

## Penerapan Algoritma C5.0 Untuk Memprediksi Tingkat Kepuasan Siswa Terhadap Kinerja Guru MAN Simalungun

Kherina Surya Ningsih<sup>1</sup>, Ilka Zufria<sup>2</sup>

<sup>1,2</sup>Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara

<sup>1</sup>kherinasry@gmail.com \*, <sup>2</sup>ilkazufria@uinsu.ac.id

### Abstract

*The role of teachers is crucial in shaping the educational landscape, which includes formal education at various levels. This responsibility includes nurturing, guiding, assessing and evaluating students. Taking inspiration from a foundational verse, this study investigates the profound influence of Quranic principles, using the life of Prophet Muhammad as a model, in shaping teacher professionalism. This relationship serves as a foundational aspect in understanding the role of educators. Identified challenges in teacher performance at Madrasah Aliyah Negeri Simalungun significantly impact student satisfaction and learning outcomes, with factors such as unclear explanations and lack of enthusiasm prominent. Using advanced techniques, this research utilizes data mining, specifically the C5.0 algorithm, to predict student satisfaction levels. This algorithm analyzed a data set consisting of 70% training data and 30% testing data. The results showed an accuracy rate of 77.78%, precision of 87.50%, recall of 87.50%, specificity of 0%, and F1 score of 87.50% on the first try. In the second experiment, with 80% training data and 20% testing data, the accuracy was 83.33%, precision 83.33%, recall 100%, specificity 0%, and F1-score 90.90%. This research underscores the importance of making informed decisions to improve the quality of education. By predicting and addressing factors that affect student satisfaction, educators can implement targeted improvements in teaching methods. In addition, this research is also in line with and demonstrates compliance with the National Education System Act. This quantitative analysis provides valuable insights for educators and policymakers in their ongoing efforts to improve and optimize teacher-student dynamics.*

**Keywords:** C5.0, Data Mining, Prediction, Teacher

### Abstrak

Peran guru sangat penting dalam membentuk lanskap pendidikan, yang mencakup pendidikan formal di berbagai tingkatan. Tanggung jawab ini mencakup pengasuhan, bimbingan, penilaian, dan evaluasi siswa. Mengambil inspirasi dari sebuah ayat yang mendasar, penelitian ini menyelidiki pengaruh mendalam dari prinsip-prinsip Alquran, dengan menggunakan kehidupan Nabi Muhammad sebagai model, dalam membentuk profesionalisme guru. Hubungan ini berfungsi sebagai aspek dasar dalam memahami peran pendidik. Tantangan yang teridentifikasi dalam kinerja guru di Madrasah Aliyah Negeri Simalungun secara signifikan berdampak pada kepuasan dan hasil belajar siswa, dengan faktor-faktor seperti penjelasan yang tidak jelas dan kurangnya antusiasme yang menonjol. Dengan menggunakan teknik-teknik canggih, penelitian ini menggunakan data mining, khususnya algoritma C5.0, untuk memprediksi tingkat kepuasan siswa. Algoritma ini menganalisis kumpulan data yang terdiri dari 70% data pelatihan dan 30% data pengujian. Hasilnya menunjukkan tingkat akurasi 77,78%, presisi 87,50%, recall 87,50%, spesifisitas 0%, dan skor F1 87,50% pada percobaan pertama. Pada percobaan kedua, dengan 80% data pelatihan dan 20% data pengujian, akurasi sebesar 83,33%, presisi 83,33%, recall 100%, spesifisitas 0%, dan F1-score 90,90%. Penelitian ini menggarisbawahi pentingnya pengambilan keputusan yang tepat untuk meningkatkan kualitas pendidikan. Dengan memprediksi dan menangani faktor-faktor yang mempengaruhi kepuasan siswa, para pendidik dapat mengimplementasikan perbaikan yang ditargetkan dalam metode pengajaran. Selain itu, penelitian ini juga sejalan dengan dan menunjukkan kepatuhan terhadap Undang-Undang Sistem Pendidikan Nasional. Analisis kuantitatif ini memberikan wawasan yang berharga bagi para pendidik dan pembuat kebijakan dalam upaya berkelanjutan mereka untuk memperbaiki dan mengoptimalkan dinamika guru-siswa.

**Kata kunci:** C5.0, Guru, Penambangan Data, Prediksi

©This work is licensed under a Creative Commons Attribution - ShareAlike 4.0 International License

### 1. Pendahuluan

Guru adalah pendidik profesional yang tanggung jawab utamanya mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan formal, pendidikan dasar, dan pendidikan menengah [1]. Undang-Undang Sistem Pendidikan Nasional Nomor 20 Tahun 2003 memerintahkan bahwa pembelajaran adalah suatu proses perhubungan antara peserta didik dan pendidik dengan menggunakan aset pembelajaran dalam iklim pembelajaran. Jadi pendidik harus memenuhi

kemampuan sesuai dengan tingkat siswa yang diajari, ilustrasi yang diajar, dan pengaturan serta aturan lainnya. Namun untuk mencapai tujuan pembelajaran, guru harus mampu menguasai sumber belajar, media pembelajaran, dan teknologi [2]. Namun, tidak semua guru di dunia terutama di Negara kita sendiri memiliki kinerja yang baik dan terkadang ada beberapa guru yang memiliki kualitas belajar yang tidak baik terhadap peserta didik [3]. Sebagaimana proses pembelajaran siswa/i yang nantinya akan menunjukkan kemampuan dan perkembangan mereka [4]. Hal itu tidak hanya dilihat dari fasilitas yang

disediakan akan tetapi juga guru yang kompeten dan memiliki kemampuan dibidangnya akan mempengaruhi kualitas peserta didik yang diajarkannya [5].

Madrasah Aliyah Negeri Simalungun (MAN Simalungun) adalah salah satu satuan pendidikan dengan jenjang MA di Kerasaan 1, Kec. Pematang Bandar, Kab. Simalungun, Sumatera Utara. MAN Simalungun adalah salah satu Madrasah yang mengedepankan kualitas imtaq dan iptek, oleh karena itu kemajuan dari Madrasah ini tentunya tidak lepas dari kontribusi dan peran dari berbagai pihak terutama guru yang selayaknya bersinergi demi meningkatnya kualitas Madrasah.

