

# Penerapan Algoritma Decision Tree Untuk Menentukan Faktor-faktor yang Mempengaruhi Kelulusan Siswa di MTs Misykatul Ulum

Naqibuzzahidin<sup>\*1</sup>, Zaehol Fatah<sup>2</sup>, A. Hamdani<sup>3</sup>

<sup>1,2,3</sup>Teknologi Informasi, Sains dan Teknologi, Universitas Ibarahimy

Email: <sup>\*1</sup>[naqibuzzahidin22@gmail.com](mailto:naqibuzzahidin22@gmail.com), <sup>2</sup>[zaeholfatah@gmail.com](mailto:zaeholfatah@gmail.com), <sup>3</sup>[dan.kidz88@gmail.com](mailto:dan.kidz88@gmail.com)

(Naskah masuk: 22 Juni 2025, diterima untuk diterbitkan: 20 Oktober 2025)

**Abstrak:** Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor utama yang memengaruhi kelulusan siswa di MTs Misykatul Ulum dengan menerapkan algoritma Decision Tree. Latar belakang dari penelitian ini adalah pentingnya pemanfaatan data pendidikan dalam proses pengambilan keputusan yang lebih objektif dan terukur, khususnya dalam meningkatkan mutu akademik siswa. Penelitian menggunakan metode data mining dengan pendekatan Knowledge Discovery in Database (KDD), yang mencakup tahapan pemahaman masalah, seleksi data, preprocessing, pemodelan, evaluasi, dan implementasi sistem. Data yang digunakan berupa nilai rata-rata Ujian Akhir Madrasah dan tingkat kehadiran siswa. Model klasifikasi dibangun menggunakan pustaka Scikit-learn dan diimplementasikan ke dalam aplikasi web berbasis Streamlit. Hasil evaluasi model menunjukkan performa klasifikasi yang sangat baik dengan akurasi sebesar 100%, presisi dan recall sebesar 1.00. Sistem ini memungkinkan pihak sekolah melakukan prediksi kelulusan secara otomatis, menyimpan hasil ke database SQLite, serta menyediakan visualisasi pohon keputusan dan grafik analitik. Penelitian ini menunjukkan bahwa penerapan algoritma Decision Tree efektif dalam membantu sekolah menganalisis dan mengevaluasi faktor-faktor kelulusan siswa. Namun, generalisasi model masih perlu divalidasi dengan jumlah data yang lebih besar.

**Kata Kunci** – Penambangan data, Algoritma Pohon Keputusan, Kelulusan

## Implementation of Decision Tree Algorithm to Determine Factors Affecting Student Graduation at MTs Misykatul Ulum

**Abstract:** This study aims to identify the key factors influencing student graduation at MTs Misykatul Ulum by implementing the Decision Tree algorithm. The background of this research lies in the importance of utilizing educational data to support more objective and measurable decision-making processes, particularly in improving student academic performance. The study employs data mining techniques using the Knowledge Discovery in Database (KDD) approach, which includes problem understanding, data selection, preprocessing, modeling, evaluation, and system implementation. The dataset consists of students' average scores on final exams and attendance percentages. A classification model was developed using the Scikit-learn library and implemented into a web-based application using Streamlit. Evaluation results showed excellent classification performance with 100% accuracy, and both precision and recall reaching 1.00. The system enables schools to automatically predict student graduation, store results in a SQLite database, and provide decision tree visualization and analytical graphs. This research demonstrates that applying the Decision Tree algorithm is effective in helping schools analyze and evaluate graduation-related factors. However, further validation with a larger dataset is recommended to ensure better model generalization.

**Keywords** – Data mining, Algoritma Decision Tree, Graduation

### 1. PENDAHULUAN

Pendidikan merupakan salah satu faktor penting dalam pembangunan sumber daya manusia yang berkualitas. Pendidikan tidak hanya berfungsi sebagai sarana untuk meningkatkan pengetahuan dan keterampilan, tetapi juga sebagai landasan dalam membentuk karakter dan kepribadian individu. Keberhasilan pendidikan menjadi salah satu indikator utama dalam menentukan kualitas sumber daya manusia suatu bangsa. keberhasilan pendidikan sering diukur

dari berbagai aspek, salah satunya adalah tingkat kelulusan siswa. Tingkat kelulusan siswa mencerminkan sejauh mana lembaga pendidikan berhasil dalam memberikan pengajaran serta kemampuan siswa dalam menyerap dan memahami materi yang diajarkan[1].

Algoritma Decision Tree merupakan salah satu metode dalam data mining yang banyak digunakan untuk klasifikasi dan prediksi. Metode ini menyajikan hasil dalam bentuk pohon keputusan yang mudah dipahami, sehingga memudahkan pengguna dalam melakukan interpretasi. Dengan menerapkan algoritma Decision Tree, data kelulusan siswa di MTs Misykatul Ulum dapat dianalisis untuk mengidentifikasi faktor-faktor utama yang memengaruhi kelulusan, seperti nilai ujian akhir madrasah dan tingkat kehadiran siswa[2].

