

**ANALISIS *DATA MINING* UNTUK MEMPREDIKSI PRESTASI AKADEMIK
SISWA MENGGUNAKAN ALGORITMA J48**

SKRIPSI

**Disusun untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana**



Disusun Oleh :

**ARIS AFRIANTO NUGROHO
2000018220**

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN**

2024

ANALYSIS TO PREDICT STUDENTS' ACADEMIC ACHIEVEMENT USING J48 ALGORITHM

S1 THESIS

Submitted as a Partial Fulfilment of Requirements to Obtain Bachelor's Degree



Written by:

ARIS AFRIANTO NUGROHO
2000018220

**INFORMATICS STUDY PROGRAM
FACULTY OF INDUSTRIAL TECHNOLOGY
UNIVERSITAS AHMAD DAHLAN**

2024

**ANALISIS *DATA MINING* UNTUK MEMPREDIKSI PRESTASI AKADEMIK
SISWA MENGGUNAKAN ALGORITMA J48**

SKRIPSI



Disusun Oleh:

ARIS AFRIANTO NUGROHO
2000018220

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN**

2024

LEMBAR PERSETUJUAN PEMBIMBING

SKRIPSI

**ANALISIS *DATA MINING* UNTUK MEMPREDIKSI PRESTASI AKADEMIK
SISWA MENGGUNAKAN ALGORITMA J48**

Dipersiapkan dan disusun oleh:

**ARIS AFRIANTO NUGROHO
2000018220**

**Program Studi S1 Informatika
Fakultas Teknologi Industri
Universitas Ahmad Dahlan**

Telah disetujui oleh:

Pembimbing



Lisna Zahrotun, S.T., M.Cs.

NIPM. 19840911 200909 011 1058758

LEMBAR PENGESAHAN

SKRIPSI

ANALISIS DATA MINING UNTUK MEMPREDIKSI PRESTASI AKADEMIK SISWA MENGUNAKAN ALGORITMA J48

Dipersiapkan dan disusun oleh:

ARIS AFRIANTO NUGROHO
2000018220

Telah dipertahankan di depan Dewan Penguji
pada tanggal 19 Juni 2024
dan dinyatakan telah memenuhi syarat

Susunan Dewan Penguji

Ketua : Lisna Zahrotun, S.T., M.Cs.

Penguji 1 : Fiftin Noviyanto, S.T., M.Cs.

Penguji 2 : Ir. Nur Rochmah D. P. A., S.T., M.Kom.

Yogyakarta, 3 Juli 2024

Dekan, Fakultas Teknologi Industri
Universitas Ahmad Dahlan



Prof. Dr. Ir. Siti Jamilatun, M.T.

NIPM. 19660812 199601 011 0784324

LEMBAR PERNYATAAN KEASLIAN

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Aris Afrianto Nugroho
NIM : 2000018220
Prodi : Informatika
Judul TA/Skripsi : Analisis *Data Mining* Untuk Memprediksi Prestasi Akademik
Siswa Menggunakan Algoritma J48

Dengan ini saya menyatakan bahwa Laporan Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar Ahli Madya/Kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 21 Juni 2024

Mengetahui,
Dosen Pembimbing



Lisna Zahrotun, S.T., M.Cs.
NIPM. 19840911 200909 011 1058758

Yang menyatakan,



Aris Afrianto Nugroho
2000018220

PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : Aris Afrianto Nugroho
NIM : 2000018220
Email : aris2000018220@webmail.uad.ac.id
Program Studi : S1 Informatika
Fakultas : Teknologi Industri
Judul Tesis : Analisis *Data Mining* Untuk Memprediksi Prestasi Akademik Siswa Menggunakan Algoritma J48

Dengan ini Saya menyerahkan hak sepenuhnya kepada Perpustakaan Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tesis elektronik sebagai berikut (beri tanda pada kotak):



Saya (~~mengijinkan~~/~~tidak mengijinkan~~)* karya tersebut diunggah ke dalam aplikasi Repository Perpustakaan Universitas Ahmad Dahlan.

Demikian pernyataan ini Saya buat dengan sebenarnya.

Yogyakarta, 21 Juni 2024

Yang Menyatakan

Aris Afrianto Nugroho

Mengetahui,

Dosen Pembimbing Skripsi

Lisna Zahrotun, S.T., M.Cs.

MOTTO

“Prosesnya memang tidak mudah tetapi pada akhirnya membuat tidak berhenti mengucapkan

Alhamdulillah”

"Bentuk takdirmu sendiri, dan jangan biarkan keputusan orang lain mengubahnya.”

