

Implementasi Algoritma C4.5 untuk Klasifikasi Penerima Rumah Bantuan Di Kampung Kotalintang

Mahdi¹, Muhammad Arhami², Khairunnisa Atami³, Renol Herwinsyah⁴

^{1,2,3} *Jurusan Teknologi Informasi dan Komputer Politeknik Negeri Lhokseumawe*

⁴ *Jurusan Tata Niaga Politeknik Negeri Lhokseumawe*

¹mahdi@pnl.ac.id, ²muhammad.arhami@pnl.ac.id, ³khairunnisaatamii.99@gmail.com, ⁴Reynold_maldini@yahoo.com

Abstrak— Kotalintang merupakan salah satu desa/kelurahan yang ada di Kecamatan Kualasimpang, Kabupaten Aceh Tamiang, Provinsi Aceh. Sebanyak 383 jumlah rumah tangga yang tergolong Keluarga Pra Sejahtera di Desa Kotalintang beberapa diantaranya memiliki rumah yang tidak layak huni. Untuk membantu mengatasi permasalahan rumah tidak layak huni pemerintah memberikan program rumah bantuan yaitu program yang diberikan oleh pemerintah pusat melalui Pemerintah Kementerian Pekerjaan Umum dan Perumahan Rakyat (PUPR) untuk memberikan rumah layak huni kepada Masyarakat Berpenghasilan Rendah (MBR). Program rumah bantuan tersebut diberikan dengan kuota terbatas, sedangkan jumlah masyarakat yang membutuhkan lebih besar sehingga harus dilakukan proses seleksi. Proses seleksi penerima rumah bantuan di Kampung Kotalintang masih dilakukan secara musyawarah sehingga hasil yang diperoleh tidak optimal dan terdapat penerima yang tidak tepat sasaran, berakibat pada saat pemeriksaan oleh Inspektorat ditemukan ketidaksesuaian hasil seleksi penerima rumah bantuan tersebut. Untuk mengatasi permasalahan tersebut dibangun sebuah sistem untuk mengklasifikasikan penerima rumah bantuan di Kampung Kotalintang dengan menggunakan metode decision tree untuk menghasilkan keputusan yang lebih optimal dan dapat membantu pihak desa/kelurahan dalam mengklasifikasikan masyarakat penerima rumah bantuan. Sistem yang dibangun berbasis web menggunakan parameter penentu yaitu status pernikahan, penghasilan, status kepemilikan tanah, status kepemilikan rumah, jenis atap, kondisi atap, jenis lantai, kondisi lantai, jenis dinding dan kondisi dinding. Data yang digunakan adalah data penerima rumah bantuan tahun 2019 s/d 2022 sebanyak 185 data. 125 data yang dijadikan sebagai data latih dan dilakukan 5 kali pengujian untuk setiap 37 data. Penelitian ini berhasil dibangun sistem klasifikasi penerima rumah bantuan menggunakan metode decision tree dengan menghasilkan rata-rata nilai akurasi sebesar 88,11% dan menunjukkan bahwa fitur-fitur yang ada telah berfungsi sesuai yang diharapkan.

Kata kunci— Rumah Bantuan, Klasifikasi, Pohon Keputusan.

Abstract— Kotalintang is one of the villages in the District of Kua-lasimpang, Aceh Tamiang Regency, Aceh Province. A total of 383 households belonging to Pre-Prosperous Families in Kotalintang Village, some of them have houses that are not suitable for habitation. To help overcome the problem of uninhabitable houses, the government provides a housing assistance program, namely a program provided by the central government through the Government of the Ministry of Public Works and Public Housing (PUPR) to provide livable houses to Low-Income Communities (MBR). The housing assistance program is provided with a limited quota, while the number of people in need is greater so a selection process must be carried out. The selection process for the recipients of the aid housing in Kampung Kotalintang is still carried out by deliberation so that the results obtained are not optimal and there are recipients who are not on target, resulting in an inspection by the Inspectorate finding discrepancies in the results of the selection of the recipients of the aid housing. To overcome this problem, a system was built to classify the beneficiaries of the aid housing in Kampung Kotalintang by using the decision tree method to produce more optimal decisions and to assist the village/kelurahan in classifying the recipients of the aid housing. The web-based system uses determining parameters, namely marital status, income, land ownership status, house ownership status, roof type, roof condition, floor type, floor condition, wall type and wall condition. The data used is data on recipients of housing assistance from 2019 to 2021 as many as 185 data. 125 data were used as training data and tested 5 times for every 35 data. This study succeeded in building a classification system for house beneficiaries using the decision tree method by producing an average accuracy value of 88,11% and showing that the existing features function as expected.