MTs Misykatul Ulum sebagai lembaga pendidikan Islam menaruh perhatian besar terhadap keberhasilan akademik siswa. Namun, masih terdapat kendala dalam mengidentifikasi faktor utama yang berkontribusi terhadap kelulusan siswa secara sistematis. Oleh karena itu, penerapan algoritma Decision Tree diharapkan dapat membantu pihak sekolah dalam menganalisis dan menentukan faktor-faktor yang paling berpengaruh terhadap kelulusan siswa. Dengan adanya sistem berbasis data ini, pihak sekolah dapat mengambil langkah-langkah strategis untuk meningkatkan kualitas pendidikan dan memastikan tingkat kelulusan yang lebih baik[3].

Selain itu, penggunaan algoritma Decision Tree dalam dunia pendidikan tidak hanya bermanfaat bagi sekolah, guru dan siswa. Guru dapat memperoleh wawasan lebih mendalam tentang bagaimana faktor-faktor tertentu memengaruhi performa siswa, sementara siswa dapat memahami aspek-aspek yang perlu ditingkatkan untuk mencapai kelulusan yang lebih baik. Dengan demikian, hasil penelitian ini diharapkan dapat menjadi panduan dalam meningkatkan mutu pendidikan di MTs Misykatul Ulum[4].

Dengan kemajuan teknologi yang terus berkembang, data yang sebelumnya hanya digunakan untuk keperluan administratif kini dapat dianalisis lebih mendalam untuk mendukung pengambilan keputusan yang lebih efektif. Penerapan algoritma Decision Tree dalam sistem pendidikan membuka peluang baru untuk meningkatkan proses evaluasi dan perencanaan akademik[5].

Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor utama yang mempengaruhi kelulusan siswa di MTs Misykatul Ulum serta mengevaluasi efektivitas algoritma Decision Tree dalam analisis data kelulusan. Dengan memahami faktor-faktor tersebut, pihak sekolah dapat mengembangkan strategi yang lebih efektif dalam mendukung kesuksesan akademik siswa. Oleh karena itu, penelitian ini memiliki nilai penting dalam meningkatkan kualitas pendidikan berbasis data dan teknologi.

## 2. METODE PENELITIAN

### 2.1. Tahapan Penelitian

Penelitian lapangan adalah metode penelitian yang dilakukan dilokasi atau lingkungan alami dari subjek yang diteliti, bukan di laboratorium atau tempat yang sudah diatur secara khusus. Tujuan utama dari penelitian lapangan adalah untuk mengumpulkan data yang mendetail dan konsektual dari sumbernya langsung, sehingga peneliti dapat memahami fenomena atau perilaku dalam konteks yang sebernarnya.[6]

#### 2.1.1. Teknik Pengumpulan Data

Tahapan pada pengumpulan data ini sebagai berikut:

##### 1. Observasi

Metode observasi merupakan teknik pengumpulan data yang digunakan untuk memperoleh informasi secara langsung dari siswa MTs Misykatul Ulum, khususnya terkait nilai ujian dan tingkat kehadiran. Pengamatan dilakukan secara sistematis dan disertai pencatatan terhadap data yang relevan dengan faktor-faktor yang diduga memengaruhi kelulusan. Melalui observasi ini, peneliti

dapat mengumpulkan data nyata di lapangan sebagai dasar dalam penerapan algoritma Decision Tree untuk menganalisis pola kelulusan siswa..[7] Melalui observasi ini, peneliti mengumpulkan dokumen-dokumen terkait yang diperlukan untuk analisis faktor-faktor yang mempengaruhi kelulusan siswa

## 2. Wawancara

Wawancara adalah metode pengumpulan data yang dilakukan melalui interaksi langsung antara peneliti dan responden.[8] Dalam hal ini guru atau staf administrasi MTs Misykatul Ulum, dengan tujuan memperoleh informasi yang dibutuhkan melalui proses tanya jawab. Proses ini memungkinkan peneliti untuk menggali data secara lebih mendalam mengenai hal-hal yang berkaitan dengan faktor-faktor yang memengaruhi kelulusan siswa, seperti kebijakan penilaian dan catatan kehadiran.

