

Penggunaan Algoritma ID3 Untuk Klasifikasi Data Calon Peserta Didik

Use of ID3 Algorithm for Data Classification of Prospective Students

Agus Sifaunajah^{a,1}, Riesca Dewi Wahyuningtyas^{b,2}

^{a,b}Program studi Sistem Informasi, Fakultas Teknologi Informasi

Universitas KH. Abdul Wahab Hasbullah, Jl.Garuda No.9 Tambakberas Jombang

¹agus.syifa85@gmail.com, ²dewiriesca@gmail.com

ABSTRAK

Penerapan sistem informasi dalam berbagai proses kehidupan sehari-hari mengakibatkan adanya penumpukan data. Klasifikasi menjadi salah satu cara yang digunakan untuk memanfaatkan penumpukan data, dengan klasifikasi kita dapat melakukan analisa data dengan tujuan membentuk model untuk menggambarkan informasi yang terkandung dalam data. Dalam Penelitian kali ini, algoritma ID3 dipergunakan untuk melakukan klasifikasi data calon peserta didik. Dalam penelitian ini model penelitian kualitatif digunakan untuk membangun kerangka menarik kesimpulan. Aplikasi Rapid Miner digunakan sebagai tool bantuan melakukan validasi proses klasifikasi. Data yang diolah merupakan data yang didapatkan dari survey melalui media google form. Hasil dari penelitian ini diperoleh sebuah model klasifikasi dengan menggunakan Algoritma Iterative Dichotomizer 3 (id3) yang dapat dikembangkan ke dalam sebuah sistem informasi penerimaan calon peserta didik. Hasil dari penelitian ini didapatkan klasifikasi data peserta didik "tidak mampu" atau "mampu" pada yayasan pondok pesantren bahrul ulum jombang dalam bentuk pohon keputusan. Dari pohon keputusan itu didapatkan data bahwa 50,72 % (35 orang) dalam klasifikasi tidak mampu, dan 49,27 % (34 orang) termasuk klasifikasi mampu dari total keseluruhan 69 data calon peserta didik.

Kata Kunci : Sistem informasi, klasifikasi, kualitatif, Algoritma id3.

ABSTRACT

The application of information systems in various processes of daily life results in the accumulation of data. Classification is one of the methods used to take advantage of the accumulation of data. With classification, we can perform data analysis to form a model to describe the information contained in the data. In this study, the ID3 algorithm was used to classify prospective student data. In this research, a qualitative research model is used to build a framework for concluding. The Rapid Miner application is used as a tool to help validate the classification process. The processed data is data obtained from a survey through the google form media. The results of this study obtained a classification model using the Iterative Dichotomizer Algorithm 3 (id3), which can be developed into an information system for the admission of prospective students. The results of this study obtained data classification of students "cannot afford" or "able" at the foundation of the boarding school Bahrul Ulum Jombang in the form of a decision tree. From the decision tree, it was found that 50.72% of the data were classified as poor, and 49.27% was classified as capable of a total of 345 data on prospective students.

Keywords : Information systems, classification, qualitative, Algorithm id3

Info Artikel :

Disubmit: 08 Agustus 2021

Direview: 03 November 2021

Diterima : 10 Mei 2022

Copyright © 2022 – CSRID Journal. All rights reserved.

1. PENDAHULUAN

Penerapan sistem informasi dalam proses penerimaan peserta didik mengakibatkan adanya penumpukan data. Sistem informasi dapat diartikan sebagai sekumpulan elemen yang saling berhubungan atau terintergasi, dengan mengumpulkan dan menyebarluaskan informasi dengan tujuan dapat menunjang proses pengambilan keputusan dan tujuan lainnya, baik pada orang maupun suatu parlemen [1].

Terdapat beberapa fungsionalitas dari *data mining*, diantaranya analisis asosiasi antar data, melakukan klasifikasi data, klastering data dan lainnya. fungsionalitas yang digunakan ialah klasifikasi data. Klasifikasi merupakan teknik analisa data bertujuan untuk menghasilkan model-model serta menggambarkan kelas-kelas dari data tersebut [2]. Model-model yang digunakan untuk menyusun kelas-kelas yang terkandung pada data disebut classifier [3].

Yayasan Pondok Pesantren Bahrul Ulum merupakan yayasan yang terdapat di Kab. Jombang. Yayasan ini membawahi banyak unit lembaga pendidikan. Salah satu data pendukung yang digunakan pengurus yayasan dalam mengambil keputusan adalah klasifikasi data calon peserta didik dilingkungan yayasan. Algoritma ID3 akan memproses klasifikasi data calon peserta didik di lingkungan Yayasan. Dengan klasifikasi diharapkan dapat menemukan pola klasifikasi para calon peserta didik yang akan mendukung dalam pengambilan keputusan oleh pengurus yayasan terkait [4].

