

**ANALISIS KEPUASAN KONSUMEN HINET BATAM
MENGUNAKAN METODE ALGORITMA C4.5**

SKRIPSI



**Oleh:
Sherly Maisa Putri
161510070**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2020**

**ANALISIS KEPUASAN KONSUMEN HINET BATAM
MENGUNAKAN METODE ALGORITMA C4.5**

SKRIPSI

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



**Oleh:
Sherly Maisa Putri
161510070**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN KOMPUTER
UNIVERSITAS PUTERA BATAM
2020**

SURAT PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini saya:

Nama : Sherly Maisa Putri
NPM : 161510070
Fakultas : Teknik dan Komputer
Program Studi : Sistem Informasi

Menyatakan bahwa skripsi yang saya buat dengan judul:

ANALISIS KEPUASAN KONSUMEN HINET BATAM MENGUNAKAN METODE ALGORITMA C4.5

Adalah hasil karya sendiri dan bukan “duplikasi” dari karya orang lain. Sepengetahuan saya, di dalam naskah skripsi ini tidak terdapat karya ilmiah atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis dikutip di dalam 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 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, 08 Februari 2020

Sherly Maisa Putri

161510070

ANALISIS KEPUASAN KONSUMEN HINET BATAM MENGUNAKAN METODE ALGORITMA C4.5

SKRIPSI
Untuk memenuhi salah satu syarat
guna memperoleh gelar Sarjana

Oleh
Sherly Maisa Putri
161510070

Telah disetujui oleh Pembimbing pada tanggal
seperti tertera di bawah ini

Batam, 08 Februari 2020

Sasa Ani Arnomo, S.Kom., M.SI.
Pembimbing

ABSTRAK

Kepuasan konsumen merupakan salah satu tujuan dari perusahaan dalam memberikan pelayanan kepada konsumennya baik perusahaan jasa maupun non jasa. Salah satu perusahaan jasa penyedia layanan akses data internet yang berkomitmen meningkatkan kualitas produk, fasilitas dan kualitas pelayanan terhadap kepuasan konsumen adalah Hinet Batam. Penelitian ini bertujuan untuk menganalisa kepuasan konsumen Hinet Batam dengan menggunakan proses *data mining* metode algoritma C4.5 dan mengukur tingkat kepuasan konsumen dari hasil pohon keputusan yang didapatkan. Variabel penilaian mencakup kualitas produk, fasilitas, pelayanan dan harga dengan hasil keputusan puas atau tidak puas. Hasil dari pengolahan metode *data mining* dengan menggunakan algoritma C4.5 ternyata mampu mengklasifikasi tingkat kepuasan konsumen. Terbukti dengan tingkat akurasi = 78.72%, *Root Mean Squared Error* (RMSE) = 0.4123, dan *Mean Absolute Error* (MSE) = 0.2408. Hasil penelitian ini bisa dijadikan acuan untuk meningkatkan kepuasan konsumen serta untuk mempertahankan loyalitas konsumen dalam menggunakan produk paket data layanan internet dan meningkatkan penjualan.

Kata kunci: *Data Mining*, Klasifikasi, Algoritma C4.5, Pohon Keputusan, Kepuasan Konsumen

ABSTRACT

Consumer satisfaction is one of the goals of the company in providing services to consumers both service and non-service companies. One of the internet data access service provider companies that is committed to improving product quality, facilities and service quality to customer satisfaction is Hinet Batam. This study aims to analyze the satisfaction of Hinet Batam consumers by using the data mining process C4.5 algorithm method and measure the level of customer satisfaction from the decision tree results obtained. The assessment variables include product quality, facilities, services and prices with the results of the decision satisfied or not satisfied. The results of processing data mining methods using the C4.5 algorithm were able to classify the level of customer satisfaction. Proven with an accuracy rate of 78.72%, Root Meas Squared Error (RMSE) = 0.4123, and Mean Absolute Error (MSE) = 0.2408. The results of this study can be used as a reference to improve consumer satisfaction and to maintain customer loyalty in using internet service data package products and increasing sales.

Keywords: *Data Mining, Classification, Algorithm C4.5, Decision Tree, Customer Satisfaction*

KATA PENGANTAR

Alhamdulillah, segala puji dan syukur atas rahmat dan karuniaNya serta kehadiran Allah Azza Wa Jalla yang telah melimpahkan segala ilmu, 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 di 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 Muhammad Rasid Ridho, S.Kom., M.SI., selaku Ketua Program Studi Sistem Informasi Universitas Putera Batam.
3. Bapak Nopriadi, S.Kom., M.Kom., selaku Pembimbing Akademik pada Program Studi Sistem Informasi di Universitas Putera Batam.
4. Bapak Sasa Ani Arnomo, S.Kom., M.SI., selaku Pembimbing Skripsi pada Program Studi Sistem Informasi Universitas Putera Batam.
5. Bapak / Ibu Dosen, seluruh Staff dan Civitas Universitas Putera Batam, yang telah memberi banyak pengetahuan pada penulis.

6. Bapak Fatria, selaku Supervisor Sales Hinet Batam yang telah memberikan persetujuan dan rekomendasi untuk melakukan penelitian ini.
7. Ibu Ani, selaku manajemen pada Hinet Batam yang telah menyetujui, memberikan informasi data dan bekerjasama dalam menyelesaikan skripsi ini.
8. Konsumen Hinet Batam yang sudah meluangkan waktu untuk mengisi kuesioner penelitian
9. Kedua orang tua penulis, Syafruddin dan Nurlina, S.H yang tak pernah henti selalu memberikan do'a, dukungan dan motivasi untuk menyelesaikan skripsi ini.
10. Kepada Saudara penulis, Irsyadi Rahmad Nur, Rossi Silvia, Abdul Wahid Nur memberikan motivasi beserta dukungan menyelesaikan skripsi ini.
11. Teman-teman Sistem Informasi angkatan 2016 memberikan semangat dan motivasi dalam penyelesaian skripsi ini.
12. Teman-teman alumni sekolah maupun teman kalangan luar: Sri Julianifah Sihotang, Restu Dinata Sari, Livia Oktora, Hendrawan Wardana memberikan semangat dan motivasi dalam penyelesaian skripsi ini.

Semoga Allah SWT membalas kebaikan dan selalu mencurahkan hidayah serta taufik-Nya, Aamiin.

Batam, 08 Februari 2020

Sherly Maisa Putri

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
SURAT PERNYATAAN ORISINALITAS	iii
HALAMAN PENGESAHAN	iiiv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
DAFTAR RUMUS	xiii
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Identifikasi Masalah.....	4
1.3. Pembatasan Masalah.....	4
1.4. Perumusan Masalah	5
1.5. Tujuan penelitian	5
1.6. Manfaat Penelitian	6
1.6.1. Aspek Teoritis.....	6
1.6.2. Aspek Praktis	6
BAB II TINJAUAN PUSTAKA	7
2.1. Teori Dasar	7
2.1.1. <i>Knowledge Discovery in Database</i>	7
2.1.2. <i>Data mining</i>	9
2.1.3. Kepuasan Konsumen	19
2.1.4. <i>Software WEKA 3.8.3</i>	22
2.2. Penelitian Terdahulu	28
2.3. Kerangka Pemikiran	34
2.4. Hipotesis Penelitian	34

BAB III METODE PENELITIAN	35
3.1. Desain Penelitian	35
3.2. Operasional Variabel	40
3.3. Populasi dan Sampel.....	42
3.3.1. Populasi	42
3.3.2. Sampel	42
3.4. Teknik Pengumpulan Data	45
3.5. Metode Analisa Data	45
3.6. Lokasi dan Jadwal Penelitian.....	46
3.6.1. Lokasi Penelitian	46
3.6.2. Jadwal Penelitian	47
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	48
4.1. Analisa Data.....	48
4.1.1. Data Hasil Kuesioner Kepuasan Konsumen.....	48
4.1.2. Menentukan Atribut Akar.....	58
4.1.3. Menyusun <i>Tree</i> dari Perhitungan <i>Node</i> 1	64
4.1.4. Menyusun <i>Tree</i> dari Perhitungan <i>Node</i> 1.3	68
4.1.5. <i>Rule</i> Hasil dari Perhitungan <i>Entropy</i> dan <i>Gain</i>	72
4.1.6. Pengujian Menggunakan <i>Software</i> WEKA	73
4.1.7. <i>Rule</i> Hasil dari Pengujian WEKA	74
4.1.8. Uji <i>Validation</i>	75
4.2. Pembahasan	76
BAB V KESIMPULAN DAN SARAN	78
5.1. Kesimpulan.....	78
5.2. Saran	79

