

**ANALISIS POLA BELAJAR SISWA MENGGUNAKAN DATA
MINING DENGAN ALGORITMA DECISION TREE C4.5
UNTUK MENINGKATKAN HASIL BELAJAR
BAHASA INGGRIS**

SKRIPSI

Untuk Memenuhi Persyaratan Memperoleh Gelar Sarjana (S1) Program Studi
Sistem Informasi Fakultas Sains dan Teknologi Universitas Labuhanbatu



OLEH :

**VIVI INDRIANI HASIBUAN
2209100134**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS LABUHANBATU
RANTAU PRAPAT
2026**

LEMBAR PERSETUJUAN SKRIPSI

JUDUL SKRIPSI : ANALISIS POLA BELAJAR SISWA
MENGUNAKAN DATA MINING DENGAN
ALGORITMA DECISION TREE C4.5 UNTUK
MENINGKATKAN HASIL BELAJAR BAHASA
INGGRIS

NAMA : VIVI INDRIANI HASIBUAN

NPM : 2209100134

PRODI : SISTEM INFORMASI

Disetujui pada tanggal : 02 April 2026

Pembimbing I



Maras Nasution, S.Kom., M.Kom

NIDN. 0130039001

Pembimbing II



Asriani Hasibuan, S.Pd., M.Hum

NIDN. 0116078802

LEMBAR PENGESAHAN SKRIPSI

JUDUL SKRIPSI : ANALISIS POLA BELAJAR SISWA
MENGUNAKAN DATA MINING DENGAN
ALGORITMA DECISION TREE C4.5 UNTUK
MENINGKATKAN HASIL BELAJAR BAHASA
INGGRIS

NAMA : VIVI INDRIANI HASIBUAN
NPM : 2209100134
PRODI : SISTEM INFORMASI

Telah Diuji dan Dinyatakan Lulus Dalam Uji Sarjana Pada
Tanggal 02 April 2026

TIM PENGUJI

Pembimbing I (Ketua)

Nama : Marnis Nasution, S.Kom., M.Kom
NIDN : 0130039001

Pembimbing II (Anggota)

Nama : Asriani Hasibuan, S.Pd., M.Hum
NIDN : 0116078802

Penguji (Anggota)

Nama : Angga Putra Juledi, S.Kom., M.Kom
NIDN : 0119079401

Tanda Tangan

()

()

()

Rantauprapat, 02 April 2026

Dekan
Fakultas Sains dan Teknologi



Assoc. Prof. Dr. Iwan Purnama, S.Kom., M.Kom
NIDN: 012029202

Ka. Program Studi
Sistem Informasi



Budianto Bangun, S.Sos., M.Kom
NIDN: 0124047003

PERNYATAAN

JUDUL SKRIPSI : ANALISIS POLA BELAJAR SISWA
MENGUNAKAN DATA MINING DENGAN
ALGORITMA DECISION TREE C4.5 UNTUK
MENINGKATKAN HASIL BELAJAR BAHASA
INGGRIS

NAMA : VIVI INDRIANI HASIBUAN

NPM : 2209100134

PRODI : SISTEM INFORMASI

Dengan ini penulis menyatakan bahwa skripsi ini disusun sebagai syarat untuk memperoleh gelar Sarjana pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Labuhanbatu adalah hasil karya tulis penulis sendiri. Semua kutipan maupun rujukan dalam penulisan skripsi ini telah penulis cantumkan sumbernya dengan benar sesuai dengan ketentuan yang berlaku. Jika ada kemudian hari ternyata ditemukan seluruh atau sebagian skripsi ini bukan hasil karya penulis atau plagiat, penulis bersedia menerima sanksi pencabutan gelar akademik yang disandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Rantauprapat, 02 April 2026