2. METODE PENELITIAN

A. Metode Kualitatif

Penelitian kualitatif merupakan penelitian yang bersifat deskriptif yang cenderung menggunakan analisis, proses, makna dan perspektif. Sedangkan subjek lebih ditonjolkan dalam penelitian kualitatif. Pada penelitian ini alur pengambilan kesimpulan menggunakan alur kualitatif. Landasan teori dipergunakan untuk pemandu supaya fokus dalam penelitian agar sesuai dengan fakta yang ada di lapangan. Penelitian kualitatif mulai dikenal sejak tahun 1960-an dan disebut dengan metode alternatif (*alternative method*). Metode tersebut tidak memakai pertanyaan yang rinci, melainkan dimulai dengan pertanyaan umum, lalu semakin meruncing dan detail. Dalam penelitian kualitatif, partisipan diperlakukan sebagai subjek sehingga partisipan beranggapan dirinya berharga karena informasi dari mereka sangat bermanfaat.

B. Pengambilan Data

Pengambilan data pada penelitian ini, peneliti menerapkan cara pengambilan data menggunakan sistem survey dengan memanfaatkan media aplikasi *google form*. Alasan penggunaan aplikasi *google form* agar mempermudah responden berpartisipasi dalam penelitian. Data yang didapatkan akan diolah lebih lanjut menggunakan proses klasifikasi data calon peserta didik.

C. Alur Kerja Algoritma ID3

Algoritma ID3 merupakan salah satu algoritma yang digunakan untuk membuat atau membangun sebuah decision tree atau pohon keputusan [1][5]. Algoritma ini ditemukan oleh J. Ross Quinlan pada tahun 1979, dengan memanfaatkan Teori Informasi milik Shanon. ID3 merupakan singkatan dari *Iterative Dichotomiser 3* [6]. Dalam prosesnya, *decision tree* memakai struktur *hierarki* untuk pembelajaran *supervised*. Alur kerja dari *decision tree* dimulai dari *root node* hingga *leaf node* yang dilakukan secara rekursif, dimana setiap percabangannya menyatakan suatu kondisi yang harus dipenuhi dan disetiap ujung pohon menyatakan kelas dari suatu data. Proses yang berjalan dalam *decision tree* ialah dengan cara merubah bentuk data yang berupa tabel dirubah menjadi model pohon (*tree*), selanjutnya model pohon diubah menjadi rule atau aturan yang dihasilkan.

a) Entropy dan Gain

Entropy merupakan formula yang digunakan sebagai alat menghitung homogenitas dari sebuah sample atau contoh [5]. Entropy merupakan suatu ukuran ketidakpastian yang diasosiasikan dengan variabel acak yang memiliki nilai antara 0-1 sebagai representasi dari informasi yang diketahui. Dengan mengetahui nilai entropy, akan bisa diketahui seberapa informatif dari node yang ada. Berikut ini merupakan persamaan yang digunakan untuk mencari nilai entropy dari suatu data :

$$Entropy(S) = -\sum_{i=1}^c p_i \log_2 p_i \quad (1)$$

Dimana:

S = Himpunan (*dataset*) kasus

K = Banyaknya partisi S

P_j = Probabilitas yang didapat dari penjumlahan nilai Y_a dibagi jumlah total kasus

Setelah didapatkan nilai *entropy*, pemilihan *attribut* dilakukan dengan memperhatikan nilai *information gain terbesar*. *Information Gain* merupakan penguatan informasi yang digunakan sebagai ukuran pemilihan atribut untuk dijadikan akar dalam pohon keputusan. Untuk menentukan *information gain*, digunakan rumus sebagai berikut :

$$Gain(A) = Entropy(S) - \sum_{i=1}^k \frac{|S_i|}{|S|} \times Entropy(S_i) \quad (2)$$

Dimana :

S = Ruang (data) sample yang digunakan untuk training

A = Atribut

$|S_i|$ = Jumlah sample untuk nilai V

$|S|$ = Jumlah seluruh sample data

Entropi $|S_i|$ = Entropi untuk sample – sample yang memiliki nilai i

b) Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan metode yang dipakai sebagai prosedur penalaran supaya mendapatkan jawaban permasalahan yang dihadapi[3]. Pohon yang terbentuk tidak selalu berupa pohon biner. Apabila dataset yang dimiliki menggunakan 2 macam nilai kategorikal maka bentuk pohon yang didapatkan dapat berupa pohon biner. Namun, jika fiturnya berisi lebih dari 2 macam nilai kategori atau menggunakan tipe data numerik maka bentuk pohon yang didapatkan biasanya tidak berupa pohon biner.

