



Klasifikasi Pemilihan Siswa untuk Rekomendasi Beasiswa Menggunakan Algoritma C4.5 di MA Tahfidh Annuqayah

Nilam Ramadhani^{1*}, Ach Muhyil Umam², Abd. Wahab Syahroni³, Sholeh Rachmatullah⁴, Wildona Zumam⁵

^{1,2,3,4} Program Studi Informatika, Fakultas Teknik, Universitas Madura, ⁵ Program Studi Pendidikan Bahasa Inggris, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Madura Jl. Raya Panglegur km 3.5 Kec. Tlanakan, Kab. Pamekasan, Jawa Timur 69371, Indonesia.

*Penulis Korespondensi, Email: nilam_ramadhani@unira.ac.id

Abstrak—Sebagai instansi sekolah yang berbeda dilingkungan pondok pesantren Annuqayah, MA Tahfidh Annuqayah mengedepankan Pendidikan Kitab dan Hafalan Al-qur’annya. MA Tahfidh Annuqayah ikut andil sebagai instansi sekolah yang merekomendasikan berbagai model beasiswa berupa bantuan biaya pendidikan maupun siswa yang berprestasi. Beasiswa yang di rekomendasikan ada dibawah naungan Kementerian Agama Republik Indonesia. Salah satunya adalah beasiswa full di Al-Azar Kairo Mesir. Banyaknya calon siswa yang akan diklasifikasi sebagai calon penerima rekomendasi beasiswa menjadikan proses seleksi memakan waktu yang lama, sehigga diperlukan system dalam rekomendasinya. Sistem yang dirancang ini akan menerapkan algoritma klasifikasi data mining C4.5 terhadap data yang telah ada pada proses seleksi dan rekomendasi penerima beasiswa. Algoritma C4.5 ini dipilih karena menghasilkan pohon keputusan berdasarkan nilai dari tiap-tiap atribut yang telah ditentukan berdasarkan persyaratan beasiswa. Sistem yang dirancang menggunakan 6 atribut yaitu nilai raport, prestasi, nilai ujian test, baca kitab, hafalan Al-Qur’an dan bahasa arab. Pemrosesan data training menghasilkan pohon keputusan klasifikasi yang bisa dipakai sebagai bahan rekomendasi. Evaluasi yang dilakukan dalam pengujian ini didapatkan nilai akurasi sebesar 97,5 %.

Kata Kunci: Algoritma C4.5; Beasiswa; Data Mining; Klasifikasi; Pohon Keputusan.

Abstract- As a different school institution within Annuqayah Islamic Boarding school, MA Tahfidh Annuqayah prioritizing Book Education and memorize the Al-Qur’an. MA Tahfidh Annuqayah takes part as a school institution that recommends various scholarship models in the form of tuition assistance and outstanding students. The recommended scholarships are under the auspices in the Ministry of Religious Affairs Republic Indonesia. One of them is a full scholarship at Al-Azar Cairo Egypt. The number of prospective students who will be classified as candidates for scholarship recommendations makes the selection process take a long time, so a system is needed in the recommendation. This designed system will apply the C4.5 data mining classification algorithm to existing data in the selection and recommendation process for scholarship recipients. The C4.5 algorithm was chosen because it produces a decision tree based on the value of each attribute that has been determined based on the scholarship requirements. The system is designed using 6 attributes, namely report card scores, achievements, test scores, reading books, memorize the Al-Qur’an and Arabic language. Processing training data produces a classification decision tree that can be used as a recommendation material. The evaluation carried out in this test obtained an accuracy value of 97,5 %.

Keywords: C4.5 Algorithm; Scholarship; Data Mining; Classification; Decision Tree.

1. PENDAHULUAN

MA Tahfidh Annuqayah adalah lembaga pendidikan tingkat SLTA sederajat yang berada dibawah naungan Yayasan Annuqayah Pondok Pesantren Annuqayah Guluk-guluk Sumenep Jawa Timur. MA Tahfidh Annuqayah juga melakukan program mendapatkan beasiswa yang orientasinya ke luar negeri. Beasiswa sebagai bentuk apresiasi yang diberikan dengan tujuan agar proses pembelajaran dapat dilanjutkan pada jenjang yang lebih tinggi[1]. Penerima tidak perlu lagi memikirkan tagihan biaya pendidikan. Mereka akan lebih fokus untuk belajar serta berusaha mendapatkan nilai dan prestasi yang tinggi. MA Tahfidh Annuqayah menawarkan beasiswa dari Kemenag dan beasiswa dari NU.

Dari sinilah MA Tahfidh Annuqayah mengambil kesempatan untuk para siswanya yang berprestasi agar bisa mendapatkan beasiswa yang terbaik, sehingga sekolah mempersiapkan setiap tahunnya rekomendasi siswa yang di ambil dari data siswa kelas akhir. Proses seleksi dimulai dari data hasil akademik siswa selama menempuh proses kegiatan belajar mengajar dan berbagai data nilai lainnya. Data siswa tersebut dimanfaatkan dan diolah kembali untuk menjadi sebuah pengetahuan dan informasi yang bermanfaat sebagai bahan pertimbangan dalam memprediksi rekomendasi dari klasifikasi pemilihan siswa[2]. Karena berdasarkan

banyaknya data siswa, perlu dilakukan analisa untuk mengetahui informasi penting berupa pengetahuan baru (Knowledge Discovery)[3].