Keywords— Home Assistance, Classification, Decision Tree.

I. PENDAHULUAN

Kotalintang merupakan salah satu desa/kelurahan yang ada di Kecamatan Kualasimpang, Kabupaten Aceh Tamiang, Provinsi Aceh. Jumlah 383 rumah tangga yang tergolong Keluarga Pra Sejahtera di Desa Kotalintang masih ada beberapa diantaranya memiliki rumah yang tidak layak huni.

Setiap tahun desa Kotalintang mendapatkan kuota bantuan rumah dari pemerintah sebanyak 30 hingga 50 rumah, sedangkan masyarakat yang membutuhkan bantuan rumah tersebut lebih besar dibandingkan kuota yang diberikan.

Program rumah bantuan merupakan program yang diberikan oleh pemerintah pusat melalui Pemerintah

Kementerian Pekerjaan Umum dan Perumahan Rakyat (PUPR) untuk memberikan rumah layak huni kepada Masyarakat Berpenghasilan Rendah (MBR) (Peraturan Menteri Pekerjaan Umum dan Perumahan Rakyat Nomor 39 Tahun 2015 tentang Pedoman Pelaksanaan Bantuan Stimulasi Perumahan Swadaya).

Pihak Desa memiliki prosedur untuk menentukan penerima program bantuan yang diberikan oleh pemerintah, salah satunya adalah melakukan seleksi bagi setiap warga berdasarkan yang ditetapkan sesuai dengan Pasal 11 Peraturan Menteri PUPR RI Nomor 07/PRT/M/2018 tentang Bantuan Stimulan dan Perumahan Swadaya yaitu warga negara Indonesia yang sudah berkeluarga, memiliki atau menguasai

tanah dengan alas hak yang sah, belum memiliki rumah atau memiliki dan menempati satu-satunya rumah dengan kondisi tidak layak huni, belum pernah memperoleh BSPS atau bantuan pemerintah untuk program perumahan, berpenghasilan paling banyak sebesar upah minimum daerah.

Proses seleksi untuk memenuhi kuota tersebut memiliki beberapa kendala diantaranya :

1. Waktu yang dibutuhkan lebih lama untuk melakukan seleksi karena masih dilakukan secara biasa, walaupun menggunakan perangkat komputer namun masih terbatas pada penggunaan aplikasi Microsoft Excel.
2. Penilaiannya masih bersifat subjektif, dimana yang mendapatkan rumah bantuan tersebut merupakan saudara dekat dari aparat desa.
3. Seleksi penerima rumah bantuan masih dilakukan secara musyawarah sehingga hasil seleksi yang diperoleh tidak optimal dan terdapat penerima yang tidak tepat sasaran, berakibat pada saat pemeriksaan oleh Inspektorat ditemukan ketidaksesuaian hasil seleksi penerima rumah bantuan tersebut.

Berdasarkan permasalahan diatas untuk menyelesaikan persoalan-persoalan tersebut maka dibutuhkan suatu sistem klasifikasi menggunakan metode *decision tree* yang dapat digunakan untuk membantu pihak desa dalam melakukan seleksi penerima rumah bantuan.

Klasifikasi merupakan jenis analisa data yang dapat menentukan kelas label dari data sampel yang ingin di klasifikasi. Klasifikasi menemukan hubungan antara atribut masukan dan atribut target yang dapat dikatakan sebagai Metode supervised learning. Tujuan klasifikasi untuk meningkatkan keakuratan hasil yang diperoleh dari data [1].

