



**PENERAPAN ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENJURUSAN SISWA**

SKRIPSI

**BRAMEFIO QIBRAN HUSAINI
181410109**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BINA DARMA PALEMBANG
2023**



**PENERAPAN ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENJURUSAN SISWA**

**BRAMEFIO QIBRAN HUSAINI
181410109**

**Skripsi ini diajukan sebagai syarat memperoleh gelar
Sarjana Komputer/AhliMadya***

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BINA DARMA PALEMBANG
2023**

HALAMAN PENGESAHAN

**PENERAPAN ALGORITMA DECISION TREE UNTUK KLASIFIKASI
PENJURUSAN SISWA**

Oleh:

**BRAMEFIO QIBRAN HUSAINI
181410109**

Telah diajukan sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada Program Studi Sistem Informasi

Pembimbing


Dr. H. Jemakmun, M.Si

**Palembang
Fakultas Sains Teknologi
Universitas Bina Darma**
Dekan


Dr. Tata Sutabri, S.Kom, MMSI, MKM

HALAMAN PERSETUJUAN

Skripsi Berjudul "**PENERAPAN ALGORITMA DECISION TREE UNTUK KLASIFIKASI PENJURUSAN SISWA**" Oleh **BRAMEFIO QIBRAN HUSAINI (181410109)** telah dipertahankan didepan komisi pengaji

Komisi Pengaji

1. Ketua : DR. H. Jemakmun, M.Si
2. Anggota : Andri, M.Cs
3. Anggota : Linda Atika, M.Kom

Mengetahui,
Program Studi Sistem Informasi
Fakultas Sains Teknologi
Universitas Bina Darma
Ketua,



Nita Rosa Damayanti., M.Kom., Ph.D

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini :

Nama : BRAMEFIO QIBRAN HUSAINI

Nim : 181410109

Dengan ini menyatakan bahwa:

1. Skripsi yang menjadi karya tulis saya adalah asli, serta belum pernah saya ajukan untuk mendapat gelar akademik (sarjana), baik di Universitas Bina Darma atau di perguruan tinggi lainnya;
2. Karya tulis yang saya buat adalah murni suatu gagasan, rumusan dan penelitian yang saya lakukan dengan arahan dari tim pembimbing;
3. Dalam karya tulis ini, saya tidak mencantumkan karya serta pendapat yang sudah ditulis atau dipublikasikan oleh orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar rujukan;
4. Saya bersedia skripsi ini di periksa keasliannya dengan menggunakan Plagiarism Checkers dan diunggah ke internet, sehingga memungkinkan untuk diakses secara daring;
5. Surat pernyataan ini saya buat dengan sebenar-benarnya dan apabila terbukti ditemukan penyimpangan atau ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan peraturan dan perundang-undangan yang berlaku.

Demikianlah surat pernyataan yang saya buat ini agar dapat dipergunakan sebagaimana mestinya.

Palembang,
Yang Membuat Pernyataan,



BRAMEFIO QIBRAN HUSAINI
181410109

MOTTO DAN PERSEMBAHAN

Motto

“ Be proud of yourself, even though there are those who don't like you sometimes they hate because they can't be like you”

Persembahan

“Dan dia bersama kamu di mana kamu berada. Dan Allah maha melihat apa yang kamu kerjakan”
(QS. Al Hadid : 4)

“Ilmu itu tidak didapatkan dari jasad yang santai”
(HR Muslim)

“Hal yang paling menyakitkan di dunia adalah ketika kita tidak bisa membahagiakan orang yang kita sayangi”
(Abdhy)

“Hidup ini seperti secangkir kopi. Dimana pahit dan manis melebur, bertemu dalam kehangatan”
(Dee Lestari)

“Skripsi ini aku persembahkan untuk Ayahanda Alm.Suyono dan Ibunda Eni Darwanti , adik-adikku, my-girlfriend, saudaraku dan keluarga besarku serta orang-orang yang aku cintai”

ABSTRACT

SMA N 1 Air Kumbang is a school that has computerized data storage, one of which is student majors data. The majors for SMA N 1 Air Kumbang students are carried out at the beginning of the school year in accordance with the 2013 curriculum which has 2 majors, namely Natural Sciences and Social Sciences, which aims to channel the abilities and interests of these students. In determining student majors, so far the process of classifying majors has not been implemented and the accuracy is also unknown. This study aims to classify student majors and apply data mining methods using the C4.5 decision tree algorithm and also the Rapidminner application. This classification uses 6 attributes as control attributes which are the average grades of junior high school report cards which include subjects B. Indonesian, B. English, Mathematics and Science, as well as recommendations for junior high school counseling teachers. Based on tests carried out on students in the 2022/2023 school year with a data sample of ± 400 data with the attributes of the recommendations of Middle School Guidance and Guidance teachers that influence student major decisions. The results obtained were an accuracy of 84.27% and a classification error of 15.73%. From the results of the accuracy of 84.27% is in the range of 0.80-0.90 and these results are included in the Good Classification category. Thus it can be concluded that determining majors using data mining decision tree algorithm C4.5 can speed up and accurately make major decisions for SMA N 1 Air Kumbang students.

