

**IMPLEMENTASI ALGORITMA C4.5 UNTUK
MEMPREDIKSI PENJURUSAN
SISWA SMAK YOS SUDARSO**

SKRIPSI



Oleh:
Ricky
140210066

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2018**

**IMPLEMENTASI ALGORITMA C4.5 UNTUK
MEMPREDIKSI PENJURUSAN
SISWA SMAK YOS SUDARSO**

SKRIPSI

Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana



Oleh:
Ricky
140210066

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2018**

PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (sarjana, dan/atau magister), baik di Universitas Putera Batam maupun di perguruan tinggi lain.
2. Skripsi ini adalah murni gagasan, rumusan, dan penelitian saya sendiri, tanpa bantuan pihak lain, kecuali arahan pembimbing.
3. Dalam skripsi ini tidak terdapat karya atau pendapat yang telah ditulis atau dipublikasikan orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di perguruan tinggi.

Batam, 04 Agustus 2018
Yang membuat pernyataan,

Materai Rp 6.000

Ricky
140210066

**IMPLEMENTASI ALGORITMA C4.5 UNTUK
MEMPREDIKSI PENJURUSAN
SISWA SMAK YOS SUDARSO**

**Oleh:
Ricky
140210066**

**SKRIPSI
Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini**

Batam, 4 Agustus 2018

**Koko Handoko, S.Kom., M.Kom.
Pembimbing**

ABSTRAK

Pendidikan merupakan aspek penting dalam kehidupan manusia. Pada dunia pendidikan, saat siswa mencapai tingkat menengah atas, siswa akan dihadapkan dengan penjurusan. Penjurusan merupakan suatu proses pemilihan jurusan perkuliahan yang dilakukan siswa-siswi pada masa sekolah tingkat menengah atas (SMA). Penjurusan dibagi menjadi IPA dan IPS yang bertujuan agar siswa-siswi dapat mendapatkan pendidikan yang sesuai dengan kemampuan dan minat mereka. Penjurusan yang dilakukan secara manual bersifat tidak efektif dan efisien mengakibatkan siswa-siswi gagal dalam mendapatkan pelajaran yang tepat. Pemanfaatan *data mining* sangat membantu dalam menentukan penjurusan siswa tingkat menengah atas. Selain dapat bekerja dengan efektif, pemanfaatan *data mining* bersifat lebih tepat dibandingkan dengan sistem manual yang dilakukan oleh guru. Penjurusan dengan menggunakan *data mining* dapat menggunakan banyak aspek sebagai variabel penelitian seperti nilai akademis, nilai IQ, minat siswa dan masih banyak lagi yang dapat mempengaruhi penjurusan. Salah satu algoritma yang digunakan dalam *data mining* adalah algoritma C4.5. Algoritma C4.5 adalah algoritma klasifikasi pohon keputusan yang berfungsi dengan menggunakan cara kerja yaitu mengolah data yang ada, mengambil simpul lalu membuat peraturan atau *rule* yang nantinya akan menghasilkan rekomendasi penjurusan yang tepat untuk siswa-siswi dalam bentuk pohon keputusan. Dengan penerapan *data mining* algoritma C4.5 ini akan dapat membantu pihak sekolah dalam menentukan penjurusan yang tepat untuk siswa tingkat menengah atas sehingga siswa-siswi dapat mendapatkan pendidikan yang tepat.

Kata Kunci : Penjurusan, *Data Mining*, Pohon Keputusan, Algoritma C45

ABSTRACT

Education is an important aspect of human life. In the world of education, when students reach the upper middle level, students will be faced with majors. Majoring is a process of choosing the course majors undertaken by students in high school. The majors are divided into IPA and IPS which aims to enable students to get an education that suits their abilities and interests. Manually administered majors are ineffective and efficient and result in the students failing to get the right lessons. Utilization of data mining is very helpful in determining the majors of upper middle level students. In addition to working effectively, the utilization of data mining is more appropriate than the manual system performed by the teacher. Majoring by using data mining can use many aspects as research variables such as academic value, IQ score, student interest and many more that can influence the majors. One of the algorithms used in data mining is the C4.5 algorithm. Algorithm C4.5 is a decision tree classification algorithm that works by using the work that is processing existing data, take a knot and then create a rule or rule that will produce the right direction for student recommendations in the form of decision trees. With the implementation of data mining algorithm C4.5 will be able to assist the school in determining the right direction for high school students so that students can get the right education.

Keywords : *Majoring, Data Mining, Decision Tree, C4.5 algorithm*

KATA PENGANTAR

Segala puji dan syukur saya panjatkan kehadirat Tuhan Yang Maha Esa sebab rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Implementasi Algoritma C4.5 Untuk Memprediksi Penjurusan Siswa SMAK Yos Sudarso Batam” ini.

Penulis menyadari bahwa skripsi ini tidak akan selesai tanpa bantuan dan bimbingan serta dorongan moral dari berbagai pihak yang telah membantu penulis baik secara langsung ataupun tidak langsung. Untuk itu, penulis menyampaikan ucapan terima kasih kepada :

1. Rektor Universitas Putera Batam.
2. Dekan Fakultas Teknik dan Komputer Universitas Putera Batam.
3. Ketua Program Studi Teknik Informatika Universitas Putera Batam.
4. Koko Handoko, S.Kom., M.Kom. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
5. Dosen dan Staff Universitas Putera Batam.
6. Bapak Stanislaus K. Teluma Kepala Sekolah SMAK Yos Sudarso yang telah memberikan dukungannya.
7. Keluarga yang selalu mendoakan dan memberi semangat.
8. Teman-teman seangkatan yang berjuang bersama dan saling memberi semangat dan motivasi.
9. Serta pihak-pihak lain yang tidak dapat disebutkan satu per satu.