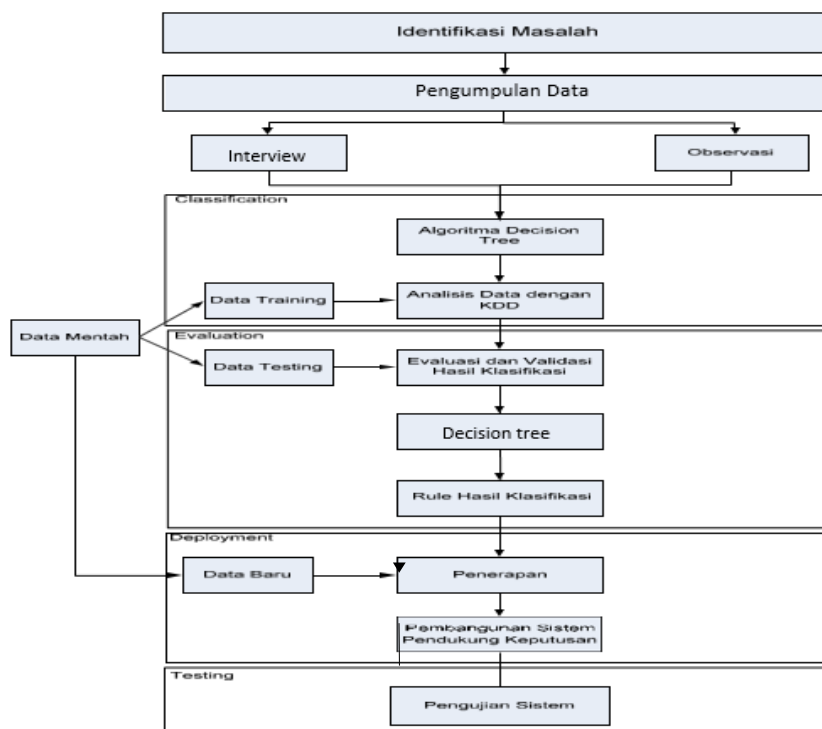
Proses rekomendasi siswa sejauh ini masih dilakukan secara manual yaitu dengan menginput satu persatu data siswa ke dalam satu file kemudian dilakukan sorting data. Hal ini seringkali menimbulkan beberapa permasalahan, antara lain membutuhkan waktu yang lama dan ketelitian yang tinggi. Prosedur pengolahan data yang dilakukan meliputi kegiatan pengumpulan data, pengelompokan, pencocokan data dengan biodata siswa, perkiraan siswa penerima, dan menyusun laporan. Sehingga pemberian beasiswa dilakukan cukup lama serta hasilnya tidak tepat sasaran. Ada siswa yang memang prestasi namun tidak diberikan rekomendasi beasiswa oleh pihak sekolah. Selain itu, transparansi dan ketidakjelasan metodologi yang digunakan dalam proses komputasi rekomendasi beasiswa serta kurang tepatnya penyaluran beasiswa juga menjadi salah satu permasalahan. Oleh karena itu dibutuhkan suatu sistem yang dapat membantu dalam proses pengambilan keputusan rekomendasi penerima beasiswa berdasarkan kriteria yang ditentukan secara cepat dan tepat sasaran.

Salah satu teknologi yang dapat digunakan adalah data mining. Karena konsep yang digunakan untuk menemukan pengetahuan atau informasi dan sebagai proses untuk memecahkan masalah dengan menganalisis data yang telah ada dalam database[4]. Bentuk dasar dan teknik dari analisis data yang sering digunakan dalam data mining adalah klasifikasi. Terdapat beberapa model klasifikasi, diantaranya yaitu Decision Tree, Algoritma C4.5, Statistical Analysis, Naive Bayes, K-Nearest Neighbor, Support Vector Machine[5].

Penerapan algoritma C4.5 dalam melakukan prediksi sudah memiliki akurasi yang cukup baik dalam melakukan penyeleksian terhadap data dengan prosedur yang terstruktur dengan baik melalui beberapa tahapan diantaranya proses klasifikasi dengan data mining[6]. Algoritma C4.5 banyak digunakan dalam klasifikasi karena mudah diinterpretasikan, cepat dan memiliki akurasi serta presisi tinggi. Dengan algoritma C4.5, akan didapatkan sebuah pohon keputusan yang mudah dipahami dan mudah dimengerti[7].

2. METODOLOGI PENELITIAN

Dalam penelitian ini, metodologi yang dilakukan seperti yang disajikan pada gambar 1[8].



Gambar 1. Kerangka Kerja Penelitian

2.1 Tahapan Penelitian

Pada penelitian ini dibutuhkan beberapa tahapan yang perlu dilakukan. Tahapan tersebut antara lain sebagai berikut[9][10].

2.1.1. Pengumpulan data

Merupakan proses mengumpulkan informasi atau fakta-fakta yang relevan dari berbagai sumber untuk tujuan analisis, penelitian, atau pengambilan keputusan. Ini melibatkan pengumpulan, pencatatan, dan penyimpanan data dalam bentuk yang dapat digunakan dan diolah. Metode pengumpulan data yang digunakan dalam aplikasi klasifikasi pemilihan siswa menggunakan dua cara yaitu observasi dan wawancara. Dimana proses pengumpulan data ini dilakukan untuk memperoleh suatu informasi yang dibutuhkan dalam rangka mencapai tujuan penelitian.

a. Teknik Penelitian Langsung (Observasi)

Data yang diperlukan untuk proses klasifikasi siswa yang berprestasi yaitu berupa penilaian terhadap siswa tersebut. Diantara penilaiannya disajikan seperti pada tabel 1.

Tabel 1. Data Atribut Penilaian

No	Atribut	Jenis Data
1	Nilai Raport	Integer
2	Nilai Test Ujian	Integer
3	Prestasi	Integer
4	Baca kitab kuning	Integer
5	Hafalan alquran	Integer
6	Bahasa arab	Integer
7	Keterangan (Rekomendasi/Tidak Rekomendasi)	Binominal

b. Teknik Wawancara (Interview)

Wawancara merupakan teknik pengumpulan data yang dilakukan melalui tatap muka dan tanya jawab langsung antara peneliti terhadap nara sumber atau sumber data. Penulis dalam hal ini telah melakukan wawancara kepada kepala pimpinan MA Tahfidh Annuqayah. Adapun hasil wawancara yang berkaitan dengan penelitian ini adalah sebagai berikut:

Sistem Informasi yang masih belum berjalan dengan baik berupa website resmi instansi sekolah di MA Tahfidh Annuqayah, dalam mempermudah menginformasikan berbagai informasi yang ada di instansi/sekolah kepada siswa atau publik. Sebelumnya dalam pemilihan kualifikasi beasiswa siswa masih dipilih secara manual karena memang belum ada sistem untuk menentukan siswa sebagai kualifikasi beasiswa tahunan. Faktor pendukung dalam klasifikasi pemilihan siswa yaitu dimulai dari data raport siswa, prestasi, nilai test, baca kitab, hafalan alquran dan fasih dalam berbahasa arab.