DAFTAR PUSTAKA

LAMPIRAN

Lampiran 1. Pendukung Penelitian

Lampiran 2. Daftar Riwayat Hidup

Lampiran 3. Surat Keterangan Penelitian

DAFTAR GAMBAR

Gambar 2. 1 Tahap-tahap Data Mining	12
Gambar 2. 3 Pohon Keputusan (<i>Decision Tree</i>).....	18
Gambar 2. 4 Data Pra Proses	23
Gambar 2. 5 Halaman <i>Welcome</i> WEKA	24
Gambar 2. 6 <i>Design Preprocess</i> WEKA.....	24
Gambar 2. 7 <i>Open File</i>	25
Gambar 2. 8 Tampilan Setelah Terhubung.....	25
Gambar 2. 9 Hasil Keputusan.....	26
Gambar 2. 10 Tampilan Pohon Keputusan.....	27
Gambar 2. 11 Kerangka Pemikiran	34
Gambar 3. 1 Desain Penelitian	36
Gambar 3. 2 Kantor Hinet Batam	46
Gambar 4. 1 Grafik Lokasi Konsumen.....	49
Gambar 4. 2 Grafik Produk	49
Gambar 4. 3 Grafik Kualitas Produk	50
Gambar 4. 4 Grafik Fasilitas	51
Gambar 4. 5 Grafik Pelayanan (<i>Emphaty</i>).....	51
Gambar 4. 6 Grafik Pelayanan (<i>Responsiveness</i>).....	52
Gambar 4. 7 Pelayanan (<i>Reliability</i>).....	52
Gambar 4. 8 Grafik Harga	53
Gambar 4. 9 Grafik Hasil	53
Gambar 4. 10 Pohon Keputusan <i>Node 1.1</i>	64
Gambar 4. 11 Pohon Keputusan <i>Node 1.2</i>	68
Gambar 4. 12 Pohon Keputusan <i>Node 1.3</i>	72
Gambar 4. 13 Hasil Pohon Keputusan WEKA	74
Gambar 4. 14 <i>Cross Validation</i> algoritma C4.5	75

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu.....	28
Tabel 3. 1 Operasional Variabel.....	40
Tabel 3. 2 Sampel / <i>Data Set</i> Penelitian.....	43
Tabel 3. 3 Jadwal Penelitian.....	47
Tabel 4. 1 Data Hasil Kuesioner Kepuasan Konsumen.....	54
Tabel 4. 2 Penyederhanaan Data Kuesioner.....	56
Tabel 4. 3 <i>Set Data</i>	58
Tabel 4. 4 Hasil Perhitungan <i>Entropy</i> Total.....	58
Tabel 4. 5 Atribut dan Nilai-Nilai.....	59
Tabel 4. 6 <i>Node 1</i>	63
Tabel 4. 7 <i>Set Data</i>	64
Tabel 4. 8 Hasil Perhitungan <i>Entropy</i> Total.....	65
Tabel 4. 9 <i>Node 1.2</i>	67
Tabel 4. 10 <i>Set Data</i>	69
Tabel 4. 11 Tabel Hasil Perhitungan <i>Entropy</i> Total.....	69
Tabel 4. 12 <i>Node 1.3</i>	71

DAFTAR RUMUS

Rumus 2. 1 Perhitungan <i>Gain</i>	16
Rumus 2. 2 Perhitungan Nilai <i>Entropy</i>	16
Rumus 2. 3 Nilai Akurasi.....	18

BAB I

PENDAHULUAN

1.1. Latar Belakang

Internet bukanlah hal yang asing di era globalisasi sekarang ini. Perkembangan teknologi komunikasi dan informasi yang pesat, menjadikan internet sebagai salah satu kebutuhan pokok bagi masyarakat. Internet di masyarakat tentu tidak bisa berdiri dengan sendirinya, diperlukan dukungan dari perusahaan *Internet Service Provider* (ISP), sebagai penyedia sarana dan prasarana pendukung akses internet. Salah satu ISP di Indonesia adalah PT Berca Hardayaperkasa. Hinet merupakan layanan data akses internet pita lebar/*broadband* nirkabel yang menggunakan teknologi berbasis *4G LTE* dengan bantuan perangkat penghubung gelombang komunikasi *Base Transceiver Station* (BTS) dari PT Berca Hardayaperkasa yang saat ini beroperasi di 8 wilayah Indonesia yaitu, Denpasar, Makassar, Pekanbaru, Batam, Medan, Palembang, Pontianak, dan Balikpapan.

Hinet Batam berdiri sejak Januari 2019 dan memiliki 4 BTS (*Base Transceiver Station*) di 4 wilayah yang ada di kota Batam yaitu Sagulung, Batu Aji, Tiban dan Bengkong. Dengan terbatasnya keberadaan BTS maka pengguna layanan akses data internet Hinet Batam masih terbatas yaitu masyarakat yang hanya berada tidak lebih dari 100meter dari lokasi BTS. Sebagai penyedia layanan

akses data internet yang baru hadir di kota Batam tekanan kompetisi, ketersediaan sarana dan prasarana serta jaringan yang tersebar menjadikan Hinet Batam memiliki tingkat daya saing yang tinggi dalam memenangi persaingan pasar industri telekomunikasi khususnya yang ada di kota Batam. Kondisi semacam ini memaksa penyedia layanan akses data internet Hinet Batam untuk melakukan strategi yang tepat untuk meningkatkan serta mempertahankan produk layanan akses data internet di pasaran.

Dengan banyaknya penyedia layanan akses data internet yang sudah mendapatkan *branding* di kota Batam maka perlu kiranya Hinet Batam mengenal seperti apa kelebihan ataupun kekurangan produknya di pasaran terlebih adanya keluhan dari konsumen dan juga adanya konsumen yang beralih ke *provider* lain. Masalah keluhan dari konsumen di antaranya, kualitas layanan jaringan seperti sinyal yang terkadang bermasalah, kecepatan internet yang lambat, fasilitas *customer service* yang kurang memadai, serta kualitas pelayanan yang kurang responsif (Spv. Sales Hinet Batam). Oleh karenanya, perlu dilakukan riset mengenai kepuasan konsumen untuk mempertahankan pengguna layanan akses data internet Hinet Batam bahkan untuk meningkatkan pengguna layanan akses data internet Hinet Batam.

Setiap konsumen memiliki tingkat kepuasan yang berbeda-beda. Bila sebuah produk sesuai dengan yang diharapkan tentunya akan mendapatkan kepuasan tersendiri bagi konsumen sehingga akan terus menggunakannya, sedangkan bila produknya tidak sesuai dengan yang diharapkan tentu akan ada rasa ketidakpuasan. Jadi kepuasan konsumen mencakup antara tingkat

kepentingan dan kinerja atau hasil yang dirasakan serta evaluasi purna beli alternatif yang dipilih dapat memberikan hasil yang sama atau melampaui harapan konsumen.