## 3. Studi Pustaka

Studi pustaka merupakan teknik pengumpulan data dengan menelaah berbagai literatur ilmiah yang relevan sebagai dasar teoritis. Metode ini dilakukan melalui pengumpulan dan penelaahan dokumen-dokumen seperti buku, jurnal, artikel ilmiah, dan sumber daring yang berkaitan dengan algoritma Decision Tree, data mining, serta faktor-faktor yang memengaruhi kelulusan siswa. Studi pustaka ini bertujuan untuk memperkuat landasan teori dan mendukung analisis dalam penelitian, dengan langkah-langkah yang mencakup pencarian, seleksi, dan pengkajian sumber informasi yang kredibel dan sesuai dengan fokus penelitian.[9]

## 4. Data Sekunder

Selain data yang diperoleh melalui observasi dan wawancara, penelitian ini juga memanfaatkan data sekunder yang bersumber dari pihak internal MTs Misykatul Ulum. Data ini mencakup informasi yang telah terdokumentasi sebelumnya oleh sekolah, seperti nilai rata-rata siswa, tingkat kehadiran, dan status kelulusan. Data tersebut diperoleh dari pihak sekolah, serta telah tersimpan dalam arsip administrasi sekolah. Pemanfaatan data sekunder ini bertujuan untuk mendukung proses analisis klasifikasi menggunakan algoritma Decision Tree. Karena bentuknya sudah terstruktur dan rapi, data ini sangat membantu dalam tahapan preprocessing, pelatihan model, dan evaluasi akurasi prediksi kelulusan.

### 2.2. Metode Pengembangan Sistem

Penelitian ini menggunakan pendekatan Knowledge Discovery in Database (KDD) dalam proses pengembangan sistem klasifikasi kelulusan siswa. KDD adalah proses sistematis untuk menemukan pola atau pengetahuan yang berguna dari data. Tahapan KDD yang digunakan dalam penelitian ini adalah sebagai berikut:

#### a. Pemahaman Masalah

Tahap ini bertujuan untuk memahami kebutuhan sekolah dalam mengetahui faktor-faktor yang memengaruhi kelulusan siswa secara objektif. Fokus utama adalah dua atribut penting yaitu nilai rata-rata dan kehadiran siswa.

#### b. Seleksi Data

Data yang digunakan merupakan data sekunder yang diperoleh dari MTs Misykatul Ulum, yaitu:

- Nilai rata-rata Ujian Akhir Madrasah (UAM)
- Presentase kehadiran Siswa
- Status kelulusan

#### c. Preprocessing Data

Data yang diperoleh dibersihkan dan disiapkan agar layak untuk digunakan dalam proses pemodelan. Proses ini mencakup:

- Penanganan data kosong
- Normalisasi data
- Transformasi nilai kedalam format numerik (contoh : 1 = lulus, 0 = tidak lulus)

#### d. Pemodelan

Pada tahap ini, algoritma Decision Tree diterapkan untuk membentuk model klasifikasi. Atribut-atribut yang sudah disiapkan digunakan untuk mempelajari pola yang berpengaruh terhadap kelulusan siswa. Model dibangun menggunakan pustaka Scikit-learn pada bahasa pemrograman Python.

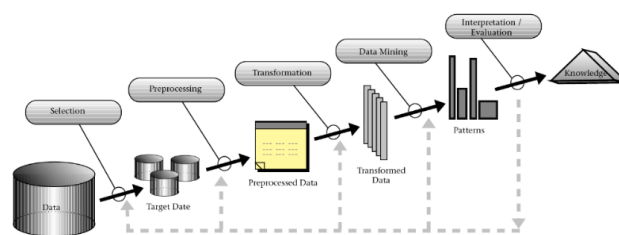
e. Evaluasi Model

Model dievaluasi menggunakan metrik:

- Akurasi: tingkat prediksi benar secara keseluruhan
- Presisi: proporsi prediksi "Lulus" yang benar
- Recall: proporsi siswa "Lulus" yang berhasil dikenali
- F1-Score: keseimbangan antara presisi dan recall

f. Implementasi Sistem

Model yang telah dilatih kemudian diimplementasikan ke dalam sistem berbasis web menggunakan Streamlit. Sistem ini memungkinkan pihak sekolah memasukkan nilai dan kehadiran siswa, lalu mendapatkan hasil klasifikasi kelulusan secara otomatis dan visual.[10]



Gambar 1. Knowledge Discovery in Database

## 2.3. Landasan Teori

### 2.3.1. Data Mining

Data mining adalah istilah yang merujuk pada proses menemukan pengetahuan tersembunyi dalam database. Proses ini dilakukan secara semi otomatis dengan memanfaatkan berbagai teknik, seperti statistik, matematika, kecerdasan buatan, dan pembelajaran mesin. Tujuannya adalah untuk mengekstraksi serta mengidentifikasi informasi atau pengetahuan yang potensial, berguna, dan bermanfaat dari kumpulan data besar yang tersimpan dalam database.[11]

### 2.3.2. Decision Tree

Pohon keputusan adalah salah satu metode dalam pembelajaran mesin yang digunakan untuk memetakan suatu observasi kedalam kelas atau kategori tertentu berdasarkan atribut yang dimilikinya. Model ini divisualisasikan dalam bentuk struktur pohon, di mana daun mewakili klasifikasi atau keputusan akhir, sedangkan cabang menunjukkan hubungan antar atribut yang berkontribusi dalam proses klasifikasi.[10]

Beberapa hal penting mengenai pohon keputusan:

- Termasuk dalam kategori pembelajaran terawasi dan dapat diterapkan untuk menyelesaikan masalah regresi maupun klasifikasi.
- Menggunakan struktur pohon untuk memecahkan masalah, di mana setiap simpul daun berisi kelas hasil klasifikasi dan setiap simpul internal mewakili atribut yang diuji.
- Pohon keputusan juga bisa digunakan untuk merepresentasikan fungsi Boolean dari atribut diskrit.

