

Aplikasi Penentuan Penerima Beasiswa Menggunakan Algoritma C4.5

¹Abdurraghib Segaf Suweleh, ²Dyah Susilowati, ³Hairani

^{1,2,3} Fakultas Teknik dan Desain, Universitas Bumigora
abdurraghibsegaf@gmail.com, dyah.bumigora@gmail.com, hairani@universitasbumigora.ac.id

Abstrak

Pada proses penentuan beasiswa sering muncul permasalahan mengenai tidak adanya perhitungan pasti untuk menentukan penerima beasiswa yang berhak. Hal ini mengakibatkan pelaksanaan seleksi beasiswa membutuhkan waktu yang relatif lama. Implementasi aplikasi yang dapat memprediksi calon penerima beasiswa yang menggunakan teknik *data mining*, dapat menjadi salah satu solusi alternatif untuk mengatasi permasalahan tersebut. Metode penelitian yang digunakan yaitu metode *waterfall* dengan tahap yaitu analisa kebutuhan, perancangan diagram alur dan *interface*, implementasi menggunakan PHP dan MySQL, dan pengujian menggunakan metode *black box*. Data yang digunakan untuk pengujian merupakan data mahasiswa sebanyak 125 data. Hasil yang dicapai dari pengujian tersebut yaitu diketahuinya tingkat akurasi implementasi algoritma C4.5 pada proses penentuan penerima beasiswa mencapai 92%, spesifisitas 92.3%, dan sensitifitas 91.6%. Kesimpulan dari penelitian ini adalah algoritma C4.5 berhasil diimplementasikan dalam proses klasifikasi penerima beasiswa berdasarkan pada tingkat akurasi yang mencapai 92% dan hasil pengujian sistem yang valid pada setiap modul menggunakan metode *black box*.

Kata Kunci : Algoritma C4.5, Data Mining, Black Box, Beasiswa

Abstract

In the process of determining the scholarship, problems often arise regarding the absence of a precise calculation to determine the right recipient of the scholarship. This problem results in the relatively long scholarship selection process. Implementation of applications that can predict prospective scholarship recipients using data mining techniques, can be an alternative solution to overcome these problems. The research method used is the waterfall method with stages: needs analysis, designing flowcharts and interfaces, implementation using PHP and MySQL, and testing using the black box method. The data used for testing is 125 student data. The results achieved from the test are knowing the accuracy of the implementation of the C4.5 algorithm in the process of determining scholarship recipients reaching 92%, specificity 92.3%, and sensitivity 91.6%. The conclusion of this study is the C4.5 algorithm was successfully implemented in the classification process of scholarship recipients based on the accuracy level that reached 92% and the results of a valid system test on each module using the black box method.

Keywords : C4.5 Algorithm, Data Mining, Black Box Method, Scholarship

I. PENDAHULUAN

Program beasiswa bertujuan untuk meningkatkan akses dan pemerataan kesempatan belajar bagi seluruh rakyat Indonesia, mengurangi jumlah pelajar yang putus sekolah atau kuliah karena tidak mampu membayar biaya pendidikan studi, serta meningkatkan prestasi dan motivasi pelajar dalam menempuh pendidikannya [1]. Beasiswa tidak hanya ditawarkan bagi mahasiswa yang berprestasi, namun juga bagi mahasiswa kurang mampu agar dapat membantu mahasiswa tersebut dalam menempuh pendidikannya.

Proses penentuan keputusan penerima beasiswa di Universitas Bumigora, didasarkan pada 4 kriteria yaitu pendapatan orang tua, jumlah sks, ipk, dan jumlah tanggungan keluarga. Adanya pendaftar beasiswa yang banyak yang contohnya saja pendaftar beasiswa pada tahun 2019 adalah

sebanyak 110 orang menyebabkan para pengambil keputusan kesulitan dalam menentukan penerima beasiswa secara cepat dan tepat. Untuk itu, dalam proses penerimaan beasiswa diperlukan suatu sistem yang dapat memprediksi penerima beasiswa yang bertujuan untuk mengurangi subyektifitas dalam pengambilan keputusan dan meningkatkan efektivitas proses penentuan penerima beasiswa. Selain itu dengan adanya sistem penentuan beasiswa akan lebih tepat sasaran untuk mahasiswa yang benar-benar membutuhkan dan dengan adanya sistem juga dapat mengurangi tingkat kesalahan penentuan kebijakan.

Untuk menyelesaikan permasalahan tersebut, para peneliti sebelumnya menggunakan konsep *data mining*. Seperti penelitian mengenai perancangan sistem pendukung keputusan untuk prediksi penerima beasiswa menggunakan metode

neural network backpropagation yang dilakukan oleh Ade Pujiyanto [3], juga penelitian mengenai implementasi algoritma K-Nearest Neighbor sebagai pendukung keputusan klasifikasi penerima beasiswa PPA dan BBM oleh Sumarlin [4], serta penelitian mengenai *development of decision support system determining the student as scholarship awardees by fuzzy multi attribute decision making* (FMADM) oleh Saragih Hoga [5]. Ketiga penelitian diatas menyelesaikan permasalahan dalam prediksi penerima beasiswa, akan tetapi metode klasifikasi yang digunakan berbeda-beda. Untuk penelitian kali ini, penulis akan menyelesaikan permasalahan dalam prediksi penerima beasiswa menggunakan algoritma C4.5 dengan melakukan *preprocessing dataset* terlebih dahulu.