Berdasarkan masalah di atas penelitian kepuasan konsumen tersebut akan dianalisis menggunakan pendekatan *data mining* dengan algoritma C4.5 untuk menentukan pola kepuasan konsumen. *Data mining* atau disebut juga sebagai penambangan data yang berguna untuk menemukan pola atau relasi antar data secara otomatis. Dengan menggunakan metode pohon keputusan algoritma C4.5 yang bisa menampilkan hasil perhitungan berupa pohon keputusan yang sangat mudah dipahami dan dimengerti, yang mana nantinya akan ditampilkan adalah tingkat kepuasan atau ketidakpuasan yang dirasakan oleh konsumen. Sehingga hal tersebut bisa memberikan gambaran kepada pimpinan perusahaan untuk bisa membuat strategi ke depannya baik untuk mempertahankan konsumen atau pun menambah jumlah konsumen pengguna layanan akses data internet. Untuk memenuhi kebutuhan-kebutuhan di atas, banyak cara yang dapat ditempuh. Salah satunya adalah menggunakan kuesioner *online* dengan memberikan seperangkat pertanyaan yang telah disediakan pilihan jawaban untuk dipilih oleh objek penelitian dengan menggunakan beberapa indikator yaitu, kualitas produk, fasilitas, pelayanan, dan harga. Oleh karena itu, maka penulis berkeinginan menuangkan penelitian ke dalam bentuk skripsi dengan judul “ANALISIS KEPUASAN KONSUMEN HINET BATAM MENGGUNAKAN METODE ALGORITMA C4.5”

1.2. Identifikasi Masalah

Berlandaskan konteks di atas, penulis menyimpulkan adanya permasalahan sebagai berikut:

1. Kecepatan sinyal akses data internet yang dirasakan konsumen belum sesuai harapan konsumen sehingga terjadinya keluhan dari konsumen.
2. Fasilitas *customer service* yang kurang memadai.
3. Kualitas pelayanan yang kurang responsif ketika konsumen menghubungi untuk menyampaikan keluhan.

1.3. Pembatasan Masalah

Supaya penelitian ini lebih terstruktur dan tidak melewati dari pokok permasalahan maka peneliti membatasi penelitian ini diantaranya:

- 1 Survey kuesioner *online* hanya dilakukan pada konsumen dalam data penjualan periode 29 Januari 2019 hingga 5 Agustus 2019.
- 2 Penelitian ini dilakukan dengan menggunakan *data mining* dengan klasifikasi pohon keputusan.
- 3 Pengujian pada penelitian ini menggunakan *software* WEKA 3.8.3.

1.4. Perumusan Masalah

Berlandaskan indikasi di atas maka permasalahan dapat dirunutkan sebagai berikut:

1. Bagaimana cara menggunakan *data mining* dengan algoritma C4.5 untuk mengetahui kepuasan konsumen Hinet Batam?
2. Bagaimana pohon keputusan yang dihasilkan dalam mengukur tingkat kepuasan konsumen?

1.5. Tujuan penelitian

Berlandaskan perumusan di atas, maka tujuan yang akan dicapai pada penelitian ini adalah:

1. Untuk mengetahui proses *data mining* dengan metode algoritma C4.5 pada kepuasan konsumen Hinet Batam.
2. Untuk mengetahui pohon keputusan yang dihasilkan dalam mengukur tingkat kepuasan konsumen.

1.6. Manfaat Penelitian

Manfaat penelitian merupakan suatu yang signifikan dari sebuah penelitian yang bisa dimanfaatkan dan digunakan setelah penelitian. Manfaat yang ingin dicapai dari hasil penelitian ini adalah:

1.6.1. Aspek Teoritis

Digunakan untuk acuan materi dalam dunia pendidikan seperti sebagai media ekspansi ilmu pengetahuan dan informasi tentang analisis kepuasan konsumen dengan metode data mining algoritma C4.5.

1.6.2. Aspek Praktis

- a. Bagi peneliti adalah untuk mengetahui hasil dari analisis kepuasan konsumen menggunakan metode *data mining* algoritma C4.5.
- b. Bagi pembaca adalah sebagai bahan tambahan ilmu pengetahuan terutama dalam aspek ilmu *data mining*, juga sebagai acuan untuk pihak-pihak yang berkepentingan misalnya kemudahan yang didapatkan oleh pihak perusahaan dalam pengelolaan dan pemanfaatan hasil penelitian.

BAB II

TINJAUAN PUSTAKA

2.1. Teori Dasar

2.1.1. *Knowledge Discovery in Database*

Data mining juga dikatakan *Knowledge Discovery in Database* (KDD) membentuk aktivitas dengan cakupan pengumpulan, pemakaian, historis yang akan menghasilkan kesesuaian, pola atau hubungan pada *set data* yang berukuran besar (Hariyadi, 2016). *Knowledge Discovery in Database* (KDD) merupakan teknik atau cara pencarian yang mengimplikasikan pelaksanaan berbagai metode algoritma guna memanipulasi data, membentuk model dari data, juga memanipulasi model (Elisa, 2018).

Hal-hal yang penting pada teknik KDD sebagai berikut (Sijabat, 2015):

1. Data selection

- a. Menemukan himpunan data target, penentuan himpunan data, maupun memfokuskan dengan menggunakan *subset* variabel atau sampel data, di mana pada teknik ini akan dilakukan proses *discovery*.
- b. Pemilihan seleksi data pada kumpulan data operasional penting diterapkan sebelum proses pencarian informasi pada KDD dilakukan. Hasil dari data penyeleksian yang nantinya diterapkan guna pemrosesan *data mining*, disimpan pada sebuah *file*, yang dipisah dari basis data operasional.

2. *Pre-processing/Cleaning*

- a. Pembersihan data sebagai operasi dasar misalnya melakukan penghapusan *noise*.
- b. Sebelum memasuki tahapan *data mining* dilakukan, maka penting diterapkannya terlebih dahulu tahapan penghapusan(*cleaning*) terhadap data yang menjadi fokus KDD.
- c. Tahapan penghapusan(*cleaning*) diantaranya membuang duplikasi data, memeriksa data tidak konsisten, juga memperbaiki kesalahan pada data, misalnya kesalahan pada cetak (tipografi).
- d. Melakukan tahapan *enrichment*, ialah tahapan “memperkaya” data yang sudah ada dan data atau informasi lainnya yang relevan juga dibutuhkan guna KDD, misalnya data ataupun informasi eksternal.

3. *Transformation*

- a. Pencarian fitur-fitur yang berguna untuk mempresentasikan data bergantung kepada *goal* yang ingin dicapai.
- b. Merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses ini merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data mining*

- a. Pemilihan tugas *data mining*: pemilihan *goal* dari proses KDD misalnya klasifikasi, *regresi*, *clustering*, dll.
- b. Pemilihan algoritma *data mining* untuk pencarian (*searching*).

- c. Proses *data mining* yaitu proses mencari pola atau informasi menarik dalam data terpilih menggunakan teknik atau metode tertentu. Teknik, metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.
5. *Interpretation/Evaluation*
- a. Penerjemahan pola-pola yang dihasilkan dari *data mining*.
 - b. Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan.
 - c. Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.1.2. Data mining

Data mining ialah sebuah tahapan penganalisaan data secara otomatis. Data yang digunakan ialah data yang sangat besar, dan sukar dilakukan secara manual. Dengan tujuan supaya menghasilkan hubungan atau pola yang informatif (Faisal, 2019).

Data mining juga berkaitan dalam pencarian aturan-aturan atas data yang tersembunyi. *Data mining* saat ini sebagai teknik yang sangat berpengaruh dalam perobahan sebuah data menjadi informasi yang bermanfaat. Sebab banyak dipakai pada berbagai bidang misalnya promosi, pengecekan, pelacakan hingga penemuan ilmiah. Sudah digunakan bertahun-tahun pada perusahaan bisnis, ilmuan, juga

pemerintahan sebagai penyaringan jumlah data misalnya histori penerbangan, sensus penduduk, juga minimarket *scanner* data yang membentuk suatu laporan riset pasar (Sijabat, 2015).

Data mining bertujuan sebagai penyederhanaan informasi dari sekumpulan data besar. Penyederhanaan informasi dilaksanakan berlandaskan metode *data mining* yang nantinya akan dipakai. Untuk melakukan penyederhanaan informasi menggunakan beberapa metode dan algoritma *data mining* diantaranya, metode klasterisasi, metode estimasi, metode klasifikasi, metode prediksi, dan metode asosiasi (Takalapeta, 2018).

Jadi menurut kesimpulan penulis *data mining* ialah pola/aturan pencarian informasi dalam data yang berjumlah besar dan tersembunyi untuk bisa menghasilkan suatu informasi yang bermanfaat. Dengan menggunakan metode-metode yang sudah ada bergantung dengan tujuan dari penggunaan data mining tersebut.

