

**ANALISIS ALGORITMA C4.5 UNTUK
MENENTUKAN TINGKAT PROFIT USAHA
PADA PT MEGACIPTA SEJATI**

SKRIPSI



**Oleh:
Harina
161510006**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

**ANALISIS ALGORITMA C4.5 UNTUK
MENENTUKAN TINGKAT PROFIT USAHA
PADA PT MEGACIPTA SEJATI**

SKRIPSI

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



**Oleh:
Harina
161510006**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2021**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Harina
NPM/NIP : 161510006
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa “**Skripsi**” yang saya buat dengan judul:

Analisis Algoritma C4.5 Untuk Menentukan Tingkat Profit Usaha Pada PT Megacipta Sejati

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, didalam naskah Skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka.

Apabila ternyata di dalam naskah Skripsi ini dapat dibuktikan terdapat unsur-unsur PLAGIASI, saya bersedia naskah Skripsi ini digugurkan dan gelar akademik yang saya peroleh dibatalkan, serta diproses sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya tanpa ada paksaan dari siapapun.

Batam, 20 Januari 2021



Harina
161510006

**ANALISIS ALGORITMA C4.5 UNTUK
MENENTUKAN TINGKAT PROFIT USAHA PADA PT
MEGACIPTA SEJATI**

Oleh
Harina
161510006

SKRIPSI

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

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

Batam, 20 Januari 2021



Rika Harman, S.Kom., M.SI.
Pembimbing



Universitas Putera Batam

ABSTRAK

Perkembangan asrama di Batam dari tahun ke tahun menunjukkan peningkatan yang signifikan. Banyak orang yang datang dari kota yang jauh dan banyak juga orang yang meninggalkan kampung halaman untuk datang dan mencari kerja di kota Batam sehingga tidak memiliki tempat tinggal membuat banyak perusahaan yang sedang mencari pekerja memberikan fasilitas tempat tinggal seperti asrama untuk para pekerja. PT Megacipta Sejati merupakan perusahaan yang bergerak dibidang sewa *dormitory* kepada perusahaan yang memberikan fasilitas tempat tinggal seperti asrama untuk para pekerjanya. Selama ini perusahaan memiliki sekumpulan data terkait pencapaian kinerja perusahaan yang telah menghabiskan banyak biaya namun data tersebut belum dapat dimanfaatkan dengan baik. Melalui penerapan teknologi data mining, sekumpulan data yang belum dimanfaatkan tersebut akan diproses dan menghasilkan *new knowledge* (pengetahuan baru) yang menjadi sesuatu yang berharga dan berguna bagi perusahaan terutama dalam peningkatan *profit*. Berdasarkan permasalahan tersebut, saya sebagai peneliti tertarik untuk meneliti analisa algoritma C4.5 untuk menentukan tingkat *profit* usaha pada PT Megacipta Sejati. Data Mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Data Mining sendiri memiliki beberapa teknik salah satunya klasifikasi. Teknik klasifikasi terdiri beberapa metode, dan *decision tree* adalah bagian dari metode klasifikasi. Kemudian metode *decision tree* memiliki algoritma, algoritma C4.5 adalah salah satu dari algoritma yang memiliki *decision tree*. Hasil dari penelitian ini akan menghasilkan rule-rule keputusan *profit* dan kerugian perusahaan.

Kata Kunci: Algoritma C4.5, *Data Mining*, Pohon Keputusan, *Profit*.