Dalam kegiatan belajar mengajar ada beberapa faktor yang menyebabkan ketidakpuasan peserta didik terhadap kinerja guru, diantaranya adalah guru yang tidak menjelaskan materi dengan baik, sering terlambat masuk kelas, tidak mampu berkomunikasi dengan baik, tidak rapi dalam berpakaian dan kurang bersemangat dalam mengajar. Dengan adanya permasalahan tersebut maka pengaruh bagi peserta didik adalah timbul rasa bosan rasa bosan dan kurang berminat ketika tiba waktu hari guru tersebut mengajar sehingga proses pembelajaran menjadi kurang efektif.

Dari permasalahan diatas, penulis melakukan upaya untuk memprediksi tingkat kepuasan siswa terhadap kinerja guru MAN Simalungun. Pelaksanaan prediksi kepuasan siswa terhadap guru akan dilakukan dengan cara memberikan kuisioner kepada siswa/I MAN Simalungun dengan menerapkan teknik *data mining* yang nantinya akan menggunakan *algoritma C5.0* sebagai metode yang efisien dalam memprediksi tingkat kepuasan siswa terhadap kinerja guru di MAN Simalungun. Algoritma C5.0 dianggap tepat dan mampu untuk menyelesaikan penelitian ini dengan beberapa alasan yang relevan. Pertama-tama, algoritma C5.0 merupakan perluasan dari algoritma C4.5, yang sudah dikenal sebagai salah satu algoritma klasifikasi yang efektif dalam menangani kumpulan data besar [6]. Dengan demikian, penggunaan C5.0 memberikan keunggulan dalam hal kecepatan memori dan efisiensi, sehingga cocok untuk penelitian yang melibatkan data peserta didik di Madrasah Aliyah Negeri Simalungun (MAN Simalungun). Dengan harapan setelah informasi terolah dan penambangan informasi diterapkan dengan menggunakan perhitungan C5.0, maka para pendidik akan lebih mampu mengambil pilihan dalam mempersiapkan hal-hal apa saja yang dirasa perlu ditingkatkan sehingga siswa merasa puas dengan pemaparan yang disampaikan oleh pendidik. yang telah diberikan.

Penelitian ini memiliki tiga tujuan utama. Pertama, untuk memberikan deskripsi yang komprehensif tentang tingkat kepuasan siswa terhadap kinerja guru di MAN Simalungun. Dengan memahami secara mendalam tingkat kepuasan siswa, penelitian ini

diharapkan dapat mengidentifikasi aspek-aspek tertentu yang perlu perhatian lebih dari pihak pendidik. Kedua, penelitian ini bertujuan untuk memprediksi tingkat kepuasan siswa terhadap kinerja guru dengan menggunakan algoritma C5.0. Prediksi ini menjadi landasan untuk merancang strategi perbaikan yang lebih terarah, membantu guru dan pihak sekolah dalam meningkatkan kualitas pembelajaran. Selanjutnya, penelitian ini juga berkomitmen untuk mengukur akurasi hasil prediksi sebagai aspek penting dalam validasi metode yang digunakan.

## 2. Metode Penelitian

### 2.1. Studi Pendahuluan

Studi pendahuluan merupakan studi yang dilakukan untuk mempertajam arah studi utama yang merupakan tahap awal dalam melakukan penelitian dan pengembangan model [7]. Dalam rangka mendapatkan wawasan yang komprehensif mengenai kondisi kinerja guru di MAN Simalungun, peneliti melakukan wawancara dengan total 75 responden, terdiri dari satu guru dan beberapa siswa. Wawancara dengan guru bertujuan untuk memahami perspektif dan pengalaman mereka dalam memberikan pengajaran, serta kendala atau tantangan yang mungkin mereka hadapi. Wawancara ini memberikan perspektif internal yang berharga terkait faktor-faktor yang dapat memengaruhi kinerja guru.

Sementara itu, wawancara dengan siswa, yang melibatkan sebanyak 75 orang, bertujuan untuk merinci tingkat kepuasan siswa terhadap kinerja guru. Pendapat dan pengalaman siswa menjadi landasan utama dalam mengevaluasi kualitas pengajaran. Jumlah responden yang signifikan ini diharapkan dapat mencakup beragam sudut pandang dan memberikan gambaran yang lebih akurat tentang dinamika interaksi antara guru dan siswa di MAN Simalungun. Hasil dari wawancara ini akan menjadi kontribusi penting dalam memahami faktor-faktor yang mungkin memengaruhi tingkat kepuasan siswa dan membentuk dasar penelitian lebih lanjut menggunakan algoritma C5.0.

### 2.2. Studi Literatur

Studi literatur atau studi kepustakaan merupakan kegiatan yang wajib dilakukan dalam sebuah penelitian yang tujuannya adalah mengembangkan aspek teoritis maupun aspek manfaat praktis. Studi kepustakaan dilakukan oleh setiap peneliti dengan tujuan utamanya yaitu mencari dasar pijakan/fondasi untuk memperoleh dan membangun landasan teori, kerangka berpikir, dan menentukan dugaan sementara atau disebut juga dengan hipotesis penelitian. Studi literatur yang digunakan oleh peneliti berupa jurnal 5 tahun terakhir, beberapa artikel dan juga buku yang berkaitan dalam penelitian ini.

### 2.3. Identifikasi Masalah Awal

Identifikasi masalah awal adalah mendefinisikan masalah penelitian, dalam tahap ini peneliti melakukan klasifikasi faktor – faktor kinerja guru yang nantinya akan menjadi bahan acuan dalam penelitian untuk menentukan tingkat kepuasan siswa terhadap kinerja guru dengan hasil prediksi puas, netral, dan tidak puas.

### 2.4. Pengumpulan Data

Penambahan informasi adalah pengalaman pendidikan PC (AI) yang memanfaatkan setidaknya satu strategi pembelajaran dengan menyelidiki desain informasi untuk mendapatkan data berharga. Penelitian ini menerapkan teknik data mining, khususnya metode klasifikasi berbasis pohon keputusan. Algoritma C5.0 merupakan salah satu algoritma pohon keputusan. Pohon keputusan adalah teknik klasifikasi sekumpulan objek atau data yang menggunakan representasi pohon [8].

Dalam tahap identifikasi masalah penelitian, peneliti melakukan klasifikasi faktor-faktor kinerja guru sebagai langkah awal dalam memahami aspek-aspek yang dapat memengaruhi tingkat kepuasan siswa. Selanjutnya, data untuk penelitian ini diperoleh melalui teknik survei dan observasi.

#### a. Survery (Kuesioner)

Peneliti merancang kuesioner yang berisi pertanyaan-pertanyaan terkait faktor-faktor kinerja guru yang telah diidentifikasi. Kuesioner ini disebarkan kepada siswa di MAN Simalungun. Responden diminta memberikan tanggapan mereka terhadap setiap faktor kinerja guru dengan memberikan penilaian skala tertentu, seperti skala Likert. Pertanyaan-pertanyaan tersebut dapat mencakup aspek seperti kemampuan menjelaskan materi, kehadiran yang tepat waktu, kemampuan berkomunikasi, keteraturan berpakaian, dan semangat dalam mengajar.