2.1.2.1 Tahapan *Data Mining*

a. Pembersihan data (*data cleaning*)

Merupakan tahapan untuk menghilangkan *noise* dan data yang inkonsisten.

b. Integrasi data (*data integration*)

Menggabungkan data dari beragam *database* pada suatu *database* yang baru.

c. Seleksi data (*data selection*)

Data yang ada pada *database* sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari *database*.

d. Transformasi data (*data transformation*)

Pengubahan data atau pengabungan data ke dalam aturan yang sesuai pada proses *data mining*.

e. Proses *mining* (*process mining*)

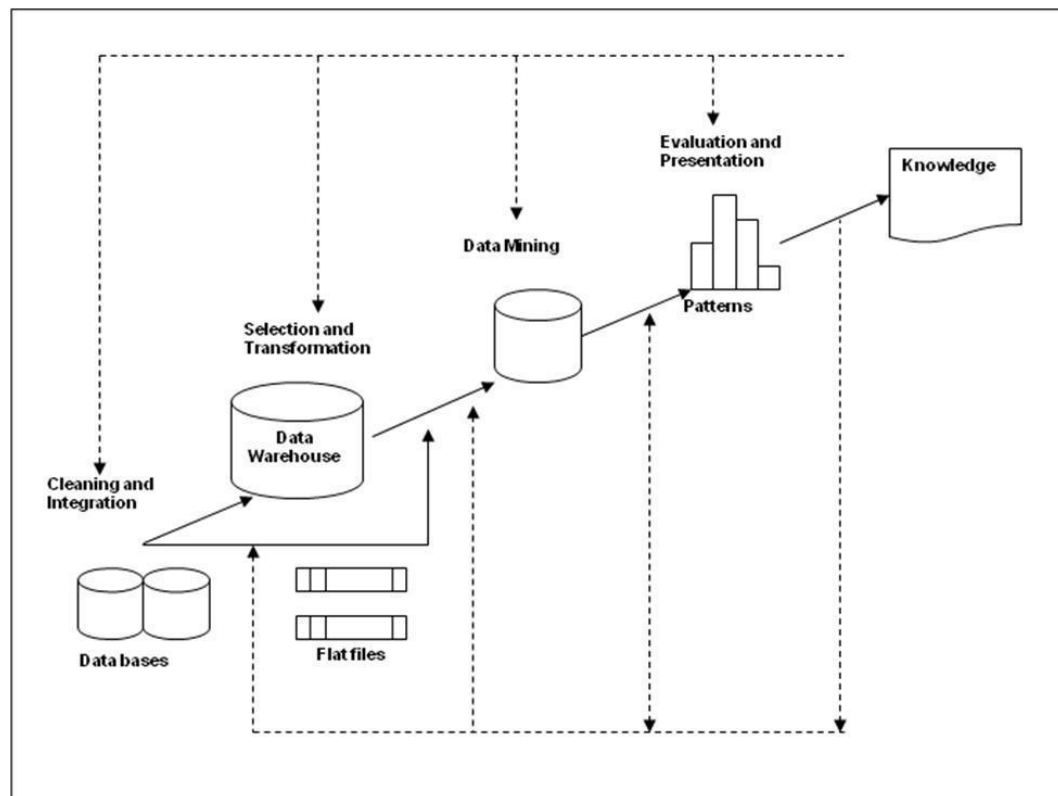
Merupakan tahapan penemuan pengetahuan dari data yang tersembunyi.

f. Evaluasi pola (*pattern evaluation*)

Untuk mendapatkan pola yang menarik ke dalam *knowledge based* yang didapatkan bisa bermanfaat.

g. Presentasi pengetahuan

Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil *data mining* (Hariyadi, 2016).



Gambar 2. 1 Tahap-tahap Data Mining

2.1.2.2 Manfaat *Data mining*

Manfaat *data mining* dilandaskan berdasarkan dua sudut pandang, yaitu sudut pandang komersial dan sudut pandang keilmuan. Dari sudut pandang komersial, pemanfaatan *data mining* dapat digunakan untuk menangani meledaknya volume data, dengan menggunakan teknik komputasi dapat digunakan untuk menghasilkan informasi-informasi yang dibutuhkan yang merupakan *asset* yang dapat meningkatkan daya saing suatu institusi. Sedangkan manfaat *data mining* dilihat berdasarkan sudut pandang keilmuan, dimanfaatkan

sebagai penganalisisan juga penyimpanan data dan memiliki sifat sangat besar juga nyata (Vulandari, 2017).

2.1.2.3 Fungsi Data mining

Fungsi *data mining* terbagi atas dua landasan yaitu deskriptif dan prediktif. Secara deskriptif, *data mining* sebagai pencarian aturan-aturan supaya bisa dimengerti oleh manusia dalam mendetailkan karakteristik data. Secara prediktif, *data mining* sebagai pembentukkan suatu bentuk ilmu yang dapat dipakai pada proses prediksi (Suyanto, 2017).

Berlandaskan dari teori fungsi *data mining* dapat dikategorikan sebagai berikut (Vulandari, 2017):

1. *Assosiation*, suatu tahapan untuk mendapatkan aturan asosiasi antara sebuah kombinasi *item* pada sebuah waktu.
2. *Sequence*, suatu tahapan untuk mendapatkan aturan asosiasi antara sebuah kombinasi *item* pada sebuah waktu dan digunakan melebihi satu periode.
3. *Clustering*, suatu tahapan pengelompokkan sejumlah data / obyek ke dalam kelompok data hingga tiap kelompok memiliki data yang sama.
4. *Classification*, merupakan suatu tahapan pencarian model atau fungsi yang mendetailkan atau memilah konsep dan kelas data, dengan tujuan supaya bisa memprediksi kelas dari sebuah objek yang labelnya tidak ada.
5. *Regression*, merupakan suatu tahapan pemetaan data pada sebuah nilai prediksi.

6. *Forecasting*, merupakan suatu tahapan perkiraan nilai prediksi berlandaskan pola-pola dari dalam kumpulan data.
7. *Solution*, merupakan suatu tahapan menemukan akar permasalahan dan *problem solving* dari masalah bisnis yang dihadapi atau setidaknya menjadi pengambilan keputusan untuk informasi.

2.1.2.4 Klasifikasi

Klasifikasi ialah tahapan penemuan pola atau fungsi-fungsi yang menjelaskan dan membagi kelas data satu dengan lainnya, agar bisa digunakan untuk prediksi dari data yang belum memiliki kelas tertentu (Febriyanto, Handoko, & Aisyah, 2018). Pada metode klasifikasi memiliki kategori variabel target. Metode/model yang pernah digunakan oleh peneliti dalam menuntaskan penelitian klasifikasi diantaranya pohon keputusan, jaringan saraf tiruan, dll (Faisal, 2019).

Kegunaan dari model klasifikasi diantaranya (Elisa, 2018):

- a. Model Deskriptif digunakan sebagai alat ilustrasi untuk memisahkan dari obyek yang berasal dari kelas yang berbeda.
- b. Mode Prediktif digunakan sebagai prediksi label sebuah kelas dari *record* yang belum teridentifikasi.

2.1.2.5 Algoritmas C4.5

Algoritma C4.5 ialah kelompok dari algoritma *decision tree*. Algoritma ini memiliki masukan berupa *samples* dan *training samples*. *Samples* merupakan *field-field* data yang akan dimanfaatkan sebagai parameter dalam melakukan klasifikasi data. Sedangkan *training samples* seperti data contoh yang akan digunakan untuk membangun sebuah *tree* (struktur pohon) yang sudah diuji kebenarannya. Algoritma C4.5 ialah hasil perkembangan dari ID3 (*Iterative Dichotomiser 3*). Beberapa hal yang merupakan perbaikan dari algoritma ID3 ke algoritma C4.5 (Takalapeta, 2018):

1. Mampu mengatasi *missing value*
2. Mampu mengatasi *data kontinyu*
3. *Pruning*
4. Adanya aturan

Pada penggunaan algoritma C4.5 terlihat berbagai tahapan yaitu perubahan bentuk data pada tabel menjadi model pohon kemudian perubahan model pohon menjadi aturan (*rule*) dan terakhir penyederhanaan *rule*. Secara umum, algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut:

- a. Menghitung jumlah data, jumlah data berlandaskan dari anggota atribut hasil menggunakan persyaratan tertentu. Pada tahapan awal persyaratannya masih kosong.
- b. Memilih atribut sebagai *Node*.
- c. Membuat cabang untuk tiap-tiap anggota dari *Node*.

