

**ALGORITMA C4.5 UNTUK MENENTUKAN
TINGKAT KELAYAKAN MOTOR BEKAS YANG
AKAN DIJUAL DI AMAZONE MOTOR**

SKRIPSI



**Oleh:
Ervina
151510058**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2019**

**ALGORITMA C4.5 UNTUK MENENTUKAN
TINGKAT KELAYAKAN MOTOR BEKAS YANG
AKAN DIJUAL DI AMAZONE MOTOR**

SKRIPSI

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



**Oleh:
Ervina
151510058**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
TAHUN 2019**

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 di kemudian 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, 01 Agustus 2019
Yang membuat pernyataan,

Ervina
151510058

**ALGORITMA C4.5 UNTUK MENENTUKAN TINGKAT
KELAYAKAN MOTOR BEKAS YANG AKAN DIJUAL DI
AMAZONE MOTOR**

**Oleh
Ervina
151510058**

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

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

Batam, 01 Agustus 2019

**Erlin Elisa, S.Kom., M.Kom.
Pembimbing**

ABSTRAK

Kendaraan pribadi adalah hal yang sangat penting bagi masyarakat Indonesia saat ini. Sebagian besar masyarakat Indonesia memilih sepeda motor sebagai kendaraan penunjang aktivitas sehari-hari mereka. Banyak juga masyarakat yang membeli sepeda motor bekas dengan berbagai alasan. Sehingga, Menjual sepeda motor bekas adalah bisnis yang cukup menjanjikan saat ini. Tetapi, tidak semua motor bekas dapat dijual begitu saja harus dilakukan penyeleksian yang menentukan tingkat kelayakannya tidak memiliki standar atau acuan yang pasti, hal ini menyebabkan banyak motor bekas yang sudah dijual malah dibawa kembali oleh pelanggan ke Amazone Motor dan menyampaikan komplain bahwa terjadi kerusakan-kerusakan yang tidak seharusnya. Hal ini menyebabkan kerugian dalam hal waktu, tenaga maupun biaya. Pada tugas akhir ini, dilakukan sebuah penelitian untuk menentukan tingkat kelayakan sepeda motor bekas yang akan dijual di Amazone Motor dengan menggunakan penerapan salah satu metode dalam teknik *datamining* yaitu metode algoritma C4.5. *Datamining* dapat membantu dalam menggali informasi dari tumpukkan data, informasi yang didapatkan dapat digunakan untuk mencapai tujuan yang kita inginkan. Pada algoritma C4.5 untuk memperoleh informasi dengan melakukan perhitungan *entropy* lalu *gain* tertinggi untuk dibentuk menjadi pohon keputusan. Hasil yang didapatkan dalam metode ini berupa model pohon keputusan. Hasil dari penelitian ini akan menghasilkan *rule* untuk keputusan standar atau acuan dalam penilaian tingkat kelayakan motor bekas yang akan dijual. Standar atau acuan yang dihasilkan dari penelitian ini menunjukkan bahwa faktor-faktor utama yang mempengaruhi penilaian tingkat kelayakan motor bekas di Amazone Motor adalah Mesin, aki dan rangka.

Kata Kunci: *datamining* , pohon keputusan, algoritma C4.5, kelayakan, motor bekas

ABSTRACT

Private vehicles are very important for Indonesian people nowadays. Most Indonesians prefer motorbikes as vehicles to support their daily activities. Many people also buy used motorbikes for various reasons. therefore, Selling used motorbikes is a promising business at this time. However, not all used motorcycles can be sold directly, firstly must be do a selection that determines the feasibility level does not have a definite standard or reference, this causes many used motorcycles to be brought back by customers to Amazone Motor and complaints that are various damage has occured. This causes losses in terms of time, energy or costs. In this final project, a study was conducted to determine the feasibility of used motorbikes to be sold at Amazone Motorbike using one of the methods in datamining techniques, namely the algorithm C4.5 method. Datamining can help collect information from data stack, information obtained that can be useful for this research. In algorithm C4.5 we can get the information by calculating entropy then get the highest to form Decision Tree. The results obtained in this method consist of a decision tree model. The results of this study will produce rules for standard or reference decisions in assessing the feasibility of used motorcycles to be sold. The standards or references produced from this study indicate the main factors that influence the feasibility of used motorcycles in Amazone Motor are engines, batteries and frames of the motorbikes.

Keywords: *Datamining , decision tree, algorithm C4.5, feasibility, used motorbikes*

KATA PENGANTAR

Puji 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 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. Dekan fakultas teknik dan komputer Universitas Putera Batam.
3. Ketua Program Studi Sistem Informasi Universitas Putera Batam.
4. Ibu Erlin Elisa, S.Kom., M.Kom selaku pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam.
5. Dosen dan Staff Universitas Putera Batam.
6. Pemilik Amazone Motor yang telah mengizinkan pengumpulan data penelitian.
7. Orang tua dan keluarga yang selalu memberikan semangat.
8. Teman-teman perkuliahan yang telah membantu dan memberikan informasi yang berguna

Semoga Tuhan membalas kebaikan dan selalu mencurahkan hidayah serta taufikNya, Amin.

Batam, 01 Agustus 2019

Ervina

DAFTAR ISI

	Halaman
HALAMAN SAMPUL DEPAN	i
HALAMAN JUDUL	ii
HALAMAN PERNYATAAN	iii
HALAMAN PENGESAHAN	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR RUMUS	xii
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang Penelitian	1
1.2. Identifikasi Masalah	3
1.3. Pembatasan Masalah	4
1.4. Perumusan Masalah	4
1.5. Tujuan Penelitian	5
1.6. Manfaat Penelitian	5
1.6.1 Manfaat teoritis	5
1.6.2 Manfaat praktis.....	6
BAB II TINJAUAN PUSTAKA	7
2.1. Teori Dasar.....	7
2.1.1 Sepeda motor.....	7
2.1.2 <i>Knowledge Discovery in Database (KDD)</i>	7
2.1.3 <i>Datamining</i>	10
2.1.4 Klasifikasi	15
2.1.5 Pohon keputusan	16
2.1.6 Algoritma C4.5.....	19
2.1.7 <i>Waikato Environment for Knowledge Analysis (WEKA)</i>	20
2.2. Penelitian Terdahulu	21
2.3. Kerangka Pemikiran.....	25
2.4. Hipotesis.....	26
BAB III METODE PENELITIAN	27
3.1. Desain Penelitian.....	27
3.2. Operasional Variabel.....	29
3.3. Populasi dan Sampel	29
3.3.1 Populasi.....	29

3.3.2	Sampel.....	29
3.4.	Teknik Pengumpulan Data.....	30
3.5.	Metode Analisis Data.....	31
3.6.	Lokasi dan Jadwal Penelitian.....	31
3.6.1	Lokasi penelitian.....	31
3.6.2	Jadwal penelitian.....	32
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....		33
4.1.	Analisa <i>Datamining</i> untuk Prediksi Kelayakan Motor Bekas.....	33
4.2.	Melakukan Pra-proses Data.....	33
4.3.	Analisa Data.....	35
4.3.1	Pembersihan data.....	35
4.3.2	Seleksi data.....	35
4.3.3	Transformasi data.....	35
4.3.4	Pohon keputusan.....	40
4.3.5	Pengujian menggunakan <i>WEKA</i>	58
4.4.	Hasil Penelitian.....	64
BAB V SIMPULAN DAN SARAN.....		65
5.1.	Simpulan.....	65
5.2.	Saran.....	66
DAFTAR PUSTAKA.....		67
LAMPIRAN		