Keyword : Datamining, Classification, Decision Tree, Algorithm C4.5

ABSTRAK

SMA N 1 Air Kumbang merupakan sekolah yang penyimpanan datanya telah terkomputerisasi, salah satunya data penjurusan siswa. Penjurusan siswa SMA N 1 Air Kumbang dilakukan pada awal tahun pelajaran sesuai dengan kurikulum 2013 yang memiliki 2 jurusan yaitu IPA dan IPS, yang bertujuan untuk menyalurkan kemampuan dan minat bakat siswa tersebut. Dalam menentukan penjurusan siswa selama ini belum diterapkannya proses klasifikasi jurusan dengan akurasi nya juga belum diketahui. Penelitian ini bertujuan untuk mengklasifikasi penjurusan siswa dan menerapkan metode data mining menggunakan algoritma *decision tree C4.5* dan juga aplikasi Rapidminner. Hasil dari penelitian ini yaitu nilai *accuracy*, dan *classification error* dari setiap kelas dilihat dari tingkat kemampuan pemrosesan sistem dalam mencari tingkat keputusan dalam menentukan jurusan bagi siswa SMA Negeri 1 Air Kumbang, dengan nilai *accuracy* yang didapatkan dari prediksi untuk mengukur tingkat akurasi adalah sebesar 84.27% dan *classification error* sebesar 15.73%. Dari hasil akurasi sebesar 84.27% berada pada nilai rentang 0,80-0,90 dan hasil ini masuk kategori ***Good Classification***, maka peneliti menyimpulkan bahwa hasil prediksi jurusan siswa menggunakan RapidMiner dengan metode Algoritma C45 dapat dikatakan baik. Demikian dapat disimpulkan bahwa menentukan penjurusan dengan menggunakan data mining algoritma *decision tree C4.5* dapat mempercepat dan akurat dalam pengambilan keputusan penjurusan siswa SMA N 1 Air Kumbang.

Kata kunci : Data Mining, Klasifikasi, Decision Tree, Algoritma C4.5

KATA PENGANTAR



Alhamdulillah puji syukur saya ucapkan atas kehadirat Allah SWT sebab atas rahmat dan karunia-Nya, skripsi ini dapat saya selesaikan untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan di program studi sistem informasi Universitas Bina Darma, Palembang. Skripsi ini jelas masih jauh dari kata sempurna, sebab adanya keterbatasan pengetahuan yang penulis miliki. Untuk itu, dalam rangka memperlengkap kesempurnaan dari penulisan karya tulis ini, penulis berharap adanya akan saran dan kritik yang bersifat membangun. Pada kesempatan yang baik ini dengan segala kerendahan hati penulis menyampaikan terima kasih sebesar-besarnya kepada :

1. Allah SWT dan junjungan Nabi Muhammad SAW, berkat Rahmat dan Ridho-nya penulis dapat menyelesaikan studi ini.
2. Kedua orang tuaku tercinta (Ayah Alm.Suyono dan Ibu Eni Darwanti), adik saya (Anggi Rangga Ainun Nadjid dan Haidar Anggoro Isnaini) yang tiada henti selalu memberikan doa, semangat, nasehat, masukan yang telah kalian berikan, terimakasih atas pengorbanan dan jerih payah kalian selama ini serta support yang tiada terhingga disetiap perjalanan hidup ini untuk menyelesaikan perkuliahan dan juga skripsi ini.
3. Ibu Dr. Sunda Ariana, M.Pd., M.M selaku Rektor Universitas Bina Darma.
4. Bapak Dr. Tata Sutabri, S.Kom. MMSI. MKM selaku Dekan Fakultas Ilmu Komputer Universitas Bina Darma.
5. Ibu Dr. Desy Misnawati, S.Sos., M.I.Kom selaku Ketua Prodi Sistem Informasi Universitas Bina Darma.
6. Bapak Dr. H. Jemakmun, M.Si selaku dosen Pembimbing atas seluruh bantuan, ide, bimbingan, doa, nasihat yang telah diberikan kepada penulis selama penyusunan skripsi hingga selesai. Semoga kebaikan selalu menyertai bapak.
7. Bapak Andri, M.Cs dan Ibu Linda Atika, M.Kom selaku dosen pembimbing atas seluruh bantuan ide, bimbingan, doa dan nasihat yang telah diberikan kepada penulis selama penyusunan skripsi hingga selesai. Semoga kebaikan selalu menyertai bapak dan ibu.

8. Partner penulis dalam segala hal yaitu my girlfriend (Riza Fadhilah) sebagai support system dengan mendoakan serta memberikan nasehat kepada penulis dari awal hingga selesai pengskripsian.
9. Semua saudaraku, seluruh teman juga sahabat yang telah saling memberikan dorongan, saran serta masukan yang membangun. Sukses selalu dimanapun kalian berada.

