGAN ALGORITMA C4.5 DAN NAIVE BAYES UNTUK MENENTUKAN PENERIMA BANTUAN BEDAH RUMAH DI KABUPATEN TANGGAMUS.”** Proses seleksi penerimaan bantuan bedah rumah dan *approved* kapan dana diberikan dan tanggal eksekusi renovasi rumah menggunakan algoritma *Naive Bayes Classifier* dan *Algoritma C45* salah satu metode data mining yang menggunakan perhitungan probabilitas dan sudah diuji dengan *Rapidminer* dengan menguji 70 data training dan 10 data uji diperoleh hasil presisi 100%, recall 100% dan akurasi 100% (Nugroho et al., 2021)

Penelitian sebelumnya yang terkait dengan tema serupa adalah penelitian dengan judul “Penentuan Kelayakan Penerima Bantuan Renovasi Rumah Warga Miskin Menggunakan Naive Bayes” tingkat akurasi dalam penentuan bantuan renovasi rumah warga miskin, pengujian dilakukan dengan menerapkan metode Naive Bayes. Dari hasil penghitungan dengan metode Naive Bayes yang menggunakan bantuan tools WAKA didapat hasil 90%. Metode naive bayes dapat membantu dalam pengambilan keputusan untuk penentuan bantuan, sedangkan 10% tidak dapat membantu pengambilan keputusan (Yul Dewi Marta & Nurlitasari, 2021). Penelitian berikutnya dengan judul “Penerapan Metode Naive Bayes Terhadap Bantuan Langsung Tunai Di Desa Baleturi Nganjuk” membahas tentang Penentuan Bantuan BLT menggunakan metode Naive Bayes sehingga dapat diketahui keputusan yang mendapat BLT (Fitriani, 2020)

2. KERANGKA TEORI

2.1 Program Bedah Rumah

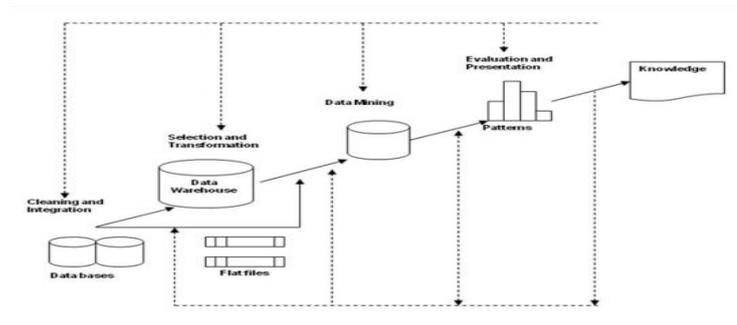
Program bedah rumah untuk memenuhi kebutuhan dasar terutama perumahan sebagai tempat tinggal, melalui peningkatan kesejahteraan sosial masyarakat dan perbaikan atau rehabilitasi rumah tidak layak huni (bedah rumah), sehingga keluarga miskin dapat menempati rumah yang layak huni dalam lingkungan yang sehat dan sejahtera.

Dengan terpenuhinya salah satu kebutuhan dasar berupa rumah yang layak huni, diharapkan tercapai ketahanan keluarga. Rumah yang baik adalah rumah yang sehat atau sering disebut layak huni, yang harus diupayakan keberadaannya, kebutuhan rumah yang layak huni diharapkan sebagai upaya mencapai ketahanan keluarga, sebaliknya jika tidak terpenuhi akan menimbulkan permasalahan, seperti keterlantaran ataupun permasalahan kesejahteraan sosial keluarga (Fitriani, 2020) (Ibrahim Irawan, Gani Hilmansyah, 2022) (Aziz, 2020)

2.2. Datamining

Data Mining adalah kegiatan menemukan pola yang menarik dari data yang jumlahnya besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lain (Ibrahim Irawan, Gani Hilmansyah, 2022)

Data Mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan dan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting dan menarik dari data yang terdapat pada basis data. Data mining digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut Knowledge Discovery Database (KDD) (Paul V.M.; et al., 2021)



Gambar 2.1 Tahapan *Knowledge Discovery Database* (KDD) (Ginting et al., 2022)