- d. Periksa apakah nilai *entropy* dari anggota *node* ada yang bernilai nol. Jika ada, tentukan daun yang terbentuk. Jika seluruh nilai *entropy* anggota *node* adalah nol, maka proses pun dihentikan.
- e. Jika ada anggota *node* yang memiliki nilai *entropy* lebih besar dari nol, ulangi lagi proses dari awal dengan *node* sebagai syarat sampai semua anggota dari *node* bernilai nol. *Node* ialah atribut yang memiliki nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk melakukan perhitungan nilai *gain* suatu atribut menggunakan rumus seperti yang tertera dalam persamaan berikut (Faisal, 2019):

$$Gain(S, A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} * Entropy(S_i) \quad \text{Rumus 2. 1 Perhitungan Gain}$$

Keterangan:

S : himpunan kasus

A : atribut

N : jumlah partisi atribut A

|S_i| : jumlah kasus pada partisi ke-i

|S| : jumlah kasus dalam S

Perhitungan nilai *entropy* dapat dilihat dalam persamaan berikut:

$$Entropy(s) = \sum_{i=1}^n - p_i * \log_2 p_i \quad \text{Rumus 2. 2 Perhitungan Nilai Entropy}$$

Keterangan:

S :himpunan kasus

A : Fitur

n : Jumlah

p_i : proporsi dari S_i terhadap S

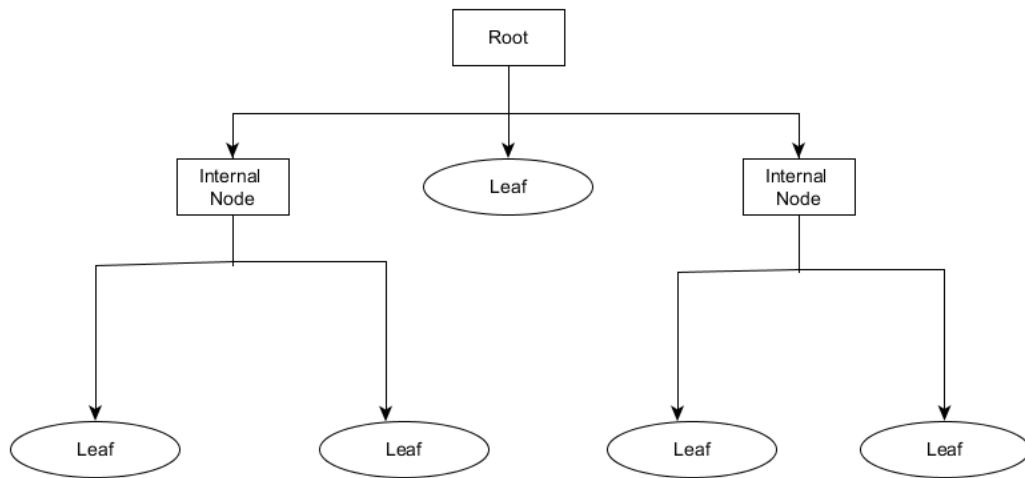
2.1.2.6 Pohon Keputusan (*Decision Tree*)

Pohon keputusan merupakan suatu metode klasifikasi yang sangat populer karena mudah untuk dimengerti oleh manusia. Pohon keputusan merupakan model prediksi yang memiliki struktur berhirarki atau struktur pohon. Konsep dari pohon keputusan dengan melakukan pengubahan data menjadi pohon keputusan dan memiliki aturan-aturan dari keputusan.

Manfaat utama dari pemakaian pohon keputusan ialah kemampuannya untuk mem-*break down* tahapan pengambilan keputusan dari yang rinci menjadi lebih *simple* sehingga pengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan. Pohon keputusan digunakan juga untuk mengeksplorasi data dan menemukan hubungan yang tersembunyi (Faisal, 2019).

Pada pohon keputusan (*decision tree*) terdapat tiga jenis *node*, yaitu (Sijabat, 2015):

1. *Root Node*, merupakan *node* yang paling atas. Pada *node* ini tidak ada input dan bisa tidak mempunyai *output* atau mempunyai output lebih dari satu.
2. *Internal Node*, merupakan *node* percabangan. Pada *node* ini hanya terdapat satu input dan mempunyai *output* minimal dua.
3. *Leaf Node* atau *Terminal Node*, merupakan *node* akhir. Pada *node* ini hanya terdapat satu input dan tidak mempunyai *output*.



Gambar 2. 2 Pohon Keputusan (*Decision Tree*)

2.1.2.7 Uji *Cross Validation*

Cross Validation ialah suatu teknik untuk memvalidasi atau menilai akurasi suatu model yang dihasilkan melalui *data set* tertentu. Salah satu teknik *cross validation* yang sering digunakan yaitu *k-fold cross validation*. Dalam teknik *k-fold cross* dilakukan sejumlah K-kali ujicoba, di mana masing-masing uji coba tersebut menggunakan data partisi ke-K sebagai data uji coba , di mana masing-masing uji coba tersebut menggunakan data partisi ke-K sebagai data uji dan sisa partisi lainnya sebagai data latih. Untuk mengetahui nilai akurasi pada uji *k-fold cross validation* menggunakan persamaan berikut (Febriana, Furqon, & Rahayudi, 2018):

$$\text{Akurasi} = \frac{\text{Jumlah data benar}}{\text{Jumlah data uji}} \times 100\%$$

Rumus 2. 3 Nilai Akurasi

2.1.3. Kepuasan Konsumen

Kepuasan konsumen mencakup perbedaan antara tingkat kepentingan dan kinerja atau hasil yang dirasakan, serta merupakan evaluasi purna beli dimana alternatif yang dipilih sekurang-kurangnya dapat memberikan hasil sama atau melampaui harapan. Sedangkan ketidakpuasan dapat terjadi apabila hasil yang diperoleh tidak memenuhi harapan yang diinginkan konsumen (Takalapeta, 2018).

Seorang konsumen menilai tingkat kepuasannya yaitu dengan melihat baik atau tidaknya kualitas dari sebuah pelayanan. Konsumen akan membandingkan pelayanan yang mereka terima. Pada umumnya kepuasan tiap konsumen berbeda-beda, tingkat kepuasan yang berbeda ini adalah indikator yang baik untuk mengukur tingkat kualitas pelayanan yang diberikan. Ada salah satu cara untuk meningkatkan kepuasan konsumen yaitu berusaha untuk menjalin hubungan yang baik dengan konsumen dan juga mau mendengar semua keluhan konsumen tentang pelayanan yang diberikan kepada kepuasan tersebut. Konsumen yang merasa puas dengan pelayanan yang diberikan maka konsumen akan datang kembali dan konsumen yang merasa puas akan berbagi pengalaman tentang pelayanan yang diterima kepada konsumen lain (Febriarini & Astuti, n.d.).

Terdapat beberapa hal penting yang ada dalam kepuasan konsumen, diantaranya (Nareswari & T, 2016):

1. Kualitas produk, yang merupakan performansi dari produk atau jasa yang ditawarkan.
2. Harga, dimana perusahaan menentukan harga sesuai dengan sensitifitasnya.

3. Kualitas jasa, dimana perusahaan / pihak penyedia layanan jasa memberikan atau menawarkan jasa sesuai dengan layanan yang ada.
4. Emosi (*emotional factor*), merupakan dimensi yang relatif penting dimana perusahaan / pihak penyedia layanan jasa harus mampu memberikan rasa percaya diri serta kebanggaan tersendiri kepada para pelanggannya terhadap produk atau jasa yang digunakan.

Berdasarkan penelitian di atas peneliti menyimpulkan bahwa kepuasan konsumen adalah ketika kebutuhan konsumen terpenuhi atas produk yang digunakan dan pelayanan yang diberikan. Kepuasan konsumen akan menentukan loyalitas seorang konsumen untuk bertahan menggunakan produk dan pelayanan yang diberikan.