2.1.2 Analisis Data

Analisis data yang digunakan untuk proses klasifikasi menggunakan Knowledge Discovery in Databases (KDD) yang terdiri dari sembilan langkah. Dimulai dari tahap pemahaman data yang akan digunakan hingga tahap terciptanya sebuah pengetahuan tentang klasifikasi beasiswa dan penggunaannya. KDD sendiri diartikan sebagai proses terorganisir untuk mengidentifikasi pola dalam data yang besar dan kompleks dimana pola data tersebut ditemukan yang bersifat sah, baru, dan dapat bermanfaat serta dapat dimengerti[11].

Berikut adalah beberapa tahapan yang dilakukan dalam menganalisis sebuah data:

a. Developing and understanding of the application domain

Pada tahap ini dilakukan pengembangan pemahaman domain aplikasi untuk persiapan langkah awal untuk memahami apa yang harus dilakukan dengan banyak keputusan (tentang transformasi, algoritma, representasi, dan lain-lain).

b. Seleksi Data (Selection)

Memilih dan menciptakan satu set data yang akan digunakan untuk penelitian harus ditentukan. Langkah ini antara lain mencari tahu data apa yang tersedia, memperoleh data tambahan yang diperlukan, dan mengintegrasikan semua data untuk penemuan pengetahuan ke dalam satu set data, termasuk atribut yang akan dipakai untuk proses tersebut. Data yang digunakan yaitu data siswa kelas akhir (XII) tahun pelajaran 2022-2023 sejumlah 325 record.

c. Data Cleaning

Pada tahap ini kehandalan data ditingkatkan dengan membersihkan data yang tidak lengkap (missing value) dan data tidak benar (noise). Data siswa yang ada diseleksi terdapat 10 data yang nilainya kosong (missing value) dan 15 data yang tidak benar (noise). Sehingga data yang akan dipakai untuk klasifikasi sejumlah 300 record.

d. Data Transformasi

Pada proses penerapan klasifikasi, terdapat 6 atribut yang akan ditransformasi. Diantaranya adalah data raport siswa, nilai test siswa, data prestasi siswa, baca kitab, hafalan alquran dan nilai bahasa arab siswa. Hasil transformasi data disajikan pada tabel 2.

Tabel 2. Data Tranformasi Tiap Atribut			
No	Atribut	Grade	Range
1	Nilai Raport	Excellent	80-100
		Good	65-79.99
		Pass	55-64.99
		Weak	40-54.99
2	Pretasi	Memuaskan	>=3
		Baik	2
		Cukup	1
		Tidak Ada	0
3	Nilai Test Ujian	A	91-100
		B	75-90
		C	65-74
		D	55-64
4	Baca Kitab	E	0-54
		Fasih	75-100
		Lancar	65-74
		Kurang Lancar	51-64
5	Hafalan Alquran	Tidak Bisa	0-50
		Istimewa	>=3
		Sempurna	2
		Kurang Sempurna	1
6	Bahasa Arab	Tidak Hafal	0
		C2	85-100
		C1	71-84
		B2	65-70
		B1	61-65
		A2	51-60
		A1	0-50

e. Data Mining

Setelah pemilihan jenis data mining yang akan digunakan yaitu klasifikasi, maka selanjutnya menentukan algoritma klasifikasi yang akan digunakan. Pada penelitian ini algoritma yang dipilih untuk komparasi adalah C4.5.

f. Interpretasi/Evaluation

Dalam tahap ini dilakukan evaluasi dan menafsirkan pola yang didapatkan dari hasil algoritma yang dipakai untuk mengetahui aturan, kehandalan, dan lain-lain. Evaluasi dilakukan dengan menerapkan pola yang didapat dari proses sebelumnya terhadap data testing yang disediakan. Evaluasi dilakukan dengan confusion matrix. Kemudian pada tahap interpretasi menggunakan pengetahuan yang diperoleh dari proses data mining untuk penerapan pada pembuatan program atau aplikasi. Pengetahuan klasifikasi siswa

penerima beasiswa diterapkan pada data baru untuk membuat klasifikasi siswa penerima beasiswa dengan pohon keputusan dan rule klasifikasi.

3. HASIL DAN PEMBAHASAN

Pada tabel 3 disajikan cuplikan data traning yang digunakan dalam klasifikasi pemilihan siswa untuk mendapatkan beasiswa dengan menggunakan algoritma C.45.

Tabel 3. Cuplikan data training yang digunakan

NIS	Nilai Raport	Prestasi	Nilai Ujian	Baca Kitab	Hafalan Al-quran	Bahasa Arab	Keterangan
1	80	tidakada	50	80	kurangsempurna	50	Tidakrekomendasi
2	80	memuaskan	60	50	kurangsempurna	50	Tidakrekomendasi
3	80	memuaskan	75	75	istimewa	80	Rekomendasi
4	60	baik	60	60	tidakhafal	50	Tidakrekomendasi
5	80	memuaskan	80	75	sempurna	80	Rekomendasi
6	50	tidakada	50	50	tidakhafal	50	Tidakrekomendasi
7	50	memuaskan	80	80	istimewa	75	Rekomendasi
8	50	baik	60	50	sempurna	50	Tidakrekomendasi
9	50	tidakada	60	50	kurangsempurna	80	Tidakrekomendasi
10	80	baik	80	80	istimewa	80	Rekomendasi
11	80	baik	80	75	istimewa	80	Rekomendasi
12	80	memuaskan	60	80	kurangsempurna	50	Tidakrekomendasi
13	60	tidakada	50	50	tidakhafal	50	Tidakrekomendasi
14	50	tidakada	75	50	sempurna	50	Tidakrekomendasi
15	50	tidakada	75	50	kurangsempurna	60	Tidakrekomendasi

Hasil dari keseluruhan record data training yang didapatkan,disajikan pada tabel 4.

Tabel 4. Jumlah Data Total dan Target Atribut

	Jumlah	Rekomendasi	Tidak Rekomendasi
Total	300	124	176

Pada tabel 4 merupakan hasil jumlah total data traning sehingga ditemukan jumlah record rekomendasi dan yang tidak rekomendasi.