"Di antara nikmat terindah adalah memiliki orang-orang yang peduli di sekitar kita, hargailah mereka sebagaimana mereka menghargaimu."

“Orang lain ga akan bisa paham *struggle* dan masa sulitnya kita, yang mereka ingin tahu hanya

bagian *success stories*. Berjuanglah untuk diri sendiri walaupun gak ada yang tepuk tangan.

Kelak diri kita di masa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini.

Tetap berjuang yaa!!!”

HALAMAN PERSEMBAHAN

Dengan rasa syukur kepada Allah SWT atas segala rahmat dan karunia-Nya, saya persembahkan skripsi ini kepada orang-orang yang sangat berarti dalam hidup saya. Skripsi ini merupakan hasil dari kerja keras dan dedikasi yang tidak terlepas dari dukungan, doa, dan kasih sayang yang saya terima selama ini.

Skripsi ini saya persembahkan kepada :

1. **Allah SWT**, yang telah memberikan rahmat, hidayah, dan kekuatan dalam setiap langkah perjalanan hidup dan studi saya. Segala puji dan syukur hanya kepada-Mu, ya Allah.
2. **Orang tua** saya yang sangat saya cintai, **Bapak Muryanto** dan **Ibu Ngatminah**. Terimakasih yang sebesar-sebesarnya atas segala doa, kasih sayang, dukungan, dan pengorbanan yang tidak ternilai harganya. Semoga Allah SWT selalu memberikan perlindungan dan memberikan kebahagiaan kepada kalian. Aamiin.
3. Kakak saya, **Sartika Novitasari** dan **Susanto** serta Adik saya, **Shabira Naura Ilmira** yang sangat saya sayangi, yang selalu memberikan dukungan dan motivasi serta menjadi sumber inspirasi dan kebahagiaan. Terimakasih atas kebersamaan dan kasih sayang yang kalian berikan.
4. **Ibu Lisna Zahrotun, S.T., M.Cs.**, yang telah meluangkan waktunya untuk memberikan arahan, bimbingan serta dukungan sehingga skripsi ini dapat terselesaikan. Terimakasih atas ilmu yang telah Ibu berikan.
5. Kepala SMPN 2 Jogonalan, **Bapak Drs. Jaka Siswanta**, yang telah memberikan kesempatan kepada saya untuk melakukan penelitian. Saya mengucapkan terimakasih sehingga penelitian ini dapat berjalan dengan lancar.

6. Sahabat-sahabat saya, **Arif Wicaksono Gegadang Putro, Ahmad Fajar, Rafif Nugie Indrasta**, yang selalu ada di sisi saya dalam suka dan duka. Terimakasih atas kebersamaan dan canda tawa selama ini. Semoga kesuksesan dan lindungan Allah SWT senantiasa membersamai kita. Aamiin.
7. Teman saya **Hanif Mahmudy Ibrahim**, yang telah memberikan motivasi untuk menyelesaikan skripsi ini. Semoga pertemanan kita terus terjaga dan menjadi berkah bagi kita kedepannya.
8. **Teman-teman seperjuangan**, yang telah bersama-sama menjalani proses belajar dan perjuangan ini. Terimakasih atas kerja sama dan semangat yang diberikan selama masa studi.
9. **Teman-teman angkatan 2020** yang tidak bisa saya sebutkan satu persatu, terimakasih atas waktu dan kebersamaannya dalam 4 tahun ini.
10. **Almamater tercinta**, yang telah memberikan kesempatan dan fasilitas belajar untuk berkembang. Terimakasih atas lingkungan akademik yang mendukung dan penuh inspirasi.

KATA PENGANTAR

Puji dan syukur penulis panjatkan atas kehadiran Allah S.W.T, atas limpahan rahmat dan hidayah-Nya penulis dapat menyelesaikan skripsi yang berjudul “Analisis *Data Mining* Untuk Memprediksi Prestasi Akademik Siswa Menggunakan Algoritma J48” sebagai syarat untuk menyelesaikan studi jenjang strata satu Program Studi Informatika Universitas Ahmad Dahlan.