DAFTAR TABEL

	Halaman
Tabel 3.1 Jadwal Penelitian	32
Tabel 4.1 <i>Format</i> Data Kendaraan	34
Tabel 4.2 Klasifikasi Atribut Mesin	36
Tabel 4.3 Klasifikasi Atribut Rangka	36
Tabel 4.4 Klasifikasi Atribut <i>Body</i>	36
Tabel 4.5 Klasifikasi Atribut Aki	37
Tabel 4.6 Klasifikasi Atribut Aksesoris	37
Tabel 4.7 Klasifikasi Atribut Tahun Pembuatan	38
Tabel 4.8 <i>Format</i> Data Kelayakan Motor Bekas	39
Tabel 4.9 Hasil Perhitungan <i>Node</i> 1	45
Tabel 4.10 Hasil Perhitungan <i>Node</i> 2	51
Tabel 4.11 Hasil Perhitungan <i>Node</i> 3	56

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Tahapan <i>knowledge discovery in database (KDD)</i>	9
Gambar 2.2 Ilustrasi Definisi <i>Datamining</i>	10
Gambar 2.3 Fase <i>Cross Industry Process for Datamining</i>	14
Gambar 2.4 Pengelompokan Teknik Klasifikasi	16
Gambar 2.5 Konsep Dasar Pohon Keputusan	18
Gambar 2.6 Kerangka Pemikiran	25
Gambar 3.1 Desain Penelitian	27
Gambar 4.1 Pohon Keputusan <i>Node 1</i>	46
Gambar 4.2 Pohon Keputusan <i>Node 2</i>	51
Gambar 4.3 Pohon Keputusan <i>Node 3</i>	57
Gambar 4.4 Data Uji <i>WEKA.xls</i>	58
Gambar 4.5 Data Uji <i>WEKA.arff</i>	59
Gambar 4.6 Menu Awal <i>WEKA 3.8.3</i>	60
Gambar 4.7 <i>Open File</i> pada <i>WEKA Explorer</i>	60
Gambar 4.8 Pemilihan Variabel pada <i>WEKA</i>	61
Gambar 4.9 <i>Tab Classify</i> pada <i>WEKA</i>	61
Gambar 4.10 Pemilihan Model <i>Classify</i>	62
Gambar 4.11 <i>Classify Output</i> pada <i>WEKA</i>	62
Gambar 4.12 <i>Visualize Tree</i> Pada <i>WEKA</i>	63
Gambar 4.13 Hasil <i>Visualize Tree WEKA</i>	63

DAFTAR RUMUS

	Halaman
Rumus 2.1 Perhitungan <i>Gain</i>	19
Rumus 2.2 Perhitungan <i>entropy</i>	19
Rumus 4.1 Perhitungan Nilai Interval	37

DAFTAR LAMPIRAN

Lampiran 1 Pendukung penelitian

Lampiran 2 Daftar Riwayat Hidup

Lampiran 3 Surat Keterangan Penelitian

BAB I

PENDAHULUAN

1.1. Latar Belakang Penelitian

Kendaraan pribadi saat ini sudah dapat disebut sebagai kebutuhan yang penting bagi masyarakat Indonesia. Menurut data Badan Pusat Statistik pada tahun 2018 banyaknya kendaraan pribadi di Indonesia adalah mobil penumpang pada 2016 terdapat 14.580.666 unit dan pada 2017 15.493.068 unit sedangkan sepeda motor pada 2016 105.150.082 unit dan 2017 113.030.793 unit. Berdasarkan data tersebut dapat dilihat bahwa peningkatan jumlah mobil dalam 1 tahun adalah 912.402 unit sedangkan, peningkatan jumlah sepeda motor adalah 7.880.711. Sehingga, dapat disimpulkan sepeda motor di Indonesia lebih banyak 7 kali lipat dibandingkan mobil dan juga peningkatan sepeda motor pada 2016 sampai 2017 lebih banyak 8 kali lipat dibandingkan peningkatan mobil.

Di Kepulauan Riau sendiri banyak kendaraan pribadi berdasarkan data Badan Pusat Statistik pada tahun 2018 terdapat sebanyak 176.764 unit mobil penumpang pada 2016 dan 184.140 pada 2017 unit sedangkan terdapat 1.163.715 unit sepeda motor dan 1.198.781 unit pada 2017. Berdasarkan data diatas dapat dilihat bahwa peningkatan jumlah mobil dalam 1 tahun adalah 7.376 unit sedangkan, peningkatan jumlah sepeda motor adalah 35.066. Sehingga, dapat disimpulkan peningkatan sepeda motor pada 2016 sampai 2017 lebih banyak hampir 5 kali lipat dibandingkan peningkatan jumlah mobil. Dengan peningkatan yang sangat banyak dalam 1 tahun

menandakan tingkat permintaan sepeda motor ini cukup tinggi. Dengan penambahan jumlah unit sepeda motor tentunya tidak semua masyarakat membeli sepeda motor baru, banyak masyarakat yang membeli sepeda motor bekas dengan berbagai alasan. Dengan adanya permintaan maka bermunculan tempat penjualan motor bekas saat ini. Salah satunya adalah Amazone Motor.

Amazone motor merupakan usaha yang bergerak dibidang jual beli motor bekas sejak 2009 beralamat di Komplek Taman Sentosa Indah Blok E No. 12 Sungai Panas, Batam, Kepulauan Riau dalam 1 bulan dapat menjual sekitar 20 sampai 30 unit motor bekas. Dengan adanya penjualan tentunya ada juga pemasukan motor bekas ini yang sebelum dijual tentunya harus dicek kelayakannya untuk dijual. Tetapi, selama usaha ini dijalankan tidak pernah dilakukan uji kelayakan motor bekas yang akan dijual secara mendalam ataupun membuat acuan atau standar mengenai kelayakan motor bekas yang akan dijual. Sehingga, banyak pelanggan yang sudah membeli motor bekas menyampaikan komplain tentang kendaraan yang sudah dibeli setelah beberapa waktu menggunakannya. Komplain yang disampaikan seperti, mesin yang tidak bisa dihidupkan, suara mesin yang keras saat dihidupkan, dan banyak masalah kendaraan lainnya. Tentunya hal ini menyebabkan kerugian dalam hal biaya, waktu, tenaga, dan citra Amazone Motor sendiri. Dengan adanya masalah tersebut maka peneliti ingin memberikan sebuah solusi terhadap masalah yang dihadapi oleh Amazone Motor selama ini yaitu dengan menganalisa data kendaraan yang akan dijual dengan menggunakan salah satu metode dalam teknik *datamining* yaitu algoritma C4.5.