Entropi adalah ukuran ketidakpastian atau ketidakteraturan dalam suatu himpunan data yang terkait dengan variabel acak. Nilai entropi berkisar antara 0 hingga 1, di mana semakin tinggi entropinya, semakin besar ketidakpastian dalam data tersebut. Konsep ini digunakan dalam teori informasi untuk mengukur seberapa banyak informasi yang dimiliki suatu sistem. Entropi dapat dihitung menggunakan rumus berikut:

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

Dimana:

1.  $S$  adalah himpunan data yang di partisi berdasarkan suatu atribut.
2.  $N$  adalah jumlah kelas dalam data.
3.  $p_i$  adalah probabilitas suatu kelas  $i$  dalam dataset.

### 2.3.3. Machine Learning

Metode machine learning merupakan salah satu cabang dari kecerdasan buatan dan juga bagian dari bidang ilmu komputer yang berperan penting dalam pengolahan data secara otomatis. Pendekatan ini berfokus pada pemanfaatan data historis dan algoritma tertentu untuk meniru cara manusia belajar dari pengalaman. Dengan terus menerus menerima dan menganalisis data baru, sistem machine learning mampu meningkatkan kemampuan dan akurasi dalam mengenali pola, membuat prediksi, serta mengambil keputusan secara lebih tepat. Oleh karena itu, semakin baik dan tepat algoritma machine learning yang digunakan dalam suatu sistem, maka semakin tinggi pula kualitas hasil analisis dan keputusan yang dihasilkan oleh sistem tersebut.[12]

### 2.3.4. Kelulusan

Kelulusan adalah tahap akhir dalam proses pengelolaan peserta didik yang menjadi indikator bahwa seorang siswa telah berhasil menyelesaikan seluruh rangkaian program pendidikan yang ditempuhnya. Proses ini mencakup evaluasi terhadap pencapaian kompetensi, penyelesaian kurikulum, serta pemenuhan syarat-syarat akademik dan administratif yang ditetapkan oleh lembaga pendidikan. Dengan demikian, kelulusan tidak hanya menandai berakhirnya masa belajar di satuan pendidikan tertentu, tetapi juga menjadi bukti bahwa peserta didik telah siap untuk melanjutkan ke jenjang pendidikan yang lebih tinggi atau memasuki dunia kerja sesuai dengan kemampuan yang dimilikinya.[13]

### 2.3.5. Klasifikasi

Klasifikasi adalah metode yang digunakan untuk mengelompokkan data tertentu ke dalam kelas yang sudah ditentukan berdasarkan pola atau atribut yang dimiliki oleh data tersebut. Teknik ini semakin populer karena mampu menangani berbagai jenis data yang lebih luas dibandingkan dengan metode regresi. Tujuan utama klasifikasi adalah memprediksi kelas atau label pada data yang belum dikenal dengan memanfaatkan informasi dari data yang sudah di ketahui sebelumnya.[14]

## 3. HASIL DAN PEMBAHASAN

Implementasi sistem klasifikasi kelulusan siswa dilakukan dengan mengikuti tahapan data mining yang telah dirancang dalam metode penelitian. Penelitian ini menggunakan data siswa MTs Misykatul Ulum tahun ajaran 2024/2025 dengan jumlah total 13 data siswa. Atribut yang digunakan meliputi nilai rata-rata Ujian Akhir Madrasah (UAM), persentase kehadiran, dan status kelulusan siswa.

NO	Nama Siswa	Nilai Rata-Rata	Kehadiran %	Kelulusan
1	DURROTUL HASANAH	86.08	98,00%	Lulus
2	HILMAH WILDATUL INAYAH	71.08	85,00%	Lulus
3	INDAH ROMADHANI	74.46	88,00%	Lulus
4	KHOIROTUN NAFI'AH	80.08	92,00%	Lulus
5	ANA MUTIAH	80.46	93,00%	Lulus
6	MAISYAROH NABILA ARIFIN	91.92	100,00%	Lulus
7	MOH. ALI WAFI	88.54	99,00%	Lulus
8	MUHAMMAD LUTFI	75.33	85,00%	Lulus
9	NOVITA PURNAMA SARI	72.77	86,00%	Lulus
10	OLIVIA DURROTIS ZAKIYA	85.83	96,00%	Lulus
11	PUSPITA SARI	89.31	98,00%	Lulus
12	SITI KHOTIMAH	70.92	84,00%	Lulus
13	SULAIMAN	91.92	98,00%	Lulus

Gambar 2. Data Siswa

Dalam penelitian ini, seluruh siswa pada dataset asli dinyatakan lulus. Oleh karena itu, untuk memungkinkan pembelajaran klasifikasi oleh algoritma Decision Tree, ditambahkan sejumlah data simulasi yang merepresentasikan siswa tidak lulus dengan memperhatikan batas nilai minimum kelulusan dan tingkat kehadiran.