Yang Membuat Pernyataan ,



VIVI INDRIANI HASIBUAN

2209100134

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pola belajar siswa dalam mata pelajaran Bahasa Inggris menggunakan teknik Data Mining dengan algoritma Decision Tree C4.5. Data yang digunakan meliputi variabel absensi, kelas tambahan, penggunaan media digital, dan nilai siswa. Metode penelitian yang digunakan mengikuti tahapan Knowledge Discovery in Database (KDD), yaitu pengumpulan data, preprocessing, pengolahan data, serta evaluasi model. Hasil penelitian menunjukkan bahwa faktor absensi, penggunaan media digital, dan keikutsertaan dalam kelas tambahan memiliki pengaruh signifikan terhadap hasil belajar siswa, dengan absensi sebagai faktor paling dominan. Model yang dihasilkan berupa pohon keputusan mampu menggambarkan hubungan antar variabel secara jelas dan mudah dipahami. Evaluasi model menunjukkan tingkat akurasi yang cukup baik, yaitu sekitar 83% hingga 85%, sehingga model dinilai mampu melakukan klasifikasi dengan baik. Hasil analisis ini dapat dimanfaatkan sebagai dasar dalam merancang strategi pembelajaran yang lebih efektif, adaptif, dan berbasis data guna meningkatkan hasil belajar Bahasa Inggris siswa.

Kata kunci: Data Mining, Decision Tree C4.5, Pola Belajar, Hasil Belajar, Bahasa Inggris

ABSTRACT

This study aims to analyze students' learning patterns in English subjects using Data Mining techniques with the Decision Tree C4.5 algorithm. The data used include attendance, additional classes, use of digital media, and student grades. The research method follows the Knowledge Discovery in Database (KDD) stages, including data collection, preprocessing, data processing, and model evaluation. The results show that attendance, the use of digital media, and participation in additional classes significantly influence students' learning outcomes, with attendance being the most dominant factor. The resulting model in the form of a decision tree is able to clearly describe the relationships between variables and is easy to interpret. The model evaluation shows a fairly good accuracy level of approximately 83% to 85%, indicating that the model performs well in classification. The findings can be utilized as a basis for developing more effective, adaptive, and data-driven learning strategies to improve students' English learning outcomes.

Keywords: Data Mining, Decision Tree C4.5, Learning Patterns, Learning Outcomes, Englis

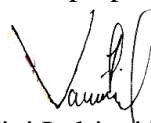
KATA PENGANTAR

Dengan menyebut nama Allah SWT Yang Maha Pengasih lagi Maha Penyayang, penulis mengucapkan puja dan puji syukur ke hadirat-Nya sehingga Skripsi penelitian berjudul “ANALISIS POLA BELAJAR SISWA MENGGUNAKAN DATA MINING DENGAN ALGORITMA DECISION TREE C4.5 UNTUK MENINGKATKAN HASIL BELAJAR BAHASA INGGRIS” ini dengan baik

dan tepat waktu. Penulis menyampaikan terima kasih kepada seluruh pihak yang telah memberikan dukungan, khususnya kepada:

1. Bapak Alm. Dr. H. Amarullah Nasution, MBA selaku pendiri Yayasan Universitas Labuhanbatu
2. Bapak Dr. Halomoan Nasution, S.H., M.H selaku ketua yayasan Universitas Labuhanbatu.
3. Bapak Rektor Universitas Labuhanbatu, Bapak Assoc, Prof. Ade Parlaungan Nasution, Ph.D
4. Bapak Assoc. Prof. Dr. Iwan Purnama, S.Kom., M.Kom selaku Dekan Fakultas Sains dan Teknologi.
5. Bapak Budianto Bangun, S.Sos., M.Kom selaku Kepala Program Studi Sistem Informasi
6. Ibu Marnis Nasution, S.Kom., M.Kom selaku Dosen Pembimbing I.
7. Ibu Asriani Hasibuan, S.pd., M.Hum selaku Dosen Pembimbing II.
8. Bapak Angga Putra Juledi, S.Kom., M.Kom selaku Dosen Penguji.