3.1 Mencari Nilai Entropy

Dalam menghitung nilai Entropy[12] dari total data sebanyak 300 record, diketahui status rekomendasi ada 124 record dan status tidak rekomendasi ada 176 record. Sehingga Entropy nilai akar total yang di dapat melalui perhitungan sesuai persamaan (1). Hasil dari perhitungan tersebut disajikan di tabel 5.

$$\begin{aligned}
 Entropy(S) &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= (- \frac{124}{300} \cdot \log_2 (\frac{124}{300})) + (- \frac{176}{300} \cdot \log_2 (\frac{176}{300})) \\
 &= 0.978217666
 \end{aligned}
 \tag{1}$$

Tabel 5. Hasil Data Entropy Total

	Jumlah	Rekomendasi	Tidak Rekomendasi	Entropy
Total	300	124	176	0.978217666

Kemudian setelah jumlah total dan entropy total sudah ditemukan, selanjutnya mencari nilai entropy pada atribut nilai raport, dimana total record pada atribut nilai raport seperti pada tabel 6.

Tabel 6. Total Atribut Nilai Raport

		Jumlah	Rekomendasi	Tidak Rekomendasi	Entropy
Total		300	124	176	0.908178347
	Excellent	102	69	33	
	Good	66	47	19	
	Pass	63	4	59	
	Weak	69	4	65	

Pada proses mencari nilai entropy terhadap atribut nilai raport dengan kriteria excellent berjumlah 102 record, rekomendasi 69 record dan tidak rekomendasi berjumlah 33 record, sehingga diperoleh perhitungan berikut:

$$\begin{aligned}
 Entropy(S) &= \sum_{i=1}^n - p_i * \log_2 p_i \\
 &= \left(- \frac{69}{102} \cdot \log_2 \left(\frac{69}{102} \right) \right) + \left(- \frac{33}{102} \cdot \log_2 \left(\frac{33}{102} \right) \right) \\
 &= \mathbf{0.90817834}
 \end{aligned}$$

Proses perhitungan diatas di ulang sampai selesai sesuai jumlah data berikutnya, sehingga diperoleh hasil data entropy seperti pada tabel 7.

Tabel 7. Data Entropy Atribut Nilai Raport

		Jumlah	Rekomendasi	Tidak Rekomendasi	Entropy
Total		300	124	176	0.978217666
	Excellent	102	69	33	0.908178347
	Good	66	47	19	0.86596535
	Pass	63	4	59	0.341153893
	Weak	69	4	65	0.319337387

3.2 Mencari Nilai Gain

Dalam mencari nilai Gain pada atribut nilai raport yakni dengan menggunakan rumus sebagai berikut[7]:

$$\begin{aligned}
 Gain(S,A) &= Entropy(S) - \sum_{i=1}^n \frac{|s_i|}{|S|} * Entropy(S_i) \\
 &= (0.978217666) - \left(\left(\frac{102}{300} \right) \cdot 0.908178347 \right) + \left(\left(\frac{66}{300} \right) \cdot 0.86596535 \right) \\
 &\quad + \left(\left(\frac{63}{300} \right) \cdot 0.341153893 \right) + \left(\left(\frac{69}{300} \right) \cdot 0.319337387 \right) \\
 &= \mathbf{0.333834734}
 \end{aligned} \tag{2}$$

Dari perhitungan di atas dapat diperoleh nilai gain dari atribut nilai raport yang disajikan pada tabel 8.

Tabel 8. Data Gain Atribut Nilai Raport

		Jumlah	Rekomendasi	Tidak Rekomendasi	Entropy	Gain
Total		300	124	176	0.978217666	
	Excellent	102	69	33	0.908178347	
	Good	66	47	19	0.86596535	0.333834734
	Pass	63	4	59	0.341153893	
	Weak	69	4	65	0.319337387	

3.3 Split Information

Selanjutnya menghitung SplitInfo pada atribut Nilai raport berdasarkan jumlah kasus. Perhitungan split information menggunakan rumus (3)[5].

$$\begin{aligned}
 SplitInfo(S,A) &= \sum_{i=1}^n - \frac{S_i}{s} \log_2 \frac{S_i}{s} \\
 &= - \left(\frac{102}{300} \cdot \log_2 \left(\frac{102}{300} \right) \right) \\
 &\quad - \left(\frac{66}{300} \cdot \log_2 \left(\frac{66}{300} \right) \right) \\
 &\quad - \left(\frac{63}{300} \cdot \log_2 \left(\frac{63}{300} \right) \right) \\
 &\quad - \left(\frac{69}{300} \cdot \log_2 \left(\frac{69}{300} \right) \right) \\
 &= 1.970237959
 \end{aligned} \tag{3}$$

Dari hasil perhitungan diperoleh nilai SplitInformasi pada atribut nilai raport yang disajikan pada tabel 9.

Tabel 9. Data SplitInfo Atribut Nilai Raport

Jumlah	Rekomendasi	Tidak Rekomendasi	Entropy	Gain	Split Info
300	124	176	0.978217666		
102	69	33	0.908178347		
66	47	19	0.86596535	0.333834734	1.970237959
63	4	59	0.341153893		
69	4	65	0.319337387		

3.4 Gain Ratio

Sebelum mencari gain ratio[13], terlebih dahulu mengulang perhitungan pada semua atribut sampai selesai. Kemudian menghitung nilai gain ratio pada Nilai Raport sebagai nilai yang digunakan sebagai pemecahan node pada pohon keputusan. Persamaan yang digunakan untuk menghitung gain ratio seperti pada rumus (4).

$$\begin{aligned}
 Gain\ Ratio &= \frac{Gain(S,A)}{SplitInfo(S,A)} \\
 &= \frac{1.970237959}{0.169438789} \\
 &= 0.169438789
 \end{aligned} \tag{4}$$