Ada peribahasa “Tiada gading yang tak retak”, penulis menyadari bahwa skripsi ini jauh dari kata sempurna, sebab itu kritik dan saran yang bersifat

membangun sangat diterima oleh penulis. Penulis juga mohon maaf jika ada kesalahan kata dalam penulisan.

Akhir kata, penulis berharap semoga skripsi ini dapat bermanfaat untuk pembaca yang akan menjadikan skripsi ini menjadi salah satu referensi.

Batam, Agustus 2018

Penulis

DAFTAR ISI

	Halaman
HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL	ii
PERNYATAAN	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
DAFTAR RUMUS	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang Penelitian	1
1.2. Identifikasi Masalah	3
1.3. Pembatasan Masalah	3
1.4. Perumusan Masalah	4
1.5. Tujuan Penelitian	4
1.6. Manfaat Penelitian	4
1.6.1. Aspek Teoritis (Keilmuan)	5
1.6.2. Aspek Praktis (Gunalaksana)	5
BAB II KAJIAN PUSTAKA	6
2.1. <i>Knowledge Discovery in Database (KDD)</i>	6
2.2. <i>Data Mining</i>	9
2.3. Metode Data Mining	11
2.3.1. Metode Prediksi	15
2.3.2. Algoritma C4.5	16
2.3.3. Pohon Keputusan	18

2.3.4.	<i>Rule Based</i>	19
2.4.	<i>Software</i> Pendukung	19
2.5.	Penelitian Terdahulu	20
2.6.	Kerangka Pemikiran.....	22
BAB III METODOLOGI PENELITIAN		24
3.1.	Desain Penelitian	24
3.2.	Teknik Pengumpulan Data.....	26
3.3.	Operasional Variabel	27
3.4.	Metode Analisis dan Rancangan Sistem.....	30
3.5.	Lokasi dan Jadwal Penelitian.....	31
BAB IV HASIL DAN PEMBAHASAN		33
4.1.	Analisa Data.....	33
4.1.1.	Seleksi Data	33
4.1.2.	<i>Cleaning</i> Data	34
4.1.3.	Tranformasi Data	34
4.1.4.	Perhitungan Algoritma.....	37
4.2.	Hasil Pengujian	53
BAB V SIMPULAN DAN SARAN		60
5.1.	Simpulan	60
5.2.	Saran	60
DAFTAR PUSTAKA		
DAFTAR RIWAYAT HIDUP		
SURAT KETERANGAN PENELITIAN		
SURAT BALASAN KETERANGAN PENELITIAN		
LAMPIRAN		

DAFTAR TABEL

	Halaman
Tabel 3.1 Tabel Waktu Penelitian	30
Tabel 4.1 Pembagian Kelas Nilai PR	33
Tabel 4.2 Pembagian Kelas Nilai PO	33
Tabel 4.3 Pembagian Kelas Nilai PB	34
Tabel 4.4 Pembagian Kelas Nilai PV	34
Tabel 4.5 Sampel Penelitian	35
Tabel 4.6 Hasil Perhitungan <i>Node 1</i>	43
Tabel 4.7 Hasil Perhitungan <i>Node 2</i>	45
Tabel 4.8 Hasil Perhitungan <i>Node 3</i>	47

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Tahapan pada KDD	7
Gambar 2.2 <i>Flowchart</i> KDD	8
Gambar 2.3 Model Metode <i>Data Mining</i>	13
Gambar 2.4 <i>Flowchart</i> Algoritma C4.5	16
Gambar 2.5 Kerangka Pemikiran	22
Gambar 3.1 Desain Penelitian	23
Gambar 3.2 Rancangan Sistem.....	28
Gambar 4.1 Pohon Keputusan <i>Node 1</i>	44
Gambar 4.2 Pohon Keputusan <i>Node 2</i>	46
Gambar 4.3 Pohon Keputusan <i>Node 3</i>	48
Gambar 4.4 Pohon Keputusan	49
Gambar 4.5 Hubungan Proses	50
Gambar 4.6 Kotak Dialog <i>Parameters Read Excel</i>	51
Gambar 4.7 Kotak Dialog <i>Data Import Wizard</i>	52
Gambar 4.8 Step 4 of 4 Data Import Wizard.....	53
Gambar 4.9 Proses <i>Cross Validation</i>	53
Gambar 4.10 Pohon Keputusan <i>RapidMiner</i>	54
Gambar 4.11 Deskripsi Pohon Keputusan <i>RapidMiner</i>	55
Gambar 4.12 Deskripsi Ketepatan <i>RapidMiner</i>	56

DAFTAR RUMUS

	Halaman
Rumus 2.1 Perhitungan Entropi	17
Rumus 2.2 Perhitungan <i>Gain</i>	17
Rumus 4.1 Perhitungan Sampel	31
Rumus 4.2 Perhitungan Interval.....	32

DAFTAR LAMPIRAN

Lampiran 1. Populasi Penelitian

Lampiran 2. Hasil *Cleaning* dan Transformasi Data Populasi

Lampiran 3. Perhitungan *Node 2*

Lampiran 4. Perhitungan *Node 3*