Rantauprapat, Maret 2026



Vivi Indriani Hasibuan

DAFTAR ISI

LEMBAR PERSETUJUAN SKRIPSI.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
PERNYATAAN.....	iii
ABSTRAK.....	iv
ABSTRACT.....	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Dan Manfaat Penelitian.....	4
1.4.1 Tujuan Penelitian.....	4
1.4.2 Manfaat Penelitian.....	5
1.5 Tinjauan Umum Objek Penelitian.....	6
1.6 Sistematika Penulisan.....	7
BAB II LANDASAN TEORI.....	9
2.1 Data.....	9
2.2. Knowledge Discovery in Database.....	11
2.2.1 Data Mining.....	12
2.3 Algoritma Decision Tree C4.5.....	13
2.3.1 Langkah – Langkah Algoritma Decision Tree C4.5.....	16
2.4 Alat Bantu Aplikasi.....	18
2.4.1 RapidMiner.....	18
2.4.2 Microsoft Excel.....	20
2.5 Pola Belajar.....	21
2.6 Hasil Belajar.....	23

2.7 Penelitian Terdahulu.....	25
BAB III METODELOGI PENELITIAN	28
3.1 Kerangka Penelitian	28
3.2 Populasi dan Sampel	29
3.2.1 Populasi	29
3.2.2 Sampel	29
3.3 Lokasi dan Waktu Penelitian	30
3.3.1 Lokasi Penelitian.....	30
3.3.2 Waktu Penelitian.....	30
3.4 Variabel Penelitian	30
3.5 Langkah – Langkah Pengelolaan Algoritma C.45.....	31
3.5.1 Pengumpulan Data	31
3.5.2 Preprocessing Data.....	33
3.5.3 Metrik Evaluasi.....	47
BAB IV HASIL DAN PEMBAHASAN	52
4.1 Analisis Data	52
4.2 Pengumpulan Data.....	52
4.3 Preprocessing Data.....	56
4.4 Pengolahan Data Menggunakan Rapidminer	63
4.4.1 Input Data.....	63
4.4.2 Perancangan dan Penambahan Operator pada RapidMiner.....	64
4.5 Pohon Keputusan	66
4.6 Analisis Faktor-Faktor yang Mempengaruhi Hasil Belajar	67
4.7 Pemanfaatan Hasil Analisis Untuk Strategi Pembelajaran	68
4.8 Hasil Evaluasi.....	69
BAB V KESIMPULAN DAN SARAN.....	71
5.1 Kesimpulan	71
5.2 Saran.....	72
DAFTAR PUSTAKA	73

DAFTAR TABEL

Tabel 2.1. Penelitian Terdahulu	25
Tabel 3.1 Variabel penelitian	31
Tabel 3.2 Sampel Data Mentah	32
Tabel 3.3 Dataset Hasil Preprocessing.....	33
Tabel 3.4 Data Training.....	34
Tabel 3.5 Data Testing.....	36
Tabel 3.6 Perhitungan Entropy, Information Gain, Split Information, dan Gain Ratio.....	43
Tabel 3.7 Root Node 1 Absensi	44
Tabel 3.8 Node 1.1 Media Digital	45
Tabel 3.9 Hasil prediksi data testing.....	48
Tabel 3.10 Confusion Matrix.....	48
Tabel 4.1 Data Sampel Penelitian.....	53
Tabel 4.2 Data Training	56
Tabel 4.3 Data Testing.....	59

DAFTAR GAMBAR

Gambar 2.1. Knowledge Discovery in Database (KDD)	11
Gambar 2.2. Tampilan Awal Aplikasi Rapidminer	19
Gambar 2.3. Tampilan Awal Microsoft Excel	20
Gambar 3.1. Kerangka Penelitian.....	31
Gambar 3.2. Pohon Keputusan.....	46
Gambar 4.1. Input Data	63
Gambar 4.2. Penambahan Operator Read Excel, Set Role dan Cross Validation ...	64
Gambar 4.3. Penambahan Operator Decision Tree, Apply Model dan Performance.....	65
Gambar 4.4. Pohon Keputusan.....	66
Gambar 4.5. Hasil Evaluasi	69