### 3.1. Import Library dan Data

Dalam penelitian ini, seluruh siswa pada dataset asli dinyatakan lulus. Oleh karena itu, untuk memungkinkan pembelajaran klasifikasi oleh algoritma Decision Tree, ditambahkan sejumlah data simulasi yang merepresentasikan siswa tidak lulus berdasarkan batas nilai kelulusan dan tingkat kehadiran yang ditetapkan.

```
1 import streamlit as st
2 import pandas as pd
3 import numpy as np
4 import sqlite3
5 from sklearn.tree import DecisionTreeClassifier
6 from sklearn import tree
7 import matplotlib.pyplot as plt
8 from sklearn.metrics import classification_report, accuracy_score, confusion_matrix
9
10
11 |
12 |
13 | data = {
14 |     'nilai_rata_rata': [78, 65, 85, 60, 88, 70, 75, 76, 74, 80],
15 |     'kehadiran': [95, 80, 90, 70, 98, 85, 80, 82, 79, 90],
16 |     'kelulusan': [1, 0, 1, 0, 1, 0, 1, 1, 0, 1]
17 | }
18 | df = pd.DataFrame(data)
19 | X = df[['nilai_rata_rata', 'kehadiran']]
20 | y = df['kelulusan']
```

Gambar 3. Kode Import Data

### 3.2. Preprocessing dan Training

Tahapan preprocessing dan training merupakan langkah penting dalam membangun model klasifikasi kelulusan siswa menggunakan algoritma Decision Tree. Data yang digunakan terdiri atas tiga atribut, yaitu nilai rata-rata ujian akhir madrasah (nilai rata-rata), persentase kehadiran (kehadiran), dan status kelulusan (kelulusan). Seluruh data disusun dalam bentuk DataFrame menggunakan pustaka pandas.

Langkah-langkah *preprocessing* yang dilakukan meliputi:

- Pembuatan Data Frame**  
Data dikonversi ke dalam struktur DataFrame agar dapat digunakan dalam pelatihan model. Semua atribut berupa data numerik, sehingga tidak diperlukan konversi lanjutan.
- Pemisahan Fitur dan Label**  
Atribut nilai rata-rata dan kehadiran digunakan sebagai fitur input (X), sedangkan kelulusan digunakan sebagai label output (y).
- Pembentukan Model**

Model klasifikasi dibentuk menggunakan algoritma DecisionTreeClassifier dari pustaka scikit-learn, dengan parameter `criterion='entropy'` untuk pengukuran pemisahan data terbaik dan `random_state=42` guna memastikan reproduibilitas. Berikut adalah cuplikan kode yang merepresentasikan tahap preprocessing dan training:

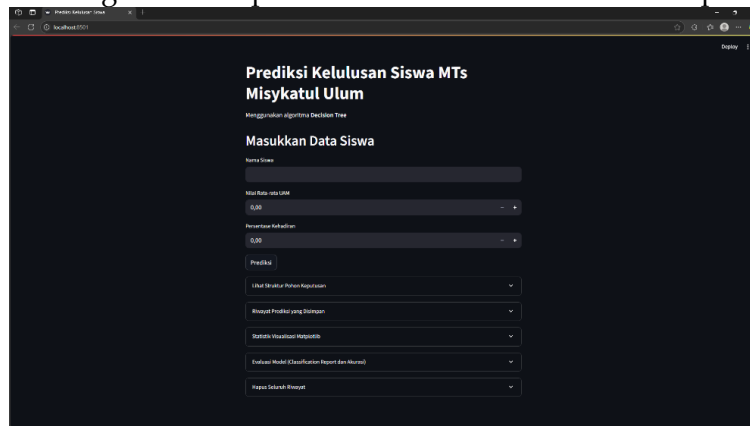
```
12 data = {  
13     'nilai_rata_rata': [78, 65, 85, 60, 88, 70, 75, 76, 74, 80],  
14     'kehadiran': [95, 80, 90, 70, 98, 85, 80, 82, 79, 90],  
15     'kelulusan': [1, 0, 1, 0, 1, 0, 1, 1, 0, 1]  
16 }  
17 df = pd.DataFrame(data)  
18 X = df[['nilai_rata_rata', 'kehadiran']]  
19 y = df['kelulusan']  
20 model = DecisionTreeClassifier(criterion='entropy', random_state=42)  
21 model.fit(X, y)
```

Gambar 4. Kode Pembentukan Model

Setelah proses pelatihan selesai, model siap digunakan untuk melakukan prediksi terhadap data baru serta dievaluasi untuk mengukur kinerja klasifikasinya. Tahap ini menjadi dasar sebelum integrasi ke dalam sistem berbasis web.