Data mining merupakan proses dalam mendapatkan informasi berguna yang berasal dari basis data yang besar [6]. Dengan bantuan *data mining*, pengguna dapat mengekstraksi informasi tersembunyi yang berharga yang dapat membantu dalam pengambilan keputusan atau digunakan untuk membuat prediksi [7].

Algoritma C4.5 adalah algoritma *data mining* yang digunakan untuk melakukan klasifikasi dan estimasi [8]. Algoritma C4.5 termasuk dalam algoritma induksi *decision tree* yang dikembangkan oleh J. Ross Quinlan yang dikenal dengan nama *ID3 (Iterative Dichotomiser 3)* [9].

Hasil klasifikasi yang dihasilkan oleh algoritma *decision tree* cukup baik bila dibandingkan dengan algoritma – algoritma yang lain [10]. Akurasi dan pemodelan yang dimiliki oleh algoritma C4.5 khususnya lebih baik daripada algoritma yang lain didalam keluarga *decision tree* ini serta memiliki tingkat kesalahan yang lebih rendah [11].

Oleh sebab itu maka penulis berinisiatif membuat aplikasi penentuan penerima beasiswa ini. Dengan adanya aplikasi ini, beasiswa dapat diberikan kepada penerima beasiswa yang tepat berdasarkan klasifikasi, kualitas, dan kompetensi penerima [5].

II. METODOLOGI

2.1. Konsep Beasiswa

Murniasih (2010) mengemukakan beasiswa adalah bentuk penghargaan yang diberikan kepada individu agar dapat melanjutkan pendidikan ke jenjang yang lebih tinggi. Penghargaan tersebut bisa berbentuk akses tertentu pada suatu instansi atau penghargaan berupa bantuan keuangan.

Lahinta (2009) mengemukakan beasiswa adalah pemberian berupa bantuan keuangan yang

diberikan kepada perorangan yang bertujuan untuk digunakan demi keberlangsungan pendidikan yang ditempuh. Beasiswa dapat diberikan oleh lembaga pemerintah, perusahaan ataupun yayasan. KBBI (Kamus Besar Bahasa Indonesia) mengemukakan beasiswa adalah tunjangan uang yang diberikan kepada pelajar atau mahasiswa sebagai bantuan biaya belajar. Dari beberapa definisi tersebut, dapat dinyatakan pengertian beasiswa yaitu pemberian bantuan keuangan kepada perorangan untuk keberlangsungan pendidikan yang ditempuhnya karena ketidakmampuan atau prestasinya.

2.2. Konsep Data Mining

Data mining merupakan proses dalam mendapatkan informasi berguna yang berasal dari basis data yang besar [6]. Kata *mining* sendiri berarti usaha untuk mendapatkan barang berharga yang berjumlah sedikit dari material dasar yang berjumlah besar. Karena itu data mining sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (*artificial intelligent*), *machine learning*, statistik dan *database*.

Decision tree adalah *flow-chart* berbentuk pohon, dimana tiap *internal node* menunjukkan sebuah *test* pada sebuah atribut, tiap cabang menunjukkan hasil dari test, dan *leaf node* menunjukkan *class-class* atau *class distribution*.

Algoritma C4.5 adalah algoritma *data mining* yang digunakan untuk melakukan klasifikasi dan estimasi [8]. Algoritma C4.5 termasuk dalam algoritma induksi *decision tree* yang dikembangkan oleh J. Ross Quinlan yang dikenal dengan nama *ID3 (Iterative Dichotomiser 3)* [9].

Decision tree memang populer dan sering digunakan dalam klasifikasi karena memiliki hasil yang cukup baik jika dibanding algoritma lainnya [10]. Ananda & Wibisono (2014) menjelaskan bahwa algoritma C4.5 memiliki tingkat akurasi dan pemodelan waktu yang lebih baik daripada algoritma lainnya dan memiliki tingkat kesalahan yang lebih rendah.

Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut menurut Kusriani (2009):

1. Pilih atribut sebagai akar.
2. Buat cabang untuk tiap-tiap nilai.
3. Bagi kasus dalam cabang.
4. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.



Gambar 1 Flowchart Algoritma C4.5 [1]

Berikut ini akan dijelaskan secara lebih detail algoritma C4.5 menggunakan flowchart yang disajikan pada gambar 1.

Entropy merupakan distribusi probabilitas dalam teori informasi dan diadopsi kedalam algoritma C4.5 untuk mengukur tingkat homogenitas distribusi kelas dari sebuah himpunan data (dataset). Semakin tinggi tingkat entropy dari sebuah data maka semakin homogen distribusi kelas pada data tersebut. Perhitungan information gain menggunakan persamaan 1, sedangkan entropy menggunakan persamaan 2.