Palembang, 10 April 2023



BRAMEFIO QIBRAN HUSAINI
181410109

DAFTAR ISI

HALAMAN PENGESAHAN.....	Error! Bookmark not defined.
HALAMAN PERSETUJUAN.....	Error! Bookmark not defined.
SURAT PERNYATAAN.....	Error! Bookmark not defined.
MOTTO DAN PERSEMBAHAN	Error! Bookmark not defined.
ABSTRACT.....	Error! Bookmark not defined.
ABSTRAK.....	Error! Bookmark not defined.
KATA PENGANTAR.....	Error! Bookmark not defined.
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN	Error! Bookmark not defined.
1.1 Latar Belakang.....	Error! Bookmark not defined.
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian.....	4
1.5 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 Landasan Teori.....	7
2.1.1 Data Mining.....	7
2.1.2 Decision Tree	Error! Bookmark not defined.
2.1.3 Algoritma C45.....	17
2.1.4 Klasifikasi Algoritma C45 (Kurva ROC)	18
2.1.5 Rapidminer	20
2.2 Penelitian Terdahulu.....	22
BAB III METODOLOGI PENELITIAN	25
3.1 Data Selection	25

3.1.1 Penjelasan Atribut Data	26
3.2 Preprocessing	27
3.3 Transformation	Error! Bookmark not defined.
3.3.1 Menentukan atribut akhir yang akan di proses	30
3.3.2 Inisialisasi atribut indikator	31
3.4 Penerapan Algoritma Decision Tree C45.....	31
BAB IV HASIL DAN PEMBAHASAN	37
4.1 Hasil	37
4.1.1 Data.....	37
4.1.2 Atribut.....	39
4.2 Pembahasan.....	39
4.2.1 Pembentukan Model	39
4.2.2 Pengujian Tingkat Kesalahan (<i>accuracy</i>)	47
4.2.3 Pengujian <i>Classification Error</i>	47
4.2.4 Pengujian Tingkat Kesalahan (<i>root mean square error</i>)	48
4.2.5 <i>Confusion Matrix</i>	49
BAB V KESIMPULAN DAN SARAN	50
5.1 Kesimpulan	50
5.2 Saran.....	51
DAFTAR PUSTAKA	52

DAFTAR TABEL

Tabel 2.1 Kategori Nilai Kurva Roc	19
Tabel 3.1 Penjelasan Atribut Data	31
Tabel 3.2 Atribut Akhir	35
Tabel 3.3 Daftar Calon Mutakhir	37
Tabel 3.4 Daftar Calon Mutakhir	38
Tabel 3.5 Nilai Candidate Split Purity Left (PL) dan Purity Right (PR)	38
Tabel 3.6 Label Jurusan	38
Tabel 3.7 Perhitungan Nilai Q (s/t)	39
Tabel 3.8 Perhitungan Nilai PL dan PR	40
Tabel 3.9 Data Uji	41
Tabel 3.10 Data Hasil Uji Perhitungan	41
Tabel 4.1 Data Atribut	44

DAFTAR GAMBAR

Gambar 2.1 Tahapan Dalam KDD	10
Gambar 2.2 Struktur Pohon Keputusan	15
Gambar 3.1 Data Siswa	31
Gambar 3.2 Dataset Hasil Integrasi Data	32
Gambar 3.3 Dataset Missing Value	33
Gambar 3.4 Hasil Cleaning Data Siswa	34
Gambar 3.5 Atribut Akhir untuk di Proses Lanjut.....	35
Gambar 3.6 Hasil Inisialisasi	36
Gambar 3.7 Data Latih Penelitian	37
Gambar 3.8 Pohon Keputusan.....	40
Gambar 4.1 Excel Hasil Proses Pengolahan Data	42
Gambar 4.2 Excel Hasil Data yang akan dilakukan Uji	43
Gambar 4.3 Memasukan Data Pembentukan Model	44
Gambar 4.4 Hasil Data Pembentukan Model	45
Gambar 4.5 Membagi Data Menggunakan Split Data	45
Gambar 4.6 Membagi Data Training Training dan Testing	46
Gambar 4.7 Data Training 80%	47
Gambar 4.8 Data Testing 20%	47
Gambar 4.9 Penggunaan Operator Decision Tree C45	48
Gambar 4.10 Pengaturan Parameter Operator Decision Tree C45	48
Gambar 4.11 Penggunaan Operator Apply Model	49
Gambar 4.12 Penggunaan Operator Performance	49
Gambar 4.13 Hasil Akhir Analisa Model	50
Gambar 4.14 Hasil Analisis Prediksi Data Decision Tree C45	51
Gambar 4.15Hasil Pohon Keputusan.....	51
Gambar 4.16 Nilai Accuracy	52

Gambar 4.17 Hasil Classification Error	53
Gambar 4.18 Nilai Root Mean Square Error	53
Gambar 4.19 Confusion Matrix	54