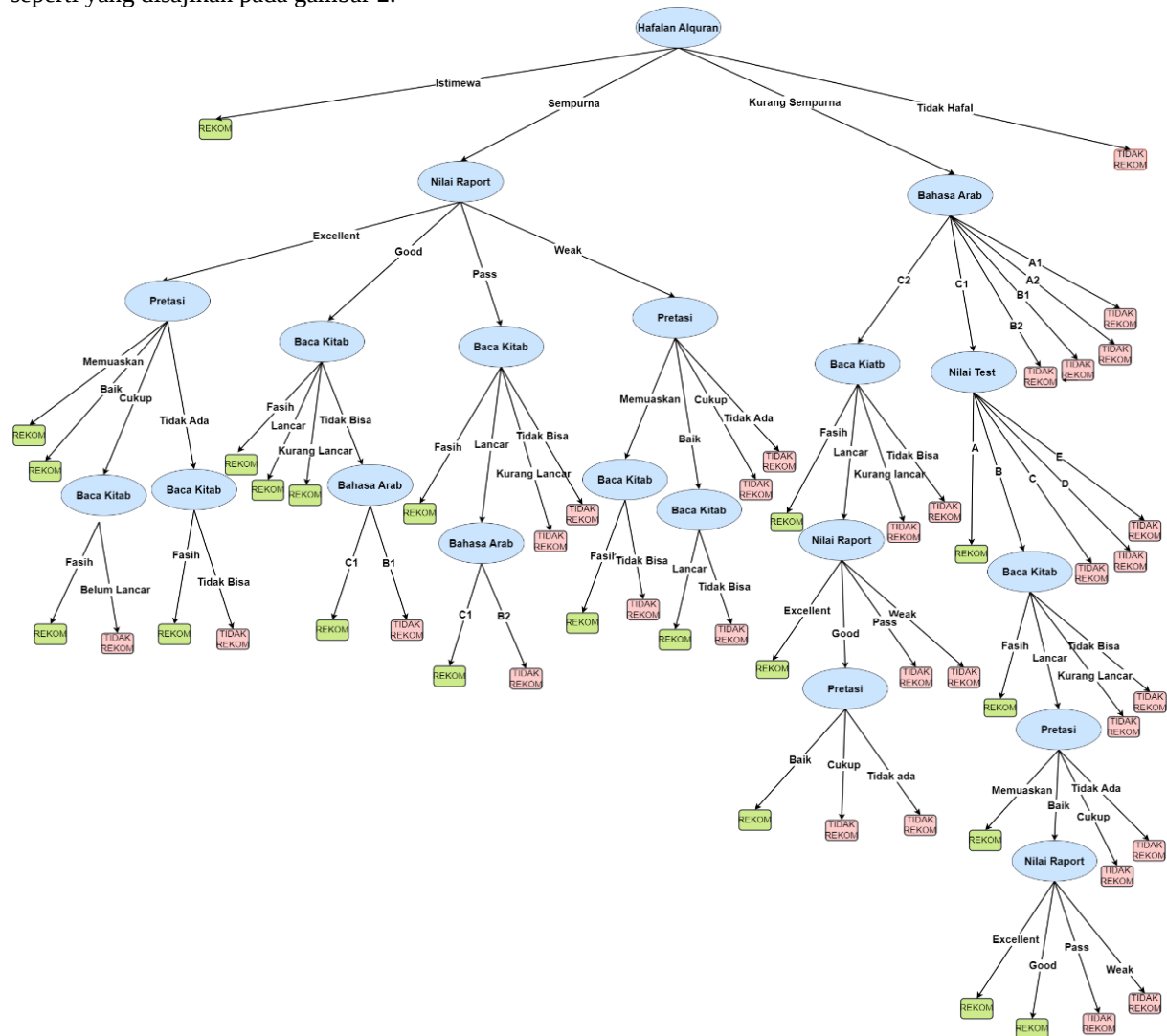
Dari hasil perhitungan diperoleh nilai SplitInformasi pada atribut nilai raport yang disajikan pada tabel 10.

Tabel 10. Data Gain Ratio Atribut Nilai Raport

	Entropy	Gain	Split Info	Gain Ratio
Total	0.978217666			
Nilai Raport	0.908178347	0.333834734	1.970237959	0.169438789
	0.86596535			
	0.341153893			
	0.319337387			

3.5 Pohon Keputusan

Dari hasil pemecahan node, dapat dibentuk pohon keputusan[14] dengan data rule klasifikasi sebanyak 15 aturan seperti yang disajikan pada gambar 2.



Gambar 2. Hasil Generate Pohon Keputusan

Berikut hasil dari aturan yang didapat :

Rule Aturan

- R1 IF (hafalan='istimewa') THEN Label = Rekomendasi
- R2 IF (hafalan='sempurna') AND (raport<=60) AND (kitab<=60) THEN Label = Tidakrekomendasi
- R3 IF (hafalan='sempurna') AND (raport<=60) AND (kitab>60) AND (prestasi='baik') THEN Label = Rekomendasi
- R4 IF (hafalan='sempurna') AND (raport<=60) AND (kitab>60) AND (prestasi='memuaskan' OR prestasi='tidakada') AND (prestasi='tidakada') THEN Label = Tidakrekomendasi
- R5 IF (hafalan='sempurna') AND (raport<=60) AND (kitab>60) AND (prestasi='memuaskan' OR prestasi='tidakada') AND (prestasi='memuaskan') THEN Label = Rekomendasi
- R6 IF (hafalan='sempurna') AND (raport>60) AND (prestasi='tidakada') AND (kitab<=60) THEN Label = Tidakrekomendasi
- R7 IF (hafalan='sempurna') AND (raport>60) AND (prestasi='tidakada') AND (kitab>60) THEN Label = Rekomendasi
- R8 IF (hafalan='sempurna') AND (raport>60) AND (prestasi='memuaskan' OR prestasi='baik') THEN Label = Rekomendasi
- R9 IF (hafalan='kurangsempurna') AND (bahasa<=60) THEN Label = Tidakrekomendasi
- R10 IF (hafalan='kurangsempurna') AND (bahasa>60) AND (ujian<=60) THEN Label = Tidakrekomendasi

- R11 IF (hafalan='kurangsempurna') AND (bahasa>60) AND (ujian>60) AND (prestasi='tidakada') THEN Label = Tidakrekomendasi
- R12 IF (hafalan='kurangsempurna') AND (bahasa>60) AND (ujian>60) AND (prestasi='memuaskan' OR prestasi='baik') AND (raport<=60) THEN Label = Tidakrekomendasi
- R13 IF (hafalan='kurangsempurna') AND (bahasa>60) AND (ujian>60) AND (prestasi='memuaskan' OR prestasi='baik') AND (raport<=60) THEN Label = Tidakrekomendasi
- R14 IF (hafalan='kurangsempurna') AND (bahasa>60) AND (ujian>60) AND (prestasi='memuaskan' OR prestasi='baik') AND (raport>60) AND (kitab>60) THEN Label = Rekomendasi
- R15 IF (hafalan='tidakhafal') THEN Label = Tidakrekomendasi

3.6 Hasil Pengujian

Dari proses mining dengan 15 data rule terhadap data traning,data testing yang digunakan disajikan pada tabel 11.