Sifatnya yang flexible membuat agar metode ini atraktif, khususnya karena memberikan keuntungan berupa visualisasi saran dalam bentuk pohon keputusan yang menjadikan prosedur prediksinya dapat diamati.

c) Karakteristik Data

Karakteristik data adalah tipe data, tingkat akurasi serta tingkat presisi, usia, rentang waktu, kelengkapan, tingkat keringkasan, relevansi, serta kemudahan untuk mengakses sumber data. Dalam penelitian ini, karakteristik data yang digunakan terdiri atas beberapa bagian berikut ini:

i. Class Target

Class taraget adalah atribut yang dijadikan sasaran dalam proses klasifikasi, class target ini sering juga disebut dengan label atau atribut target . Adapun class target yang diinginkan dalam penelitian ini adalah kategori keluarga mampu atau keluarga tidak mampu yang mana penentuan kategori ini didasarkan pada kategori yang ditetapkan oleh Badan Pusat Statistik (BPS).

ii. Syarat – syarat dari class target antara lain [7]:

1. keluarga Mampu keluarga memliki kategori sebagai berikut :

- Gaji Ayah diatas Rp. 1.000.000 - 5.000.000 per bulan
- Gaji ibu diatas Rp. 500.000 - Rp.3000.000 per bulan
- Perkerjaan Ayah dan Ibu
- Tidak Mempunyai saudara atau mempunyai saudara tetapi cuma 1
- Mempunyai kendaraan dengan harga cash
- Saudara sudah kerja atau sedang menempuh pendidikan tinggi

2. Keluarga Tidak Mampu

- Gaji Ayah diatas Rp. 500.000 - Rp. 1.000.000
- Gaji Ibu diatas Rp.0 – Rp. 500.000
- Perkerjaan Ayah dan Ibu
- Mempunyai saudara lebih dari 3

- Mempunyai kendaraan motor saja dengan melakukan pembayaran secara kredit
 - Ada yang masih kerja atau yang masih menempuh pendidikan
- iii. Atribut Dari Variabel
- Variable dari pekerjaan terdiri dari kategori pekerjaan ayah , pekerjaan ibu
 - Variabel dari gaji terdiri dari gaji ayah dan gaji ibu
 - Variabel jumlah saudara terdiri dari berapa saudara seperti satu , dua , tiga atau empat saudara
- iv. Identifikasi Data
- Identifikasi data merupakan proses dilakukannya pemeriksaan mengenai kualitas dan kelengkapan data. Proses identifikasi ini meliputi nilai-nilai yang hilang atau kosong pada data.
- d) Persiapan Data (Data Preparation)
- Pada tahap ini mencakup semua data, data tersebut akan dipersiapkan terlebih dahulu dan akan dimasukkan kedalam alat pemodelan, data tersebut diolah dari awal. Data diolah menggunakan tool bantu Ms.Excel dengan menetapkan parameter class target terlebih dahulu atau atribut variable dan mempelajari data tersebut sehingga diharapkan data tersebut dapat memiliki pengetahuan yang lebih pada data yang diolah sehingga dapat membantu untuk melakukan pemilihan data yang digunakan untuk proses data mining. Dalam melakukan *data preparation* yang perlu dilakukan ada beberapa tahapan, antara lain :
- 1) Data Sampling

Data sampling adalah data yang akan digunakan pada proses data mining. Adapun data yang dikelola antara lain :

 - a) Data pekerjaan ayah dan pekerjaan ibu
 - b) Data gaji ayah dan data gaji ibu
 - c) Data jumlah saudara
 - d) Kelayakan = (Mampu , Tidak Mampu)
 - 2) Seleksi Data