#### b. Oservasi

Selain survei, peneliti melakukan observasi langsung terhadap kegiatan belajar mengajar di MAN Simalungun. Observasi mencakup pengamatan terhadap interaksi guru-siswa, teknik pengajaran yang diterapkan, keteraturan dan kedisiplinan guru, serta ekspresi siswa selama pembelajaran. Observasi ini membantu mendapatkan pemahaman yang lebih mendalam mengenai dinamika kelas dan faktor-faktor yang mungkin tidak terungkap melalui kuesioner.

Data yang diperoleh dari kedua teknik ini akan menjadi dasar untuk mengidentifikasi pola-pola dan hubungan antara faktor-faktor kinerja guru dengan tingkat kepuasan siswa. Selanjutnya, data ini akan diolah menggunakan teknik data mining, dengan algoritma C5.0 sebagai metode utama untuk memprediksi tingkat kepuasan siswa terhadap kinerja guru di MAN Simalungun.

### 2.5. Analisis

Dalam tahapan analisis ini terbagi kedalam dua bagian yaitu.

#### a. Analisis Data

Analisis data adalah interaksi yang melakukan hal-hal seperti memeriksa, membersihkan, mengubah, dan mendemonstrasikan informasi yang bertekad untuk mencapai tujuan.

#### b. Analisis Kebutuhan

Pada tahap analisis kebutuh terbagi menjadi analisis kebutuhan fungsional dan analisis kebutuhan non-fungsional. Proses kegiatan yang akan dilaksanakan dalam suatu sistem digambarkan dalam analisis fungsional, beserta persyaratan yang diperlukan agar sistem dapat berfungsi dengan baik dan sesuai dengan kebutuhan. Pemeriksaan kebutuhan utilitarian ini mencakup penyelidikan prasyarat informasi dan tampilan kerangka. Sedangkan kajian non-utilitarian menggambarkan prasyarat kerangka kerja yang berpusat pada sifat-sifat sosial kerangka, termasuk pemrograman, peralatan, dan kebutuhan klien kerangka sebagai bahan untuk membedah kekurangan dan kebutuhan yang harus dipenuhi dalam perencanaan kerangka yang akan dilaksanakan.

7 aturan yang digunakan dalam penelitian ini dapat dilihat pada Tabel 1 di bawah:

Tabel 1. Tabel Kriteria Penelitian

No	Kriteria	Komponen
1	Evaluasi Pengajaran	1. Kemampuan menjelaskan materi 2. Kejelasan komunikasi 3. Keadilan dalam penilaian
2	Interaksi Guru Dengan Siswa	1. Frekuensi pertanyaan siswa 2. Kualitas tanggapan guru, terhadap pertanyaan 3. Guru merespon kebutuhan dan masalah siswa
3	Ketersediaan bahan Belajar	1. Penilaian terhadap ketersediaan 2. Kualitas materi pembelajaran yang disediakan oleh guru
4	Kejelasan Tujuan Pembelajaran	1. Guru menjelaskan tujuan pembelajaran kepada siswa 2. Guru memberikan panduan yang jelas terhadap apa yang dibayangkan
5	Keberagaman Pendekatan Pengajaran	1. Pengguna multimedia 2. Metode aktif 3. Penggunaan teknologi
6	Kemampuan Motivasi	1. Pujian guru terhadap siswa 2. Penghargaan yang diberikan kepada siswa 3. Dukungan emosional
7	Responsif Terhadap Kebutuhan Siswa	1. Kebutuhan individual siswa baik dalam pembelajaran 2. Kebutuhan individual siswa baik dalam pribadi

## 2.6. Implementasi Algoritma C5.0

Algoritma C5.0 merupakan algoritma penyempurnaan dari algoritma C4.5. Prosesnya hampir sama dengan algoritma C4.5, namun algoritma C5.0 memiliki keunggulan dibandingkan algoritma C4.5 [9].

Algoritma C5.0 merupakan perpanjangan dari ID3, serta perpanjangan dari algoritma C4.5. Perhitungan C5.0 adalah perhitungan karakterisasi yang masuk akal untuk koleksi informasi yang sangat besar. Dari segi kecepatan dan efisiensi memori, algoritma C5.0 lebih unggul dibandingkan C4.5. Ukuran rasio gain digunakan saat memilih atribut yang akan diproses dalam algoritma C5.0. Setiap node dengan gain rasio tertinggi dipilih sebagai induk untuk node berikutnya dengan menggunakan ukuran gain rasio. Usaha kerja pembuatan pohon pada perhitungan C5.0 seperti membuat pohon pada C4.5. Proses penghitungan entropi dan penguatan adalah salah satu contoh persamaan ini. Di C5.0, ia akan terus menghitung rasio keuntungan menggunakan keuntungan dan entropi yang ada jika berhenti di titik penghitungan keuntungan di C4.5 [10].

Dijelaskan didalam [11] bahwa model berbasis aturan digunakan dalam algoritma C5.0, yang memudahkan untuk melihat aturan di pohon keputusan. Algoritma C5.0 juga dapat menangani nilai yang hilang. Ini merupakan keunggulan C5.0 yang dinilai lebih baik dibandingkan perhitungan lainnya. Cara kerja pengembangan pohon pada perhitungan C5.0 secara praktis setara dengan penataan pohon pada perhitungan C4.5. Perhitungan entropi dan perolehan informasi dimasukkan dalam persamaan ini. Jika penghitungan C4.5 terhenti hingga penghitungan perolehan data, maka penghitungan C5.0 dilanjutkan dengan memastikan proporsi kenaikan menggunakan penjumlahan dan entropi saat ini [12]. Pemilihan kualitas dalam perhitungan ini menggunakan proporsi peningkatan dimana karakteristik dengan nilai proporsi penambahan terbesar akan dipilih sebagai induk untuk hub berikutnya.