### 3.3. Visualisasi dan Implementasi Sistem

Setelah model klasifikasi kelulusan berhasil dilatih menggunakan algoritma Decision Tree, langkah selanjutnya adalah mengimplementasikan model tersebut dalam sistem berbasis web. Framework Streamlit digunakan untuk membangun antarmuka pengguna yang interaktif dan ringan, serta memudahkan guru atau operator sekolah dalam melakukan prediksi kelulusan siswa.



Gambar 5. Antarmuka

Sistem memungkinkan pengguna untuk memasukkan data siswa berupa:

- Nama siswa
- Nilai rata-rata UAM
- Presentase kehadiran



Gambar 6. Input dan Prediksi

Setelah data diinput, sistem akan memprosesnya menggunakan model Decision Tree yang telah dilatih, lalu menampilkan hasil prediksi dalam bentuk label.

Sistem juga menampilkan tingkat probabilitas dari prediksi menggunakan metode predict\_proba dari scikit-learn.

#### Penyimpanan ke Database SQLite

Untuk mendokumentasikan hasil prediksi, sistem terhubung dengan basis data lokal SQLite. Tiap hasil prediksi disimpan ke dalam tabel hasil-prediksi.



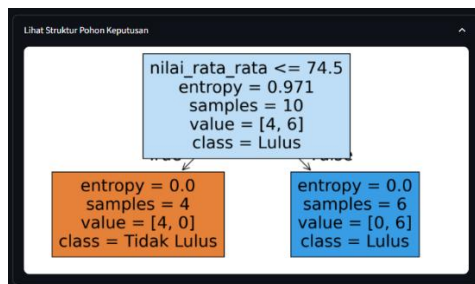
ID	Nama	Nilai Rata-rata	Kehadiran (%)	Hasil
0	22 NAQIBUZZAHIDIN	75	80	LULUS
1	21 NAQIBUZZAHIDIN	71	80	TIDAK LULUS
2	20 NAQIBUZZAHIDIN	70	80	TIDAK LULUS

Gambar 7. Penyimpanan Hasil Prediksi

Hal ini memungkinkan pihak sekolah untuk melakukan pelacakan riwayat prediksi serta mengevaluasi hasil-hasil sebelumnya secara kolektif.

#### Visualisasi Pohon Keputusan

Sebagai bentuk interpretabilitas model, sistem menyediakan fitur untuk menampilkan struktur pohon keputusan yang dibangun. Visualisasi ini divisualisasikan menggunakan fungsi plot\_tree dari pustaka matplotlib dan sklearn.tree, dengan pengkodean warna berdasarkan kelas hasil.



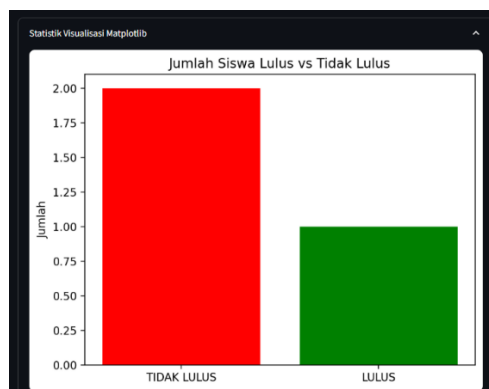
Gambar 8. Visualisasi Pohon Keputusan

#### Statistik dan Grafik Visualisasi

Sistem juga menyertakan dua bentuk grafik utama:

##### 1. Grafik Batang

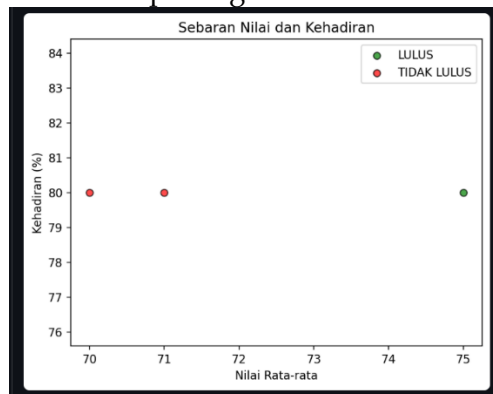
Menampilkan perbandingan jumlah siswa yang diklasifikasikan sebagai lulus dan tidak lulus.



Gambar 9. Grafik Batang

##### 2. Grafik Sebaran

Menampilkan penyebaran nilai dan kehadiran terhadap hasil prediksi, dengan pewarnaan berbeda untuk setiap kategori.