2.1.3.1 Kualitas Produk

Produk adalah segala sesuatu yang dapat ditawarkan perusahaan ke pasar untuk diamati, disukai, dan dibeli sehingga dapat memuaskan keinginan dan kebutuhan konsumen (Zahrah, 2017). Semua konsumen menginginkan produk yang tidak hanya murah, namun juga berkualitas. Dengan menciptakan produk yang berkualitas, perusahaan akan memperoleh kepercayaan serta menciptakan kepuasan konsumen (Insyroh & Setyowati, 2018). Kualitas produk merupakan performansi dari produk atau jasa yang digunakan (Nareswari & T, 2016).

2.1.3.2 Pelayanan

Kualitas pelayanan merupakan tingkatan kondisi baik buruknya sajian yang diberikan oleh perusahaan / penyedia jasa layanan. Dengan meningkatkan kualitas layanan maka konsumen akan memiliki sikap loyalitas terhadap produk yang digunakan (Rangga & Nalendra, 2018).

Kualitas pelayanan yang terdapat pada industri *Internet Service Provider* (ISP) yaitu (Nareswari & T, 2016):

- a. *Information quality & website information support*. Informasi yang diberikan harus lengkap dan detail namun tetap harus singkat, padat, dan jelas.
- b. *Security & privacy* merupakan salah satu dimensi yang penting dari kualitas jasa pada industri ISP yang harus disediakan guna menjaga data/identitas diri dari konsumennya.
- c. *Customer service & technical support* merupakan bagian yang penting dari penggunaan layanan jasa *provider* setiap harinya.

2.1.3.3 Fasilitas

Fasilitas merupakan sumber daya fisik yang harus ada sebelum suatu jasa dapat ditawarkan kepada konsumen. Fasilitas dapat berupa segala sesuatu yang memudahkan konsumen dalam memperoleh kepuasan. Karena suatu bentuk jasa tidak bisa dilihat, tidak bisa dicium dan tidak bisa diraba maka aspek wujud fisik menjadi penting sebagai ukuran dari pelayanan. Konsumen yang ingin mencari kenyamanan selama proses menunggu pelaksanaan layanan jasa akan lebih

merasa nyaman bila fasilitas yang digunakan oleh pelanggan dibuat nyaman dan menarik (Moha & Loindong, n.d.).

2.1.3.4 Harga

Harga adalah sejumlah uang yang harus dibayar oleh konsumen untuk memperoleh produk. Penetapan harga harus disesuaikan dengan keinginan kebutuhan dan kemampuan konsumen agar perusahaan dapat menarik minat dan mampu mempertahankan kesetiaan konsumen. Selain itu, Penetapan harga dapat memperoleh laba (Insyroh & Setyowati, 2018). Harga juga seringkali dijadikan sebagai bahan pertimbangan bagi konsumen dalam melakukan pembelian (Devi & Hoyyi, 2015).

2.1.4. Software WEKA 3.8.3

Implementasi dan pengujian dari penelitian ini memakai *Software WEKA 3.8.3*. Dengan pengujian data memakai *software WEKA 3.8.3* peneliti hendak melakukan perbandingan dari hasil pengolahan data secara manual dengan memakai sebuah *software* atau perangkat dari sebuah aplikasi yang sudah ada.

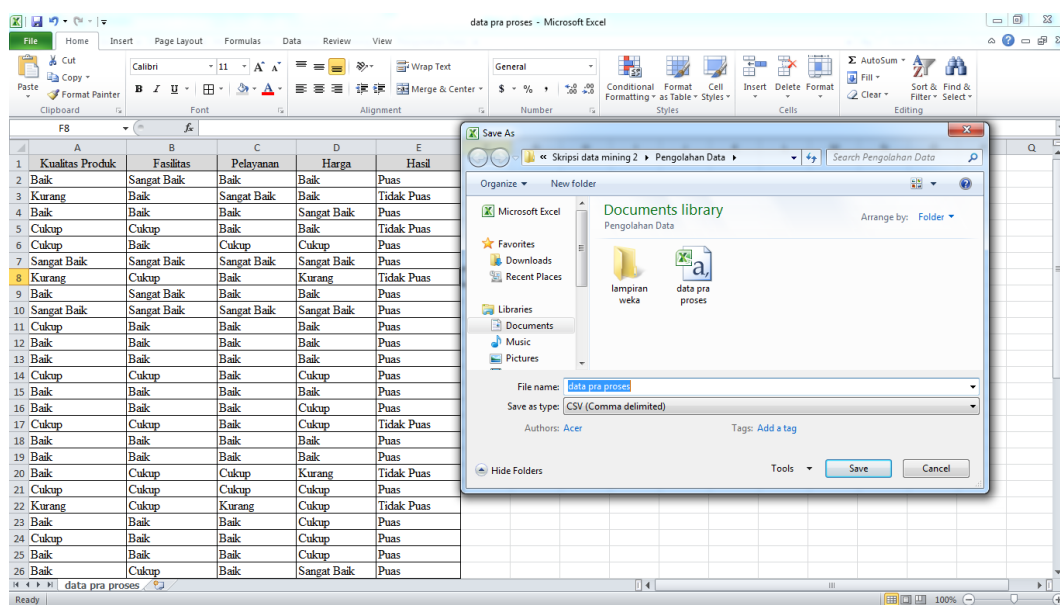
Waikato Environment for Knowledge Analysis (WEKA) ialah suatu perangkat lunak yang mengimplementasikan algoritma *machine learning* untuk dilakukannya berbagai proses yang bersangkutan dengan sistem temu pada informasi atau *data mining*. *Software WEKA* menyediakan teknik klasifikasi

dengan metode algoritma *decision tree J48*. Teknik klasifikasi dan algoritma yang dipakai disebut *Classification* (Fitriani, Putera, & Siahaan, 2016).

Proses pengujian dengan aplikasi WEKA:

1. Proses Persiapan *Data Set*

Sebagai *data set* yang digunakan untuk pengujian dengan algoritma C4.5 dengan persiapan data dari *excel CSV*.