ABSTRACT

The development of boarding houses in Batam from year to year shows a significant increase. Many people come from distant cities and many people leave their hometowns to come and look for work in Batam, so they do not have a place to live. Many companies are looking for workers to provide housing facilities such as dormitories for workers. PT Megacipta Sejati is a company engaged in the rental of dormitory to companies that provide housing facilities such as dormitories for their workers. During this time the company has a set of data related to the achievement of company performance which has cost a lot but the data cannot be utilized properly. Through the application of data mining technology, a collection of untapped data will be processed and produce new knowledge that becomes valuable and useful for companies, especially in increasing profits. Based on these problems, I as a researcher are interested in researching the c4.5 algorithm analysis to determine the level of business profit at PT Megacipta Sejati. Data Mining is mining or finding new information by looking for certain patterns or rules from a large amount of data that is expected to overcome these conditions. Data Mining itself has several techniques, one of which is classification. The classification technique consists of several methods, and the decision tree is part of the classification method. Then the decision tree method has an algorithm, C4.5 algorithm is one of the algorithms that has a decision tree. The results of this study will produce rules for the company's profit and loss decisions.

Keywords: *Algorithm C4.5, Data Mining, Decision Tree, Profit.*

KATA PENGANTAR

Dengan mengucapkan puji dan syukur kepada Tuhan yang Maha Esa yang telah melimpahkan segala rahmat dan karunianya, sehingga penulis dapat menyelesaikan skripsi yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Sistem Informasi Universitas Putera Batam.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Karena itu, kritik dan saran akan senantiasa penulis terima dengan senang hati.

Dengan segala keterbatasan, penulis menyadari pula bahwa skripsi ini takkan terwujud tanpa bantuan, bimbingan, dan dorongan dari berbagai pihak. Untuk itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih kepada:

1. Ibu Dr. Nur Elfi Husda, S.Kom., M.SI, selaku Rektor Universitas Putera Batam.
2. Bapak Welly Sugianto, S.T., M.M, selaku Dekan Fakultas Teknik Dan Komputer Universitas Putera Batam.
3. Bapak Muhammad Rasyid Ridho, S.Kom., M.SI., selaku Ketua Program Studi Sistem Informasi Universitas Putera Batam.
4. Bapak Rika Harman, S.Kom., M.SI., sebagai pembimbing skripsi yang telah menyediakan waktu dalam memberikan bimbingan dan petunjuk sampai selesainya skripsi ini.
5. Seluruh dosen dan staff Universitas Putera Batam.
6. Dan pihak-pihak yang telah memberikan dukungannya hingga tersusunnya penelitian ini yang tidak dapat disebutkan oleh penulis satu persatu.

Semoga Tuhan YME membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Amin.

Batam, 20 Januari 2021

Penulis



Universitas Putera Batam

DAFTAR ISI

	Halaman
HALAMAN SAMPUL	
HALAMAN JUDUL	
HALAMAN PERNYATAAN ORISINALITAS	
HALAMAN PENGESAHAN	
ABSTRAK	v
<i>ABSTRACT</i>	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR RUMUS	xii
BAB I PENDAHULUAN	
1.1 Latar Belakang Penelitian	1
1.2 Identifikasi Masalah	6
1.3 Pembatasan Masalah	6
1.4 Perumusan Masalah	7
1.5 Tujuan Penelitian	8
1.6 Manfaat Penelitian	8
1.6.1 Manfaat Teoritis	9
1.6.2 Manfaat Praktis	9
BAB II TINJAUAN PUSTAKA	
2.1 Teori	11
2.1.1 Teori Umum	11
2.1.1.1 <i>Knowledge Discovery in Database (KDD)</i>	11
2.1.1.2 <i>Data Mining</i>	14
2.1.1.3 Pengelompokan <i>Data Mining</i>	16
2.1.1.4 Klasifikasi	20
2.1.1.5 Pohon Keputusan	22
2.1.1.6 Algoritma C4.5	25
2.1.2 Teori Khusus	29
2.1.2.1 Profit	29
2.1.2.2 <i>Waikato Environment for Knowledge Analysis (WEKA)</i>	30
2.2 Kerangka Berpikir	31
2.3 Hipotesis Penelitian	32
2.4 Penelitian Terdahulu	33
BAB III METODE PENELITIAN	
3.1 Desain Penelitian	36
3.2 Objek Penelitian	38