Perhitungan algoritma C5.0 secara umum memiliki rumus berikut ini :

Untuk mencari nilai Entropy :

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

S = Himpunan Kasus

n = Jumlah Partisi S

$p_i$  = Proporsi dari  $S_i$  terhadap S

Untuk mencari nilai information gain :

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

Keterangan :

S = Himpunan Kasus

A = Atribut

n = Jumlah Partisi Atribut A

$|S_i|$  = Jumlah Kasus Pada Partisi ke i

$|S|$  = Jumlah Kasus Dalam S

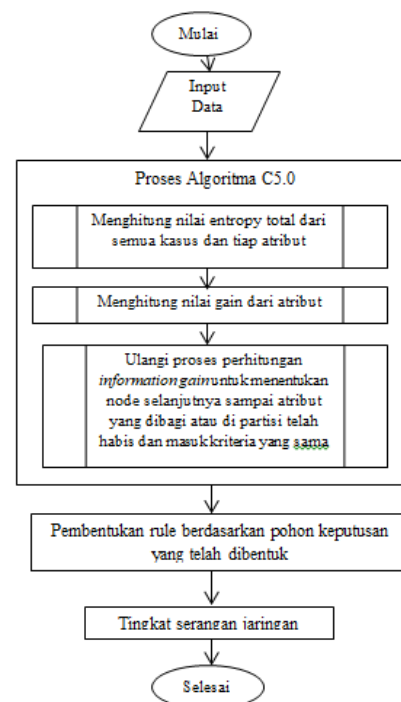
Setelah nilai entropy dan information gain di dapat, kemudian dilanjutkan dengan mencari nilai gain ratio menggunakan rumus berikut :

$$Gain\ Ratio = \frac{Information\ Gain(S,A)}{\sum_{i=1}^n Entropy(S_i)} \quad (3)$$

Ulangi proses penghitungan gain rasio hingga setiap cabang memiliki kelasnya masing-masing. Pada iterasi selanjutnya, nilai gain rasio tidak lagi dihitung menggunakan atribut yang dipilih. Anda dapat menggunakan persamaan di bawah ini untuk menentukan keakuratan aturan.:

$$Akurasi = \frac{Nilai\ Kecocokan}{Jumlah\ Seluruh\ Kejadian} * 100\% \quad (4)$$

Gambar 2 berikut merupakan tahapan pembentukan pohon algoritma C5.0:



Gambar 1. Tahapan Pembentukan Pohon Algoritma C5.0

Tahap ini membentuk model pohon pilihan untuk mengenali contoh-contoh penting dalam kumpulan data dengan melakukan perhitungan C5.0.

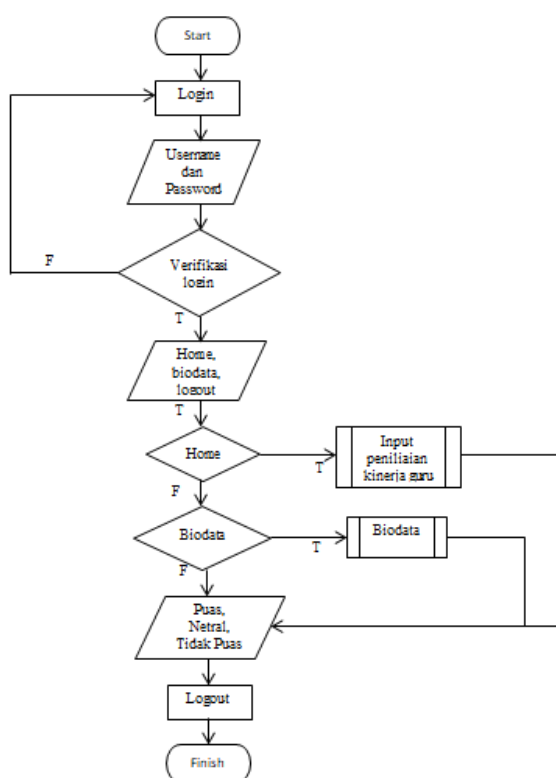
Adapun proses pembentukan pohon keputusan dijelaskan sebagai berikut [8] :

- Tentukan total perolehan informasi kasus;

- b. Tentukan perolehan informasi dan entropi dari setiap kemungkinan posisi kerusakan atribut.
- c. Menghitung nilai Entropi dan Gain absolut dari setiap posisi rincian kualitas yang mungkin, Gain yang paling tinggi akan dipilih sebagai hub akar.
- d. Buat cabang dari setiap kelas hub root. Percabangan berakhir dan berubah menjadi simpul daun jika atribut Entropi diatur ke nol. Node daun dipilih berdasarkan suara terbanyak.
- e. Jika cabang mengarah ke node internal, Anda harus mengikuti prosedur yang sama seperti sebelumnya.
- f. Siklus tersebut dilakukan hingga klasifikasi karakteristik tidak dapat dipisahkan lagi.

## 2.7. Desain

Dari informasi dan data yang telah dikumpulkan maka tahap berikutnya adalah pembuatan desain, desain dibuat terhadap tampilan halaman dari sebuah situs terkait dengan desain UI/UX sebuah website. Tampilan flowchart struktur program diilustrasikan pada Gambar 2 berikut :



Gambar 2. Flowchart Struktur Program

## 2.8. Penerapan

Pada tahap penerapan ini nantinya akan dilakukan proses prediksi kinerja guru dengan hasil prediksi puas, netral, tidak puas. Untuk sampel data yang diambil sebanyak 75 siswa kelas 11 jurusan Matematika dan Ilmu Pengetahuan Alam. Pada tahap ini juga data akan diolah menggunakan aplikasi

berbasis website yang telah dirancang sebelumnya untuk membantu proses prediksi.

## 2.9. Pengujian

Pada tahap pengujian, fokus utama dilakukan terhadap sistem website untuk memastikan bahwa fungsi-fungsi yang telah diimplementasikan berjalan dengan baik dan sesuai dengan harapan. Pengujian sistem ini mencakup berbagai aspek, di antaranya:

### a. Uji Fungsionalitas

Pengujian ini dilakukan untuk memastikan setiap fitur dan fungsi yang telah diimplementasikan berjalan sesuai dengan spesifikasi yang telah ditetapkan. Termasuk di dalamnya adalah verifikasi bahwa tombol-tombol berfungsi dengan benar, formulir dapat diisi dan disubmit dengan sukses, serta navigasi antarmuka pengguna berjalan lancar.

### b. Uji Kinerja

Uji kinerja dilakukan untuk mengidentifikasi sejauh mana website dapat menangani beban penggunaan yang tinggi. Pengujian ini melibatkan pemantauan respons waktu, kecepatan akses, dan stabilitas sistem saat ada banyak pengguna yang mengakses dan menggunakan website secara bersamaan.

Jika dalam tahap pengujian terdapat indikasi bug atau ketidaksesuaian dengan harapan, maka dilakukan iterasi dan debugging untuk memperbaiki masalah tersebut. Proses ini berulang hingga diperoleh hasil pengujian yang mendekati kata sempurna, di mana website dapat beroperasi secara optimal dan sesuai dengan tujuan yang telah ditetapkan.