Teknik *datamining* dengan algoritma C4.5 pada penelitian (Wira & Putra, 2016) yang meneliti tentang menentukan tingkat kelayakan motor bekas yang akan dijual menyatakan *datamining* adalah rangkaian proses dalam menganalisis data dari sudut pandang berbeda dan membentuk data tersebut menjadi informasi yang berguna sehingga pengguna dapat memahami substansi hubungan antara data. Sedangkan, pada penelitian yang pernah dilakukan oleh (Elisa, 2017) yang meneliti tentang pengidentifikasian faktor-faktor penyebab kecelakaan kerja dengan teknik *datamining* dan metode klasifikasi yaitu algoritma C4.5 mengatakan bahwa algoritma C4.5 adalah algoritma untuk klasifikasi data yang memiliki atribut-atribut numerik dan kategorial. menunjukkan bahwa faktor-faktor yang menyebabkan kecelakaan kerja adalah pekerja dan cara kerja, lingkungan tempat kerja, dan alat pelindung diri. Dengan menggunakan metode klasifikasi yaitu algoritma C4.5 nantinya diharapkan dapat digunakan untuk membantu pengambilan keputusan tentang kelayakan motor bekas yang akan dijual.

Berdasarkan uraian latar belakang diatas, maka saya sebagai peneliti tertarik untuk meneliti melalui skripsi saya yang berjudul “Algoritma C4.5 untuk menentukan tingkat kelayakan motor bekas yang akan dijual di Amazone Motor”.

1.2. Identifikasi Masalah

Berdasarkan uraian latar belakang diatas, maka peneliti mengidentifikasi beberapa masalah sebagai berikut:

1. Sering mendapatkan komplain pelanggan terhadap kendaraan yang sudah

dibeli.

2. Tidak adanya acuan atau standar dalam menentukan tingkat kelayakan kendaraan yang akan dijual.

1.3. Pembatasan Masalah

Untuk lebih terarahnya pembahasan penelitian ini, maka peneliti menentukan batasan masalah yang akan diteliti yaitu:

1. Penelitian ini dilakukan di Amazone Motor dengan menganalisa data motor bekas yang akan dijual.
2. Menganalisa data motor bekas dengan metode yang ada dalam *datamining* yaitu algoritma C4.5 dan menguji kebenaran hasil dengan *software WEKA 3.8.3*.

1.4. Perumusan Masalah

Berdasarkan identifikasi masalah tersebut, maka peneliti merumuskan masalah sebagai berikut:

1. Bagaimana teknik *datamining* dengan metode algoritma C4.5 menganalisa kelayakan motor bekas yang akan dijual pada Amazone Motor?
2. Bagaimana hasil pohon keputusan yang terbentuk dari hasil perhitungan *entropy* dan *gain* pada metode algoritma C4.5?

1.5. Tujuan Penelitian

Berdasarkan pada uraian masalah diatas, adapun tujuan dari pelaksanaan penelitian ini adalah untuk:

1. Untuk menganalisa kelayakan motor bekas yang akan dijual Amazone Motor sehingga bisa mengurangi resiko mendapatkan komplain dari pelanggan.
2. Untuk memperoleh keputusan tentang acuan atau standar kelayakan motor bekas yang akan dijual dengan metode dalam *datamining* yaitu algoritma C4.5.

1.6. Manfaat Penelitian

Penelitian ini diharapkan bermanfaat bagi pembaca. Manfaat dari penelitian yang dilakukan adalah sebagai berikut:

1.6.1 Manfaat teoritis

Manfaat teoritis yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Memperkuat teori tentang algoritma C4.5 dalam menentukan tingkat kelayakan.
2. Sebagai referensi pada penelitian-penelitian selanjutnya yang berhubungan dengan teknik *datamining* dengan metode algoritma C4.5.

1.6.2 Manfaat praktis

Selain manfaat teoritis, manfaat praktis yang diharapkan dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Dijadikan pedoman standar atau acuan untuk menentukan tingkat kelayakan motor bekas yang akan dijual.
2. Memberi masukan kepada Amazone Motor untuk menentukan tingkat kelayakan.
3. Meningkatkan citra bisnis Amazone Motor dalam menjual kendaraan yang layak.

BAB II

TINJAUAN PUSTAKA

2.1. Teori Dasar

2.1.1 Sepeda motor

Sepeda motor adalah salah satu jenis alat transportasi. Seiring dengan zaman yang terus berkembang menjadi modern transportasi ini menjadi salah satu kebutuhan manusia yang menginginkan mengerjakan sesuatu dengan cepat (Novera, 2014). Sepeda motor dinilai praktis oleh masyarakat dan menjadi pilihan masyarakat untuk menjalankan berbagai aktivitas sehingga tidak mengherankan jika pasar bisnis untuk penjualan sepeda motor sangat luas (Yuliana, 2013).

2.1.2 *Knowledge Discovery in Database (KDD)*

Saat ini, komputer merupakan salah satu teknologi yang berkembang sangat pesat. Dari waktu ke waktu banyak sekali teknologi baru pada komputer yang tentunya juga memperbanyak manfaat dan tujuan penggunaan komputer saat ini salah satunya adalah untuk pemrosesan data. Pemrosesan data secara komputerisasi sudah diimplementasikan pada banyak bidang pekerjaan seperti pada bidang pendidikan, pemerintahan, ataupun bisnis. Data yang akan diproses ini biasanya tersimpan dalam sebuah bank data atau bisa juga kita sebut sebagai *database*. Tapi,

dengan penggunaan *database* secara berkepanjangan data yang ada semakin menumpuk sementara data yang tersimpan tidak digunakan secara maksimal. Hal seperti inilah yang dapat kita sebut sebagai istilah *rich of data but poor of information*. Sebuah solusi untuk masalah ini adalah dengan menggunakan *knowledge discovery in database (KDD)*.

Menurut penelitian yang dilakukan oleh (Sari & Sindunata, 2014) *knowledge discovery in database (KDD)* merupakan suatu proses *non-trivial* untuk mengidentifikasi pola dalam data yang dimana pola tersebut bersifat baru, sah, dan dapat bermanfaat. *knowledge discovery in database (KDD)* merupakan sebuah proses yang terstruktur sehingga *KDD* terbagi menjadi beberapa tahap, seperti yang disebutkan dalam penelitian (Sulastri & Gufroni, 2017) tahap dalam *KDD* terdiri dari:

1. *Cleaning data*

Pada tahap ini dilakukan proses pembersihan data dari data *noise* dan tidak konsisten.

2. *Data integration*

Proses untuk menggabungkan data dari berbagai sumber yang berbeda.

3. *Data selection*

Pada tahap ini, data diseleksi dari *database* dengan memilih data yang sesuai dengan tujuan analisis.

4. *Data transformation*

Data diubah menjadi jenis data yang sesuai untuk proses *datamining* .

5. *Datamining*

Pada tahap ini, dilakukan penerapan teknik *datamining* sesuai dengan metode tertentu untuk memperoleh sebuah pola dari data.

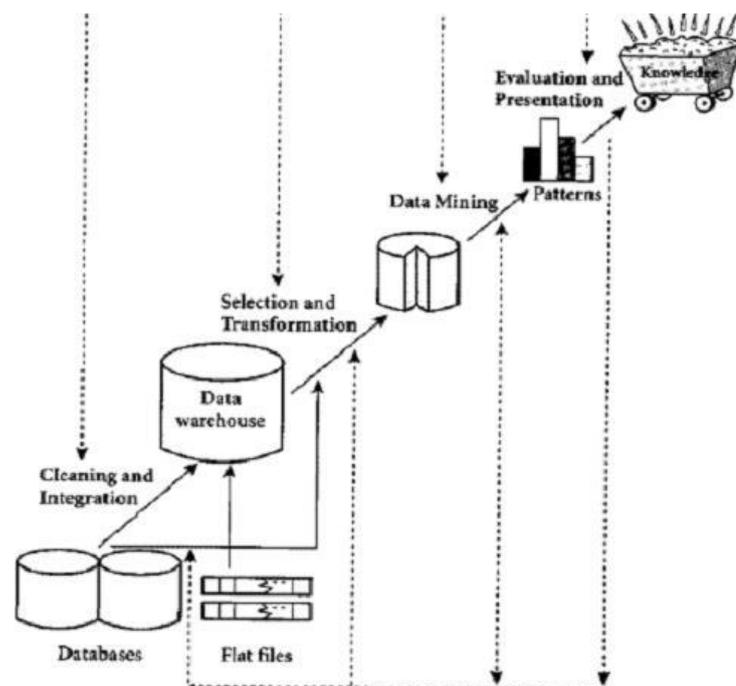
6. *Pattern evaluation*

Proses pengidentifikasian pola pada data.