Gambar 10. Grafik Sebaran

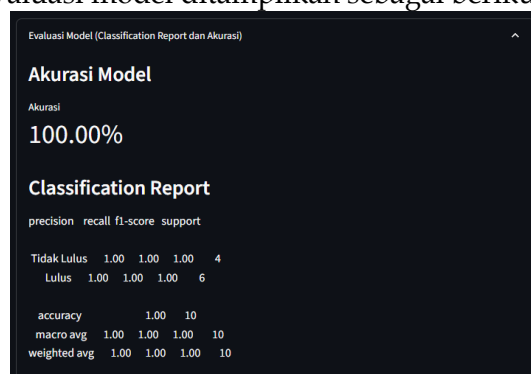
Fitur-fitur ini memberikan gambaran menyeluruh kepada pengguna dalam menganalisis hubungan antara nilai dan kehadiran terhadap kemungkinan kelulusan siswa secara visual dan intuitif.

### 3.4. Evaluasi Model

Evaluasi model dilakukan untuk mengukur sejauh mana model Decision Tree mampu mengklasifikasikan kelulusan siswa dengan akurat. Evaluasi ini dilaksanakan menggunakan metrik umum klasifikasi, yaitu akurasi, presisi, recall, dan F1-score..

Model dievaluasi menggunakan data yang sama seperti data pelatihan (karena jumlah data terbatas). Prediksi dibandingkan dengan label aktual untuk menghasilkan *classification report* dan *confusion matrix*. Proses ini dilakukan menggunakan fungsi bawaan *scikit-learn*, yaitu *accuracy\_score*, *classification\_report*, dan *confusion\_matrix*.

Berikut adalah hasil evaluasi model ditampilkan sebagai berikut:



Gambar 11. Hasil Pengukuran

Model mampu mengklasifikasikan semua data dengan benar, baik untuk kategori Lulus maupun Tidak Lulus, tanpa terjadi kesalahan klasifikasi. Nilai presisi, recall, dan F1-score untuk kedua kelas adalah sempurna (1.00). Hal ini menunjukkan bahwa model memiliki performa yang sangat optimal, setidaknya pada data yang tersedia saat ini. Namun demikian, perlu diingat bahwa ukuran dataset masih terbatas (hanya 10 data), sehingga meskipun hasil evaluasi menunjukkan akurasi 100%, hal ini perlu divalidasi lebih lanjut dengan data yang lebih besar dan beragam agar model lebih generalisasi dan tahan terhadap data baru.

## 4. KESIMPULAN

Berdasarkan hasil penelitian dan implementasi sistem prediksi kelulusan siswa menggunakan algoritma Decision Tree di MTs Misykatul Ulum, maka dapat disimpulkan beberapa hal berikut:

1. Model klasifikasi kelulusan siswa berhasil dikembangkan dengan menggunakan algoritma Decision Tree berbasis atribut nilai rata-rata Ujian Akhir Madrasah (UAM) dan persentase kehadiran siswa.
2. Proses pengembangan sistem menggunakan pendekatan Knowledge Discovery in Database (KDD) yang terdiri dari tahapan pemahaman masalah, seleksi data, preprocessing, pemodelan, evaluasi, hingga implementasi sistem berbasis web menggunakan Streamlit.
3. Model klasifikasi yang dihasilkan menunjukkan performa sangat baik dengan akurasi sebesar 100%, presisi dan recall sebesar 1.00 untuk masing-masing kelas (Lulus dan Tidak Lulus), serta tidak terjadi kesalahan klasifikasi pada data uji.
4. Implementasi sistem prediksi berbasis web memungkinkan pihak sekolah melakukan klasifikasi kelulusan siswa secara otomatis dan cepat, serta menyimpan hasil prediksi ke dalam basis data SQLite sebagai dokumentasi.
5. Visualisasi dalam sistem seperti pohon keputusan dan grafik sebaran nilai-kehadiran memudahkan pengguna dalam memahami pola kelulusan siswa, sehingga dapat dijadikan bahan pertimbangan dalam evaluasi akademik.
6. Meskipun hasil evaluasi menunjukkan akurasi yang tinggi, penelitian ini masih menggunakan jumlah data yang terbatas. Oleh karena itu, pengujian lebih lanjut dengan jumlah data yang lebih besar sangat disarankan untuk memastikan validitas dan generalisasi model.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terimakasih kepada:

1. Dosen pembimbing dan seluruh civitas akademika Universitas Ibrahimi atas bimbingan, masukan, dan dorongan yang sangat berarti dalam penyusunan jurnal ini.
2. MTs Misykatul Ulum yang telah memberikan izin, dukungan, serta menyediakan data yang dibutuhkan selama proses penelitian ini berlangsung.
3. Seluruh pihak yang tidak dapat disebutkan satu per satu yang telah membantu secara langsung maupun tidak langsung dalam proses pelaksanaan dan penyelesaian penelitian ini.

Semoga hasil dari penelitian ini dapat memberikan kontribusi positif bagi pengembangan sistem informasi pendidikan, khususnya dalam pengambilan keputusan berbasis data di lingkungan sekolah.

