

**ANALISIS ALGORITMA C4.5 UNTUK
MEMPREDIKSI PENJUALAN PADA *DEALER*
MOTOR HONDA**

SKRIPSI



**Oleh:
Nolaskus Sinaga
150210162**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

**ANALISIS ALGORITMA C4.5 UNTUK
MEMPREDIKSI PENJUALAN PADA *DEALER*
MOTOR HONDA**

SKRIPSI

**Untuk memenuhi salah satu syarat
Guna memperoleh gelar sarjana**



**Oleh:
Nolaskus Sinaga
150210162**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2020**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini saya:

Nama : Nolaskus Sinaga

Npm : 150210162

Fakultas : Teknik dan Komputer

Program Studi : Teknik Informatika

Menyatakan bahwa Skripsi yang saya buat dengan judul:

ANALISIS ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN PADA *DEALER* MOTOR HONDA 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 tertulis dikutip didalam naskah ini dan disebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata di dalam naskah skripsi ini digugurkan dan skripsi 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, 21 Februari 2020

Yang membuat pernyataan,

Nolaskus Sinaga
150210162

ANALISIS ALGORITMA C4.5 UNTUK MEMPREDIKSI PENJUALAN PADA *DEALER* MOTOR HONDA

SKRIPSI

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

**Oleh :
Nolaskus Sinaga
150210162**

**Telah disetujui oleh Pembimbing pada tanggal
seperti tertera dibawah ini**

Batam, 21 februari 2020

**Sunarsan Sitohang, S.Kom., M.TI.
Pembimbing**

ABSTRAK

Motor menjadi pilihan setiap orang sebagai andalan *transportasi* karena dari harga yang terjangkau, perawatan yang mudah serta dapat digunakan dalam jangka waktu yang cukup lama. Di Indonesia sendiri, motor yang paling laris dipasaran adalah motor merk Honda. Untuk wilayah kota batam ada banyak perusahaan yang menjual motor Honda salah satunya adalah PT. Capella Dinamik Nusantara Cabang Muka Kuning. Wilayah penjualan sangat mempengaruhi terhadap penjualan motor selain ke showroom langsung, beberapa tempat ramai seperti mall dapat digunakan sebagai promosi penjualan motor karna tempat yang strategis dan mudah dijangkau masyarakat. Data penjualan yang banyak akan sangat berpengaruh terhadap pengambilan keputusan serta sangat sulit untuk dianalisa. Dari data penjualan yang begitu banyak, maka dilakukan metode *data mining* dengan menggunakan algoritma C4.5. algoritma C4.5 adalah salah satu *data mining* yang digunakan untuk melakukan klasifikasi dan segmentasi yang bertujuan untuk menemukan pola yang berharga dari data yang berukuran relatif besar hingga sangat besar. Dari penelitian yang dilakukan terdapat 3 root (akar) dari 5 variabel yang digunakan sebagai sampel yaitu segmen, harga, tahun produksi semester dan penjualan via. Hasil dari kegiatan *data mining* ini diharapkan mampu memberikan sebuah pohon keputusan untuk membantu pihak PT. Capella Dinamik Nusantara Cabang Muka Kuning dalam melihat pola prediksi perilaku konsumen dalam membeli motor.

Kata Kunci : Algoritma C4.5, *Data mining*, Klasifikasi, Pohon keputusan.

ABSTRACT

Motorbike is the choice of everyone as a mainstay of transportation because of the affordable price, easy maintenance and can be used in a long time. In Indonesia alone, the best-selling motorbike in the market is the Honda motorcycle brand. For the city of Batam there are many companies that sell Honda motorcycles, one of which is PT. Capella Dinamik Nusantara Yellow Front Branch. The sales area is very influential on motorcycle sales other than the direct showroom, some crowded places like malls can be used as motorcycle sales promotions because of the strategic location and easy to reach by the public. A lot of sales data will greatly affect the decision making and very difficult to analyze. From so many sales data, the data mining method is performed using the C4.5 algorithm. C4.5 algorithm is one of the data mining used to do classification and segmentation which aims to find valuable patterns from relatively large to very large data. From the research conducted there are 3 root (roots) of 5 variables used as samples, namely segment, price, semester production year and sales via. The results of data mining activities are expected to be able to provide a decision tree to help PT. Capella Dinamik Nusantara, Branch of Muka Kuning in looking at predictive patterns of consumer behavior in buying a motorcycle.

Keywords: *C4.5 Algorithm, Classification, Data mining, Decision Tree.*

KATA PENGANTAR

Puji dan syukur kepada tuhan yang maha Esa yang telah melimpahkan segala rahmat dan karuniaNya, sehingga penulis dapat menyelesaikan laporan tugas akhir yang merupakan salah satu persyaratan untuk menyelesaikan program studi strata satu (S1) pada Program Studi Teknik Informatika 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. Rektor Universitas Putera Batam.
2. Ketua Program Studi bapak Andi Maslan, S.T., M.SI.
3. Sunarsan Sitohang, S.Kom., M.TI. selaku pembimbing Skripsi pada Program Studi Teknik Informatika Universitas Putera Batam.
4. Dosen dan Staf Universitas Putera Batam.
5. Orang tua penulis yang selalu mendoakan dan memberi semangat.
6. Fatul fikri, Oda, irwansyah, Sopar sebagai sahabat yang selalu memberi support.
7. Teman-teman seangkatan yang berjuang bersama dan saling memberi semangat dan motivasi.
8. Serta berbagai pihak yang telah membantu dalam menyelesaikan skripsi ini secara langsung maupun tidak langsung.