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

dimana,

S : Himpunan kasus

A : Atribut

n : Jumlah partisi atribut A

|S_i| : Jumlah kasus pada partisi ke i

|S| : Jumlah kasus dalam S

$$Entropy(S) = \sum_{i=1}^n -P_i * \log_2 P_i \quad (2)$$

dimana,

S : Himpunan kasus

A : Fitur

n : Jumlah partisi S

p_i : Proporsi dari S_i terhadap S

Sensitivitas akan mengukur proporsi positif asli yang dikenali (diprediksi) secara benar sebagai

positif asli. Rumus perhitungannya menggunakan rumus 3. Sedangkan spesifisitas akan mengukur proporsi negatif asli yang dikenali (diprediksi) secara benar sebagai negatif asli. Rumus perhitungannya menggunakan persamaan 4.

$$Sensitivitas = \frac{TP}{TP+FN} \quad (3)$$

Keterangan :

TP : Keterangan terima yang diprediksi secara benar sebagai Keterangan terima

FN : Keterangan terima yang diprediksi secara salah sebagai Keterangan tidak terima.

$$spesifisitas = \frac{TN}{FP+TN} \quad (4)$$

Keterangan :

TN : Keterangan tidak terima yang diprediksi secara benar sebagai Keterangan tidak terima

FP : Keterangan tidak terima yang diprediksi secara salah sebagai Keterangan terima.

2.3. Rencana Pengujian Algoritma C4.5

Uji ketepatan klasifikasi algoritma C4.5 dilakukan menggunakan confusion matrix. Pengukuran akurasi dengan confusion matriks dapat dilihat pada tabel 1.

Tabel 1. Tabel Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Formulasi perhitungan yang digunakan dapat ditunjukkan pada persamaan 5.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

Keterangan :

TP = True Positive

TN = True Negative

FP = False Positive

FN = False Negative

2.4. Analisis Masalah

Adanya data yang banyak menyebabkan para pengambil keputusan kesulitan dalam menentukan penerima beasiswa secara cepat dan tepat. Seperti yang disebutkan di latar belakang, untuk menjalankan proses penilaian dengan banyak data, maka diperlukan sebuah sistem yang mampu

memprediksi guna meningkatkan efektivitas pengambilan keputusan serta mengurangi subyektivitas dalam proses pengambilan keputusan. Selain itu dengan adanya sistem penentuan beasiswa akan lebih tepat sasaran untuk mahasiswa yang benar-benar membutuhkan dan dengan adanya sistem juga dapat mengurangi tingkat kesalahan penentuan kebijakan.

Proses seleksi beasiswa di UBG tidak selalu diputuskan berdasarkan perhitungan yang pasti tetapi kebijakan dari pembuat keputusan dalam menentukan penerima beasiswa. Contohnya pada tahun 2019, jumlah pendaftar beasiswa mencapai 110 orang. Dengan banyaknya pendaftar, menyebabkan seleksi beasiswa masih mengalami kendala pada proses pengambilan keputusan untuk menentukan mahasiswa yang berhak mendapatkan beasiswa.

Dalam penyelesaian permasalahan tersebut pada penelitian sebelumnya banyak yang menggunakan teknik *data mining*. Akan tetapi, pada penelitian ini digunakan algoritma C4.5 karena algoritma C4.5 memiliki tingkat akurasi dan pemodelan waktu yang lebih baik daripada algoritma lainnya dan memiliki tingkat kesalahan yang lebih rendah.

2.5. Pengumpulan Data

- 1) Studi Pustaka
Penulis mengambil data dari berbagai sumber, antara lain : artikel, jurnal, internet, yang berhubungan dengan data mining beasiswa.
- 2) Wawancara
Penulis mewawancarai Wakil Rektor III Universitas Bumigora untuk mendapatkan keterangan tentang beasiswa di Universitas Bumigora serta untuk mendapatkan data *training*.
- 3) Data dokumen pendaftar beasiswa tahun 2015 sampai dengan tahun 2019.

2.6. Data Preprocessing

Dataset yang bersifat mentah dan kurang berkualitas sering ditemukan, misalnya terdapat nilai yang hilang, salah input nilai, dan tidak konsisten. Akibatnya perlu dilakukan prapemrosesan data terlebih dahulu. Proses pembersihan mencakup menghilangkan duplikasi data, mengisi atau membuang data yang hilang, memperbaiki data yang tidak konsisten, dan memperbaiki kesalahan ketik. Meskipun data telah disempurnakan dan dibersihkan, kualitas data dapat ditingkatkan dengan menormalisir rentang data ataupun dengan menemukan representasi data yang lebih sedikit dari data aslinya.

id	nim	nama	prodi	jml_tanggungan	penghasilan_ortu	pekerjaan_ortu	jml_sks	ipk	sertifikat	keterangan
26	1710520189	Dwi Kurnianingsih	S1 Ilmu Komputer	3	18000000	Buruh	22	3.83		NULL
27	1710520011	Dedi Awa	S1 Ilmu Komputer	2	5000000	Buruh	24	3.28	Aktif	NULL
28	1710510025	Husnul Khotimah Ni Made	S1 Ilmu Komputer	2	9000000	Petani	24	3.69	Aktif	NULL
29	1710510032	Anggrena Yalesia Chandrawaty	S1 Ilmu Komputer	4	18000000	Wiraswasta	24	3.87	Aktif	NULL
30	1710530215	Firman Anshari Adjie	S1 Ilmu Komputer	3	15300000	Pegawai Swasta	24	3.66	Aktif	NULL