7. *Knowledge presentation*

Pada tahap ini, hasil *datamining* berupa informasi yang dihasilkan dipresentasikan dan dapat digunakan oleh pemilik data.

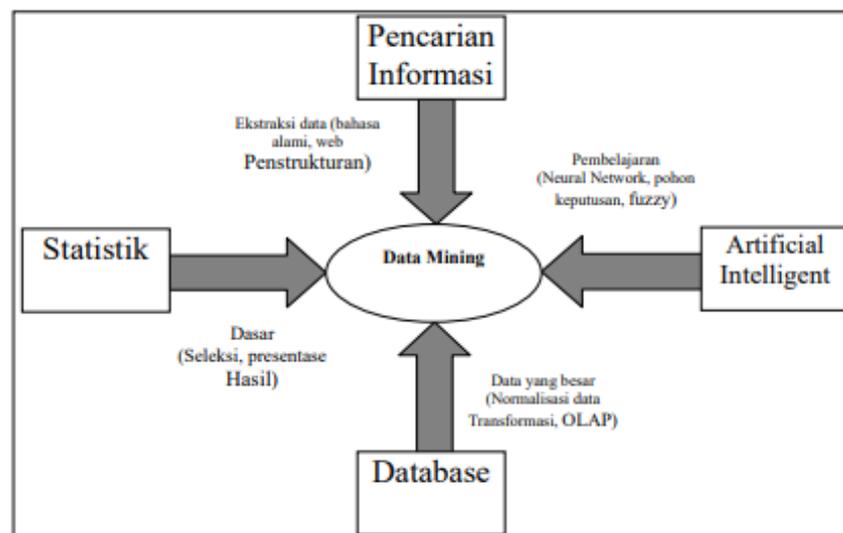
Untuk lebih jelasnya proses *knowledge discovery in database (KDD)* dapat dilihat melalui gambar 2.1 berikut.



Gambar 2.1 Tahapan *knowledge discovery in database (KDD)*

2.1.3 *Datamining*

Datamining pada penelitian (Jamhur, 2016) adalah proses yang memperkerjakan satu ataupun lebih teknik pembelajaran komputer untuk mengekstrak dan menganalisis pengetahuan secara otomatis atau serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang tidak diketahui secara manual. Definisi lain dari *Datamining* pada penelitian Eki Roziqa Maris (2017) dalam (Harman, 2018) adalah serangkaian proses dalam menemukan pengetahuan yang berharga dan tersembunyi dari sekumpulan data yang dapat digunakan untuk memprediksi yang *valid* pada berbagai bidang, baik ekonomi, sosial, budaya, pendidikan, bahkan dalam pada bidang pemerintahan.



Gambar 2.2 Ilustrasi Definisi *Datamining*

Melalui gambar 2.2. diatas dapat dilihat bahwa *datamining* memiliki akar dari berbagai bidang ilmu seperti *artificial interlligent*, statistik, *database*, *machine learning*, pencarian informasi. *Datamining* memiliki beberapa karakteristik seperti yang disebutkan dalam penelitian (Lorena, Zarman, & Hamidah, 2014)

karakteristik *datamining* adalah sebagai berikut:

- a. *Datamining* mempunyai hubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
- b. *Datamining* biasa menggunakan data yang sangat besar, biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
- c. *Datamining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Datamining juga mempunyai standar untuk proses atau strategi pemecahan masalah secara umum dari bisnis atau unit penelitian. Salah satunya adalah *Cross-Industry Process for Datamining (CRISP-DM)*. *CRISP-DM* dikembangkan oleh analisis dari beberapa industri seperti Daimler Chrysler, SPPS, NCR pada tahun 1996 (Imtiyaz, Nasrun, & Ahmad, 2015). *CRISP-DM* dibagi menjadi 6 fase berurutan yang bersifat adaptif. Fase-fase dalam *CRISP-DM* yaitu (Feblian & Daihani, 2017).

1. *Business Understanding Phase* atau Fase Pemahaman Bisnis

Hal yang dilakukan dalam fase ini terdiri dari:

- a. Penelitian secara keseluruhan atau penentuan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis.
- b. Menerjemahkan tujuan dan batasan menjadi formula dan definisi dari permasalahan *datamining*.
- c. Menyiapkan strategi awal untuk mencapai tujuan.

2. *Data Understanding Phase* atau Fase Pemahaman Data

Hal yang dilakukan dalam fase ini terdiri dari:

- a. Mengumpulkan data.
- b. Menganalisis data untuk mengenali lebih lanjut mengenai data dan pencarian awal pengetahuan yang ada didalamnya.
- c. Mengevaluasi kualitas data.

3. *Data Preparation Phase* atau Fase Pengolahan Data

Hal yang dilakukan dalam fase ini terdiri dari:

- a. Fase ini merupakan pekerjaan berat yang perlu dilaksanakan secara intensif. Dalam fase ini data yang tersedia dari awal disiapkan, karena merupakan kumpulan data yang akan digunakan untuk keseluruhan fase berikutnya.
- b. Memilih kasus dan variabel yang ingin dianalisis dan yang sesuai dengan analisis yang akan dilakukan.
- c. Melakukan perubahan pada beberapa variabel jika dibutuhkan.
- d. Menyiapkan data awal sehingga siap untuk perangkat pemodelan.

4. *Modeling Phase* atau Fase Pemodelan

Hal yang dilakukan dalam fase ini terdiri dari:

- a. Memilih teknik pemodelan yang sesuai dan mengaplikasikannya.
- b. Mengoptimalkan hasil dengan mengkalibrasi aturan model.
- c. Perlu memerhatikan bahwa beberapa teknik mungkin untuk digunakan pada permasalahan *datamining* yang sama.
- d. Proses dapat kembali kepada fase pengolahan data untuk menjadikan data kedalam bentuk yang sesuai dengan spesifikasi kebutuhan dari teknik *datamining* tertentu.