## 3. Hasil dan Pembahasan

Penelitian ini menjadi langkah awal untuk meningkatkan hubungan guru-siswa dan kualitas pendidikan dengan mengidentifikasi masalah kepuasan siswa terhadap guru. Komunikasi terbuka antara guru, siswa, dan pihak sekolah diakui sebagai solusi utama. Feedback siswa menjadi sumber informasi berharga untuk perbaikan dalam pengajaran. Pelatihan guru dan dukungan pihak sekolah diperlukan untuk meningkatkan kualitas pengajaran.

Dalam merancang sistem, data set kepuasan siswa terhadap guru terdiri dari 75 hasil riset, dengan kelas atau label "puas" dan "tidak puas". Pengelompokan ini menjadi dasar analisis sistem dengan tingkat akurasi yang optimal.

Tabel 2. Data Set Penilaian Kepuasan Siswa Terhadap Guru

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP

6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP
7	Rita Ramadhani Purba	7	6	6	4	9	6	P
...	...	...	...	...	...	...	...	...
75	Fadilah Sindiyanti	4	6	4	4	6	4	TP

Keterangan tabel :

A = Pengajaran  
B = Interaksi  
C = Bahan Ajar  
D = Tujuan  
E = Pendekatan  
F = Motivasi  
G = label  
P = Puas  
TP = Tidak Puas

### 3.1. Analisis Data

Dalam merancang sistem yang efisien, dibutuhkan informasi sesuai prasyarat kerangka kerja. Evaluasi kerangka kerja masa lalu atau saat ini memungkinkan penentuan tingkat kinerja sistem, persyaratan operasional, dan kebutuhan yang belum terpenuhi [13].

Metode validasi model klasifikasi yang digunakan adalah hold out dengan perbandingan 70:30 (data training 21, data testing 9) dan 80:20 (data training 24, data testing 6). Penelitian menunjukkan bahwa pembagian dataset dengan perbandingan 70:30 dan 80:20 memiliki kinerja baik dalam memaksimalkan akurasi model prediksi label, menjadi acuan dalam penelitian ini dengan penggunaan dataset acak berjumlah 30 seperti yang tertera pada Tabel 3, 4, 5 dan 6 dibawah.

Tabel 3. Data Set Awal Untuk Perhitungan Manual

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP
6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP
7	Rita Ramadhani Purba	7	6	6	4	9	6	P
...	...	...	...	...	...	...	...	...
21	Syahni Nazwa	7	5	6	4	9	6	P

Tabel 4. Data Testing Hold Out 70:30 Perjumlah 9

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP
6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP
7	Rita	7	6	6	4	9	6	P

8	Ramadhani Purba Bukhori Rahman Hutagaol	6	6	4	4	6	4	P
9	Lusi Amelia	6	6	5	4	6	5	P

Tabel 5. Data Training Hold Out 80:20 Berjumlah 24

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP
6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP
7	Rita Ramadhani Purba	7	6	6	4	9	6	P
...	...	...	...	...	...	...	...	...
24	Nisa Azzahra	7	7	5	5	7	7	P

Tabel 6. Data Testing Hold Out 80:20 Berjumlah 6

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP
6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP

## 1. Penerapan Metode C5.0

Sebelum melakukan proses perhitungan maka ditentukan attribut yang digunakan dalam klasifikasi pohon keputusan C5.0 untuk menentukan klasifikasi berdasarkan data yang akan dilatih dan diuji. Penilaian kepuasan siswa terhadap guru merupakan nilai yang akan dihitung disini.

### A. Perhitungan Manual

Data yang akan dicoba untuk di hitung manual berjumlah 30 data seperti tabel dibawah dimana attribute bernilai numerik bukan kategorik.

Tabel 7. Pengujian 1 Data Set 30

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa Oryza Syahputra	7	5	5	4	6	4	P
2	Java Maulana	4	6	7	4	8	4	P
3	Merri Andani	6	6	5	6	6	4	P
4	Bella Gusmita Sari	3	5	3	2	5	3	TP
5	Syifa Azzahra	3	4	2	2	5	3	TP
6	Muhammad Abdi Pratama	5	5	4	2	5	2	TP
7	Rita Ramadhani Purba	7	6	6	4	9	6	P
...	...	...	...	...	...	...	...	...
30	Fatma	6	6	6	4	7	5	P

Azzahra

Tabel diatas merupakan 30 data penilaian siswa terhadap guru untuk dilakukan perhitungan manual C5.0 selanjutnya.

B. Normalisasi

Tahap ini dilakukan untuk mengubah nilai kedalam rentang yang sama, untuk mengatasi masalah skala dalam data dan mengurangi kesalahan perhitungan maupun hasil akhir yang kurang maksimal. Rumus dalam proses normalisasi sebagai berikut :

$$z = (x - \min(x)) / (\max(x) - \min(x)) \quad (5)$$

Keterangan :

$x_1(1) = 7$ , min (kolom pengajaran) = 3, max (kolom pengajaran) = 8

$z_1(1) =$

$$= (7 - 3) / (8 - 3)$$

$$= 4 / 5$$

$$= 0,8$$

$x_2(1) = 4$ , min (kolom interaksi) = 3, max (kolom interaksi) = 8

$z_2(1) =$

$$= (4 - 3) / (8 - 3)$$

$$= 1 / 4$$

$$= 0,2$$

$x_1(2) = 5$ , min (kolom interaksi) = 4, max (kolom interaksi) = 8

$z_1(2) =$

$$= (5 - 4) / (8 - 4)$$

$$= 1 / 4$$

$$= 0,25$$

$x_2(2) = 6$ , min (kolom pengajaran) = 4, max (kolom pengajaran) = 8

$z_2(2) =$

$$= (6 - 4) / (8 - 4)$$

$$= 2 / 4$$

$$= 0,5$$

Tabel 8. Pengujian 1 Data Set 30

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa	0,8	0,25	0,5	0,5	0,4	0,4	P
	Oryza							
	Syahputra							
2	Java Maulana	0,2	0,50	1	0,5	0,8	0,4	P
3	Merri Andani	0,6	0,50	0,5	1	0,4	0,4	P
4	Bella	0	0,25	0	0	0,2	0,2	TP
	Gusmita Sari							
5	Syifa	0	0,00	0	0	0,2	0,2	TP
	Azzahra							
6	Muhammad	0,4	0,25	0,25	0	0,2	0	TP
	Abdi Pratama							
7	Rita	0,8	0,50	0,75	0,5	1	0,8	P
	Ramadhani							
	Purba							

...	...	...	...	...	...	...	...	...
30	Fatma	0,6	0,50	0,75	0,5	0,6	0,6	P
	Azzahra							

Tabel 8 diatas merupakan tabel normalisasi untuk dilakukan perhitungan manual C5.0 selanjutnya.