Dalam penyusunan skripsi tentunya tidak terlepas dari arahan serta bimbingan dari berbagai pihak mulai dari tahap persiapan hingga terselesaikannya skripsi ini. Oleh karena itu penulis mengucapkan terimakasih kepada :

1. Bapak Prof. Dr. Muchlas, M.T., selaku Rektor Universitas Ahmad Dahlan.
2. Ibu Prof. Dr. Ir. Siti Jamilatun, M.T., selaku Dekan Fakultas Teknologi Industri Universitas Ahmad Dahlan.
3. Bapak Dr. Murinto, S.Si., M.Kom., selaku Kepala Program Studi Informatika Fakultas Teknologi Industri Universitas Ahmad Dahlan.
4. Bapak Muslihudin, S.T., M.T., selaku Dosen Pembimbing Akademik yang senantiasa memberikan arahan, dorongan dan motivasi.
5. Ibu Lisna Zahrotun. S.T., M.Cs., selaku Dosen Pembimbing yang senantiasa sabar membimbing dan mengarahkan penulis.
6. Bapak dan Ibu Dosen Program Studi S1 Informatika Universitas Ahmad Dahlan.
7. Teman-teman seperjuangan serta semua pihak yang tidak dapat disebutkan satu-persatu.

Penulis menyadari bahwa karya tulis ini masih terdapat banyak kekurangan, oleh karena itu penulis mengharapkan kritik dan saran yang membangun agar laporan ini menjadi lebih baik.

Yogyakarta, Juni 2024

Aris Afrianto Nugroho

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERNYATAAN KEASLIAN	iv
PERNYATAAN PERSETUJUAN AKSES.....	v
MOTTO	vi
HALAMAN PERSEMBAHAN.....	vii
KATA PENGANTAR.....	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
DAFTAR KODE PROGRAM.....	xiv
DAFTAR LAMPIRAN	xv
DAFTAR SINGKATAN DAN ARTI LAMBANG	xvi
ABSTRAK	xvii
ABSTRACT	xviii
BAB I. PENDAHULUAN	1
1.1. Latar belakang	1
1.2. Batasan Masalah	4
1.3. Rumusan Masalah	4
1.4. Tujuan Penelitian.....	5
1.5. Manfaat Penelitian	5
BAB II. KAJIAN PUSTAKA	6
2.1. Kajian Penelitian Terdahulu	6
2.2. Landasan Teori	12
2.2.1. <i>Data Mining</i>	12
2.2.2. <i>Knowledge Discovery in Databases (KDD)</i>	13
2.2.3. Pengertian Prediksi.....	15
2.2.4. Pengertian Klasifikasi.....	16
2.2.5. Pengertian Prestasi Akademik.....	16
2.2.6. Pengertian Nilai Sumatif Tengah Semester.....	16
2.2.7. Algoritma J48.....	17
2.2.8. <i>Stratified K-Fold Cross Validation</i>	19
2.2.9. <i>Confusion Matrix</i>	21
2.2.10. Contoh Kasus.....	22
BAB III. METODOLOGI PENELITIAN	45
3.1. Objek Penelitian	45
3.2. Kebutuhan Data	45
3.3. Metode Pengumpulan Data	46
3.4. Definisi Operasional	47
3.5. Analisis Kebutuhan.....	51
3.5.1. Perangkat Lunak.....	51
3.5.2. Perangkat Keras.....	52
3.6. Tahapan Penelitian.....	52
3.6.1. Studi Literatur.....	53
3.6.2. Pengumpulan Data	54

3.6.3.	Analisis dan Persiapan Data	54
3.6.4.	Pengolahan Data Mining	54
3.6.5.	Hasil dan Pembahasan	57
3.6.6.	Kesimpulan dan Saran	57
BAB IV.	HASIL DAN PEMBAHASAN	58
4.1.	Hasil Pengumpulan Data	58
4.2.	Analisis dan Persiapan Data	59
4.2.1.	Integrasi Data	59
4.2.2.	Deskripsi Data.....	59
4.2.3.	Data Awal	69
4.2.4.	Pendeklarasian Library	72
4.2.5.	Load Data	74
4.3.	Pengolahan Data Mining	75
4.3.1.	<i>Selection Data</i>	76
4.3.2.	<i>Cleaning Data</i>	77
4.3.3.	<i>Transformation Data</i>	78
4.3.4.	<i>Modelling Pohon Keputusan Menggunakan Algoritma J48</i>	79
4.3.5.	<i>Pengujian Akurasi</i>	85
4.3.6.	<i>Interpretation</i>	86
4.4.	Hasil dan Pembahasan	88
4.4.1.	Analisis Rule.....	88
4.4.2.	Analisis Pengujian Akurasi.....	91
4.4.3.	Rekomendasi Tindakan	97
BAB V.	KESIMPULAN	101
5.1.	Kesimpulan.....	101
5.2.	Saran.....	102
DAFTAR PUSTAKA.....		103
LAMPIRAN		106
Lampiran 1.	Kuesioner Penelitian	106