Pohon keputusan merupakan sebuah struktur yang menerapkan serangkaian aturan keputusan yang digunakan untuk membagi sekumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil. Data dalam pohon keputusan memiliki atribut dan record disajikan dalam bentuk tabel. Pembentukan pohon keputusan atribut mempresentasikan suatu parameter yang dibuat sebagai kriteria. Satu atribut digunakan untuk menyatakan data solusi atau prediksi dari per item data yang disebut target atribut. Nilai-nilai yang terdapat didalam atribut disebut dengan instance. Proses pada pohon keputusan adalah mengubah bentuk data (tabel) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule [2].

Tujuan dari penelitian ini adalah membangun sistem berbasis web dengan mengimplementasikan metode Decision Tree guna menghasilkan keputusan yang lebih optimal dalam menentukan penerima rumah bantuan dan penelitian ini diharapkan dapat memberikan manfaat diantaranya :

1. Penelitian ini dapat membantu pihak desa/kelurahan dalam mengklasifikasi masyarakat penerima rumah bantuan.
2. Penelitian ini diharapkan dapat dijadikan sebagai referensi untuk penelitian selanjutnya.

Telah dilakukan penelitian terdahulu oleh Junaidi Noh pada tahun 2018 yang berjudul Implementasi Metode Profile Matching Dalam Penentuan Keputusan Penerima Bantuan Rumah Layak Huni Di Desa Trans Maudi Kota Tidore Kepulauan. Pada penelitian tersebut membangun sebuah Sistem Pendukung Keputusan (SPK) berdasarkan luas tanah, pekerjaan, umur, tingkat pendidikan dan menghasilkan perengkingan untuk setiap calon penerima.

II. METODOLOGI PENELITIAN

A. Decision Tree

Decision tree merupakan salah satu metode klasifikasi untuk melakukan pengukuran menggunakan struktur pohon yang terdiri dari kumpulan node yang dihubungkan oleh cabang-cabang dari decision root sampai ke leaf node. Metode ini digunakan untuk mengecek setiap atribut-atribut yang memungkinkan untuk menghasilkan keputusan-keputusan baru hingga pada akhirnya menemukan keputusan yang tepat (leaf node) [3].

B. Algoritma C4.5

Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang mana algoritma C4.5 memiliki keunggulan dibanding algoritma ID3 yaitu bisa mengatasi missing value, bisa mengatasi continue data, dan pruning [4]. Kelebihan lain dari algoritma C4.5 memiliki nilai akurasi yang dapat diterima, hasil pohon keputusan yang dibangun mudah diinterpretasikan, efisien dalam menangani atribut bertipe diskret dan numerik. Melalui *Decision Tree* juga akan didapat pola atau rules dari klasifikasi tersebut [5].

Secara umum Langkah-langkah untuk yang digunakan pada algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut yang dijadikan sebagai akar(node)
2. Buat cabang untuk masing-masing nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk tiap-tiap cabang hingga semua kasus pada cabang mempunyai kelas yang sama.

Atribut akar dapat ditentukan dengan menggunakan nilai Gain tertinggi dari masing-masing atribut.

Rumus menghitung nilai Entropy dapat dilihat pada rumus 1 berikut :

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

Dimana
 S : Himpunan Kasus
 n : Jumlah partisi
 pi : Jumlah yang bersolusi negative atau positif mendukung pada data sampel untuk kriteria tertentu

Rumus menghitung Gain tertera dalam Rumus 2

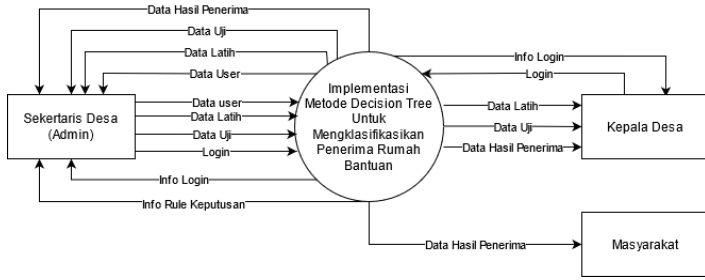
$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * \text{Entropy}(S_i) \dots\dots\dots(2)$$

Dimana
 S : Himpunan Kasus
 A : Atribut
 N : Jumlah partisi atribut A
 |Si| : Jumlah kasus pada partisi ke i
 |S| : Jumlah kasus dalam S [2].