2.4. Naive Bayes

Penomoran persamaan dilakukan secara berurutan, dengan nomor persamaan ditulis di dalam tanda kurung dan Algoritma Naive Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi yang menggunakan metode probabilitas yang sederhana berdasarkan pada teorema bayes dengan asumsi tidak ketergantungan (*independent*) yang tinggi (Ginting et al., 2022). Beberapa studi mengenai algoritma klasifikasi menunjukkan bahwa naive bayes memiliki performa yang sebanding dengan *decision tree* dan *neural network classifier* tertentu. Selain itu, metode ini juga menunjukkan akurasi dan kecepatan yang tinggi ketika digunakan dalam basis data yang berukuran besar (Kusmira, 2019).

Rumus *Naive Bayes* nya adalah:

$$P(H|X) = \frac{P(H)P(X|H)}{P(X)} \quad (2.2)$$

Keterangan:

X = data dengan kelas yang belum diketahui

H = hipotesis data X, merupakan suatu kelas yang spesifik

$P(H|X)$ = probabilitas hipotesis H berdasar kondisi X (posteriori probability)

$P(H)$ = probabilitas hipotesis H (posteriori probability)

$P(X|H)$ = probabilitas X berdasar kondisi H

$P(X)$ = probabilitas dari X

atau

$$\text{Posterior Probability} = \frac{\text{Prior Probability} \times \text{likelihood}}{\text{evidence}}$$

2.5. Algoritma C.4.5

Algoritma C4.5 adalah sebuah algoritma yang berfungsi untuk membangun decision tree (pohon keputusan). Algoritma C4.5 dan pohon keputusan merupakan dua model yang tidak terpisahkan. Algoritma C4.5 adalah salah dari satu algoritma klasifikasi yang kuat dan banyak digunakan atau diimplementasikan untuk pengklasifikasian dalam berbagai hal. Algoritma C4.5 diperkenalkan oleh J. Ross Quinlan (1996) sebagai versi perbaikan dari algoritma Iterative Dichotomiser 3 (ID3). Serangkaian perbaikan dilakukan pada algoritma ID3 mencapai puncaknya dengan menghasilkan sebuah sistem praktis dan simple yang berpengaruh untuk pembentukan pohon keputusan. Perbaikan tersebut meliputi metode untuk menangani data kontinu, mengatasi missing data, dan melakukan pemangkasan pohon (Yul Dewi Marta & Nurlitasari, 2021)

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i)$$

Keterangan:

S: Himpunan kasus

A: Data Atribut

n: Jumlah partisi di dalam atribut

|S_i|: Jumlah kasus pada partisi ke-i

|S|: Jumlah kasus

Sedangkan untuk menghitung nilai entropy dapat dihitung dengan rumus (Munandar & Munir, 2022), persamaan 2.2.

$$\text{Entropy}(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

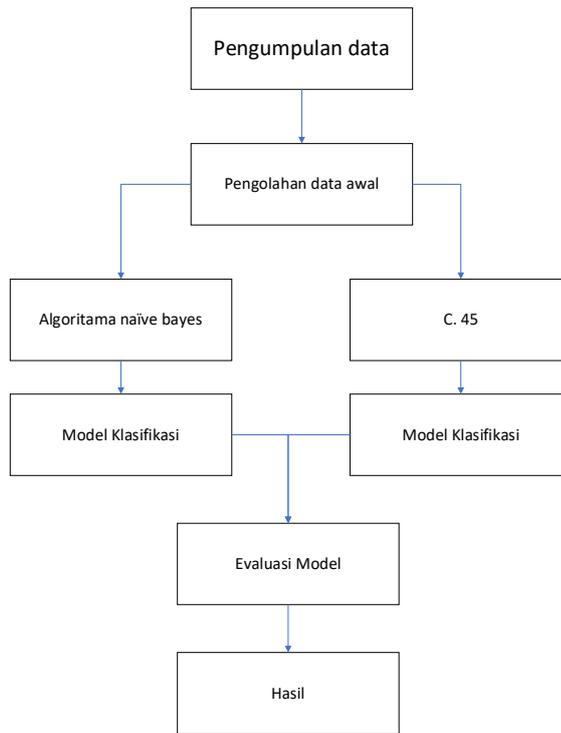
Keterangan:

S: Himpunan kasus

n: Jumlah partisi dalam atribut

p_i: Proporsi dari S_i terhadap S

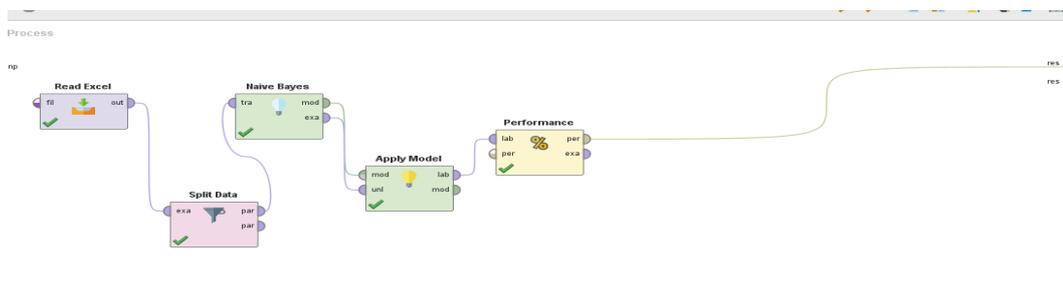
2. METODOLOGI



Gambar 3.1. Alur Penelitian (Andrianti & Firmansyah, 2020)

4. HASIL DAN PEMBAHASAN

4.1. Menggunakan Algoritma Naïve Bayes



Gambar 4.3. Proses menggunakan Algoritma Naïve Bayes Split Validation

accuracy: 82.21%

	true tidak	true ya	class precision
pred. tidak	314	56	84.86%
pred. ya	18	28	60.87%
class recall	94.58%	33.33%	

Gambar 4.4. Nilai Accuracy Algoritma Naïve Bayes. Split Validation

$$\begin{aligned}
 \text{akurasi} &= \frac{(TN + TP)}{(TN + FN + FP + TP)} \\
 &= \frac{(28 + 314)}{(28 + 18 + 56 + 314)} \\
 &= \frac{342}{416} = 0,8221
 \end{aligned}$$

akurasi = 82,2%

precision: 60.87% (positive class: ya)

	true tidak	true ya	class precision
pred. tidak	314	56	84.86%
pred. ya	18	28	60.87%
class recall	94.58%	33.33%	

Gambar 4.5. Nilai Precision Algoritma Naïve Bayes. Split Validation

$$\begin{aligned}
 \text{precision} &= \frac{TP}{(TP + FP)} \\
 &= \frac{314}{(314 + 56)} \\
 &= \frac{314}{370} = 0,84864
 \end{aligned}$$

precision = 84,86%

recall: 33.33% (positive class: ya)

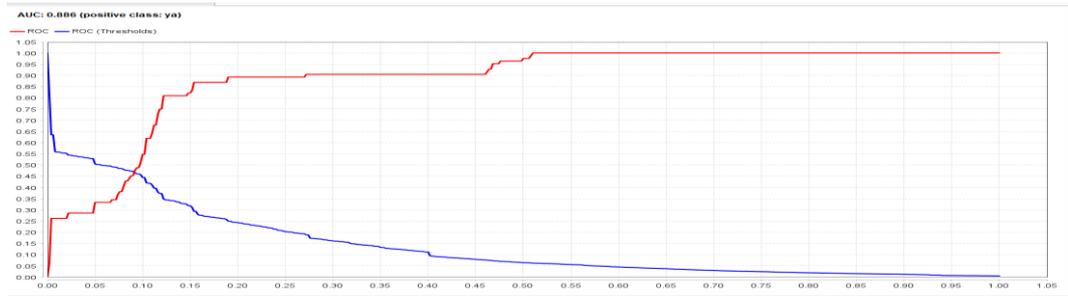
	true tidak	true ya	class precision
pred. tidak	314	56	84.86%
pred. ya	18	28	60.87%
class recall	94.58%	33.33%	

Gambar 4.6. Confusion Matrix Hasil klasifikasi Algoritma Naïve Bayes. Split Validation

$$\begin{aligned}
 \text{precision} &= \frac{TP}{(TP + FP)} \\
 &= \frac{314}{(314 + 18)} \\
 &= \frac{314}{332} = 0,9457
 \end{aligned}$$