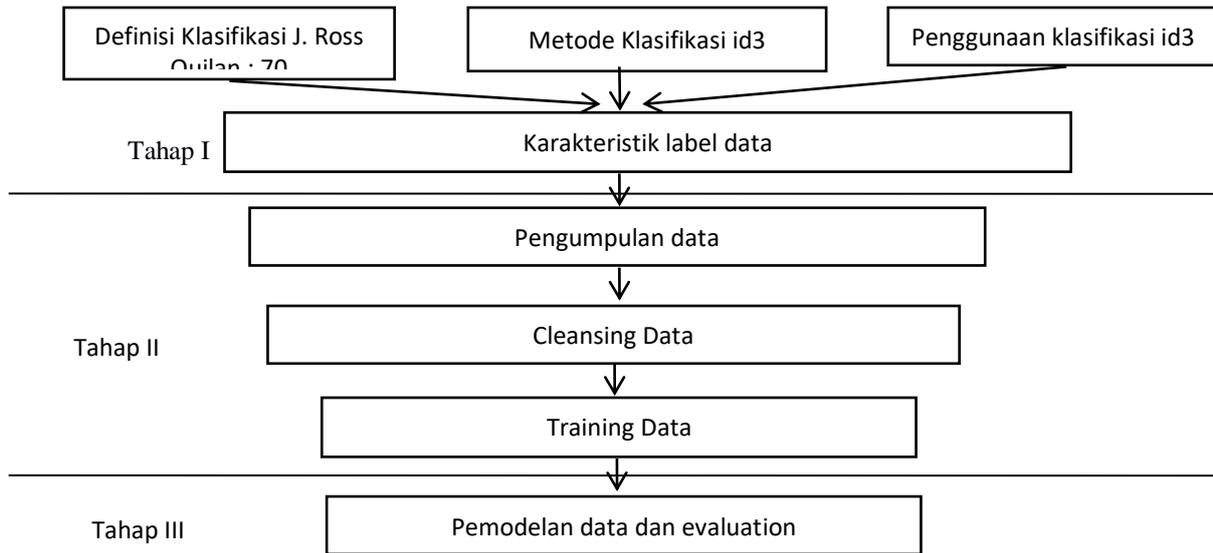
Data mentah yang didapatkan dari google form dan dilakukan proses pemilihan terhadap parameter yang akan dianalisis. Parameter yang akan diambil merupakan variable atau atribut yang akan mempengaruhi terhadap penentuan class target yang sebelumnya telah ditentukan yaitu class target “mampu” atau “tidak mampu”.
 - 3) Membersihkan Data

Tahapan ini digunakan untuk membersihkan data dari setiap data yang tidak konsisten dan noise agar data siap untuk di tahapan modeling.
 - 4) Transformasi Data

Setelah data bersih dari noise dan data-data yang tidak konsisten selanjutnya melakukan transformasi data. Transformasi ini dilakukan untuk memudahkan proses klasifikasi dan perhitungan Entropy dan Gain dalam membangun pohon keputusan. Transformasi data ini dilakukan jika dalam dataset terdapat data-data yang belum sesuai kaidah proses data mining.

D. Bagan Alur Penelitian

Alur penelitian yang digunakan secara garis besar terbagi menjadi 3 tahap. Pada tahap pertama, dilakukan pendefinisian karakteristik data dan penentuan model klasifikasi serta penentuan *class target* yang diinginkan. Selanjutnya, pada tahap kedua dilakukan pengumpulan data dan dilakukan pengolahan awal terkait anomali dan inkonsistensi data. Setelah itu dilakukan training data untuk menguji performa dari algoritma yang digunakan. Apabila training data menghasilkan informasi yang sesuai dengan yang kita harapkan maka dilanjutkan dengan membuat model klasifikasi untuk menguji data, proporsi data yang dipakai untuk training dan data yang dipakai untuk testing adalah 60 persen berbanding 40 persen. Hasil dari pemodelan tersebut kemudian dilakukan evaluasi untuk mengetahui tingkat efektifitas dari kinerja algoritma. Berikut ini adalah gambaran alur proses penelitian yang dilakukan.