C. Diagram Konteks

Sistem klasifikasi penerima rumah bantuan memiliki 3 entitas yaitu entitas Sekertaris Desa sebagai admin, entitas Kepala Desa dan entitas Masyarakat. Gambar 2 merupakan diagram konteks system klasifikasi penerima rumah bantuan di Kampung Kotalintang yang menggambarkan Sekertaris Desa dapat login ke sistem, menginputkan data latih yang digunakan, menginputkan data data uji yaitu meliputi data-data penduduk dan keterangan ekonomi penduduk, melihat hasil penerima rumah bantuan, melihat rule keputusan yang dihasilkan dan mengolah data-data user yang digunakan.

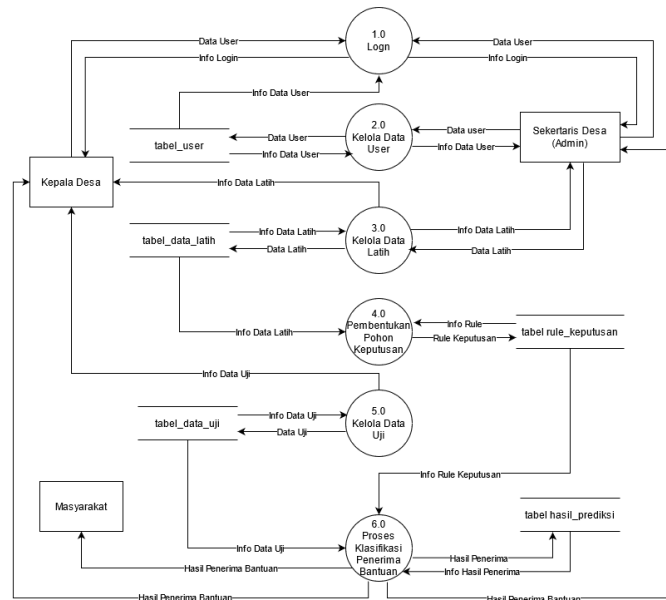
Kepala Desa dapat login dan melihat data latih, data uji dan hasil penerima rumah bantuan. Masyarakat hanya dapat melihat data hasil penerima rumah bantuan.



Gambar 1. Diagram Konteks

D. DFD (Data Flow Diagram)

DFD menggambarkan secara rinci aliran dari suatu data yang dikembangkan berdasarkan konteks diagram. Gambar 3 merupakan DFD yang telah dirancang pada Implementasi Decision Tree Untuk Mengklasifikasikan Penerima Rumah Bantuan di Kampung Kotalintang.



Gambar 2. Data Flow Diagram

Proses yang terjadi pada DFD Level 1 adalah sebagai berikut :

1. Proses Login

Pada proses login Pengguna sistem yaitu Sekertaris Desa dan Kepala Desa melakukan login terlebih dahulu untuk dapat mengakses sistem dengan memasukkan username dan password. Sistem akan melakukan pengecekan username dan password yang dimasukkan dengan username dan password yang terdapat didalam database table user. Jika username dan password sesuai maka pengguna dapat diizinkan masuk ke sistem.

2. Kelola Data User

Sekertaris Desa selaku admin dapat mengelola data user dan disimpan di table user. Data user ini akan digunakan untuk login.

3. Data Latih

Sekertaris Desa selaku admin dapat mengelola data latih dan disimpan di table data latih. Kepala Desa hanya dapat melihat data latih.