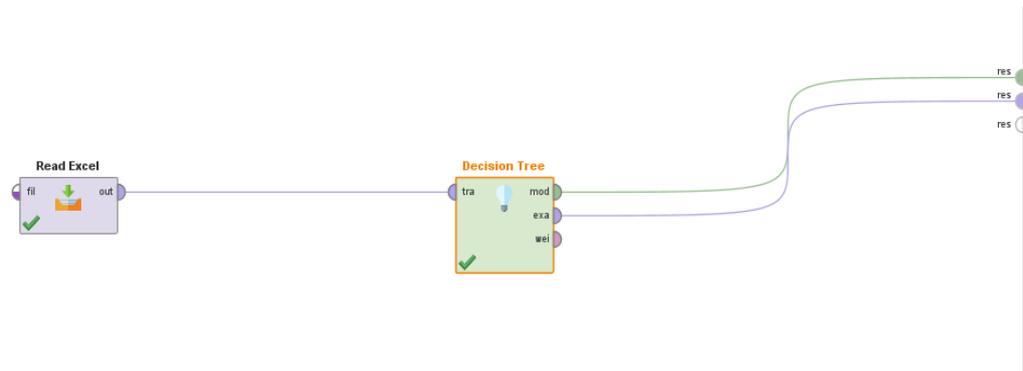
precision = 94,58%

Berikut adalah hasil kurva AUC yaitu mendapatkan nilai 0,886 seperti gambar dibawah ini:

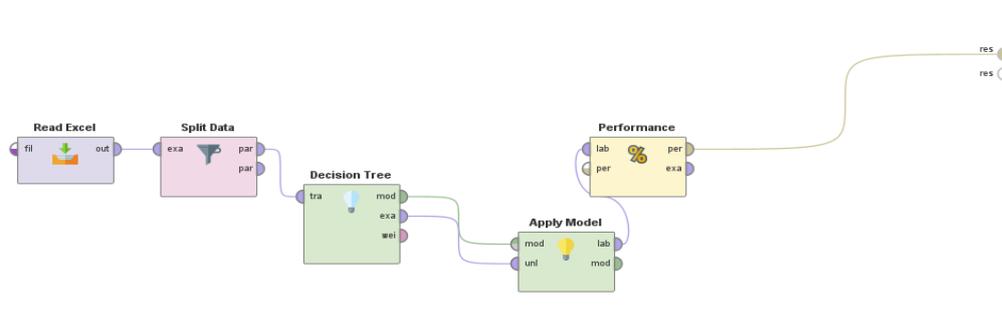


Gambar 4.7. Nilai dan Curve AUC Algoritma Naïve Bayes Split Validation

4.2. Menggunakan Algoritma Decision Tree C.45



Gambar 4.8. Proses



Gambar 4.10. Proses menggunakan Algoritma Decision Tree C. 45 Split Validation

accuracy: 96.88%

	true tidak	true ya	class precision
pred. tidak	320	1	99.69%
pred. ya	12	83	87.37%
class recall	96.39%	98.81%	

Gambar 4.11. Nilai Accuracy Algoritma Decision Tree C. 45 Split Validation

Menghitung akurasi:

Rumus:

$$\text{Akurasi} = \frac{(TN + TP)}{(TN + FN + FP + TP)}$$

$$\text{Akurasi} = \frac{(320+83)}{(320+83+1+12)} = \frac{403}{416} = 96,88\%$$

precision: 87.37% (positive class: ya)

	true tidak	true ya	class precision
pred. tidak	320	1	99.69%
pred. ya	12	83	87.37%
class recall	96.39%	98.81%	

Gambar 4.12. Precision Algorithm Decision Tree C. 45 Split Validation

Menghitung Precision, Rumus

$$P:\text{precision} = \frac{TP}{(TP+FP)}$$

$$:p(1) = \frac{320}{(320+1)} = \frac{320}{321} = 99,69\% \quad :p(0) = \frac{83}{(83+12)} = \frac{83}{95} = 87,37\%$$

recall: 98.81% (positive class: ya)

	true tidak	true ya	class precision
pred. tidak	320	1	99.69%
pred. ya	12	83	87.37%
class recall	96.39%	98.81%	