5. *Evaluation Phase* atau Fase Evaluasi

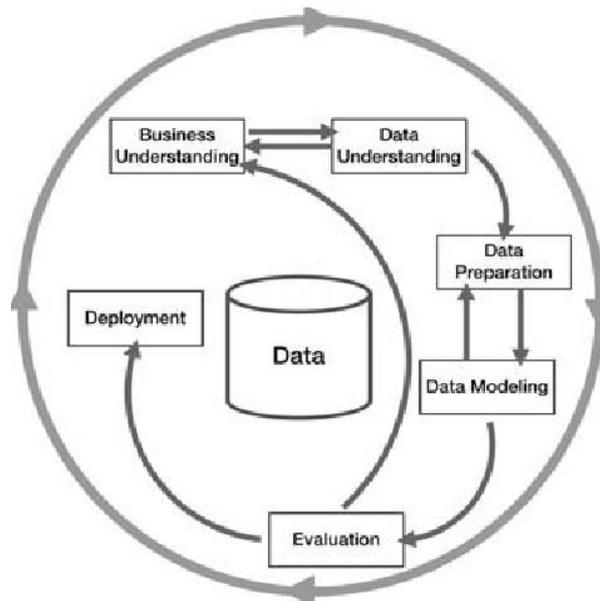
Hal yang dilakukan dalam fase ini terdiri dari:

- a. Sebelum disebarkan untuk digunakan model yang digunakan dalam fase pemodelan dievaluasi untuk mendapatkan kualitas dan efektivitas.
 - b. Menetapkan model yang memenuhi tujuan pada fase awal.
 - c. Mengevaluasi bisnis atau penelitian tertangani dengan baik dan adanya permasalahan.
 - d. Mengambil keputusan tentang penggunaan hasil dari *datamining* .
6. *Deployment Phase* atau Fase Penyebaran

Hal yang dilakukan dalam fase ini terdiri dari :

- a. Menggunakan model yang dihasilkan.
- b. Pembuatan laporan.
- c. Penerapan *datamining*.

Karena fase pada *CRISP-DM* yang berurutan sehingga fase selanjutnya pada urutan bergantung pada keluaran fase sebelumnya. Pada gambar 2.3 berikut ini akan menggambarkan secara umum mengenai urutan fase-fase dalam *CRISP-DM* dimana hubungan antar fase digambarkan dengan panah.



Gambar 2.3 Fase Cross Industry Process for Datamining

Berdasarkan penjelasan diatas dapat disimpulkan bahwa *datamining* digunakan untuk memproses data untuk mendapatkan informasi. Beberapa solusi yang dapat diselesaikan dengan *datamining* , diantaranya (Santoso, 2014):

a. Menembak target pasar

Datamining dapat melakukan klasifikasi dari model-model pembeli sesuai dengan karakteristik yang diinginkan seperti tingkat penghasilan, kebiasaan membeli, kesukaan pembeli, dan karakteristik yang lainnya.

b. Melihat pola pemakai dari waktu ke waktu

Datamining dapat menganalisa pola beli seseorang dari waktu ke waktu.

c. *Cross-Market Analysis*

Datamining dapat menganalisa hubungan antara penjualan satu produk dengan produk lainnya.

d. Profil *Customer*

Dapat menganalisis klasifikasi pelanggan melalui profil pelanggan untuk

mengetahui produk apa yang sering dibeli oleh pelanggan.

2.1.4 Klasifikasi

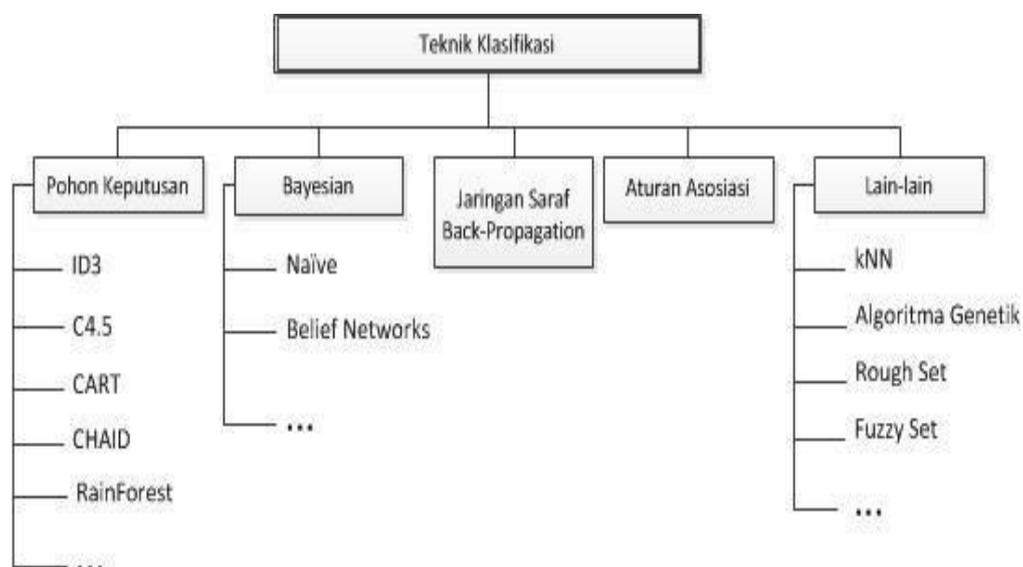
Menurut (Azwanti, 2018), klasifikasi adalah suatu proses yang menemukan properti-properti yang sama dalam sebuah himpunan objek didalam sebuah basis data dan mengklasifikasikannya ke dalam kelas-kelas berbeda menurut model klasifikasi yang ditetapkan. Tujuan klasifikasi yaitu untuk menemukan model dari *training set* yang membedakan atribut ke dalam kategori atau kelas yang sesuai, model tersebut kemudian digunakan untuk mengklasifikasikan atribut yang kelasnya belum diketahui sebelumnya (Lorena et al., 2014).

Proses klasifikasi memiliki dua tahap yaitu tahap pertama adalah proses pembelajaran dimana data pelatihan dianalisis dengan algoritma klasifikasi sehingga terbentuk aturan klasifikasi dan tahap kedua adalah proses pengujian dimana data uji digunakan untuk memperkirakan keakuratan aturan klasifikasi. Adapun model klasifikasi yang digunakan yaitu (Elisa, 2018):

1. Pemodelan Deskriptif digunakan sebagai perangkat penggambaran dari klas berbeda untuk membedakan objek-objek.
2. Pemodelan Prediktif digunakan untuk memprediksi label klas untuk *record* yang tidak dikenal atau tidak diketahui.

Menurut penelitian (Mardi, 2017), klasifikasi terdapat target variabel kategori. Contohnya penggolongan pendapatan dapat dipisahkan dalam tiga kategori yaitu pendapatan tinggi, sedang, dan rendah. Contoh lainnya dibidang bisnis dan

penelitian adalah menentukan apakah suatu transaksi kartu kredit curang atau bukan, memperkirakan apakah suatu pengajuan hipotek oleh nasabah merupakan suatu kredit baik atau buruk, mendiagnosis penyakit seorang pasien untuk mendapatkan termasuk penyakit apa, dan lain-lain. Dalam melakukan klasifikasi terdapat beberapa teknik seperti bayesian, aturan asosiasi, jaringan saraf *back-propagation*, pohon keputusan, dan lain-lain. Untuk lebih jelasnya dapat terlihat seperti gambar 2.4 berikut ini.



Gambar 2.4 Pengelompokan Teknik Klasifikasi

Setiap teknik tentu juga memiliki kelebihan dan kekurangannya masing-masing.

Dalam penelitian ini menggunakan teknik klasifikasi pohon keputusan.