4. Kelola Data Uji

Sekertaris Desa selaku admin dapat mengelola data uji dan disimpan di table data uji. Kepala Desa hanya dapat melihat data uji.

5. Pembentukan Pohon Keputusan

Proses ini akan melakukan pembentukan pohon keputusan berdasarkan data latih yang ada. Pembentukan pohon keputusan menggunakan perhitungan metode decision tree dengan algoritma C.45 dan menghasilkan rule-rule keputusan.

6. Proses Klasifikasi Penerima Rumah Bantuan

Pada proses ini akan melakukan klasifikasi data uji berdasarkan rule-rule keputusan dan sistem akan memberikan hasil penerima rumah bantuan kepada Sekertaris Desa, Kepala Desa dan Masyarakat.

E. ERD (Entity Relationship Diagram)

Gambar 3 merupakan ERD yang diterapkan untuk sistem klasifikasi penerima rumah bantuan di Kampung Kotalintang.



Gambar 3. ERD (Entity Relationship Diagram)

Pada Gambar 3 dapat dilihat bahwa ERD Sistem Klasifikasi Penerima Rumah Bantuan menggunakan entitas dusun, data latih, data uji, tabel keputusan untuk menyimpan rule-rule keputusan yang dihasilkan dari perhitungan metode decision tree.

F. Confusion Matriks

Pada data mining untuk mengukur atau ada beberapa cara untuk mengukur kinerja dari model yang dihasilkan salah satunya menggunakan confusion matriks (akurasi) Confusion

matrix adalah suatu metode yang digunakan untuk melakukan perhitungan akurasi pada konsep data mining [5]

Table 1. Model *Confusion Matriks*

Actual	Classified As	
	+	-
+	True Positives (A)	False Negatives (B)
-	False Positives (C)	True Negatives (D)

Perhitungan akurasi dengan tabel confusion matrix adalah sebagai berikut:

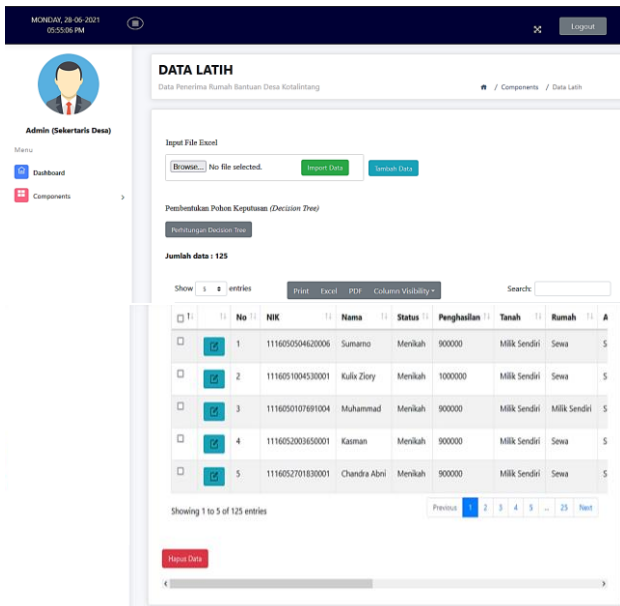
$$\text{Akurasi} = (A+D)/(A+B+C+D).$$

III. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Implementasi sistem merupakan hasil rancangan menjadi sebuah Sistem Klasifikasi Penerima Rumah Bantuan di Kampung Kotalintang Berbasis Web dengan menggunakan metode Decision Tree.