C. Holdout Validation

Merupakan strategi persetujuan silang yang lebih mudah. Kumpulan data dibagi menjadi dua bagian untuk validasi ketidaksepakatan: data untuk pengujian dan pelatihan. Seringkali, 70 hingga 80 persen data digunakan untuk pelatihan, dan sisanya digunakan untuk pengujian. Disini kita menggunakan pembagian 70:30 yaitu 70% data persiapan dan 30% data tes, dibawahnya data persiapan yang akan kita gunakan.

Tabel 9. Data Latih 70%

No	Nama Siswa	A	B	C	D	E	F	G
1	Adhwa							
	Oryza	0,8	0,25	0,5	0,5	0,4	0,4	P
	Syahputra							
2	Java Maulana	0,2	0,50	1	0,5	0,8	0,4	P
3	Merri Andani	0,6	0,50	0,5	1	0,4	0,4	P
4	Bella	0	0,25	0	0	0,2	0,2	TP
	Gusmita Sari							
5	Syifa	0	0,00	0	0	0,2	0,2	TP
	Azzahra							
6	Muhammad	0,4	0,25	0,25	0	0,2	0	TP
	Abdi Pratama							
7	Rita	0,8	0,50	0,75	0,5	1	0,8	P
	Ramadhani							
	Purba							
...	...	...	...	...	...	...	...	...
21	Syahni	0,8	0,50	0,5	0,5	1	0,8	P
	Nazwa							

Tabel 10. Uji 30%

No	Nama Siswa	A	B	C	D	E	F	G
21	Syahni	0,8	0,50	0,5	0,5	1	0,8	P
	Nazwa							
22	Dian Aini	0,6	0,50	0,25	0,5	0,4	0,8	P
	Ramadhani							
23	Ayudiah	0,6	0,50	0,25	0,5	0,2	0,4	TP
	Pratiwi							
24	Nisa Azzahra	0,8	0,75	0,5	0,75	0,6	1	P
...	...	...	...	...	...	...	...	...
30	Fatma	0,6	0,50	0,75	0,5	0,6	0,6	PP
	Azzara							

D. Perhitungan nilai rata-rata dan median

Hitung nilai rata-rata dan median untuk atribut numerik, dilakukan pemecahan dengan mengambil nilai rata-rata dan median. Dengan rumus :

$$x = x_1 + x_2 = x_3 \dots + x_n/n \quad (6)$$

Tabel 11. Nilai Rata-Rata dan Median Data Latih

	A	B	C	D	E	F
Rata-rata	0,54	0,40	0,44	0,43	0,50	0,44
Median	0,60	0,50	0,50	0,50	0,40	0,40

E. Atribut Label

Elemen yang terdapat pada C5.0 yaitu Information Gain, Entropy, Gain Ratio, sedangkan C4.5 tidak memiliki Gain Ratio. Menghitung atribut keputusan yaitu jumlah, puas dan tidak puas.

Tabel 12. Keterangan Penilaian Atribut Label

Atribut	Jumlah	Puas	TidakPuas
Label	21	15	6

Menghitung entropy untuk mengukur kecenderungan suatu kelas dari sekumpulan data. Rumus entropy:

$$\text{entropy}(S) = (-p_1) * \log_2 p_1 + (-p_2) * \log_2 p_2 \quad (7)$$

Keterangan:

$$\begin{aligned} \text{entropy} &= ((-15 / 21) * \log_2(15 / 21) + (-6 / 21) * \log_2(6 / 21)) \\ &= 0,863120569 \end{aligned}$$

Menghitung atribut numerik pengajaran yaitu jumlah, puas dan tidak puas.

Tabel 13. Keterangan Penilaian Atribut Pengajaran

Atribut	Jumlah	Puas	TidakPuas
<=0,54	6	2	4
>0,54	15	13	2
<=0,60	16	12	4
>0,60	5	3	2

Entropy digunakan untuk mengukur kecenderungan suatu kelas dari sekumpulan data.

Rumus entropy:

$$\text{Entropy}(S) = (-p_1) * \log_2 p_1 + (-p_2) * \log_2 p_2$$

Keterangan:

$$\begin{aligned} \text{entropy} &= ((-2 / 6) * \log_2(2 / 6) + (-4 / 6) * \log_2(4 / 6)) \\ &= 0,918295834 \end{aligned}$$

Information Gain adalah nilai informasi untuk mengukur tingkat keberagaman (heterogenitas) dari sekumpulan data.

Rumus information gain:

$$\text{Ig}(t) = (S) - ((P_m / P_n) * (S_m) + (Q_m / P_n) * (S_i))$$

Keterangan:

$$\begin{aligned} \text{Information gain} &= (0,8631205) - ((6 / 21) * (0,9182958) + (15 / 21) * (0,5665095)) \\ &= 0,196100683 \end{aligned}$$

Split mengacu pada pemisahan data berdasarkan nilai-nilai atribut tertentu.

Rumus split info:

$$\text{Split}(E) = -(P_m / P_n) * \log_2(P_m / P_n) - (Q_m / P_n) * \log_2(Q_m / P_n)$$

Keterangan:

$$\begin{aligned} \text{Split info} &= -(6 / 21) * \log_2(6 / 21) - (15 / 21) * \log_2(15 / 21) \\ &= 0,863120569 \end{aligned}$$

Gain Ratio dapat diartikan sebagai ukuran efektifitas atau seberapa informatif suatu atribut dalam melakukan klasifikasi data.

Rumus gain ratio:

$$\text{Gain}(A) = \text{Ig}(t) / \text{Split}(E)$$

Keterangan:

$$\begin{aligned} \text{Gain Ratio} &= 0,196100683 / 0,863120569 \\ &= 0,22719964 \end{aligned}$$

Setelah dilakukan perhitungan entropy, information gain dan gain ratio selanjut dilakukan pruning atau pemangkasan dimana node atau simpul yang akan diambil dengan nilai gain ratio tertinggi.

Tabel 14. Atribut Nilai Gain Ratio Tertinggi

Atribut	Keterangan	Atribut Nilai	Gain Ratio
Label	Rata - rata	0,40	0,577923613
Tujuan	Rata - rata	0,43	0,492399749

Disini kita mengambil 2 nilai gain ratio tertinggi untuk dilakukan pohon keputusan.