Tabel 11. Cuplikan Data Uji yang Digunakan

NIS	Nilai Raport	Nilai Prestasi	Nilai Ujian	Baca Kitab	Hafalan Al-quran	Bahasa Arab	Kelas Asli
1	80	memuaskan	50	80	sempurna	80	Rekomendasi
2	80	memuaskan	50	50	kurangsempurna	50	Tidakrekomendasi
3	80	baik	80	80	istimewa	75	Rekomendasi
4	80	memuaskan	60	80	sempurna	80	Rekomendasi
5	80	memuaskan	80	80	sempurna	75	Rekomendasi
6	60	tidakada	60	60	tidakhafal	50	Tidakrekomendasi
7	75	baik	75	75	kurangsempurna	50	Tidakrekomendasi
8	80	memuaskan	80	60	sempurna	80	Rekomendasi
9	75	tidakada	75	75	kurangsempurna	75	Tidakrekomendasi
10	75	baik	75	60	sempurna	75	Rekomendasi
11	60	tidakada	75	60	kurangsempurna	60	Tidakrekomendasi
12	50	tidakada	50	50	sempurna	50	Tidakrekomendasi
13	80	baik	75	75	sempurna	60	Rekomendasi
14	50	tidakada	50	75	tidakhafal	50	Tidakrekomendasi
15	50	tidakada	60	50	kurangsempurna	50	Tidakrekomendasi

Pada proses ini mencocokkan data asli dari data uji dengan data hasil dari proses rule klasifikasi. Dari data uji sebanyak 40 record dapat diperoleh hasil pengujian dengan akurasi persentase seperti pada tabel 12.

Tabel 12. Data Akurasi Hasil Klaifikasi Sistem

No	Hasil Akurasi Klasifikasi	Jumlah/Perentase
1	Data Uji klasifikasi	40
2	Data Tepat	39
3	Data Tidak Tepat	1
4	Akurasi	97.5%
5	Laju Error	2.5%

Pada tabel 12 merupakan hasil pengujian sistem yang menunjukkan hasil akurasi 97.5% dan laju Error 2.5% dengan banyak data uji 40 record. 39 data tepat dan 1 data tidak tepat. Untuk mengetahui tingkat akurasi digunakan rumus 5.

$$\text{Rumus} = \frac{\text{Jumlah tepat/tidak}}{\text{Jumlah Uji}} * 100 \% \quad (5)$$

Maka diperoleh tingkat akurasi data dengan data tepat sebanyak 39 dan jumlah data 40 :

$$\begin{aligned}
 &= \frac{39}{40} * 100 \% \\
 &= 0,975 * 100 \% \\
 &= 97.5 \%
 \end{aligned}$$

Kemudian juga diperoleh nilai tingkat laju error data dengan data tidak tepat sebanyak 1 dan total jumlah data 40 maka diperoleh:

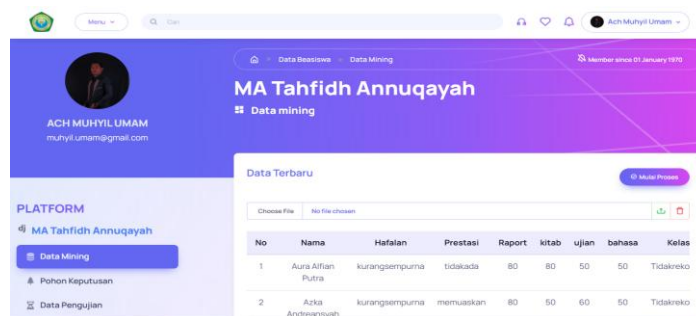
$$\begin{aligned}
 \text{Laju Error} &= \frac{\text{Jumlah tepat}}{\text{Jumlah Uji}} * 100 \% \\
 &= \frac{1}{40} * 100 \% \\
 &= 0,025 * 100 \% \\
 &= 2.5 \%
 \end{aligned} \tag{6}$$

Maka hasil dari data pengujian sistem klasifikasi diperoleh tingkat akurasi 97.5% dan laju eror 2.5%.

3.7 Tampilan System Interface

Implementasi antarmuka ini merupakan tahap dalam proses pengembangan sistem sebagai langkah penting untuk memastikan bahwa perangkat lunak atau sistem yang dikembangkan dapat digunakan dengan efisien dan sesuai dengan kebutuhan pengguna[15].

3.7.1 Halaman awal



Gambar 3. Tampilan halaman awal sistem

Pada halaman System Mining ini yakni dapat memanajemen data traning yang akan di gunakan sebagai data proses mining, dimana pada halaman ini pengguna bisa melakukan CRUD data dan proses mining.

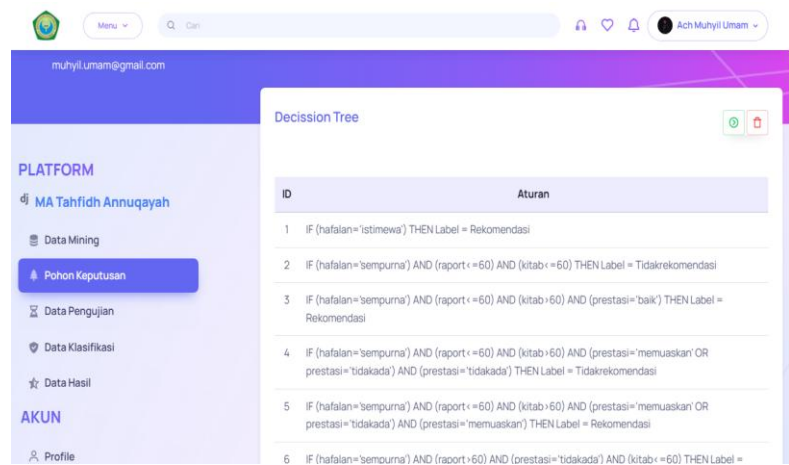
3.7.2. Hasil Proses Mining

Nilai Atribut	Jumlah data	Jumlah Rekomendasi	Jumlah Tidak rekomendasi	Entropy	Gain
Total/Terpilih	300	124	176	0.978	
hafalan="istimewa"	54	54	0	0	
hafalan="sempurna"	107	60	47	0.989	0.321
hafalan="kurang sempurna"	93	10	83	0.492	

Gambar 4. Tampilan Hasil Proses Mining

Pada halaman hasil proses mining ini menampilkan proses data traning. Pada halaman ini dapat mengetahui jalannya data, jumlah data, entropy dan gain.