1. Halaman Data Latih



Gambar 4. Halaman Data Latih

Halaman data latih merupakan halaman yang menampilkan data yang dijadikan sebagai data training atau latih. Data latih yang digunakan merupakan data penerima rumah bantuan di Kampung Kotalintang pada tahun 2018, 2019, 2020 dan 2021 yang diberikan oleh Dinas Pekerjaan Umum dan Perumahan Rakyat (PUPR) Aceh Tamiang. Data yang diberikan berupa parameter penerima rumah bantuan diantaranya :

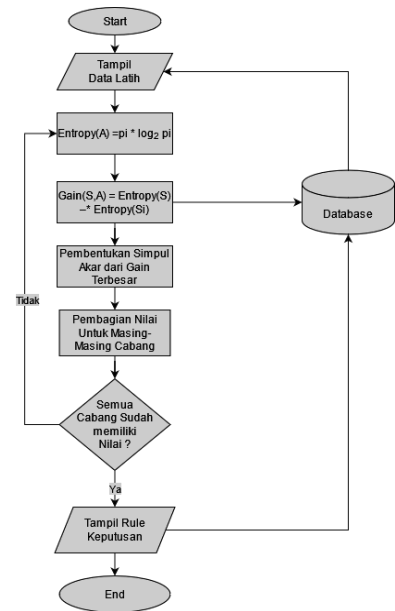
1. Status Pernikahan
2. Penghasilan
3. Status Tanah

4. Status Rumah
5. Jenis Atap
6. Kondisi Atap
7. Jenis Lantai
8. Kondisi Lantai
9. Jenis Dinding
10. Kondisi Diding

Data latih akan digunakan sebagai perhitungan untuk membentuk pohon keputusan sehingga menghasilkan rule-rule keputusan. Pada halaman ini admin dapat melakukan import data excel, tambah data, hapus dan edit.

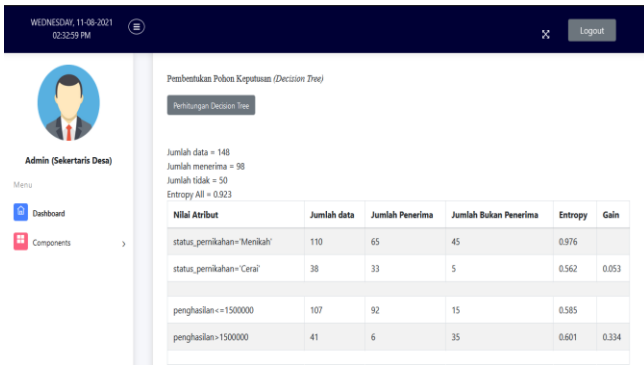
2. Halaman Pembentukan Pohon Keputusan

Alur proses implementasi metode decision tree dapat dilihat pada gambar 1. Pada bagian awal sistem menampilkan data latih yang digunakan dari database, selanjutnya melakukan proses hitung nilai entropy untuk setiap parameter yang terdapat di data latih, setelah mendapat nilai entropy, menghitung nilai gain untuk setiap parameter yang terdapat di data latih dan disimpan di database. Nilai gain terbesar akan digunakan sebagai pembentukan simpul. Setelah pembentukan akar akan dilakukan pemilihan nilai untuk masing-masing cabang dari akar yang dibentuk. Pada proses cek dilihat apakah semua cabang yang dimiliki sudah memiliki nilai. Jika ya akan tampil rule keputusan, jika tidak akan melakukan perhitungan nilai entropy Kembali sampai semua cabang memiliki nilai.



Gambar 5. Flowchart Perancangan Metode

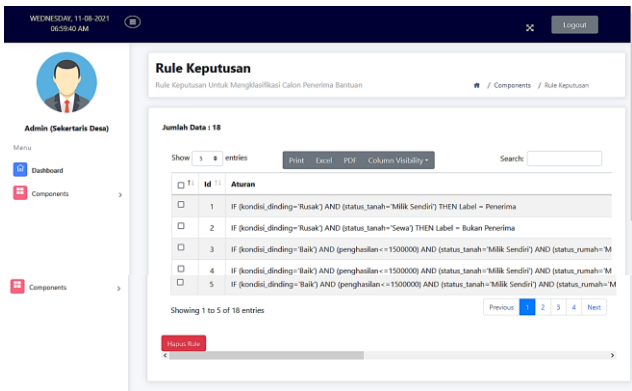
Halaman pembentukan pohon keputusan merupakan halaman yang menampilkan perhitungan algoritma C4.5 untuk membentuk suatu pohon keputusan hingga menghasilkan rule-rule keputusan untuk mengkasifikasikan penerima rumah bantuan.