3.2.1	Lokasi Penelitian.....	38
3.2.2	Jadwal Penelitian	38
3.3	Populasi dan Sampel	39
3.3.1	Populasi.....	39
3.3.2	Sampel.....	39
3.4	Variabel Penelitian.....	40
3.5	Teknik Pengumpulan Data.....	41
3.6	Model Penelitian	42
BAB IV HASIL DAN PEMBAHASAN		
4.1	Hasil	43
4.2	Pembahasan.....	45
4.2.1	Melakukan Pra Proses Data Penelitian	46
4.2.2	Analisa Data.....	48
4.2.2.1	Pembersihan Data	48
4.2.2.2	Seleksi Data	48
4.2.2.3	Transformasi Data.....	48
4.2.3	Pohon Keputusan	52
4.2.4	Keputusan Hasil Pengujian	76
4.2.5	Pengujian Menggunakan <i>WEKA</i>	77
BAB V SIMPULAN DAN SARAN		
5.1	Simpulan	86
5.2	Saran	88
DAFTAR PUSTAKA		90
LAMPIRAN		

DAFTAR GAMBAR

	Halaman
Gambar 1.1 Data Penghuni.....	3
Gambar 2.1 Proses Dalam KDD atau Knowledge Discovery in Database.....	15
Gambar 2.2 Bidang Ilmu Data Mining.....	16
Gambar 2.3 Pengelompokan Teknik Data Mining.....	20
Gambar 2.4 Pengelompokan Teknik Klasifikasi.....	22
Gambar 2.5 Pohon Keputusan Identifikasi.....	25
Gambar 2.6 Tampilan Utama <i>WEKA</i>	31
Gambar 2.7 Kerangka Berpikir.....	33
Gambar 3.1 Desain Penelitian.....	36
Gambar 4.1 Pohon Keputusan <i>Node 1</i>	61
Gambar 4.2 Pohon Keputusan <i>Node 2</i>	69
Gambar 4.3 Pohon Keputusan <i>Node 3</i>	76
Gambar 4.4 Data Variabel Pada Aplikasi Microsoft Excel.....	77
Gambar 4.5 Data Variabel Pada Aplikasi Notepad.....	78
Gambar 4.6 Tampilan Awal Aplikasi <i>WEKA</i>	79
Gambar 4.7 Tampilan Pemilihan <i>File</i> Pengujian <i>WEKA</i>	80
Gambar 4.8 Tampilan Pemilihan Atribut Pengujian <i>WEKA</i>	80
Gambar 4.9 Tampilan <i>Tab Menu Classify</i>	81
Gambar 4.10 Tampilan Pemilihan <i>Classifier</i>	82
Gambar 4.11 Tampilan Pemilihan <i>Classifier</i>	83
Gambar 4.12 Tampilan Visualisasi Pohon.....	84
Gambar 4.13 Tampilan Hasil Pengujian <i>WEKA</i>	84

DAFTAR TABEL

	Halaman
Tabel 3.1 Jadwal Penelitian	38
Tabel 4.1 <i>Format</i> Pra Proses Data Profit Usaha	47
Tabel 4.2 Biaya Tenaga Kerja	48
Tabel 4.3 Biaya Operasional	49
Tabel 4.4 Biaya Perlengkapan	49
Tabel 4.5 Biaya Perawatan	49
Tabel 4.6 Biaya Pajak	50
Tabel 4.7 Biaya Legalitas	50
Tabel 4.8 <i>Format</i> Data Akhir Profit Usaha	51
Tabel 4.9 Hasil Perhitungan Manual <i>Node 1</i>	60
Tabel 4.10 Hasil Perhitungan Manual <i>Node 2</i>	68
Tabel 4.11 Hasil Perhitungan Manual <i>Node 3</i>	75

DAFTAR RUMUS

	Halaman
Rumus 2.1 Perhitungan <i>Gain</i>	26
Rumus 2.2 Rumus Perhitungan <i>Entropy</i>	27