DAFTAR GAMBAR

Gambar 2.1. Tahapan KDD	13
Gambar 2.2. <i>Stratified K-Fold Cross Validation</i>	19
Gambar 2.3. Evaluasi <i>Confusion Matrix</i>	21
Gambar 2.4. Pohon Awal.....	34
Gambar 2.5. Pohon Setelah <i>Node</i> Cabang Pertama Ditemukan.....	37
Gambar 2.6. Pohon Keputusan Yang Terbentuk.....	40
Gambar 3.1. Tahapan Penelitian.....	53
Gambar 4.1. Data Jumlah Responden	60
Gambar 4.2. Data Skorsing.....	61
Gambar 4.3. Data Ketepatan Mengerjakan Tugas	61
Gambar 4.4. Data Ketepatan Kedatangan Siswa	62
Gambar 4.5. Data Dukungan Belajar	63
Gambar 4.6. Data Memikirkan Cita-cita.....	64
Gambar 4.7. Data Organisasi/Ekstrakurikuler.....	64
Gambar 4.8. Data Terpengaruh Lingkungan	65
Gambar 4.9. Data Kondisi Lingkungan Sekolah.....	66
Gambar 4.10. Data Kepemilikan KIP	66
Gambar 4.11. Data Kondisi Finansial Orang Tua.....	67
Gambar 4.12. Data Status Prestasi Siswa.....	68
Gambar 4.13. Hasil Pemuatan Data	75
Gambar 4.14. Hasil Seleksi Data.....	76
Gambar 4.15. Menghitung Jumlah <i>Missing Value</i> Setiap Atribut.....	77
Gambar 4.16. Hasil Pembersihan Data	78
Gambar 4.17. Hasil Transformasi Data	79
Gambar 4.18. Grafik Pengujian <i>Max_Depth</i> dan <i>StratifiedKFold Cross Validation</i>	80
Gambar 4.19. Visualisasi Pohon Keputusan.....	83
Gambar 4.20. Hasil Pengujian <i>Stratified K-Fold Cross Validation</i>	85
Gambar 4.21. Hasil Penghitungan <i>Confusion Matrix</i>	86
Gambar 4.22. Hasil Penghitungan <i>Classification Report</i>	86
Gambar 4.23. Hasil Penghitungan <i>Accuracy Score</i>	86

DAFTAR TABEL

Tabel 2.1. Perbandingan Peneliti Terdahulu	9
Tabel 2.2. Data Uji Penelitian	23
Tabel 2.3. Data Hasil Proses Seleksi	24
Tabel 2.4. Data Hasil Proses <i>Cleaning</i>	26
Tabel 2.5. Data Hasil Proses Transformasi	27
Tabel 2.6. Hasil Penghitungan <i>Entropy</i> dan <i>Gain Node</i> Awal.....	32
Tabel 2.7. Penghitungan Data <i>Node</i> Cabang Pertama	34
Tabel 2.8. Penghitungan <i>Entropy</i> Dan <i>Gain Node</i> Cabang Pertama	35
Tabel 2.9. Penghitungan <i>Entropy</i> Dan <i>Gain Node</i> Cabang Kedua Bagian I	37
Tabel 2.10. Penghitungan <i>Entropy</i> Dan <i>Gain Node</i> Cabang Kedua Bagian II	38
Tabel 3.1. Definisi Operasional.....	47
Tabel 4.1. Integrasi Data	59
Tabel 4. 2 Data Awal Penelitian.....	70
Tabel 4.3. Penghitungan <i>Precision</i>	94
Tabel 4.4. Penghitungan <i>Recall</i>	95
Tabel 4.5. Penghitungan F1-Score.....	95
Tabel 4.6. Penghitungan Support.....	96

DAFTAR KODE PROGRAM

Kode Program 4.1. Pendeklarasian <i>Library</i>	74
Kode Program 4.2. <i>Load Data</i>	75
Kode Program 4.3. Seleksi Data	76
Kode Program 4.4. Menampilkan Jumlah <i>Missing Value</i> Setiap Atribut	77
Kode Program 4.5. Pembersihan Data	78
Kode Program 4.6. Transformasi Data	79
Kode Program 4.7. Pemisahan Data Fitur Dan Target	79
Kode Program 4.8. Pengolahan Data Menggunakan Algoritma J48	81
Kode Program 4.9. Hasil Penghitungan Akurasi	85

DAFTAR LAMPIRAN

Lampiran 1. Kuesioner Penelitian	106
--	-----

DAFTAR SINGKATAN DAN ARTI LAMBANG

CM	: Confusion Matrix
KDD	: Knowledge Discovery in Databases
KIP	: Kartu Indonesia Pintar
OECD	: Organisation for Economic Co-operation and Development.
PISA	: Programme for International Student Assessment
SMK	: Sekolah Menengah Kejuruan
SMP	: Sekolah Menengah Pertama
STS	: Sumatif Tengah Semester