2.1.5 Pohon keputusan

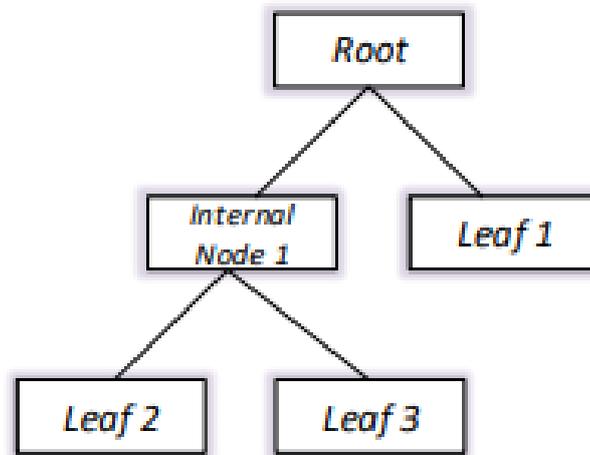
Salah satu teknik pengklasifikasian adalah pohon keputusan. Dalam penelitian (Song & Lu, 2015) menyebutkan bahwa pohon keputusan adalah alat

pendukung pengambil keputusan yang digunakan dalam penambangan data dan penelitian kecerdasan buatan dan memvisualisasikan proses pengambilan keputusan dalam bentuk struktur atau hirarki. Pengertian lainnya dari pohon keputusan merupakan metode yang melibatkan konstruksi pohon yang terdiri dari keputusan yang dihubungkan dengan cabang-cabang dari simpul akar sampai dengan *node* daun (Elisa, 2018). Konsep dasar dari suatu pohon keputusan adalah mengolah data menjadi pohon keputusan, memangkas pohon keputusan dan membuat aturan keputusan.

Pohon keputusan menggunakan representasi struktur pohon (*tree*) dimana setiap *node* mempresentasikan atribut, cabangnya mempresentasikan nilai dari atribut dan daun mempresentasikan kelas (Harman, 2018). Konsep pohon keputusan terdapat 3 jenis *node* yaitu (Khasanah, 2017):

1. *Root node* yaitu *node* paling atas, *node* ini tidak mempunyai *input* dan bisa juga tidak memiliki *output* atau mempunyai lebih dari satu *output*.
2. *Internal node* yaitu *node* percabangan, *node* ini hanya mempunyai satu *input* dan memiliki lebih dari satu *output*.
3. *Leaf node* atau terminal *node* yaitu *node* akhir, *node* hanya mempunyai satu *input* dan tidak memiliki *output*.

Untuk lebih mudah dimengerti konsep dasar pohon keputusan dapat dilihat pada gambar 2.5 dibawah ini (Pambudi, Setiawan, & Indriati, 2018):



Gambar 2.5 Konsep Dasar Pohon Keputusan

Manfaat utama menggunakan pohon keputusan dalam penelitian (Kamagi & Hansun, 2014) adalah kemampuan pohon keputusan untuk melakukan *breakdown* terhadap proses pengambilan keputusan yang kompleks menjadi lebih mudah sehingga pengambilan keputusan akan menjadi lebih menginterpretasikan solusi permasalahan. Dengan menggunakan pohon keputusan ada juga beberapa keuntungan yaitu :

1. Mudah dimengerti
2. Menyederhanakan hubungan kompleks antar variabel dengan membagi variabel *input* asli kedalam subkelompok yang signifikan
3. Pendekatan non-parametrik tanpa asumsi distribusi
4. Mudah menangani nilai-nilai yang hilang
5. Mudah menangani data tanpa melakukan transformasi data
6. Kuat untuk pecilan.

2.1.6 Algoritma C4.5

Salah satu algoritma untuk membangun pohon keputusan yang paling umum digunakan adalah algoritma C4.5. Algoritma C4.5 adalah pengembangan dari ID3, pengembangannya antara lain bisa mengatasi *missing value*, bisa mengatasi *continue data*, dan *praining*. Menurut (Rahman, 2015), algoritma C4.5 adalah suatu metode untuk membuat pohon keputusan berdasarkan data yang ada. Ide dasar dari algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai *gain* tertinggi berdasarkan nilai *entropy* atribut tersebut sebagai proses atribut klasifikasi. Untuk mencari *gain* tertinggi menggunakan rumus berikut.

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i)$$

Rumus 2.1 Perhitungan *Gain*

Dimana :

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Sementara itu, perhitungan nilai *entropy* dapat dilihat pada persamaan 2 berikut.

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i$$

Rumus 2.2 Perhitungan *entropy*

Di mana :

S : himpunan kasus

A : fitur

N : jumlah partisi S

p_i : proporsi dari S_i terhadap S

Pengujian hasil perhitungan *gain* dan *entropy* ini dapat diuji melalui *software Waikato Environment for Knowledge Analysis (WEKA)*. Didalam penelitian (Rifqo & Arzi, 2016), secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

1. Pilih atribut sebagai akar
2. Buat cabang untuk tiap-tiap nilai
3. Bagi kasus dalam cabang
4. Ulangi proses untuk setiap cabang sampai kasus pada cabang memiliki kelas yang sama.

2.1.7 Waikato Environment for Knowledge Analysis (WEKA)

Waikato Environment for Knowledge Analysis atau biasanya disingkat *WEKA* adalah paket *tools machine learning* praktis, *Waikato Environment for Knowledge Analysis (WEKA)* ini dikembangkan di Universitas Waikato yang terletak di Selandia Baru untuk penelitian, pendidikan dan berbagai aplikasi (Pujiono, Amborowati, & Suyanto, 2013). *WEKA* merupakan *software* yang bersifat *open source*, jadi siapapun dapat ikut menambahkan perangkat lunak ini (Haryati,

Sudarsono, & Suryana, 2015).

WEKA dapat digunakan untuk menyelesaikan masalah-masalah *datamining* di dunia nyata terutama klasifikasi yang menggunakan *machine learning* sebagai pendekatannya. *WEKA* mengandung *tools pre-processing* data, klasifikasi, regresi, *clustering*, aturan asosiasi, dan visualisasi. Pengujian dengan menggunakan *WEKA* lebih fleksibel dan efektif karena pohon keputusan yang terbentuk hampir sama dengan perhitungan secara manual (Azwanti, 2018).

2.2. Penelitian Terdahulu

Peneliti juga menemukan beberapa hasil penelitian terdahulu yang berkaitan dengan menganalisa menggunakan algoritma C4.5 yang dirangkum yaitu:

1. Penelitian yang berjudul “*implementation of decision trees algorithm C4.5*” yang dilakukan oleh (Chauhan & Chauhan, 2013) menunjukkan bahwa algoritma C4.5 menunjukkan hasil yang lebih akurat dan pengimplementasian pohon keputusan dengan metode algoritma C4.5 sangat mudah dan cepat. Algoritma ini lebih cocok untuk *data sets* yang mengandung tidak terlalu banyak data.
2. Penelitian yang berjudul “*Performance improvement of datamining in WEKA through GPU acceleration*” yang dilakukan oleh (Augusto & Char, 2014) menunjukkan bahwa proses yang memakan waktu banyak dapat mengadaptasi terhadap GPU. *show that set of operations that are time consuming can easily adapted to GPUs*. Ketika memparalelkan metode

penggunaan matriks, metode ini dapat beradaptasi untuk memanfaatkan *CPU* dan *GPU* sesuai dengan ukuran matriks, metode ini mempercepat proses hingga 49%.