Gambar 4.13 Confusion Matrix Hasil klasifikasi Bedah Rumah menggunakan

Algoritma Decision Tree pada Rapidminer 9.10

Menghitung Recall, Rumus: $Recall = \frac{TP}{(TP+FN)}$

$$R(1) = \frac{320}{(320+12)} = \frac{320}{332} = 96,39\% \quad R(0) = \frac{83}{(83+1)} = \frac{83}{84} = 98,81\%$$

Level 1

$$\text{Jenis dinding} = -\left(\frac{588}{117} \times \log_2 \frac{588}{117}\right) + \frac{471}{471} \times \log_2 \frac{471}{117} = 0,79882578$$

$$\text{Kayu} = -\left(\frac{104}{105} \times \log_2 \frac{104}{105}\right) + \frac{1}{105} \times \log_2 \frac{1}{104} = 0,707619511$$

$$\text{Semen} = -\left(\frac{13}{483} \times \log_2 \frac{13}{483}\right) + \frac{470}{483} \times \log_2 \frac{470}{483} = 0,17867714$$

$$\text{Gain} = -\left(\frac{103}{587}\right) \times 0,79882578 + \left(\frac{483}{587}\right) \times 0,707619511 = 0,556251206$$

Jenis Lantai

$$\text{Semen} = -\left(\frac{74}{520}x \log 2 \frac{74}{520}\right) + \frac{446}{520}x \log 2 \frac{446}{520} = 0,590250689$$

$$\text{Tanah} = -\left(\frac{43}{68}x \log 2 \frac{43}{68}\right) + \frac{25}{68}x \log 2 \frac{25}{68} = 0,948848294$$

$$\text{Gain} = -\left(\frac{73}{519}\right)x 0,590250689 + \left(\frac{68}{519}\right)x 0,948848294 = 0,089141344$$

Bahan Bakar Masak

$$\text{Gas} = -\left(\frac{6}{476}x \log 2 \frac{6}{476}\right) + \frac{470}{476}x \log 2 \frac{470}{476} = 0,097606125$$

$$\text{Kayu} = -\left(\frac{111}{112}x \log 2 \frac{111}{112}\right) + \frac{1}{112}x \log 2 \frac{1}{112} = 0,073603483$$

$$\text{Gain} = -\left(\frac{6}{519}\right)x 0,097606125 + \left(\frac{111}{519}\right)x 0,073603483 = 0,623961814$$

Jumlah Tanggungan

$$\text{Kurang Dari Lima} = -\left(\frac{103}{548}x \log 2 \frac{103}{548}\right) + \frac{445}{548}x \log 2 \frac{445}{548} = 0,697176365$$

$$\text{Lebih Dari Lima} = -\left(\frac{14}{40}x \log 2 \frac{14}{40}\right) + \frac{26}{40}x \log 2 \frac{26}{40} = 0,934068055$$

$$\text{Gain} = -\left(\frac{102}{547}\right)x 0,697176365 + \left(\frac{40}{519}\right)x 0,934068055 = 0,006743719$$

Fasilitas Mck

$$\text{Ada} = -\left(\frac{0}{470}x \log 2 \frac{0}{470}\right) + \frac{1}{470}x \log 2 \frac{1}{470} = 0$$

$$\text{Tidak Ada} = -\left(\frac{117}{118}x \log 2 \frac{117}{118}\right) + \frac{470}{118}x \log 2 \frac{470}{118} = 0,070998947$$

$$\text{Gain} = -\left(\frac{0}{470}\right)x 0 + \left(\frac{117}{470}\right)x 0,070998947 = 0,070998947$$

$$\text{Total 1.1 Fasilitas MCK Tidak ada} = -\left(\frac{117}{118}x \log 2 \frac{117}{118}\right) + \frac{1}{118}x \log 2 \frac{1}{118} = 0,07050176$$

Level 2

Jenis Dinding

$$\text{Kayu} = -\left(\frac{104}{105}x \log 2 \frac{104}{105}\right) + \frac{1}{105}x \log 2 \frac{1}{105} = 0,707619511$$

$$\text{Semen} = -\left(\frac{13}{13}x \log 2 \frac{13}{13}\right) + \frac{0}{13}x \log 2 \frac{0}{13} = 0$$

$$\text{Gain} = -\left(\frac{104}{105}\right)x 0,707619511 + \left(\frac{13}{105}\right)x 0 = 0,706$$