Batam, Februari 2020

Penulis

DAFTAR ISI

	HALAMAN
HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
PERNYATAAN	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	x
DAFTAR TABEL	xi
DAFTAR RUMUS	xii
BAB I PENDAHULUAN	
1.1. Latar Belakang	1
1.2. Identifikasi Masalah	3
1.3. Batasan Masalah	4
1.4. Perumusan Masalah	4
1.5. Tujuan Penelitian	5
1.6. Manfaat Penelitian	5
1.6.1. Manfaat Teoritis	5
1.6.1. Manfaat Praktis.....	
BAB II KAJIAN PUSTAKA	
2.1. <i>Knowledge Discovery In Database (KDD)</i>	7
2.2. <i>Data mining</i>	8
2.3. Metode Klasifikasi.....	11
2.4. <i>Software</i> Pendukung	15
2.4.1. <i>RapidMiner</i>	15
2.5. Penelitian Terdahulu	16
2.6. Kerangka Pemikiran	19
BAB III METODE PENELITIAN	
3.1. Desain Penelitian	21
3.2. Teknik Pengumpulan Data	23
3.3. Operasional Variabel	24
3.4. Metode Analisis dan Rancangan Sistem.....	26
3.5. Lokasi dan Jadwal Penelitian	28
3.5.1. Lokasi Penelitian	28
3.5.2. Jadwal Penelitian	28

BAB IV HASIL DAN PEMBAHASAN

4.1. Analisis <i>Data mining</i> Menggunakan Algoritma C4.5	29
4.1.1. Seleksi dan Pemilihan Atribut	29
4.1.2. Melakukan Pra-Proses	30
4.2. Menghitung <i>Entropy</i> dan <i>Gain Node 1</i>	33
4.2.1. Menghitung <i>Entropy</i> Total	33
4.2.2. Menghitung <i>Entropy</i> Semester	34
4.2.3. Menghitung <i>Entropy</i> Segmen.....	36
4.2.4. Menghitung <i>Entropy</i> Tahun Produksi	37
4.2.5. Menghitung <i>Entropy</i> Harga.....	39
4.2.6. Menghitung <i>Entropy</i> Penjualan Via	41
4.3. Menghitung <i>Entropy</i> dan <i>Gain Node 1.1</i>	46
4.3.1. Menghitung <i>Entropy</i> Total.....	46
4.3.2. Menghitung <i>Entropy</i> Semester	47
4.3.3. Menghitung <i>Entropy</i> Segmen.....	48
4.3.4. Menghitung <i>Entropy</i> Tahun Produksi.....	49
4.3.5. Menghitung <i>Entropy</i> Harga.....	50
4.4. Menghitung <i>Entropy</i> dan <i>Gain Node 1.2</i>	54
4.4.1. Menghitung <i>Entropy</i> Total.....	54
4.4.2. Menghitung <i>Entropy</i> Semester	55
4.4.3. Menghitung <i>Entropy</i> Segmen.....	57
4.4.4. Menghitung <i>Entropy</i> Tahun Produksi.....	58
4.4.5. Menghitung <i>Entropy</i> Harga.....	60
4.5. Hasil Pengujian.....	64

BAB V SIMPULAN DAN SARAN

5.1. Simpulan	72
5.2. Saran	73

DAFTAR PUSTAKA	74
-----------------------------	----

LAMPIRAN

Lampiran 1. Pendukung Penelitian

Lampiran 2. Daftar Riwayat Hidup

Lampiran 3. Surat Keterangan Penelitian

DAFTAR GAMBAR

	HALAMAN
Gambar 2.1 Pengelompokan Teknik Klasifikasi	12
Gambar 2.2 Kerangka Pemikiran	20
Gambar 3.1 Desain Penelitian.....	21
Gambar 3.2 Rancangan Sistem	26
Gambar 4.1 Pohon Keputusan <i>Node 1</i>	45
Gambar 4.2 Pohon Keputusan <i>Node 1.1</i>	53
Gambar 4.3 Pohon Keputusan <i>Node 1.2</i>	63
Gambar 4.4 <i>Welcome To Rapidminer Studio</i>	65
Gambar 4.5 <i>Import Data Read Excel</i>	65
Gambar 4.6 <i>Data Import Select Step 1</i>	66
Gambar 4.7 <i>Data Import Select Step 3</i>	67
Gambar 4.8 <i>Data Import Select Step 4</i>	67
Gambar 4.9 <i>Data Import Finish</i>	68
Gambar 4.10 <i>Import Operator Decesion Tree</i>	69
Gambar 4.11 Menghubungkan Operator	70
Gambar 4.12 Pohon Keputusan	70

DAFTAR TABEL

	HALAMAN
Tabel 3.1 Jadwal Penelitian	28
Tabel 4.1 Penjualan Tahun 2019	30
Tabel 4.2 Klasifikasi Semester	30
Tabel 4.3 Klasifikasi Tahun Produksi	31
Tabel 4.4 Klasifikasi Penjualan Via	31
Tabel 4.5 Klasifikasi Harga	32
Tabel 4.6 Format Akhir Data Penjualan	32
Tabel 4.7 Hasil Perhitungan <i>Node 1</i>	44
Tabel 4.8 Hasil Perhitungan <i>Node 1.1</i>	46
Tabel 4.9 Hasil Perhitungan <i>Node 1.2</i>	52
Tabel 4.10 Data Penjualan <i>Node 1.1</i>	53
Tabel 4.11 Data Penjualan <i>Node 1.2</i>	62

DAFTAR RUMUS

	HALAMAN
Rumus 2.1 <i>Entropy</i>	15
Rumus 2.2 <i>Gain</i>	15