3. Penelitian yang berjudul “*Performance evaluation of the GIS-based datamining techniques of best-first decision tree, random forest, and naive bayes tree for landslide susceptibility modeling*” yang dilakukan oleh (Chen, Zhang, Li, & Shahabi, 2018) menghasilkan bahwa *random forest model* merupakan yang paling akurat prediksinya.
4. Penelitian tentang meneliti kelayakan motor bekas dengan menggunakan metode *datamining* teknik algoritma C4.5 yang dilakukan oleh (Wira & Putra, 2016) menunjukkan bahwa aki, mesin dan aksesoris mempengaruhi penentuan tingkat kelayakan motor yang akan dijual kesimpulan ini dapat digunakan untuk membantu para penjual ataupun pembeli dalam menentukan kelayakan motor bekas yang akan dibeli ataupun dijual.
5. Penelitian tentang pengimplementasian teknik *datamining* metode algoritma C4.5 untuk mengukur tingkat resiko kredit yang dapat menentukan calon debitur dilakukan di Bank BRI Cabang Curup yang dilakukan oleh (Rifqo & Arzi, 2016) menghasilkan sebuah aplikasi sistem pendukung keputusan yang sangat membantu dalam pengambilan keputusan untuk menghindari kredit macet. Hasil penelitian ini dapat digunakan bank dan organisasi penyedia pinjaman lainnya untuk menghindari pembayaran kredit macet oleh nasabah.
6. Penelitian tentang pengimplementasian teknik *datamining* metode algoritma C4.5 untuk menentukan penerima beasiswa di STT Harapan Medan yang

dilakukan oleh (Rismayanti, 2016) menunjukkan bahwa IPK, penghasilan orang tua, dan jumlah tanggungan orang tua menjadi atribut yang berpengaruh dalam menentukan mahasiswa yang berhak menerima beasiswa dimana mahasiswa dengan IPK *Cumlaude* dan *very good* dan jumlah tanggungan orang tua paling banyak yang menerima beasiswa.

7. Penelitian tentang menentukan kelayakan penerima bantuan rehabilitasi sosial rumah tidak layak huni di pemerintah Kabupaten Solok Selatan dengan menerapkan teknik *datamining* metode algoritma C4.5 yang diteliti oleh (Zaman, 2016) menghasilkan penerima yang layak menerima bantuan rehabilitasi sosial rumah tidak layak huni adalah yang memiliki rumah dengan lantai rumah tanah, dan tidak memiliki sarana air bersih dan MCK.
8. Penelitian yang menganalisa jumlah pelanggan aktif dengan menerapkan teknik *datamining* metode algoritma C4.5 diteliti oleh (Jamhur, 2016) menunjukkan bahwa pelanggan dengan jumlah bayar tinggi, pembelian pertahun sedang, diskon rendah maka merupakan pelanggan aktif. Penelitian ini dapat membantu perusahaan dalam menyortir pelanggan yang aktif dan tidak aktif.
9. Penelitian tentang penggunaan *WEKA* untuk menganalisis kepuasan publik dalam mewujudkan *good governance* di kota Yogyakarta yang diteliti oleh (Pujiono et al., 2013) menunjukkan bahwa pelayanan yang diberikan Dinas Perizinan selama ini cukup baik sehingga pelaksanaan *good governance* dapat disimpulkan akan berjalan dengan baik.
10. Penelitian tentang memprediksi mahasiswa yang mengulang mata kuliah

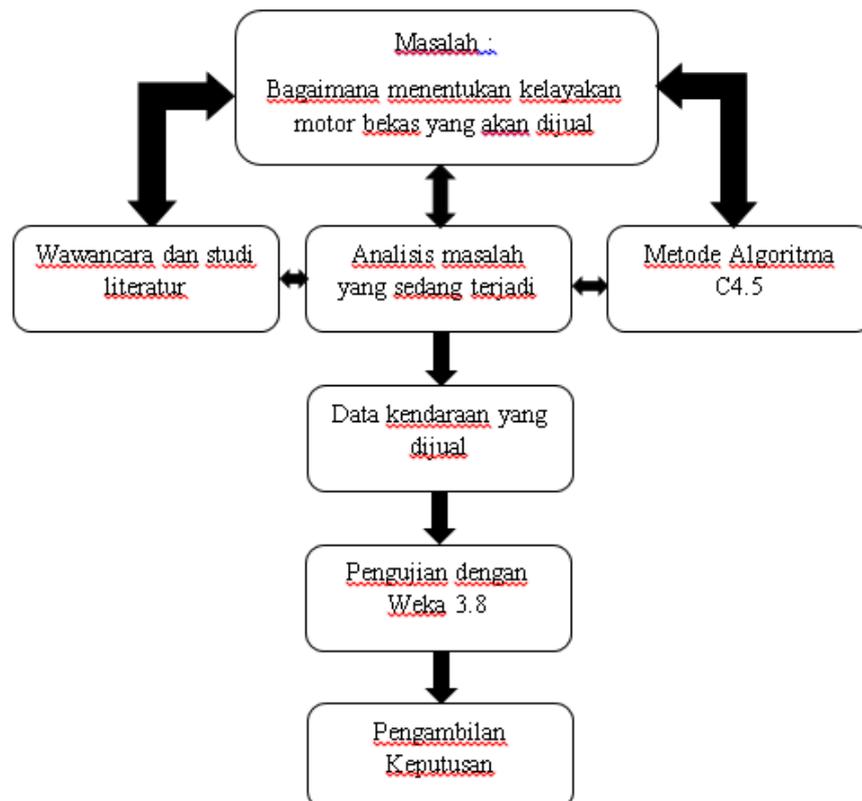
menggunakan teknik *datamining* metode algoritma C4.5 ini dilakukan di AMIK Labuhan Batu dan diteliti oleh (Azwanti, 2018) menyimpulkan bahwa algoritma C4.5 merupakan algoritma yang sangat membantu dalam mengklasifikasi data karena dapat memperoleh dengan jelas karakteristik data yang diklasifikasi baik dalam bentuk pohon keputusan maupun *if-then*.

11. Penelitian tentang memprediksi masa studi mahasiswa berdasarkan nilai akademik dengan menganalisis dan menerapkan teknik *datamining* metode algoritma C4.5 yang dilakukan oleh (Lorena et al., 2014) menghasilkan aplikasi untuk menentukan masa studi mahasiswa.
12. Penelitian tentang memprediksi profit pada perusahaan dengan menggunakan teknik *datamining* metode algoritma C4.5 yang dilakukan oleh (Elisa, 2018) menghasilkan bahwa biaya mobilisasi dan biaya tenaga kerja merupakan biaya yang paling banyak mempengaruhi profit perusahaan. Penelitian ini dapat dijadikan referensi bagi perusahaan yang ingin meningkatkan profitnya dengan cara *saving cost*.
13. Penelitian tentang memprediksi indikator *website* yang baik dengan menerapkan teknik *datamining* metode algoritma C4.5 yang dilakukan oleh (Harman, 2018) menunjukkan bahwa yang paling mempengaruhi indikator *website* yang baik adalah navigasi, selanjutnya adalah *graphic design*.
14. Penelitian tentang klasifikasi menggunakan algoritma C4.5 yang dilakukan oleh (Mardi, 2017) menghasilkan bahwa dengan menggunakan metode klasifikasi yaitu algoritma C4.5 dapat menghasilkan pohon keputusan yang dapat memberikan informasi yang dibutuhkan.