#### DAFTAR PUSTAKA

- [1] N. Nailil Amani, M. Martanto, and U. Hayati, "Penggunaan Algoritma Decision Tree Untuk Prediksi Prestasi Siswa Di Sekolah Dasar Negeri 3 Bayalangu Kidul," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 473–479, 2024, doi: 10.36040/jati.v8i1.8355.
- [2] F. R. Al-Giffary and M. Martanto, "Klasifikasi Kelulusan Siswa Tahun 2024 Menggunakan Metode Decision Tree (Studi Kasus Sma Islam Alazhar 5 Cirebon)," *J. Manajemen Inform. Jayakarta*, vol. 4, no. 2, p. 195, 2024, doi: 10.52362/jmijayakarta.v4i2.1408.
- [3] S. Kasus, U. D. Karunia, and S. Utama, "Journal Of Industrial Engineering And Technology ( Jointech ) UNIVERSITAS MURIA KUDUS Analisis Postur Kerja pada Stasiun Kerja Proses Som Kaos Kaki," *J. Ind. Eng. Technol. Univ. MURIA KUDUS*, vol. 1, no. 2, pp. 52–60, 2021.
- [4] N. K. Augustianingrum and A. C. Padmasari, "Implementasi Algoritma Decision Tree Boardgame Dwipantara sebagai Media Pembelajaran Sejarah Kerajaan untuk Meningkatkan Minat Belajar Siswa SD," *Edsence J. Pendidik. Multimed.*, vol. 2, no. 1, pp. 57–64, 2020, doi: 10.17509/edsence.v2i1.25036.
- [5] Y. Susanti, M. G. Choyyin, A. Priyatna, and S. Lestari, "Perbandingan Penerapan Algoritma Decision Tree C.45 dan Naïve Bayes Dalam Analisa Kelulusan Siswa pada SMK Swadhipa 2

- Natar Kabupaten Lampung Selatan," *J. SIMADA (Sistem Inf. dan Manaj. Basis Data)*, vol. 6, no. 2, pp. 117-123, 2023, doi: 10.30873/simada.v6i2.3772.
- [6] D. Haryanti, "Pengelolaan Kelas Inklusi Melalui Metode Belajar Bersama Alam ( MBBA ) di Sekolah Alam Bangka Belitung," vol. 7, no. 2, pp. 1-9, 2020.
- [7] P. Hasibuan, R. Azmi, D. B. Arjuna, and S. U. Rahayu, "Analisis Pengukuran Temperatur Udara Dengan Metode Observasi Analysis of Air Temperature Measurements Using the Observational Method," *ABDIMASJurnal Garuda Pengabd. Kpd. Masy.*, vol. 1, no. 1, pp. 8-15, 2023, [Online]. Available: <http://creativecommons.org/licenses/by-sa/4.0/>
- [8] Q. FAUZIAH, "Penerapan Metode Wawancara Narasumber Untuk Meningkatkan Kemampuan Menulis Teks Tanggapan," *Lang. J. Inov. Pendidik. Bhs. dan Sastra*, vol. 3, no. 2, pp. 77-83, 2023, doi: 10.51878/language.v3i2.2305.
- [9] Nur Latifah, A. Marini, and A. Maksum, "Pendidikan Multikultural di Sekolah Dasar (Sebuah Studi Pustaka)," *J. Pendidik. Dasar Nusantara*, vol. 6, no. 2, pp. 42-51, 2021, doi: 10.29407/jpdn.v6i2.15051.
- [10] A. Wanto *et al.*, "Data Mining : Algoritma dan Implementasi," *Yayasan Kita Menulis*. 2023
- [11] A. Nugraha, O. Nurdiawan, and G. Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 849-855, 2022, doi: 10.36040/jati.v6i2.5755.
- [12] I. M. Faiza, G. Gunawan, and W. Andriani, "Tinjauan Pustaka Sistematis: Penerapan Metode Machine Learning untuk Deteksi Bencana Banjir," *J. Minfo Polgan*, vol. 11, no. 2, pp. 59-63, 2022, doi: 10.33395/jmp.v11i2.11657.
- [13] F. Umi, S. Marsidin, and A. Sabandi, "Analisis Kebijakan Dan Pengelolaan terkait Peserta Didik Di Sekolah Dasar," vol. 2, no. 2, pp. 128-133, 2020.
- [14] A. Yogiarto, A. Homaidi, and Z. Fatah, "Implementasi Metode K-Nearest Neighbors (KNN) untuk Klasifikasi Penyakit Jantung," *G-Tech J. Teknol. Terap.*, vol. 8, no. 3, pp. 1720-1728, 2024, doi: 10.33379/gtech.v8i3.4495.
- [15] Bakiatus Sarifah, Rokhman Fauzi, and Iqbal Santosa, "Analisis Dan Perancangan Proses Manajemen Sistem Kontrol Internal Ti Menggunakan Kerangka Kerja Cobit 2019 Di Pt Inti (Persero)," *e-Proceeding Eng.*, vol. 70, no. 2, pp. 67-43, 2020.