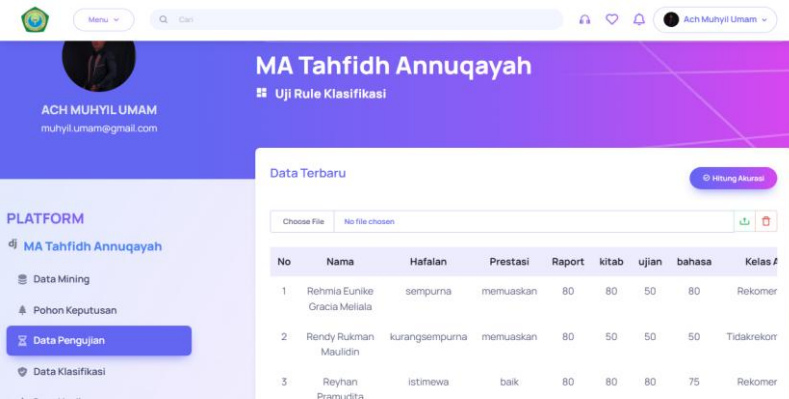
3.7.3. Halaman Rule Klasifikasi



Gambar 5. Tampilan Halaman Rule klasifikasi

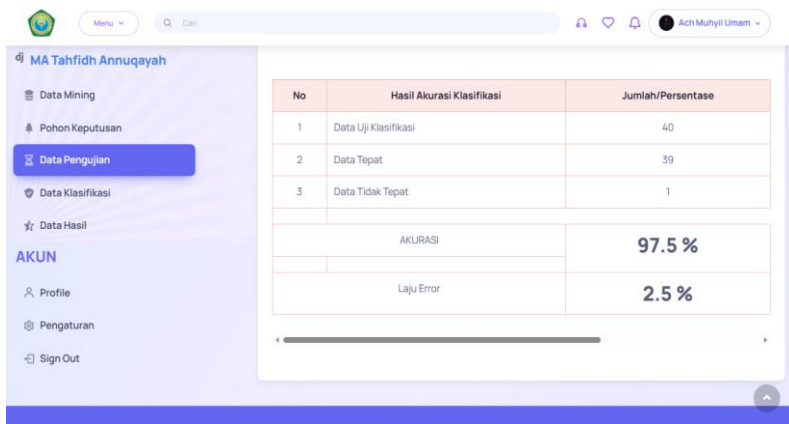
Pada halaman rule klaifikasi/pohon keputusan ini menampilkan rule/aturan dari hasil dari proses mining.

3.7.4. Halaman Pengujian



Gambar 6. Tampilan Halaman Pengujian

Pada halaman sistem pengujian,menampilkan manajemen data uji yang akan di gunakan sebagai data proses hitung akurasi.Pada halaman ini pengguna bisa melakukan CRUD data dan proses akurasi. Tampilan Hasil Akurasi Pengujian



Gambar 7. Tampilan Hasil Akurasi Pengujian

Gambar 7 menampilkan hasil akurasi data uji yang sudah diproses. Pada halaman ini dapat dilihat jumlah data tepat dan data tidak tepat, kemudian tingkat akurasi dan laju error data.

3.7.5. Halaman Cek Klasifikasi

Platform: MA Tahfidh Annuqayah

Menu: Data Mining, Pohon Keputusan, Data Pengujian, **Data Klasifikasi**, Data Hasil, AKUN (Profile, Pengaturan, Sign Out)

Data Hasil Klasifikasi

Pastikan data anda di isi dengan benar!

Your name: _____

Hafalan: ☐ istimewa ☐ sempurna ☐ kurang sempurna ☐ tidak hafal

Prestasi: ☐ memuaskan ☐ baik ☐ tidak ada

Nilai raport: _____

Nilai kitab: _____

Nilai ujian: _____

Nilai bahasa: _____

Buttons: Refresh, Cek Klasifikasi

Gambar 8. Halaman Cek Klasifikasi

Pada gambar 8, halaman cek klasifikasi ini dapat dilihat klasifikasi siswa dengan nilai data yang diminta pada form, kemudian hasilnya akan disimpan sebagai informasi klasifikasi.

3.7.6. Tampilan Hasil Rekomendasi

Menu: Data Pengujian, **Data Klasifikasi**, Data Hasil, AKUN (Profile, Pengaturan, Sign Out)

Rekomendasi

Selamat Data Siswa Atas Nama ACH MUHYIL UMAM Telah Memenuhi Persyaratan Rekomendasi Beasiswa. Konfirmasi kan kepada Siswa Jika Ingin Melanjutkannya, semangat!!

Buttons: Refresh, Cek Klasifikasi

Gambar 9. Tampilan Hasil Rekomendasi

Pada gambar 9 menampilkan informasi siswa yang sudah direkomendasi, kemudian hasilnya akan di simpan sebagai informasi klasifikasi.

3.7.7. Halaman Hasil Klasifikasi

Pada gambar 10 disajikan halaman hasil klasifikasi.Yakni dapat mengetahui semua informasi siswa yang sudah diklasifikasi.