Gambar 6. Halaman Pembentukan Pohon Keputusan

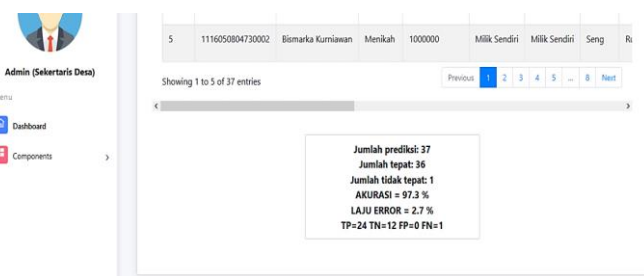
3. Halaman Rule Keputusan

Halaman rule keputusan merupakan halaman yang menampilkan rule-rule keputusan yang dihasilkan dari proses *decision tree*. Rule keputusan bersifat dinamis berdasarkan dari hasil perhitungan *decision tree*.



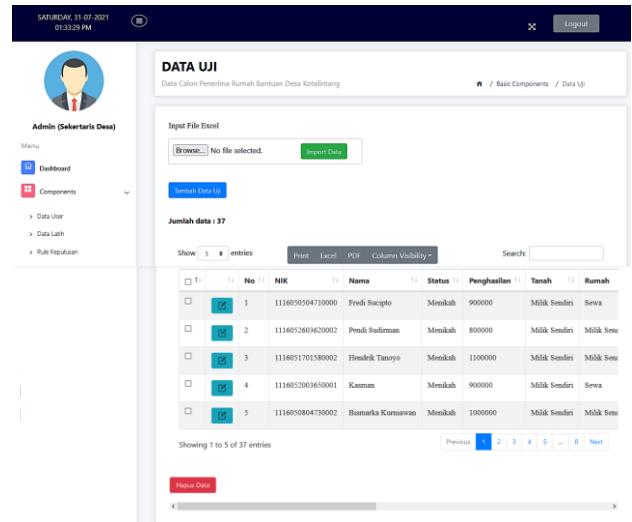
Gambar 7. Halaman Rule Keputusan

4. Halaman Data Uji



Halaman data uji merupakan halaman yang menampilkan data penerima rumah bantuan yang akan diklasifikasikan berdasarkan rule-rule keputusan. Pada halaman ini admin

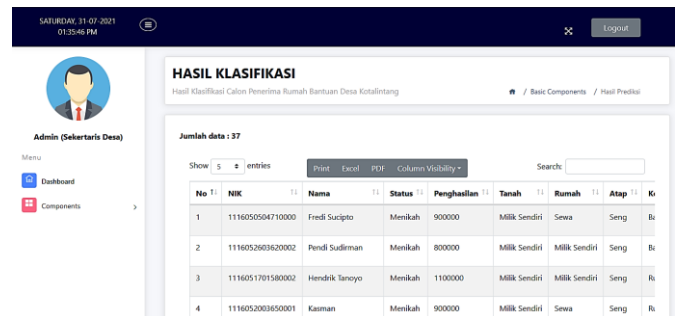
dapat melakukan import data excel, tambah data, lihat prediksi, hapus, edit data dan pencarian data. Halaman data uji dapat dilihat pada Gambar 7.



Gambar 8. Halaman Data Uji

5. Halaman Klasifikasi

Halaman hasil klasifikasi merupakan halaman yang menampilkan hasil dari proses klasifikasi pada data uji. Proses klasifikasi dilakukan dengan mencocokkan kriteria-kriteria pada data uji dengan rule keputusan yang telah dibentuk dari perhitungan *decision tree* sebelumnya. Hasil proses klasifikasi memberikan keputusan prediksi sebagai penerima atau bukan penerima. Halaman data latihan dapat dilihat pada Gambar 8.