Gambar 2 *Missing value* pada dataset

Seperti pada gambar 2 terdapat *missing value* (ditandai garis merah). Penanganan *missing value* bisa melalui beberapa cara yaitu,

- 1) *Listwise Deletion*
Listwise deletion adalah cara yang paling umum untuk menangani *missing value*. Data yang terdapat *missing value* dihapus dari analisis. Kekurangan dari cara ini adalah semakin banyak data yang dihapus akan mengurangi akurasi dari prediksi.
- 2) *Mean Substitution*
Mean substitution merupakan cara penanganan *missing value* dengan mengganti *missing value* dengan rata – rata kolom tempat *missing value* tersebut berada.
- 3) *KNN Imputation*
K-Nearest Neighbor Imputation merupakan teknik penanganan *missing value* dengan mengganti *missing value* tersebut dengan nilai tetangga terdekat data tersebut.

2.7. Analisis Kebutuhan

2.7.1. Analisis Kebutuhan Fungsional

Aplikasi akan dibagi menjadi dua bagian yaitu aplikasi untuk *client* dan aplikasi untuk *administrator*. Aplikasi *client* diharapkan akan mempunyai dan menjalankan beberapa fitur berikut:

- 1) Sistem dapat menampilkan informasi mengenai beasiswa terkait seperti waktu pendaftaran dibuka, syarat – syarat beasiswa berupa dokumen yang harus dilengkapi.
- 2) User dapat menggunakan sistem sebagai wadah untuk pendaftaran beasiswa.
- 3) User dapat berlangganan aplikasi untuk mendapatkan informasi terbaru mengenai beasiswa melalui email mereka.

Sedangkan untuk aplikasi *administrator* diharapkan akan mempunyai dan menjalankan fitur-fitur seperti berikut.

- 1) Aplikasi dapat menampilkan, menambah, mengubah, dan menghapus dataset *training*.
- 2) Aplikasi dapat menampilkan, menambah, mengubah, dan menghapus informasi mengenai beasiswa.
- 3) Aplikasi dapat mengelola data mengenai fakultas dan program studi yang tersedia dikampus.

- 4) Aplikasi dapat memprediksi atau mengklasifikasi penerima beasiswa secara otomatis menggunakan algoritma c4.5.
- 5) Panitia dapat memverifikasi data – data yang diinput pendaftar melalui aplikasi.

2.7.2. Analisis Kebutuhan non Fungsional

Kebutuhan penelitian ini mencakup kebutuhan perangkat keras, kebutuhan perangkat lunak, dan pembagian hak akses untuk pengguna.

1) Kebutuhan perangkat keras

Perangkat keras yang dibutuhkan untuk penelitian ini adalah sebuah laptop untuk membuat aplikasi dan pengerjaan laporan dengan spesifikasi *Processor* AMD Ryzen 7 2700U, 8 GB RAM, AMD Radeon Vega 10 Mobile Graphic VGA, Sistem Operasi Windows 10 Home Single Language 64-bit.

2) Kebutuhan perangkat lunak

Perangkat lunak yang dibutuhkan dalam penelitian ini adalah Microsoft Office Word Professional Plus 2016, Software ini

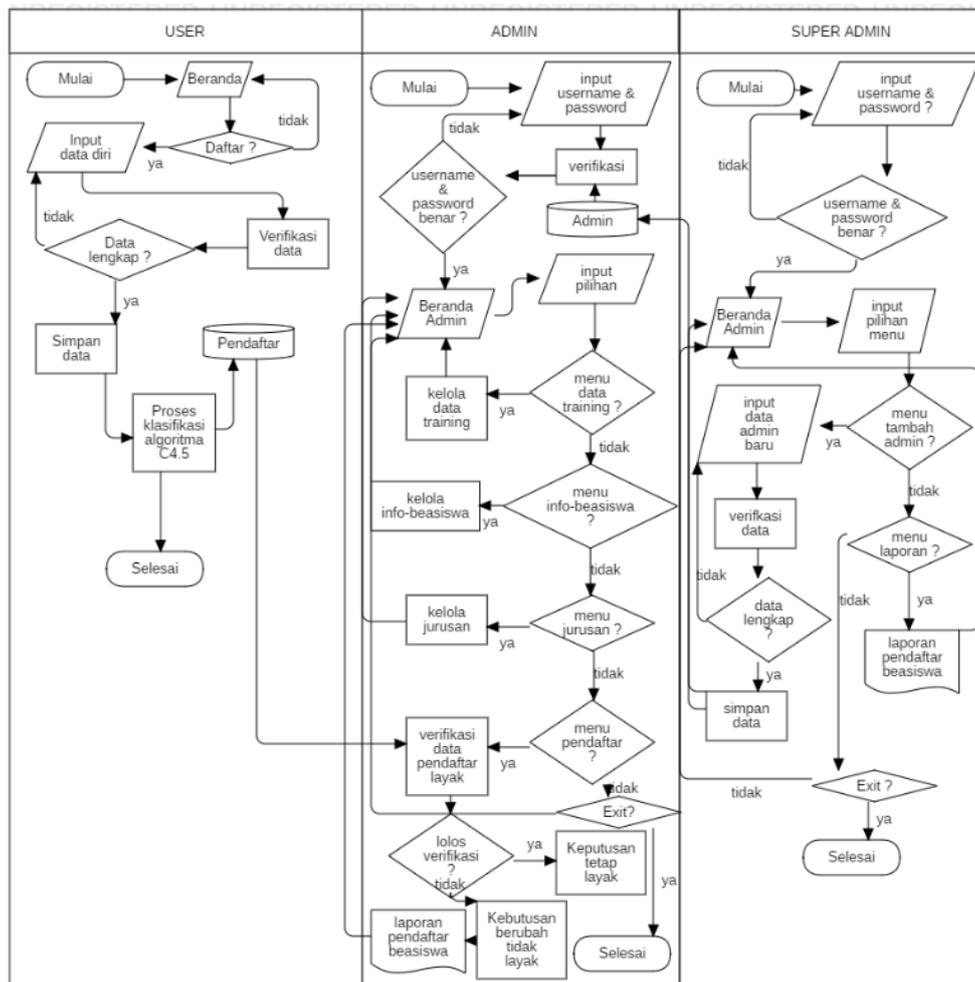
digunakan untuk mengolah laporan hasil penelitian. Microsoft Office Excel Professional Plus 2016