Tabel 15. Confusion matrix node 1

Aktual	Prediksi	
	Puas	Tidak Puas
Puas	7	1
Tidak Puas	1	0

Tabel 16. Atribut nilai gain ratio tertinggi

Atribut	Rumus	Hasil	Hasil %
Accuracy	TP+TN/Jumlah Data	0,7778	77,78 %
Precision	TP/(TP+FP)	0,8750	87,50 %
Recall	TP/(TP+FN)	0,8750	87,50 %
Specifity	TN/(TN+FP)	0	0 %
F1-Score	2 * (Recall * Precision) / (Recall + Precision)	0,8750	87,50 %

Disini nilai akurasi yaitu 78% jadi nilai akurasi dari hold out 70:30 didapat yaitu 78%. Dibawah ini contoh kedua dari hold out 80:30 dalam klasifikasi pohon keputusan.

Tabel 17. Confusion Matrix Node 1

Aktual	Prediksi	
	Puas	Tidak Puas
Puas	6	1
Tidak Puas	0	0

Tabel 18. Atribut Nilai Gain Ratio Tertinggi

Atribut	Rumus	Hasil	Hasil %
Accuracy	TP+TN/Jumlah Data	0,8333	83,33 %
Precision	TP/(TP+FP)	0,8333	83,33 %
Recall	TP/(TP+FN)	1	100 %
Specifity	TN/(TN+FP)	0	0 %
F1-Score	2 * (Recall * Precision) / (Recall + Precision)	0,9091	90,91 %

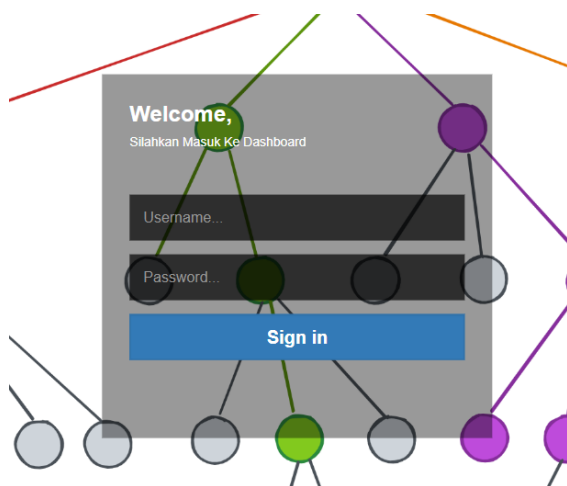


Disini nilai akurasi dari hold out data 80:20 yaitu 83,33% jadi nilai akurasi dari hold out 80:20 didapat yaitu 83,33%.

### 3.2. Rancangan Antar Muka

Berikut adalah gambaran yang akan di bangun dalam pembuatan aplikasi web yang sudah dirancang dan ditentukan, mockup merupakan gambaran kecil dalam penggunaan aplikasi yang sudah diterapkan. Tahap pengujian ini merupakan tahap yang direncanakan untuk melihat apakah setiap kapabilitas dalam framework telah berjalan sesuai rencana yang dibuat. Tahap pengujian dilakukan dengan memanfaatkan aplikasi web dengan browser internet khususnya Google Chrome. Tahap pengujian ini dilakukan dengan memanfaatkan localhost sebagai server pengujian. Hasil dari pengujian yang dilakukan adalah sebagai berikut :

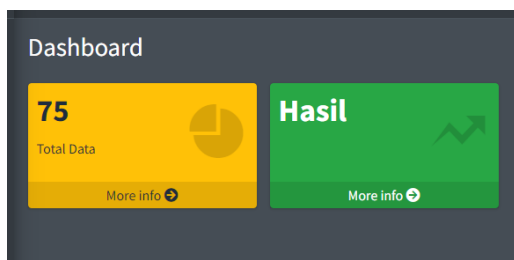
Disini user melakukan login agar bisa masuk ke aplikasi sistem klasifikasi pohon keputusan menentukan kepuasan siswa terhadap guru.



Gambar 3. Aplikasi Web Halaman Login

Gambar 3 Di atas adalah tampilan aplikasi dari halaman login, klien harus masuk untuk mengawasi akses aplikasi

Kemudian, masuk ke dashboard dan informasi penilaian siswa secara keseluruhan pada pendidik yang disimpan di awal muncul.



Gambar 4. Tampilan Total Data dan Hasil

Dashboard yang berisi data total penilaian guru terhadap kepuasan siswa dan hasil perhitungannya akan ditampilkan setelah pengguna login.

Menampilkan total data penilaian siswa terhadap guru yang sudah diinputkan diawal.

No	Nama Siswa	Pengajaran	Interaksi	Bahan Ajar	Tujuan	Pendekatan	Motivasi	Label
1	Adhwa Grysya Syahputra	7	5	5	4	6	4	Puas
2	Jaya Maulana	4	6	7	4	8	4	Puas

Gambar 5. Data Penilaian Siswa Terhadap Guru

Pada menu data penilaian guru siswa, pengguna memasukkan data dan atribut yang akan digunakan dalam proses klasifikasi pohon keputusan. Di sini, pengguna memasukkan penilaian guru berdasarkan informasi atribut yang disediakan.

Kemudian memasukkan data penilaian siswa terhadap guru yang sudah diberikan oleh MAN Simalungun kedalam aplikasi.

Gambar 6. Menyimpan Penilaian Siswa Terhadap Guru

Kemudian user melakukan input data yang sudah didapatkan dari hasil riset di MAN Simalungun.

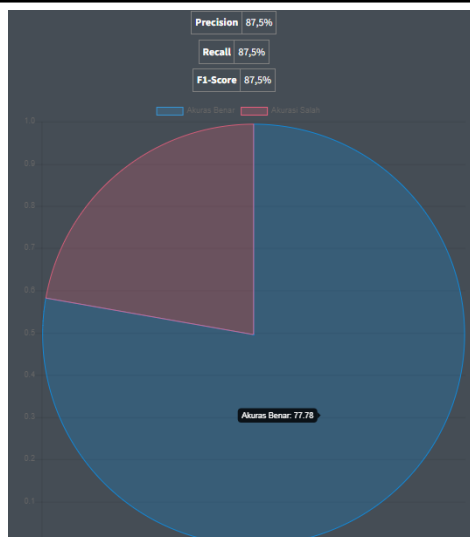
Menampilkan hasil klasifikasi pohon keputusan dengan algoritma C5.0.