Jenis Lantai

$$\text{Semen} = -\left(\frac{74}{74}x \log 2 \frac{74}{74}\right) + \frac{0}{74}x \log 2 \frac{0}{74} = 0$$

$$\text{Tanah} = -\left(\frac{43}{44}x \log 2 \frac{43}{44}\right) + \frac{1}{44}x \log 2 \frac{1}{44} = 0,156491063$$

$$\text{Gain} = -\left(\frac{74}{74}\right)x 0 + \left(\frac{43}{74}\right)x 0,156491063 = 0,7082$$

$$\begin{aligned}
 \text{Bahan Bakar Masak} &= -\left(\frac{6}{6}x \log_2 \frac{6}{6}\right) + \frac{0}{6}x \log_2 \frac{0}{6}=0 \\
 \text{Gas} & \\
 \text{Kayu} &= -\left(\frac{111}{112}x \log_2 \frac{111}{112}\right) + \frac{1}{112}x \log_2 \frac{1}{112}=0,073603483 \\
 \text{Gain} &= -\left(\frac{6}{6}\right)x0 + \left(\frac{111}{6}\right)x0,073603483=0,7059 \\
 \text{Jumlah Tanggungan} & \\
 \text{Kurang Dari Lima} &= -\left(\frac{103}{104}x \log_2 \frac{103}{104}\right) + \frac{1}{104}x \log_2 \frac{1}{104}=0,078232465 \\
 \text{Lebih Dari Lima} &= -\left(\frac{14}{14}x \log_2 \frac{14}{14}\right) + \frac{0}{14}x \log_2 \frac{0}{14}=0 \\
 \text{Gain} &= -\left(\frac{103}{104}\right)x0,078232465 + \left(\frac{0}{104}\right)x0=0,706 \\
 \text{Total 2,1 Jenis Lantai Tanah} &= -\left(\frac{43}{44}x \log_2 \frac{43}{44}\right) + \frac{1}{44}x \log_2 \frac{1}{44}=0,156496106 \\
 \text{Level 3} & \\
 \text{Kayu} &= -\left(\frac{43}{44}x \log_2 \frac{43}{44}\right) + \frac{1}{44}x \log_2 \frac{1}{44}=0,156491063 \\
 \text{Semen} &= -\left(\frac{0}{0}x \log_2 \frac{0}{0}\right) + \frac{0}{0}x \log_2 \frac{0}{0}=0 \\
 \text{Gain} &= -\left(\frac{43}{44}\right)x0,156491063 + \left(\frac{0}{0}\right)x0=0,7082 \\
 \text{Bahan Bakar Masak} & \\
 \text{Gas} &= -\left(\frac{0}{0}x \log_2 \frac{0}{0}\right) + \frac{0}{0}x \log_2 \frac{0}{0}=0 \\
 \text{Kayu} &= -\left(\frac{43}{44}x \log_2 \frac{43}{44}\right) + \frac{1}{44}x \log_2 \frac{1}{44}=0,156491063 \\
 \text{Gain} &= -\left(\frac{43}{44}\right)x0,156491063 + \left(\frac{0}{0}\right)x0=0,7082 \\
 \text{Jumlah Tanggungan} & \\
 \text{Kurang Dari Lima} &= -\left(\frac{42}{43}x \log_2 \frac{42}{43}\right) + \frac{1}{43}x \log_2 \frac{1}{43}=0,159350063 \\
 \text{Lebih Dari Lima} &= -\left(\frac{1}{1}x \log_2 \frac{1}{1}\right) + \frac{0}{1}x \log_2 \frac{0}{1}=0 \\
 \text{Gain} &= -\left(\frac{43}{44}\right)x0,156491063 + \left(\frac{0}{0}\right)x0=0,7082
 \end{aligned}$$

5. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan maka dapat diambil kesimpulan diantaranya:

1. Dengan menganalisa hasil perhitungan perbandingan dengan menggunakan algoritma Naïve Bayes dan C4.5 maka dari pengujian hasil klasifikasi bantuan bedah rumah menggunakan 594 data sampel yang telah dilakukan diperoleh nilai akurasi yang berbeda jauh, tingkat akurasi naïve bayes sebesar 82,2% sedangkan tingkat akurasi algoritma C4.5 96,88 sebesar %.