3) Hak akses pengguna

Dalam aplikasi ini pengguna dibagi menjadi tiga yaitu pertama user yaitu dapat melihat informasi beasiswa, dapat mendaftarkan dirinya ke beasiswa yang diinginkan, dapat melihat pengumuman hasil seleksi beasiswa. Kedua admin yaitu dapat mengelola *dataset training* beasiswa, dapat mengelola informasi beasiswa, Admin dapat mengelola data fakultas dan program studi. Ketiga pengawas yaitu dapat mengelola data admin, Super Admin dapat melakukan semua hal yang dilakukan admin.

2.8. Flowchart Sistem

User tidak perlu login untuk mengakses sistem. Pada saat mengakses sistem, user langsung diarahkan ke halaman beranda. Pada halaman beranda, user dapat memilih beasiswa yang ingin didaftarnya.



Gambar 3. Flowchart Sistem

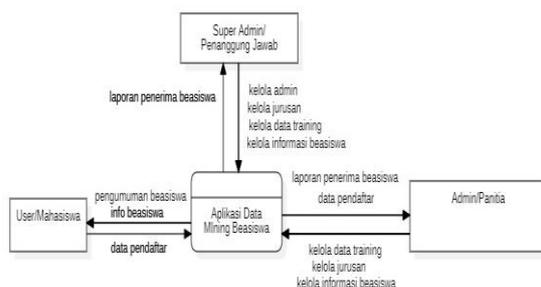
Ketika *user* mengklik tombol daftar di salah satu beasiswa, *user* akan diarahkan ke halaman pendaftaran. *User* akan diminta mengisi data pribadi, data sekolah, dan data keluarganya. Setelah itu sistem akan memvalidasi data-data tersebut seperti mengecek apakah semua data sudah diisi dan apakah sudah sesuai dengan *form* nya. Apabila lolos validasi, data-data pendaftar tersebut akan disimpan di *database* pendaftar yang kemudian akan dilakukan proses klasifikasi untuk menentukan apakah ia layak mendapat beasiswa atau tidak.

Berbeda dengan *user*, untuk mengakses aplikasi admin perlu login terlebih dahulu. Username dan password didapat melalui super admin yang mendaftarkan mereka ke sistem. Setelah *login*, admin akan dialihkan ke halaman beranda admin. Apabila admin memilih menu *data training*, maka admin dapat mengelola *data training* yang ada, sama seperti menu info-beasiswa dan jurusan. Untuk menu pendaftar, admin dapat melakukan verifikasi data yang diinputkan oleh pendaftar. Apabila data yang diinputkan sudah sesuai dengan data aslinya, maka keputusan sudah dipastikan layak, dan jika data yang diinput berbeda dengan data aslinya, maka keputusan yang semulanya layak akan berubah menjadi tidak layak.

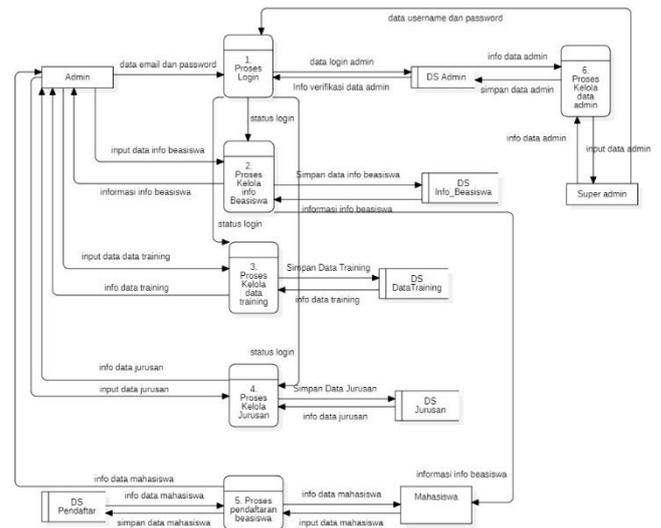
Sama seperti admin, super admin harus login terlebih dahulu untuk mengakses sistem. Peran utama seorang super admin adalah mengelola data-data admin dan melihat laporan penerima beasiswa.