15. Penelitian tentang pengimplementasian teknik *datamining* untuk memprediksi masa studi dari mahasiswa dengan menggunakan metode Algoritma C4.5 yang dilakukan di Universitas Dehasen yang terletak di Bengkulu dengan pengimplementasian sistem baru pengambilan keputusan yang sangat kompleks sebelumnya menjadi lebih mudah dan spesifik.

2.3. Kerangka Pemikiran

Didalam kerangka pemikiran ini peneliti menjelaskan secara garis besar alur logika berjalannya penelitian ini. Berikut kerangka penelitian ini:



Gambar 2.6 Kerangka Pemikiran

2.4. Hipotesis

Hipotesis adalah dugaan sementara terhadap hasil penelitian yang masih harus diuji kebenarannya. Hipotesis peneliti dalam penelitian ini yaitu:

1. Teknik *datamining* dengan algoritma C4.5 dapat menganalisis data dari data kendaraan.
2. Hasil dari analisis data dengan algoritma C4.5 dapat digunakan sebagai standar atau acuan dalam menentukan kelayakan motor bekas yang akan dijual.

BAB III METODE PENELITIAN

3.1. Desain Penelitian

Desain penelitian adalah langkah-langkah untuk memecahkan masalah yang akan dibahas. Desain penelitian pada skripsi ini yaitu sebagai berikut:



Gambar 3.1 Desain Penelitian

1. Mendeskripsikan masalah

Dalam mendeskripsikan masalah yang akan diteliti harus menentukan dahulu batasan masalah penelitian, sehingga hasil dari penelitian akan lebih membantu karena penelitian tersebut dapat lebih terarah. Sehingga, meskipun ini

merupakan langkah awal tapi, ini merupakan langkah yang sangat penting.

2. Analisa masalah

Dalam langkah ini peneliti menganalisa lebih lanjut tentang deskripsi masalah yang sudah dijelaskan sehingga dapat dipahami lagi.

3. Mempelajari literatur

Peneliti mempelajari literatur-literatur yang sekiranya berkaitan dengan penelitian untuk dipelajari lebih lanjut untuk dipergunakan dalam penelitian.

4. Mengumpulkan data

Pada tahap ini, peneliti melakukan wawancara dengan mendatangi objek penelitian dan melakukan tanya jawab kepada pemilik dan pekerja di Amazone Motor untuk mendapatkan lebih banyak data dan informasi yang dibutuhkan.

5. Analisa teknik pengolahan data dengan algoritma C4.5

Dalam langkah ini, peneliti melakukan perancangan dengan algoritma C4.5 dari model sistem sampai terbentuknya pohon keputusan dan menghasilkan *rule* kelayakan motor bekas.

6. Implementasi Algoritma C4.5

Peneliti pada tahap ini, menentukan atribut sebagai akar dan menghitung nilai *gain* atribut, membentuk pohon keputusan dan mengubah pohon keputusan menjadi *rule*.

7. Pengujian hasil

Setelah menghitung nilai *gain* melalui perhitungan dan mendapat struktur pohon keputusannya, peneliti akan melakukan pengujian terhadap hasil perancangan dan implementasi dengan menggunakan *software WEKA 3.8.3*.

3.2. Operasional Variabel

Untuk menganalisis data dalam sebuah penelitian memerlukan variabel-variabel, data yang akan diolah adalah data kendaraan yang akan dijual dengan variabel penilaian kelayakan sebagai berikut:

1. Mesin
2. Rangka
3. *Body*
4. Aki
5. Aksesoris
6. Tahun Pembuatan

Untuk variabel keputusannya adalah layak dan tidak layak.

3.3. Populasi dan Sampel

3.3.1 Populasi

Populasi pada penelitian ini adalah data transaksi penjualan motor di Amazone Motor dari periode januari 2018 sampai dengan maret 2019.

3.3.2 Sampel

Data yang akan dijadikan sampel pada penelitian ini, peneliti akan mengambil data transaksi penjualan motor secara acak sebanyak 50 data sampel dari populasi data yaitu data periode januari 2018 sampai dengan maret 2019.

3.4. Teknik Pengumpulan Data

Pengumpulan data dilakukan untuk memperoleh informasi yang dibutuhkan untuk mencapai tujuan penelitian, atau mengungkap informasi atau kondisi lokasi penelitian sesuai dengan kebutuhan penelitian. Teknik pengumpulan data adalah sebagai berikut:

1. Teknik Studi Literatur adalah pembahasan literatur pada bidang tertentu dari sebuah penelitian. Studi ini adalah gambaran singkat dari apa yang telah dipelajari, argumentasi, dan ditetapkan tentang suatu topik, dan biasanya diorganisasikan secara kronologis atau tematis. Dalam hal ini, peneliti mempelajari berbagai jurnal yang memiliki keterkaitan terhadap permasalahan yang dihadapi.
2. Teknik Observasi adalah teknik pengumpulan data dengan cara peneliti melakukan pengamatan secara langsung dilapangan, dalam hal ini peneliti terjun langsung kepada Amazone Motor untuk mengamati bagaimana proses pembelian dan penjualan motor bekas yang akan dijual.
3. Teknik Dokumentasi adalah peneliti mempelajari sumber-sumber tertulis yang ada pada lokasi penelitian yang menyangkut masalah data yang telah ada sebelumnya, atau melakukan proses penggalian terhadap data yang sudah ada tentang dan kemudian mengelompokkan data tersebut sesuai dengan kebutuhan dari penelitian yang akan dilakukan. Dalam hal ini, peneliti mengamati dan mempelajari data-data kendaraan yang ada di Amazone Motor.
4. Teknik Wawancara adalah peneliti melakukan proses tanya jawab dalam penelitian yang berlangsung secara lisan dalam mana dua orang atau lebih

bertatap muka mendengarkan secara langsung informasi-informasi atau keterangan-keterangan. Dalam hal ini, peneliti melakukan tanya jawab terhadap pemilik dan pekerja di Amazone Motor.

3.5. Metode Analisis Data

Pada subbab ini menjelaskan metode analisis data yang akan digunakan peneliti pada penelitian ini. Dalam penelitian ini peneliti menggunakan teknik *datamining* dengan metode klasifikasi yaitu algoritma C4.5. Dalam metode ini data terlebih dahulu akan diklasifikasi sesuai dengan kategorinya. Hasil klasifikasi tersebut dihitung dengan menggunakan algoritma C4.5 meliputi pencarian nilai *entropy*, *gain* untuk membentuk suatu pohon keputusan. Proses akan dilakukan berulang kali hingga *node* pada pohon keputusan tersebut tidak dapat diturunkan lagi. Dari pohon keputusan tersebut akan menggambarkan *rules* atau aturan-aturan keputusan yang akan dijadikan sebagai acuan selanjutnya dalam penilaian kelayakan motor bekas yang akan dijual di Amazone Motor.

3.6. Lokasi dan Jadwal Penelitian

3.6.1 Lokasi penelitian

Penelitian ini berlokasi di Amazone Motor yang beralamat di Komplek Taman Sentosa Indah Blok E No.12, Sungai Panas, Batam, Kepulauan Riau.

3.6.2 Jadwal penelitian

Tabel 3.1 Jadwal Penelitian

No	Kegiatan	Mar-19	Apr-19	Mei-19	Jun-19	Jul-2019
1	Bimbingan					
2	Identifikasi Masalah					
3	Studi Literatur					
4	Penentuan Judul					
5	Pengumpulan Data					
6	Pengolahan Data					
7	Pengujian Hasil					
8	Penyusunan Laporan					