No	Nama Siswa	Pengajaran	Interaksi	Bahan Ajar	Tujuan	Pendekatan	Motivasi	Label	Klasifikasi
22	Dan Aini Ramadhani	6	6	4	4	6	6	Puas	Puas
23	Ayudha Pratiwi	6	6	4	4	5	4	Tidak Puas	Puas

Gambar 7. Menampilkan Grafik Hasil Klasifikasi Pohon Keputusan C5.0

Gambar diatas merupakan tampilan grafik hasil klasifikasi pohon keputusan dengan algoritma C5.0.

Menampilkan performance vector dari algoritma C5.0 dalam klasifikasi pohon keputusan.



Gambar 8. Menampilkan Grafik Performance Vector

Gambar diatas merupakan tampilan hasil akurasi, recall, precision dan f1-score dimana penilaian diatas merupakan keakuratan dan ketepatan dalam klasifikasi data dimana data hold out 70:30 pada perhitungan diatas.

#### 4. Kesimpulan

Penelitian ini menggunakan algoritma pohon keputusan C5.0 untuk mengevaluasi tingkat kepuasan siswa terhadap guru di MAN Simalungun. Dengan pembagian data uji sebesar 30% (9 data) dan data latih sebesar 70% (21 data) pada percobaan pertama, serta pembagian data uji sebesar 20% (6 data) dan data latih sebesar 80% (24 data) pada percobaan kedua, algoritma C5.0 memberikan hasil klasifikasi yang signifikan. Pada percobaan pertama, akurasi mencapai 77,78%, dengan precision 87,50%, recall 87,50%, specificity 0%, dan f1-score 87,50%. Percobaan kedua menunjukkan peningkatan performa, dengan akurasi sebesar 83,33%, precision 83,33%, recall 100%, specificity 0%, dan f1-score 90,90%.

Hasil ini menyoroti pengaruh pembagian data latih dan uji terhadap performa model, dan menunjukkan bahwa rasio pembagian 70:30 dan 80:20 memberikan hasil yang memuaskan. Dengan menggunakan atribut dan label pada data latih yang telah dikategorikan sebagai puas dan tidak puas, algoritma C5.0 memberikan wawasan mendalam mengenai kepuasan siswa terhadap guru melalui model pohon keputusan. Pemahaman ini dapat menjadi dasar bagi pengambilan

keputusan di lingkungan pendidikan untuk meningkatkan kualitas pengajaran dan kepuasan siswa.

#### Daftar Rujukan

- [1] Dewan Perwakilan Rakyat Indonesia, *Undang-Undang (UU) tentang guru dan dosen*. Dewan Perwakilan Rakyat Indonesia, 2005.
- [2] C. H. Sancoko and R. Sugiarti, "Kinerja Guru Dan Faktor Yang Mempengaruhinya," *J. Pendidik. Rokania*, vol. 7, no. 1, pp. 1–14, 2022.
- [3] M. Muspawi, "Strategi Peningkatan Kinerja Guru," *J. Ilm. Univ. Batanghari Jambi*, vol. 21, no. 1, p. 101, 2021, doi: 10.33087/jiubj.v21i1.1265.
- [4] A. Dina, D. Yohanda, J. Fitri, and ..., "Teori Kinerja Guru Dalam Meningkatkan Mutu Pendidikan," *J. Edukasi Nonform.*, vol. 1, no. 1, pp. 149–158, 2022.
- [5] I. Asakir and F. Mahmudah, "Kreativitas dan Inisiatif Guru dalam Pengembangan Mutu Pembelajaran Online," *J. Stud. Guru dan Pembelajaran*, vol. 5, no. 1, pp. 31–40, 2022, doi: 10.30605/jsdp.5.1.2022.1541.
- [6] A. Ulfa, D. Winarso MKom, and E. Arribe MMSi, "SISTEM REKOMENDASI JURUSAN KULIAH BAGI CALON MAHASISWA BARU MENGGUNAKAN ALGORITMA C4.5 (Studi Kasus: Universitas Muhammadiyah Riau)," *Fasilkom*, vol. 10, no. 1, pp. 61–65, 2020.
- [7] N. K. W. Oktaviani and M. Putra, "Motivasi dan Disiplin Kerja Terhadap Kinerja Guru di Sekolah Dasar," *J. Imiah Pendidik. dan Pembelajaran*, vol. 5, no. 2, p. 294, 2021, doi: 10.23887/jipp.v5i2.35146.
- [8] R. N. Amalda, N. Millah, and I. Fitria, "Implementasi Algoritma C5.0 Dalam Menganalisa Kelayakan Penerima Keringanan Ukt Mahasiswa Itk," *Teorema Teor. dan Ris. Mat.*, vol. 7, no. 1, p. 101, 2022, doi: 10.25157/teorema.v7i1.6692.
- [9] M. S. Sungkar and M. T. Qurohman, "Penerapan Algoritma C5.0 Untuk Prediksi Kelulusan Pembelajaran Mahasiswa Pada Matakuliah Arsitektur Sistem Komputer," *J. Media Inform. Budidarma*, vol. 5, no. 3, p. 1166, 2021, doi: 10.30865/mib.v5i3.3116.
- [10] R. Pratiwi, M. N. Hayati, and S. Prangga, "Perbandingan Klasifikasi Algoritma C5.0 Dengan Classification and Regression Tree (Studi Kasus : Data Sosial Kepala Keluarga Masyarakat Desa Teluk Baru Kecamatan Muara Ancalong Tahun 2019)," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 14, no. 2, pp. 273–284, 2020, doi: 10.30598/barekengvol14iss2pp273-284.
- [11] A. Apriyadi, M. R. Lubis, and B. E. Damanik, "Penerapan Algoritma C5.0 Dalam Menentukan Tingkat Pemahaman Mahasiswa Terhadap Pembelajaran Daring," *Komputa J. Ilm. Komput. dan Inform.*, vol. 11, no. 1, pp. 11–20, 2022, doi: 10.34010/komputa.v11i1.7386.
- [12] N. T. Rahman, "Analisa Algoritma Decision Tree Dan Naïve Bayes Pada Pasien Penyakit Liver," *J. Fasilkom*, vol. 10, no. 2, pp. 144–151, 2020, doi: 10.37859/jf.v10i2.2087.
- [13] S. Suraji, A. C. Fauzan, and H. Harliana, "Penerapan Algoritma Decision Tree C5.0 Untuk Memprediksi Tingkat Kematian Pasien Penyakit Gagal Jantung," *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 4, no. 02, pp. 216–222, 2022, doi: 10.46772/intech.v4i02.682.