2.9. Data Flow Diagram Sistem

Ada tiga entitas utama pada sistem yaitu *user/mahasiswa*, *admin/panitia*, dan *super admin/pengawas*. *User* mengirim data – data pendaftarannya ke sistem dan menerima informasi mengenai beasiswa dari system. Admin mengelola data informasi beasiswa, data data training dan data jurusan.



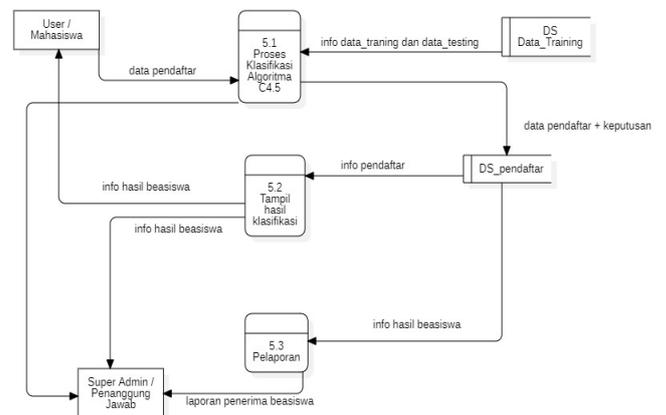
Gambar 4. DFD Level 0



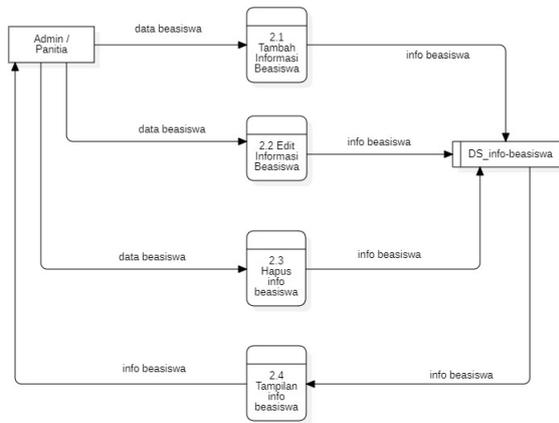
Gambar 5. DFD Level 1

Super admin melihat laporan penerima beasiswa dan mampu melakukan apa yang dapat dilakukan admin.

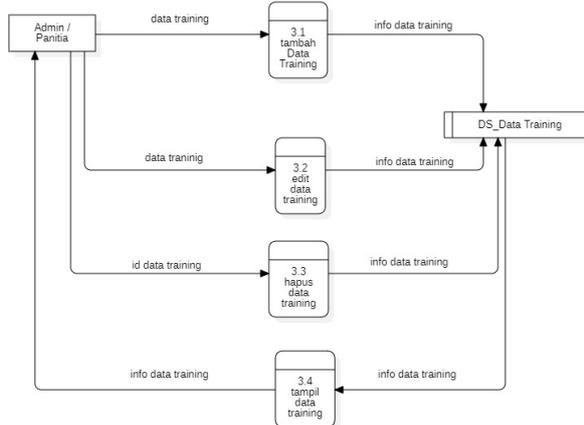
Admin dan super admin harus *login* terlebih dahulu agar dapat mengakses halaman admin. Sementara *user* tidak perlu login untuk mengakses halaman *user*. Admin dan super admin dapat mengelola informasi beasiswa yang kemudian informasi ini dikirim ke halaman *user*. Admin dan super admin juga dapat mengelola data training yang digunakan sebagai acuan dalam pengambilan keputusan. Admin dan super admin dapat mengelola jurusan. Super admin dapat mengelola admin – admin yang dapat login ke halaman admin. Untuk *user*, *user* dapat mendaftarkan dirinya untuk beasiswa.



Gambar 6. DFD Level 2 Proses Pendaftaran



Gambar 7. DFD Level 2 Proses Kelola Info Beasiswa

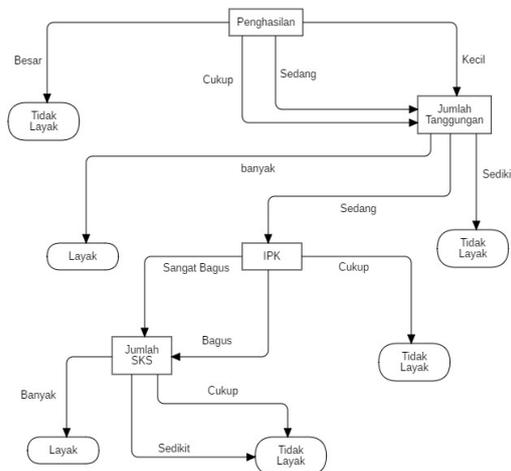


Gambar 8. DFD Level 2 Kelola Data Training

III. HASIL DAN PEMBAHASAN

3.1. Hasil Pemodelan Pohon Keputusan dan Rules

Berdasarkan dari algoritma yang telah dibuat dengan menggunakan data training telah dikumpulkan dan diolah, maka terbentuklah pohon keputusan untuk klasifikasi beasiswa PPA yang dapat dilihat pada gambar 9.