2. Nilai keakurasian tertinggi hasil klasifikasi yang ditinjau dari parameter algoritma C4.5 mendapatkan nilai akurasi yang tertinggi dibandingkan dengan algoritma Naïve Bayes.
3. Dari kedua hasil akurasi berhasil divalidasi dan di implementasi dapat mengklasifikasikan jumlah penerima bantuan rumah sehat yang telah menjadi program pemerintah pusat. Berdasarkan data training dan pengujian langkah berikutnya melakukan klasifikasi sehingga menghasilkan kesimpulan bahwa jumlah data training mempengaruhi pada presentase keakurasian.

DAFTAR PUSTAKA

- Andrianti, F., & Firmansyah, R. (2020). *Penerapan Clustering Data Kurang Mampu Di Desa Situmekar Menggunakan Algoritma K-Means*. 1(1), 88. <http://eprosiding.ars.ac.id/index.php/pti>
- Aziz, A. (2020). Implementasi Algoritma Rough Set Dan Naive Bayes Untuk Mendapatkan Rule Dalam Menyeleksi Pemohon Bantuan Fasilitas Rumah Ibadah (Studi Kasus : Pemerintah Kabupaten Pringsewu). *Jl. ZA. Pagar Alam*, 03(93), 74–83.
- Fitriani, E. (2020). Perbandingan Algoritma C4.5 Dan Naïve Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Sistemasi*, 9(1), 103. <https://doi.org/10.32520/stmsi.v9i1.596>
- Ginting, K. A., Buaton, R., Kom, M., & Syari, M. A. (2022). *Penerapan Data Mining Dalam Pengelompokan Penerimaan Bantuan Untuk UMKM dengan Metode Clustering (Studi Kasus : Kec . Salapian)*. 6(3).
- Ibrahim Irawan, Gani Hilmansyah, L. R. H. Y. (2022). Perbandingan algoritma naïve bayes dan c4.5 untuk klasifikasi bantuan rumah sehat. *Juik (Jurnal Ilmu Komputer)*, 2.
- Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naïve Bayes Clasifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 7(1), 32. <https://doi.org/10.24014/coreit.v7i1.12945>
- Kusmira, M. (2019). Penerapan Data Mining Pengajuan Pembiayaan Perumahan (Consumer Financing) Individual Menggunakan Algoritma C4.5. *Jurnal Khatulistiwa Informatika*, 7(1), 66–71. <https://doi.org/10.31294/jki.v7i1.126>
- Munandar, T. A., & Munir, A. Q. (2022). Implementasi K-Nearest Neighbor Untuk Prototype Sistem Pakar Identifikasi Dini Penyakit Jantung. *Respati*, 17(2), 44. <https://doi.org/10.35842/jtir.v17i2.457>
- Nugroho, D. A., Savitri, N., Saputra, N., Adji, T. B., Permasari, A. E., Ardiansyah, M. N., Umar, R., Sunardi, Alifah, U., Arsyah Monica Pravina, Imam Cholissodin, P. P. A., Darwis, D., Pratiwi, E. S., Pasaribu, A. F. O., Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Data Presiden Jokowi Dengan Preprocessing Normalisasi Dan Stemming Menggunakan Metode Naive Bayes Dan SVM. *Seminar Nasional Teknologi Fakultas Teknik Universitas Krisnadwipayana*, 3(1), 1–11.
- Paul V.M., Indra, G. M., Damanik, B. E., Parlina, I., & Saputra, W. (2021). *DALAM MENENTUKAN KELAYAKAN PENERIMAAN BANTUAN BEDAH RUMAH PADA DESA TIGA DOLOK Paul V . M ., Indra Gunawan , Bahrudi Efendi Damanik , Iin Parlina dan Widodo Saputra STIKOM Tunas Bangsa Pematangsiantar Abstrak Penerapan Data Mining Menggunakan Algoritma C4. 1*, 396–409.
- Yul Dewi Marta, F., & Nurlitasari, R. (2021). Implementasi Penyaluran Bantuan Langsung Tunai Dana Desa di Era Pandemi Covid-19 di Kabupaten Sigi 2020. *Jurnal Terapan Pemerintahan Minangkabau*, 1(1), 47–59. <https://doi.org/10.33701/jtpm.v1i1.1870>
- Andrianti, F., & Firmansyah, R. (2020). *Penerapan Clustering Data Kurang Mampu Di Desa Situmekar*