ABSTRAK

Indonesia masih menghadapi berbagai tantangan dalam sektor pendidikan. Salah satu permasalahan yang menonjol adalah kurangnya tindakan preventif di tingkat sekolah untuk mengatasi potensi penurunan prestasi akademik siswa. Banyak institusi pendidikan di jenjang Sekolah Menengah Pertama yang belum menerapkan tindakan proaktif sejak dini untuk mengidentifikasi siswa-siswa yang berisiko mengalami kesulitan belajar. Sehingga perlu dilakukan suatu tindakan untuk mengatasi permasalahan tersebut salah satunya dengan memprediksi prestasi siswa agar sekolah dapat menentukan strategi yang harus dipersiapkan untuk menghadapi permasalahan yang mungkin terjadi di masa yang akan datang.

Penelitian ini bertujuan untuk memprediksi prestasi akademik siswa menggunakan algoritma J48. Metode pengumpulan data yang digunakan yakni wawancara, kuesioner, dan dokumentasi sementara subjek penelitian adalah siswa kelas 7 dan 8 SMPN 2 Jogonalan tahun ajaran 2023/2024. Variabel yang digunakan dalam penelitian ini meliputi kedisiplinan, motivasi belajar, lingkungan, dan status ekonomi orang tua sebagai data fitur dan nilai STS siswa sebagai data target. Beberapa tahap terlibat dalam pembangunan model, yakni pengumpulan data, analisis dan persiapan data, pengolahan data mining yang meliputi proses *selection*, *cleaning*, *transformation*, *modelling* pohon keputusan menggunakan algoritma J48, pengujian akurasi, *interpretation*, lalu tahap pemaparan hasil dan pembahasan dan yang terakhir adalah kesimpulan dan saran. Teknik pengujian akurasi yang digunakan dalam penelitian ini adalah menggunakan *stratified k-fold cross validation* dan *confusion matrix*.

Berdasarkan hasil pengujian yang telah dilakukan terhadap model yang telah dibangun, dilakukan pengujian menggunakan *stratified k-fold cross validation* dengan 10-fold dan didapatkan hasil rata-rata akurasinya sebesar 79,66%, dimana hasil pengujian dengan menggunakan *confusion matrix* diperoleh akurasi sebesar 78%. Ini berarti 78 dari 100 prediksi yang dibuat oleh model adalah benar. Dalam hal ini membuktikan bahwa hasil prediksi dengan menggunakan algoritma J48 berhasil dan cukup efisien sehingga dapat dijadikan pedoman oleh SMPN 2 Jogonalan dalam menentukan strategi yang harus disiapkan oleh sekolah untuk menghadapi masalah yang mungkin terjadi di masa depan.

Kata Kunci : Prediksi; Prestasi Akademik; Algoritma J48;

ABSTRACT

Indonesia still faces various challenges in the education sector. One prominent issue is the lack of preventive measures at the school level to address the potential decline in students' academic performance. Many educational institutions at the junior high school level have not implemented proactive measures early on to identify students at risk of learning difficulties. So it is necessary to take action to overcome these problems, one of which is by predicting student achievement so that schools can determine strategies that must be prepared to deal with problems that may occur in the future.

This research aims to predict student academic achievement using the J48 algorithm. The data collection methods used are interviews, questionnaires, and documentation while the research subjects are 7th and 8th grade students of SMPN 2 Jogonalan in the 2023/2024 school year. The variables used in this study include discipline, learning motivation, environment, and parents' economic status as feature data and students' STS scores as target data. Several stages are involved in model building, namely data collection, data analysis and preparation, data mining processing which includes the process of selection, cleaning, transformation, decision tree modeling using j48 algorithm, accuracy testing, interpretation, then the stage of presenting results and discussion and the last is conclusions and suggestions. The accuracy testing technique used in this research is using stratified k-fold cross validation and confusion matrix.

Based on the test results that have been carried out on the model that has been built, testing using stratified k-fold cross validation with 10-fold and obtained an average accuracy of 79.66%, where the test results using confusion matrix obtained an accuracy of 78%. This means that 78 out of 100 predictions made by the model are correct. In this case, it proves that the prediction results using the J48 algorithm are successful and efficient enough so that it can be used as a guide by SMPN 2 Jogonalan in determining the strategies that must be prepared by the school to deal with problems that may occur in the future.

Keywords: Prediction; Academic Achievement; Algorithm J48;