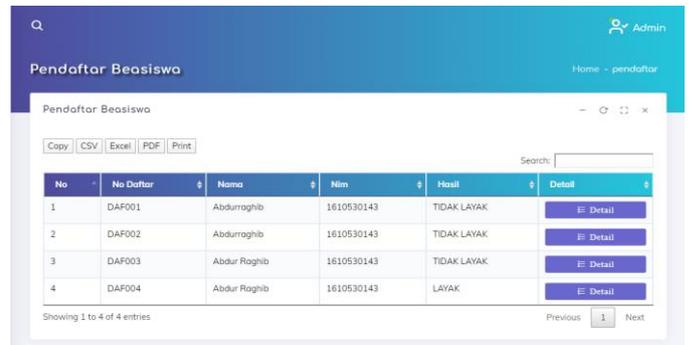
Gambar 9. Pohon Keputusan Hasil Pemodelan

Dengan melihat pohon keputusan pada gambar 15, dapat dibentuk beberapa rules :

Tabel 2. Rules Hasil Pemodelan

No	Rule	Hasil
1.	IF Penghasilan = Besar	Tidak Layak
2.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedikit	Tidak Layak
3.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedikit	Tidak Layak
4.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedikit	Tidak Layak
5.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Banyak	Layak
6.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Banyak	Layak
7.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Banyak	Layak
8.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Cukup	Tidak Layak
9.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Cukup	Tidak Layak
10.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Cukup	Tidak Layak
11.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Banyak	Layak
12.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Banyak	Layak
13.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Banyak	Layak
14.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Banyak	Layak
15.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Banyak	Layak
16.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Banyak	Layak
17.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Cukup	Tidak Layak
18.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Cukup	Tidak Layak
19.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Cukup	Tidak Layak

20.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Cukup	Tidak Layak
21.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Cukup	Tidak Layak
22.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Cukup	Tidak Layak
23.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Sedikit	Tidak Layak
24.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Sedikit	Tidak Layak
25.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Bagus AND Jumlah SKS = Sedikit	Tidak Layak
26.	IF Penghasilan = Kecil AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Sedikit	Tidak Layak
27.	IF Penghasilan = Sedang AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Sedikit	Tidak Layak
28.	IF Penghasilan = Cukup AND Jumlah Tanggungan = Sedang AND IPK = Sangat Bagus AND Jumlah SKS = Sedikit	Tidak Layak



Gambar 12. Form pendaftar beasiswa

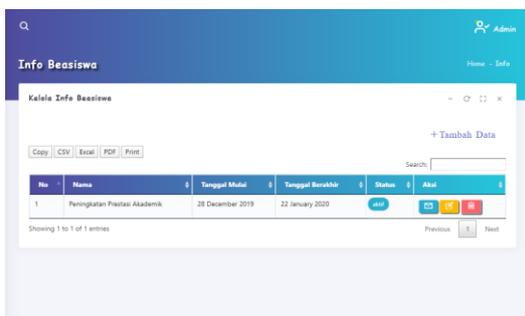


Gambar 13. Form rules dan akurasi

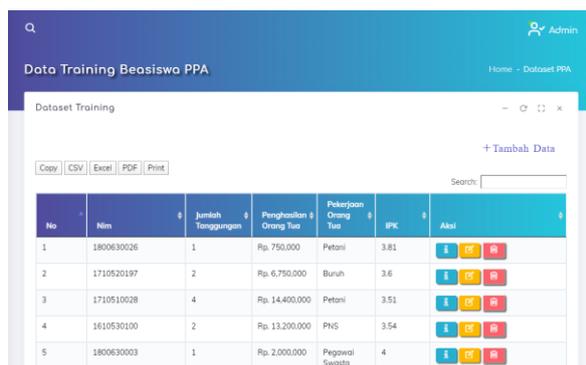


Gambar 14. Form utama user

3.2. Fitur – Fitur Aplikasi



Gambar 10. Form info beasiswa



Gambar 11. Form dataset training



Gambar 15. Form pendaftaran beasiswa

3.3. Evaluasi Hasil Pemodelan

Kesuksesan dari hasil pemodelan dapat dilihat dari besar akurasi dari klasifikasi algoritma C4.5 nya. Selain akurasi, untuk evaluasi ini juga dihitung spesifisitas dan sensitifitas nya. Untuk perhitungan akurasi menggunakan *confusion matrix*.

Dataset yang berjumlah 126 dibagi menjadi 80% untuk data training dan 20% untuk data testing. Untuk pengujian akurasi menggunakan 2 metode penanganan *missing value*. Yang pertama adalah *Listwise Deletion*. Untuk hasilnya dapat dilihat pada tabel 3.