Platform: MA Tahfidh Annuqayah

Menu: Data Mining, Pohon Keputusan, Data Pengujian, **Data Klasifikasi**, Data Hasil, AKUN (Profile, Pengaturan, Sign Out)

Hasil Klasifikasi

No	Nama	Hafalan	Prestasi	Raport	kitab	ujian	bahasa	Kelas Asil
1	Bilmy Maulidi	sempurna	baik	80	67	78	70	Rekomendasi
2	Agung Pamungkas	sempurna	tidakada	80	10	1	95	Tidakrekomendasi
3	ACH MUHYIL UMAM	sempurna	tidakada	80	100	85	95	Rekomendasi
4	ACH MUHYIL UMAM	istimewa	memuaskan	80	100	80	95	Rekomendasi
5	ACH MUHYIL UMAM	istimewa	memuaskan	80	100	80	95	Rekomendasi

Gambar 10. Halaman Hasil Klasifikasi

4. KESIMPULAN

Berdasarkan pada hasil pembahasan yang telah dijabarkan, maka dapat ditarik sebuah kesimpulan bahwa algoritma C4.5 mampu mengklasifikasi data menjadi pohon keputusan yang membantu rekomendasi beasiswa sesuai atribut yang diberikan. Hasil pengujian sistem klasifikasi siswa di MA Tahfidh Annuqayah dengan data training yang didapatkan menghasilkan nilai akurasi 97.5% dan nilai laju eror 2.5%. Hal ini menunjukkan bahwa penerapan algoritma menghasilkan keputusan klasifikasi yang baik

UCAPAN TERIMA KASIH

Penulis ingin mengucapkan terima kasih kepada Fakultas Teknik Prodi Informatika Universitas Madura, dosen dan mahasiswa serta pihak yang tidak bisa penulis sebutkan satu persatu yang telah membantu dan terlibat dalam pelaksanaan penelitian ini. Diharapkan artikel ini dapat menjadi referensi dan memberi kontribusi positif bagi pembaca.

REFERENCES

- [1] T. Ramadhan, E. Suswaini, and A. Uperiati, "Implementasi Algoritma C4.5 Dalam Klasifikasi Ketepatan Waktu Kelulusan Pada Data Mahasiswa Penerima Beasiswa Bidikmisi (Studi Kasus : Universitas Maritim Raja Ali Haji)," *SOJ (Student Online Journal)*, vol. 2, no. 1, pp. 257–270, 2021.
- [2] W. Susanto and A. Mulyani, "Analisa Algoritma C4 . 5 Terhadap Penentuan Rekomendasi," *OKTAL J. Ilmu Komput. dan Sci.*, vol. 1, no. 10, pp. 1607–1613, 2022.
- [3] N. Ramadhani and N. Fajarianto, "Sistem Informasi Evaluasi Perkuliahan dengan Sentimen Analisis Menggunakan Naïve Bayes dan Smoothing Laplace," *J. Sist. Inf. Bisnis*, vol. 02, 2020, doi: 10.21456/vol10iss2pp228-234.
- [4] N. Ramadhani, Z. Effendy, and I. Darmawan, "Penerapan Algoritma Naïve Bayes Classifier dan Fungsi Gaussian Untuk Penentuan Penjurusan Siswa Kelas X," *SMARTICS J.*, vol. 8, no. 1, pp. 14–21, 2022, [Online]. Available: <https://doi.org/10.21067/smartics.v8i1.6996>.
- [5] P. B. N. Setio, D. R. S. Saputro, and Bowo Winarno, "Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5," *Prism. Pros. Semin. Nas. Mat.*, vol. 3, pp. 64–71, 2020.
- [6] K. Provinsi, A. Carolina, and K. Ade, "Penerapan Data Mining Dengan Menggunakan Algoritma C4.5 Pada Klasifikasi Fasilitas Kesehatan Provinsi Di Indonesia," *J. Ilm. Komputasi*, vol. 19, no. 1, 2020, doi: 10.32409/jikstik.19.1.153.
- [7] Yadi, "Implementation Algorithm C4.5 Classification Of Prospective Scholarship Recipients," *J. SimanteC*, vol. 11, no. 1, pp. 27–32, 2022.
- [8] W. Z. Nilam Ramadhani, Abd Wahab Syahrone, Arin Supikar, "Penerapan Market Basket Analysis Menggunakan Metode Multilevel Association Rules dan Algoritma ML_T2L1 Pada Data Order PT. Unirama," *J. Nas. Inform. dan Teknol. Jar.*, vol. 4, no. 2, pp. 71–84, 2020, [Online]. Available: <https://jurnal.uisu.ac.id/index.php/infotekjar/article/view/2405>.
- [9] K. Kunci, "triple exponential smoothing .," vol. 10, no. 2, 2022.
- [10] A. Purwanto, "Jurnal Teknoinfo," *Tong Sampah Pint. Dengan Perintah Suara Guna Menghilangkan Perilaku Siswa Membuang Sampah Sembarangan Di Sekol.*, vol. 14, pp. 48–58, 2020, [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/article/view/336/329>.
- [11] R. Girsang, E. F. Ginting, and M. Hutasuht, "Penerapan Algoritma C4.5 Pada Penentuan Penerima Program Bantuan Pemerintah Daerah," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 449, 2022, doi: 10.53513/jursi.v1i4.5727.
- [12] A. Andriani, "Sistem Pendukung Keputusan Berbasis Decision Tree Dalam Pemberian Beasiswa (Studi Kasus: AMIK 'BSI Yogyakarta')," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2013, no. Sentika, pp. 163–168, 2013.
- [13] A. Rofik and D. R. Prehanto, "Sistem Informasi Pendukung Keputusan Penerimaan dan Monitoring Beasiswa Menggunakan Algoritma C4. 5 (Studi Kasus Universitas Negeri Surabaya)," *J. Emerg. Inf. Syst. Bus. Intell.*, vol. 01, no. 1, pp. 16–23, 2020, [Online]. Available: <https://ejournal.unesa.ac.id/index.php/JEISBI/article/view/36799>.
- [14] Y. Partogi and A. Pasaribu, "Perancangan Metode Decision Tree Terhadap Sistem Perpustakaan STMIK Kuwera," *J. Sist. Inf. dan Teknol.*, vol. 1, no. 2, pp. 20–25, 2022, doi: 10.56995/sintek.v1i2.4.
- [15] F. Dwi Meliani Achmad, Budanis, Slamet, "Klasifikasi Data Karyawan Untuk Menentukan Jadwal Kerja Menggunakan Metode Decision Tree," *J. IPTEK*, vol. 16, no. 1, pp. 18–23, 2012, [Online]. Available: <http://jurnal.itats.ac.id/wp-content/uploads/2013/06/3.-BUDANIS-FINAL-hal-17-23.pdf>.