Gambar 9. Halaman Hasil Klasifikasi

B. Hasil Pengujian

Untuk mengetahui tingkat keberhasilan dari sistem klasifikasi penerima rumah bantuan dengan metode *decision tree*, diperlukan pengujian untuk mengetahui seberapa besar nilai akurasi yang dihasilkan oleh metode *decision tree* dalam mengklasifikasikan penerima rumah bantuan di Kampung Kotalintang. Telah dilakukan 5 kali pengujian dari 185 dataset dengan menggunakan metode *confusion matrix*.

Perhitungan akurasi pengujian 1 :

TP (True Positif) → Hasil prediksi penerima, kelas asli penerima.

Jumlah = 25

TN (True Negatif) → Hasil prediksi bukan penerima, kelas asli bukan penerima.

Jumlah = 8

FP (False Positif) → Hasil prediksi penerima, kelas asli bukan penerima.

Jumlah = 4

FN (False Negatif) → Hasil prediksi bukan penerima, kelas asli penerima

Jumlah = 0

$$\begin{aligned} \text{Akurasi} &= ((TP+TN))/((TP+TN+FP+FN)) \times 100 \\ &= ((25+8))/((25+8+4+0)) \times 100 = 0,8919 \times 100 \\ &= \mathbf{89.19\%} \end{aligned}$$

Nilai akurasi untuk setiap pengujian dapat dilihat pada tabel 2.

Table 2. Nilai Akurasi Pengujian Metode

	Jumlah Data	Akurasi
Pengujian 1	37	89.19 %
Pengujian 2	37	100%
Pengujian 3	37	83.78 %
Pengujian 4	37	83.78 %
Pengujian 5	37	83.78 %
Rata-Rata		88,11%

IV. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan, maka penulis dapat menyimpulkan beberapa hal berikut :

1. Sistem klasifikasi penerimaan rumah bantuan dengan mengimplementasikan metode decision tree untuk

memberikan hasil berupa suatu prediksi penerima rumah bantuan di Kampung Kotalintang.

2. Implementasi metode decision tree pada sistem dengan cara membentuk pohon keputusan berdasarkan data latih yang digunakan, pohon keputusan yang dibentuk menghasilkan rule-rule keputusan. Berdasarkan rule keputusan tersebut akan memberikan hasil prediksi terhadap calon penerima rumah bantuan.
3. Sistem akan memberikan output atau hasil akhir berupa hasil klasifikasi sebagai penerima atau bukan penerima rumah bantuan berdasarkan rule-rule keputusan.
4. Berdasarkan pengujian yang telah dilakukan metode decision tree memberikan rata-rata nilai akurasi sebesar 88,11% untuk 5 kali pengujian dari 185 dataset.
5. Hasil pengujian sistem membuktikan bahwa semua fitur yang terdapat didalam sistem dapat berjalan sesuai dengan fungsinya.

REFERENSI

- [1] Hendrian, S. (2018). Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan. *Faktor Exacta*, 11(3), 266–274. <https://doi.org/10.30998/faktorexacta.v11i3.2777>
- [2] Mardi, Y. (2017). Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 2(2), 213–219.
- [3] Vijay Kotu and Bala Deshpande, Ph.D. Predictive Analytics and Data Mining: Concepts and Practice with RapidMiner. Elsevier. 2015.
- [4] Meilina, P. (2015). Penerapan Data Mining dengan Metode Klasifikasi Menggunakan Decision Tree dan Regresi. *Jurnal Teknologi Universitas Muhammadiyah Jakarta*, 7(1), 11–20. jurnal.ftumj.ac.id/index.php/jurtek
- [5] Selly Artaty Zega. (2014). Penggunaan Pohon Keputusan untuk Klasifikasi Tingkat Kualitas Mahasiswa Berdasarkan Jalur Masuk Kuliah <http://p2m.polibatam.ac.id/wp-content/uploads/2014/07/Selly-Artati.pdf>. 1907
- [6] Rosandy, T. (2016). *Perbandingan metode naive bayes classifier dengan metode decision tree (c4 . 5) untuk menganalisa kelancaran pembiayaan (study kasus : kspps / bmt al-fadhila*