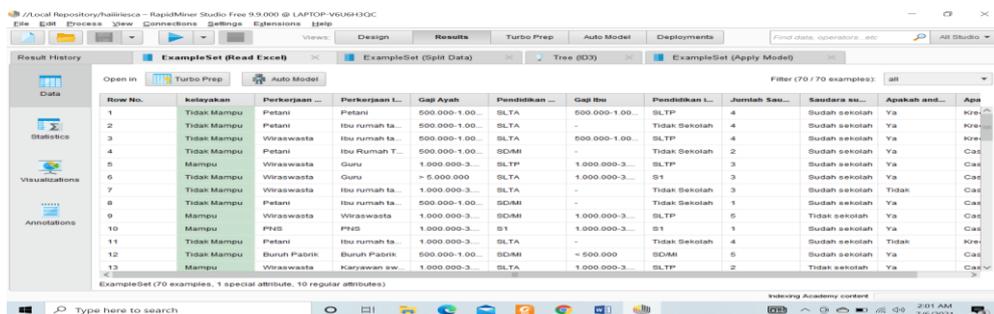


Gambar 1. Bagan alur instrument penelitian

3. HASIL DAN PEMBAHASAN

A. Data Set

Data mentah yang berasal dari survey menggunakan *google form*, selanjutnya dilakukan untuk memilih parameter yang akan dianalisis. Parameter yang diambil menggunakan class target yang sebelumnya dibuat dengan class target “mampu” dan “tidak mampu” atribut tersebut akan dijadikan parameter atau variable input. Dataset ini diolah menggunakan tool Ms.Excel lalu file ini yang akan dibaca dan diolah oleh aplikasi rapid miner. Rapid miner dapat membaca dataset dengan berbagai format file dan dapat terhubung langsung dengan database suatu sistem informasi. Berikut ini adalah tampilan interface rapid miner saat membaca dataset dari Ms. Excel.



Gambar 2. Load Data di Rapid Miner

B. Pengolahan Data (Processing)

Pada proses pengolahan data kali ini, data diolah menggunakan tool Ms.Excel dengan menerapkan perhitungan secara manual. Pada pengolahan ini, dilakukan proses menghitung guna mencari nilai *entropy* dan nilai *gain* di setiap parameter yang ada. Proses ini dilakukan secara manual untuk membuktikan proses perhitungan nilai *entropy* dan nilai *gain* berdasarkan rumus yang telah ada. Sebagai perbandingan dari perhitungan manual ini, hasilnya akan dibandingkan dengan hasil proses yang dilakukan menggunakan aplikasi rapid miner sebagai kontrolnya. Dari jumlah total data 345, nilai *gain* tertinggi diperoleh pada atribut gaji ayah. Berikut ini adalah tabel perhitungan manual dari proses mencari nilai *entropy* dan nilai *gain*.

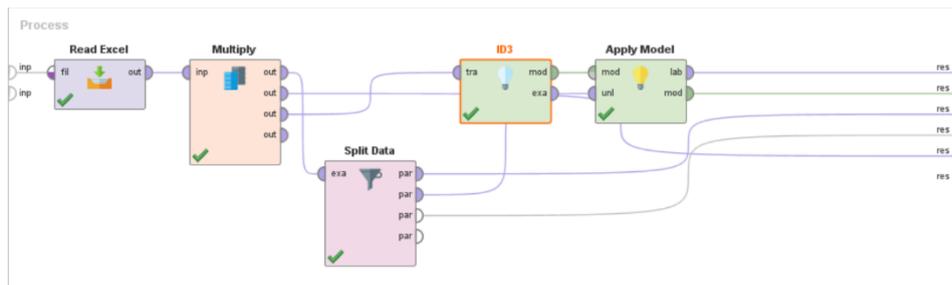
Tabel 3. Data diolah di excel

| Atribut | kategori | jumlah | Mampu | Tidak Mampu | entropy | gain |
|-----------------|---------------------|--------|-------|-------------|-------------|-------------|
| Total | | 345 | 170 | 175 | 0.999848483 | |
| Gaji Ayah | | | | | | 0.223008761 |
| | < 500.000 | 24 | 0 | 24 | 0 | |
| | 500.000-1.000.000 | 101 | 22 | 79 | 0.756170149 | |
| | 1.000.000-3.000.000 | 109 | 79 | 30 | 0.848868756 | |
| | > 5.000.000 | 111 | 13 | 42 | 0.892881302 | |
| Gaji Ibu | | | | | | 0.213204659 |
| | - | 41 | 15 | 26 | 0.947435136 | |
| | < 500.000 | 158 | 54 | 104 | 0.926504446 | |
| | 500.000-1.000.000 | 64 | 24 | 40 | 0.954434003 | |
| | 1.000.000-3.000.000 | 62 | 57 | 5 | 0.404448386 | |
| | > 5.000.000 | 19 | 19 | 0 | 0 | |
| Pekerjaan Ayah | | | | | | 0.069777051 |
| | Wiraswasta | 145 | 75 | 70 | 0.999142104 | |
| | Petani | 125 | 45 | 80 | 0.942683189 | |
| | Guru | 15 | 5 | 10 | 0.918295834 | |
| | PNS | 45 | 35 | 10 | 0.764204507 | |
| | Buruh Pabrik | 10 | 5 | 5 | 1 | |
| Pekerjaan Ibu | | | | | | 0.168671041 |
| | Wiraswasta | 50 | 35 | 15 | 0.881290899 | |
| | Petani | 60 | 15 | 45 | 0.811278124 | |
| | Guru | 30 | 15 | 15 | 1 | |
| | PNS | 15 | 15 | 0 | 0 | |
| | Buruh Pabrik | 5 | 0 | 5 | 0 | |
| | Ibu Rumah Tangga | 165 | 75 | 90 | 0.994030211 | |
| Pendidikan Ayah | | | | | | 0.02215492 |
| | SD/MI | 100 | 40 | 60 | 0.970950594 | |
| | SLTP | 60 | 30 | 30 | 1 | |
| | SLTA | 150 | 75 | 75 | 1 | |
| | S1 | 35 | 25 | 10 | 0.863120569 | |
| Pendidikan Ibu | | | | | | 0.03997623 |
| | Tidak Sekolah | 12 | 4 | 8 | 0.918295834 | |
| | SD/MI | 86 | 40 | 46 | 0.99648599 | |
| | SLTP | 70 | 30 | 40 | 0.985228136 | |
| | SLTA | 122 | 56 | 66 | 0.995148096 | |
| | S1 | 50 | 35 | 15 | 0.881290899 | |
| Kendaraan | | | | | | 0.002156212 |
| | Cash | 245 | 125 | 120 | 0.999699543 | |
| | Kredit | 100 | 45 | 55 | 0.992774454 | |
| Saudara | | | | | | 0.173041197 |
| | 1 | 65 | 50 | 15 | 0.779349837 | |
| | 2 | 70 | 45 | 25 | 0.940285959 | |