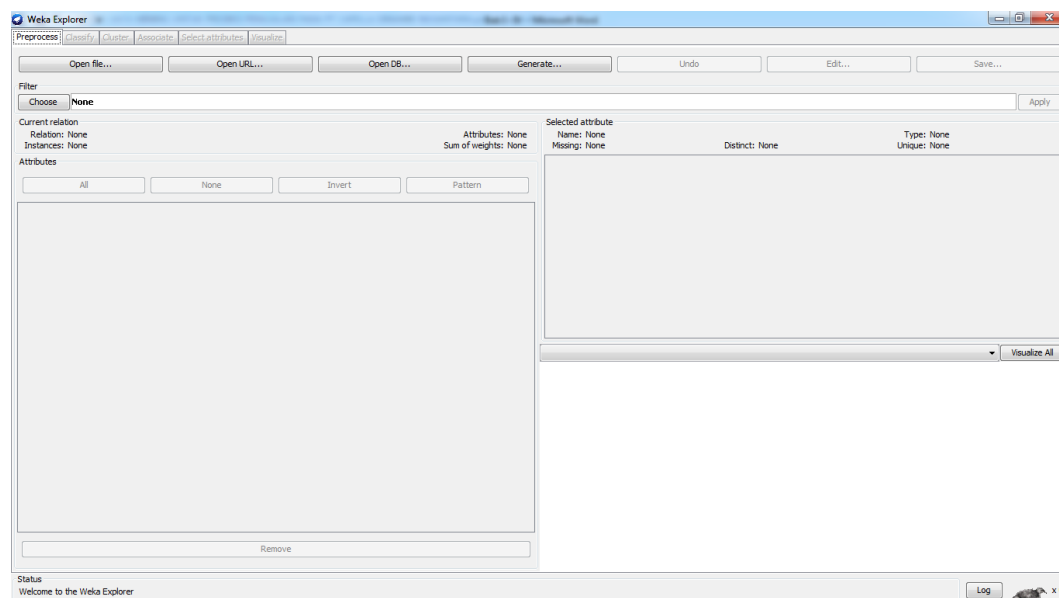
Gambar 2.3 Data Pra Proses

2. Menghubungkan Data Dengan Aplikasi WEKA



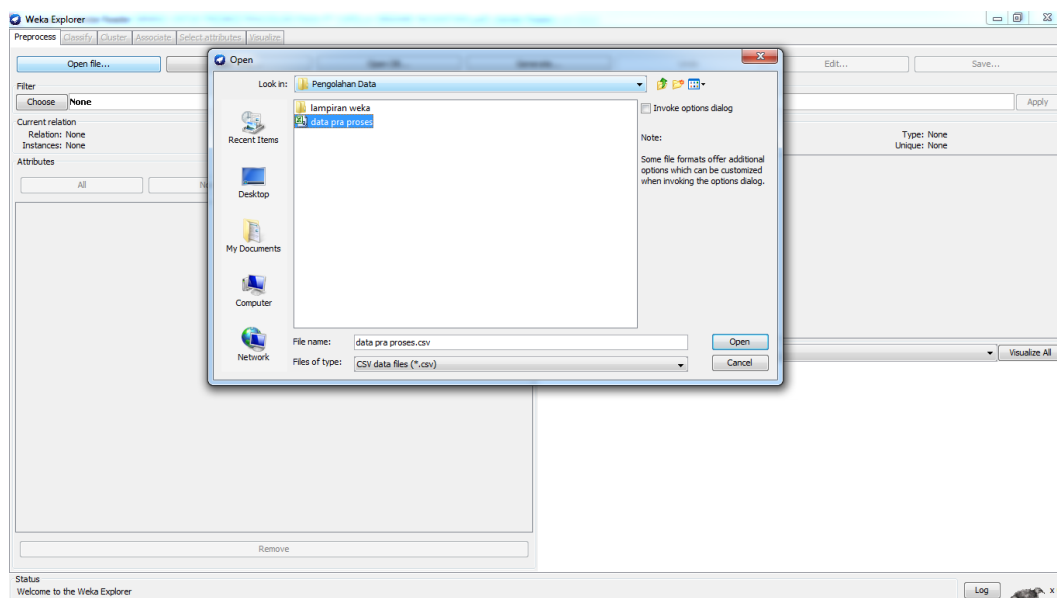
Gambar 2. 4 Halaman Welcome WEKA

3. Lalu klik *Explorer* otomatis masuk ke bagian *Design Preprocess* atau *worksheet* WEKA seperti gambar di bawah ini.



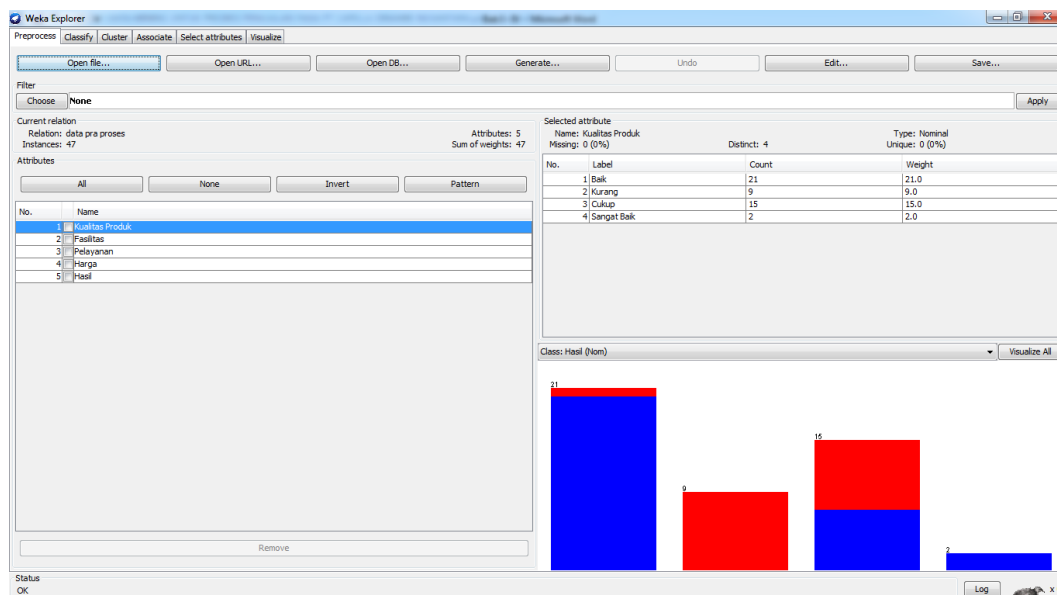
Gambar 2. 5 Design Preprocess WEKA

4. Kemudian *open file* dan pilih data yang akan diproses



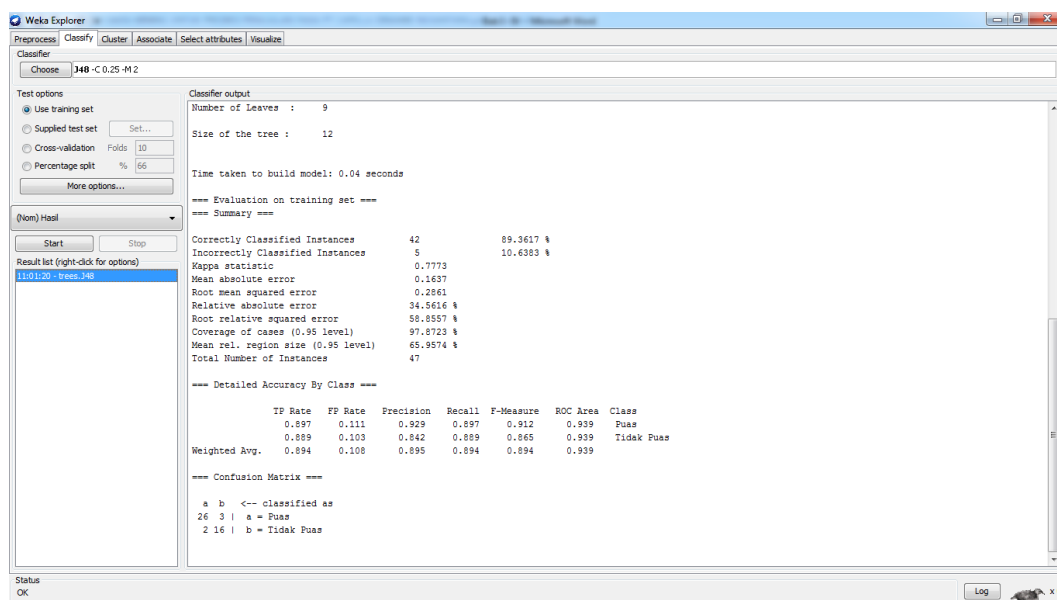
Gambar 2. 6 Open File

5. Setelah data dihubungkan maka secara otomatis akan terbentuk tampilan berikut:



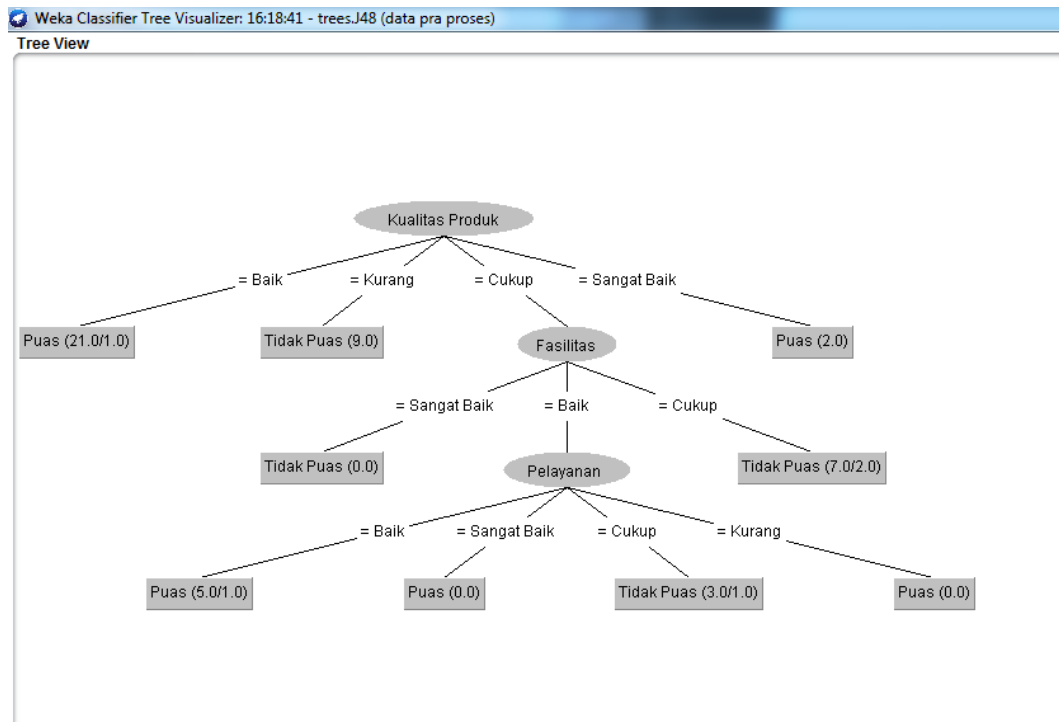
Gambar 2. 7 Tampilan Setelah Terhubung

6. Selanjutnya untuk proses penentuan pohon keputusan dengan uraian berikut ini. Langkah pertama adalah pilih *Classify* dan pilih *Choose* untuk menentukan algoritma (*trees > J48*) lalu *use training set* dan klik *button start*.



Gambar 2. 8 Hasil Keputusan Text

7. Langkah selanjutnya yaitu membuka lembaran hasil dari pohon keputusan. Klik yang ada pada *Result List > trees.J48 > Klik kanan > Visualize Tree*



Gambar 2. 9 Tampilan Pohon Keputusan

2.2. Penelitian Terdahulu

Tabel 2. 1 Penelitian Terdahulu

No	Judul Penelitian	Peneliti dan Tahun	NISN	Hasil Penelitian
1	Implementasi Algoritma C4.5 Untuk Klasifikasi Tingkat Kepuasan Pembeli <i>Online Shop</i>	(Febriyanto et al., 2018)	2407-389X	Melalui penerapan algoritma C4.5 untuk melakukan perhitungan, ternyata mampu mengklasifikasikan tingkat kepuasan pelanggan pada toko <i>online shop</i> tersebut dengan tingkat keakuratan sebesar 91%. Pohon Keputusan yang dihasilkan oleh teknik klasifikasi algoritma C4.5 dari variabel yang memiliki <i>gain</i> yang tertinggi yaitu Pelayanan, sehingga pemilik <i>online</i>

				<i>shop</i> dapat menjadikan acuan untuk memperbaiki kualitas <i>online shop</i> tersebut.
2	Klasifikasi <i>Data Mining</i> Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang	(Faisal, 2019)	2503-054X	Hasil penelitian yang dapat dicapai menggunakan algoritma C4.5 adalah <i>accuracy</i> = 95%. Penelitian ini juga sebagai acuan untuk pihak-pihak yang berkepentingan misalnya kemudahan yang didapatkan oleh pihak perusahaan dalam pengelolaan dan pemanfaatan hasil penelitian.
3	Prediksi Profit Pada Perusahaan Dengan Klasifikasi Algoritma C4.5	(Elisa, 2018)	2406-7857	Algoritma C4.5 dapat digunakan sebagai metode klasifikasi dalam memprediksi Profit Perusahaan CV.XYZ

			<p>dengan memperhatikan nilai <i>gain</i> (penguatan) tertinggi dari tujuh atribut yaitu Biaya Tenaga Ahli, Biaya, Biaya Tenaga Teknik, Biaya Tenaga Pendukung, Biaya Operasional Kantor, Biaya Peralatan, Biaya Pelaporan dan Biaya Mobilisasi.</p> <p>Pohon Keputusan yang dihasilkan oleh teknik klasifikasi Algoritma C4.5 dari variabel yang memiliki <i>gain</i> tertinggi yaitu Biaya Mobilisasi dan Biaya Tenaga Teknik, sehingga perusahaan dapat menjadikan acuan dalam pengaturan dan penggunaan dana anggaran biaya untuk</p>