Tabel 3. Hasil Evaluasi Pemodelan *Listwise Deletion*

		Predicted Class	
		Positive	Negative
Actual Class	Positive	11 (True Positive)	1 (False Negative)
	Negative	1 (False Positive)	12 (True Negative)

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + TN + FN + FP} \\
 &= \frac{11 + 12}{11 + 12 + 1 + 1} \\
 &= 0.92 \times 100\% = 92\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Sensitivitas} &= \frac{TP}{TP + FN} = \frac{11}{11 + 1} \\
 &= 0.916 \times 100\% = 91.6\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Spesifisitas} &= \frac{TN}{TN + FP} = \frac{12}{12 + 1} \\
 &= 0.923 \times 100\% = 92.3\%
 \end{aligned}$$

Untuk perhitungan akurasi dengan metode penanganan *missing value Mean Substitution* didapat hasil sebagai berikut :

Tabel 4. Hasil Evaluasi Pemodelan *Mean Substitution*

		Predicted Class	
		Positive	Negative
Actual Class	Positive	13 (True Positive)	0 (False Negative)
	Negative	3 (False Positive)	9 (True Negative)

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP + TN}{TP + TN + FN + FP} \\
 &= \frac{13 + 9}{13 + 9 + 0 + 3} = 0.88 \times 100\% \\
 &= 88\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Sensitivitas} &= \frac{TP}{TP + FN} = \frac{13}{13 + 0} = 1 \times 100\% \\
 &= 100\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Spesifisitas} &= \frac{TN}{TN + FP} = \frac{9}{9 + 3} \\
 &= 0.75 \times 100\% = 75\%
 \end{aligned}$$

IV. KESIMPULAN

4.1. Kesimpulan

Berdasarkan hasil perancangan, pengembangan, dan implementasi algoritma C4.5 dalam penentuan beasiswa, dapat disimpulkan bahwa akurasi yang didapat menggunakan cara penanganan *missing value Listwise Deletion* dalam klasifikasi penerima beasiswa mencapai 92%, spesifisitas mencapai 92.3%, dan sensitivitas mencapai 91.6%, sedangkan dengan cara penanganan *missing value Mean Substitution* akurasinya mencapai 88%, spesifisitas mencapai 75%, dan sensitivitas 100%. Kedua cara penanganan menggunakan persentase split 80% untuk data training yaitu sejumlah 100 data dan 20% data testing yaitu sejumlah 25 data. Berdasarkan hasil tersebut penanganan *Listwise Deletion* lebih tepat untuk klasifikasi penerima beasiswa. Algoritma C4.5 berhasil diimplementasikan dalam proses penentuan penerima beasiswa.

REFERENSI

- [1] Ristekdikti, "Pedoman Program Beasiswa PPA dan BBM," *Lembaga Layanan Pendidikan Tinggi Wilayah VI Jawa Tengah*, 2010. .
- [2] Friyadi, "Penerapan Metode Ahp Sebagai Pendukung Keputusan Penetapan Beasiswa," *J. Pilar Nusa Mandiri*, vol. 13(4), pp. 2527 – 6514, 2017.
- [3] A. Pujianto, K. Kusri, and A. Sunyoto, "Perancangan Sistem Pendukung Keputusan Untuk Prediksi Penerima Beasiswa Menggunakan Metode Neural Network Backpropagation," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 157, 2018.
- [4] Sumarlin, "Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM," vol. 01, pp. 52–62, 2015.
- [5] H. Saragih, M. Marbun, and B. Reza, "DEVELOPMENT OF DECISION SUPPORT SYSTEM DETERMINING THE STUDENT AS SCHOLARSHIP AWARDEES BY FUZZY MULTI ATTRIBUTE DECISION MAKING (FMADM)," pp. 90–95, 2007.
- [6] R. Rustiyan and M. Mustakim, "Penerapan Algoritma Fuzzy C Means untuk Analisis Permasalahan Simpanan Wajib Anggota Koperasi," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 2, p. 171, 2018.
- [7] N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, "A Comprehensive Survey of Data Mining Techniques on Time Series Data for

- Rainfall Prediction,” *J. ICT Res. Appl.*, vol. 11, no. 2, p. 168, 2017.
- [8] H. Amalia and E. Evicienna, “Komparasi Metode Data Mining Untuk Penentuan Proses Persalinan Ibu Melahirkan,” *J. Sist. Inf.*, vol. 13, no. 2, p. 103, 2017.
- [9] Y. I. Kurniawan, U. M. Surakarta, and N. Bayes, “Comparison of Naive Bayes and C . 45 Algorithm in Data Mining,” vol. 5, no. 4, pp. 455–464, 2018.
- [10] M. Mirqotussa’adah, M. A. Muslim, E. Sugiharti, B. Prasetyo, and S. Alimah, “Penerapan Dizcretization dan Teknik Bagging Untuk Meningkatkan Akurasi Klasifikasi Berbasis Ensemble pada Algoritma C4.5 dalam Mendiagnosa Diabetes,” *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 8, no. 2, p. 135, 2017.
- [11] D. B. Ananda and A. Wibisono, “C4.5 DECISION TREE IMPLEMENTATION IN SISTEM INFORMASI ZAKAT (SIZAKAT) TO AUTOMATICALLY DETERMINING THE AMOUNT OF ZAKAT RECEIVED BY MUSTAHIK MUSTAHIK,” *J. Inf. Syst.*, vol. 10, no. Vol 10 No 1 (2014): Jurnal Sistem Informasi (Journal of Information System), pp. 28–35, 2014.