| | | | | | | |
|--|---|-----|----|----|-------------|--|
| | 3 | 100 | 50 | 50 | 1 | |
| | 4 | 55 | 5 | 50 | 0.439496987 | |
| | 5 | 45 | 20 | 25 | 0.99107606 | |
| | 6 | 10 | 0 | 10 | 0 | |

C. Modelling

Pada proses ini, tahapan dilakukan dengan menggunakan aplikasi rapid miner. Tahapan ini adalah tahapan memproses data untuk dilakukan klasifikasi menggunakan model algoritma yang terdapat pada aplikasi rapid miner. Pada proses ini semua tahapan algoritma ID3 dilakukan dijalankan secara otomatis oleh aplikasi rapid miner. Apabila ditemukan kesalahan dalam parameter akan diberikan tanda seru dengan warna kuning. Berikut ini adalah tampilan pembuatan model dari algoritma ID3 :



Gambar 4. Modelling Algoritma ID3

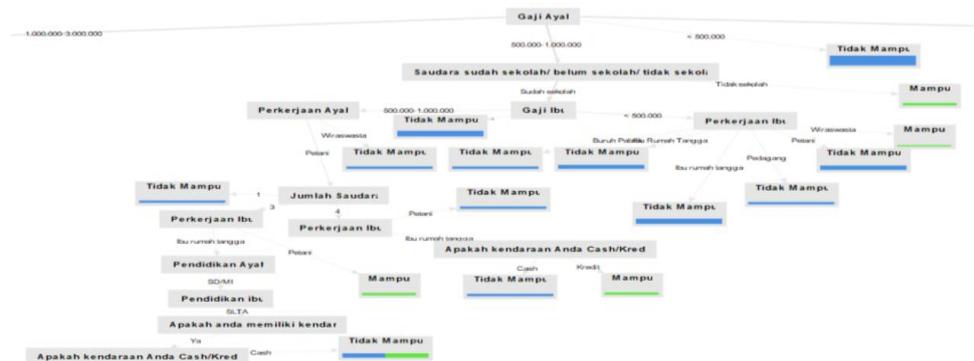
D. Pohon Keputusan (Decision Tree)

Pada tahapan ini, kita sudah sampai pada melihat hasil perhitungan dari nilai *gain*. Penyusunan pohon keputusan didasarkan dengan nilai *gain* dari tiap parameter yang telah diproses sebelumnya dan nilai *gain* terbesar akan dijadikan sebagai akar atau *root* pada pohon keputusan tersebut. Dalam kasus kali ini, parameter yang dijadikan sebagai *root* adalah gaji ayah.

Pada gambar 5 menunjukkan pembentukan pohon keputusan, yaitu atribut gaji ayah sebagai *root*. Cabang yang terbentuk berisi data kategori pada atribut ayah, yaitu gaji ayah <500.000 , gaji ayah 500.000 - 1000.000 , gaji ayah 1.000.000 – 3.000.000 , gaji ayah >5.000.000.

Sedangkan pada gambar 6 menunjukkan pembentukan cabang pohon keputusan selanjutnya yang sebelah kiri, yaitu atribut gaji ibu. Cabang yang terbentuk berisi data kategori pada atribut gaji, yaitu gaji ibu <500.000 , gaji ibu 500.000 - 1000.000 , gaji ibu 1.000.000 – 3.000.000 , gaji ibu >5.000.000.