-
- Menggunakan Algoritma K-Means*. 1(1), 88. <http://eprosiding.ars.ac.id/index.php/pti>
- Aziz, A. (2020). Implementasi Algoritma Rough Set Dan Naive Bayes Untuk Mendapatkan Rule Dalam Menyeleksi Pemohon Bantuan Fasilitas Rumah Ibadah (Studi Kasus : Pemerintah Kabupaten Pringsewu). *Jl. ZA. Pagar Alam*, 03(93), 74–83.
- Fitriani, E. (2020). Perbandingan Algoritma C4.5 Dan Naive Bayes Untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *Sistemasi*, 9(1), 103. <https://doi.org/10.32520/stmsi.v9i1.596>
- Ginting, K. A., Buaton, R., Kom, M., & Syari, M. A. (2022). Penerapan Data Mining Dalam Pengelompokan Penerimaan Bantuan Untuk UMKM dengan Metode Clustering (Studi Kasus : Kec . Salapian). 6(3).
- Ibrahim Irawan, Gani Hilmansyah, L. R. H. Y. (2022). Perbandingan algoritma naïve bayes dan c4.5 untuk klasifikasi bantuan rumah sehat. *Juik (Jurnal Ilmu Komputer)*, 2.
- Krisdiyanto, T. (2021). Analisis Sentimen Opini Masyarakat Indonesia Terhadap Kebijakan PPKM pada Media Sosial Twitter Menggunakan Naive Bayes Clasifiers. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer Dan Teknologi Informasi*, 7(1), 32. <https://doi.org/10.24014/coreit.v7i1.12945>
- Kusmira, M. (2019). Penerapan Data Mining Pengajuan Pembiayaan Perumahan (Consumen Financing) Individual Menggunakan Algoritma C4.5. *Jurnal Khatulistiwa Informatika*, 7(1), 66–71. <https://doi.org/10.31294/jki.v7i1.126>
- Munandar, T. A., & Munir, A. Q. (2022). Implementasi K-Nearest Neighbor Untuk Prototype Sistem Pakar Identifikasi Dini Penyakit Jantung. *Respati*, 17(2), 44. <https://doi.org/10.35842/jtir.v17i2.457>
- Nugroho, D. A., Savitri, N., Saputra, N., Adji, T. B., Permanasari, A. E., Ardiansyah, M. N., Umar, R., Sunardi, Alifah, U., Arsyah Monica Pravina, Imam Cholissodin, P. P. A., Darwis, D., Pratiwi, E. S., Pasaribu, A. F. O., Arsi, P., & Waluyo, R. (2021). Analisis Sentimen Data Presiden Jokowi Dengan Preprocessing Normalisasi Dan Stemming Menggunakan Metode Naive Bayes Dan SVM. *Seminar Nasional Teknologi Fakultas Teknik Universitas Krisnadwipayana*, 3(1), 1–11.
- Paul V.M.; Indra, G. M., Damanik; B. E., Parlina; I., & Saputra; W. (2021). *DALAM MENENTUKAN KELAYAKAN PENERIMAAN BANTUAN BEDAH RUMAH PADA DESA TIGA DOLOK Paul V . M ., Indra Gunawan , Bahrudi Efendi Damanik , Iin Parlina dan Widodo Saputra STIKOM Tunas Bangsa Pematangsiantar Abstrak Penerapan Data Mining Menggunakan Algoritma C4. 1*, 396–409.
- Yul Dewi Marta, F., & Nurlitasari, R. (2021). Implementasi Penyaluran Bantuan Langsung Tunai Dana Desa di Era Pandemi Covid-19 di Kabupaten Sigi 2020. *Jurnal Terapan Pemerintahan Minangkabau*, 1(1), 47–59. <https://doi.org/10.33701/jtpm.v1i1.1870>