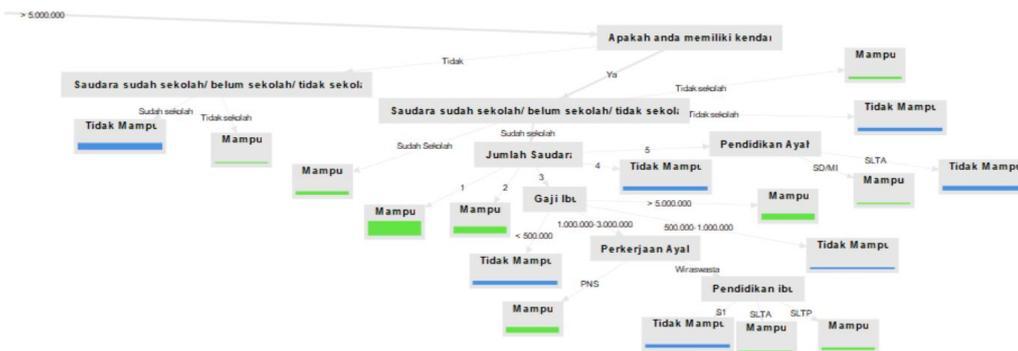
Yang terakhir gambar 7 menunjukkan pembentukan cabang pohon keputusan pada sebelah kanan, yaitu atribut saudara. Cabang yang terbentuk berisi data kategori pada atribut saudara, yaitu saudara 1, 2 , 3 , 4 , 5 , dan 6.



Gambar 5. Pohon Keputusan dengan label gaji ayah sebagai root



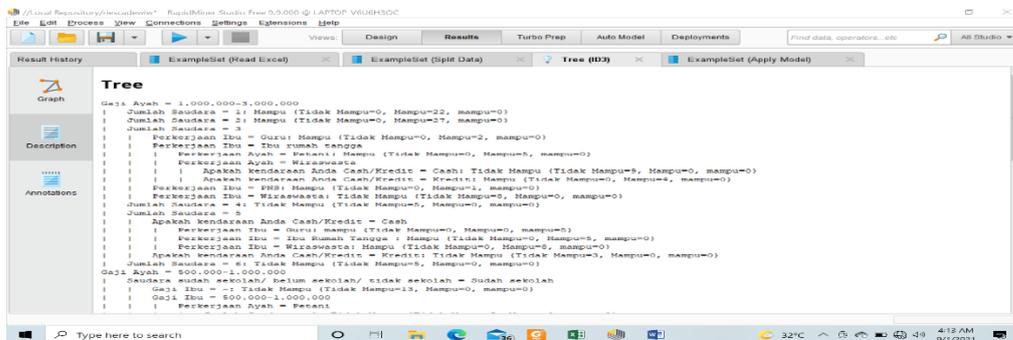
Gambar 6. Pohon Keputusan Cabang Sebelah Kiri



Gambar 7. Pohon Keputusan Cabang Sebelah Kanan

E. Rule Base

Rule base merupakan hasil dari deskripsi sebuah pohon keputusan yang berhasil dibangun. Rule base ini dapat membantu kepada peneliti dan memudahkan di dalam membaca hasil dari klasifikasi data. Berikut ini adalah rule base dari pohon keputusan yang berhasil dibangun sebagaimana pada gambar 8 berikut ini :



Gambar 8. Rule Base dari Pohon Keputusan

F. Pembahasan

Pada penelitian ini, titik tekan yang dilakukan adalah membuat model klasifikasi data calon peserta didik di lingkungan pendidikan di bawah naungan yayasan pondok pesantren bahrul ulum tambakberas jombang. Ide ini muncul pertama kali dilatarbelakangi adanya keinginan pengurus

yayasan untuk dapat melakukan monitoring dan pengaturan calon peserta didik di lembaga pendidikan pada naungan yayasan terkait.

Berdasarkan model yang telah dibangun, perbandingan antara calon peserta didik dari keluarga mampu dan tidak mampu hampir berimbang. Dari data sample yang digunakan dalam proses klasifikasi ini, terdapat 50,72% calon peserta didik kurang mampu dan 49,27 % calon peserta didik dari keluarga mampu. Hal ini menunjukkan bahwa, lembaga pendidikan di lingkungan Yayasan Pendidikan Pondok Pesantren Bahrul Ulum diminati dari berbagai kalangan. Namun, angka 50,72% harus menjadi perhatian dari pengurus yayasan terkait. Pengurus yayasan harus merumuskan kebijakan terkait pemantauan kondisi peserta didik dan merumuskan solusi-solusi untuk menekan jumlah angka dari calon peserta didik dari kalangan kurang mampu agar dapat terus melanjutkan pendidikan.

Salah satu hal lain yang bisa dilakukan adalah memastikan angka jumlah peserta didik pada lembaga pendidikan di lingkungan Yayasan Pondok Pesantren Bahrul Ulum. Ini bisa dilakukan dengan membangun sistem pendataan tersentral agar anomali data dapat ditekan semaksimal mungkin sehingga fungsi yayasan sebagai pengayom dan pengelola lembaga pendidikan dapat dimaksimalkan.

4. KESIMPULAN

Berdasarkan hasil dari pembahasan penelitian ini, dapat disimpulkan antara lain bahwa dapat dibangun sebuah model klasifikasi menggunakan Algoritma Iterative Dichotomizer(id3) pada data calon peserta didik lembaga pendidikan di lingkungan Yayasan Pondok Pesantren Bahrul Ulum Jombang berbentuk pohon keputusan. Dari pohon keputusan tersebut, dapat dideskripsikan bahwa gaji dari seorang ibu dibawah angka > 500.000 dan gaji ayah 1000.000-3000.000 yang memiliki anak lebih dari 3 maka diklasifikasikan tidak mampu, Sedangkan gaji seorang ibu diatas angka 1.000.000-3.000.000 dan gaji seorang ayah di angka >5000.000 yang memiliki anak 1, 2 atau 3 maka diklasifikasikan mampu.

Secara keseluruhan dari data yang diolah menggunakan rapidminer, didapatkan bahwa 50,72 % masuk klasifikasi tidak mampu serta 49,27 % menyatakan masuk klasifikasi mampu dari total keseluruhan 345 data peserta didik.

Penelitian ini dapat dilanjutkan dengan cara mengintegrasikan model klasifikasi ini ke dalam sebuah sistem informasi penerimaan peserta didik baru tersentral di lingkungan Yayasan Pondok Pesantren Bahrul Ulum Tambakberas Jombang agar lebih memudahkan pengurus yayasan dalam menentukan kebijakan pada lembaga pendidikan di bawah naungannya.

REFERENSI

- [1] D. K. Widiyati, M. Wati, and H. S. Pakpahan, "Penerapan Algoritma ID3 Decision Tree Pada Penentuan Penerima Program Bantuan Pemerintah Daerah di Kabupaten Kutai Kartanegara," *J. Rekayasa Teknol. Inf.*, vol. 2, no. 2, p. 125, 2018, doi: 10.30872/jurti.v2i2.1864.
- [2] B. Novianti, T. Rismawan, and S. Bahri, "Implementasi Data Mining Dengan Algoritma C4.5 Untuk Penjurusan Siswa (Studi Kasus: Sma Negeri 1 Pontianak)," *J. Coding, Sist. Komput. Untan*, vol. 04, no. 3, pp. 75–84, 2016.
- [3] M. Safii, "Implementasi Data Mining Dengan Metode Pohon Keputusan Algoritma Id3 Untuk Menentukan Status Mahasiswa," *J. Mantik Penusa*, vol. 2, no. 1, pp. 82–86, 2019.

- [4] Ardiyansyah, P. A. Rahayuningsih, and R. Maulana, "Analisis Perbandingan Algoritma Klasifikasi Data Mining Untuk Dataset Blogger Dengan Rapid Miner," *J. Khatulistiwa Inform.*, vol. VI, no. 1, pp. 20–28, 2018.
- [5] I. R. Munthe and V. Sihombing, "Klasifikasi Algoritma Iterative Dichotomizer (ID3) untuk Tingkat Kepuasan pada Sarana Laboratorium Komputer," *J. Teknol. dan Ilmu Komput. Prima*, vol. 1, no. 2, pp. 27–34, 2018, doi: 10.34012/jutikomp.v1i2.237.
- [6] F. Rahman, D. Muhammad, and I. Firdaus, "Penerapan Data Mining Metode Naïve Bayes Untuk Prediksi Hasil Belajar Siswa Sekolah Menengah Pertama (Smp)," *Al Ulum Sains dan Teknol.*, vol. 1, no. 2, pp. 76–78, 2016.
- [7] A. Sitepu, "Karakteristik Keluarga Menurut Peringkat Kemiskinan: Studi Pendahuluan untuk Perumusan Kriteria Fakir Miskin," *Informasi*, vol. 17, no. 01, pp. 48–63, 2012, [Online]. Available: <https://ejournal.kemsos.go.id/index.php/Sosioinforma/article/viewFile/930/490>.