--	--	--	---

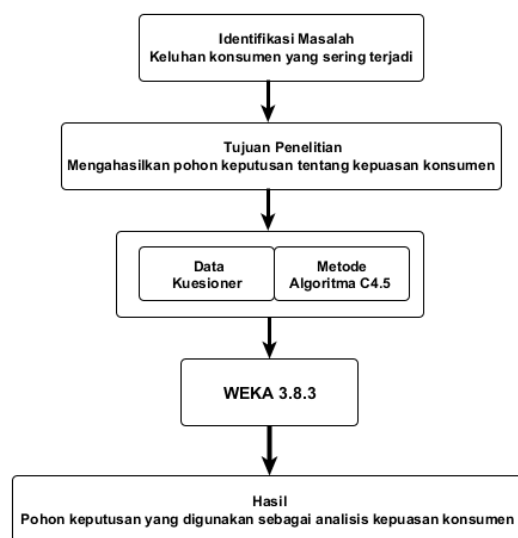
				proyek yang akan datang.
4	Uji Klasifikasi dan Validasi Data Pada Penggunaan Metode Membership <i>Function</i> Dan Algoritma C4.5 Dalam Penilaian Penerima Beasiswa	(Budiman et al., 2018)	2252-2983	Dari hasil uji validasi silang (<i>cross validation</i>) diperoleh nilai error terendah adalah pada metode Integrasi <i>Membership Function</i> dengan Algoritma C4.5, yaitu sebesar 0.2714 untuk <i>Root Mean Square Error</i> (RMSE) dan 0.132 untuk <i>Mean Square Error</i> (MAE), Sementara dari hasil Perbandingan Uji <i>Correctly/incorrectly Classified Instances</i> , diperoleh persentase 92.8571% untuk instances menggunakan metode Integrasi <i>Membership Function</i> dengan Algoritma C4.5, dimana

				persentase tersebut lebih baik dari metode kedua.
5	Penerapan <i>Data Mining</i> Untuk Pengolahan Data Siswa Dengan Menggunakan Metode <i>Decision Tree</i> (Studi Kasus : Yayasan Perguruan Kristen Andreas)	(Sijabat, 2015)	2339-210X	Hasil dari <i>data mining</i> dengan menggunakan sebuah susunan urutan <i>Decision Tree</i> merupakan sebuah susunan urutan kegiatan yang saling mendukung di dalam proses penilaian siswa sehingga lebih mudah dipahami dengan melihat tahapan-tahapan gambar pohon keputusan.
6	<i>Comparison Between WEKA and Salford System in Data Mining Software</i>	(Fitriani et al., 2016)	2393-9141	Hasil yang didapat pada penelitian ini menunjukkan bahwa data konfigurasi pada penggunaan <i>software</i> WEKA lebih mudah dan struktur pohon terlihat lebih jelas dibanding dengan <i>Salford Predictive</i>

				<i>Modeler.</i>
7	Determinan Kualitas Pelayanan, Harga dan Produk Terhadap Kepuasan Pelanggan Telkomsel Flash Di Bondowoso	(Insyroh & Setyowati, 2018)	2549-6409	<p>Hasil yang didapat pada penelitian ini menunjukkan bahwa kualitas pelayanan, produk dan harga berpengaruh positif dan signifikan terhadap kepuasan pelanggan Telkomsel Flash di Bondowoso. Hasil analisis <i>koefisien</i> determinasi menunjukkan bahwa 71,5% kepuasan pelanggan dipengaruhi oleh kualitas pelayanan, produk dan harga. Sedangkan sisanya sebesar 28,5% dipengaruhi oleh variabel lain di luar model.</p>

2.3. Kerangka Pemikiran

Kerangka pemikiran merupakan gambaran singkat yang mendetailkan alur logika dari penelitian yang telah dirumuskan. Kerangka pemikiran pada penelitian ini tertera pada gambar 2.10 berikut ini.



Gambar 2. 10 Kerangka Pemikiran

2.4. Hipotesis Penelitian

Hipotesis adalah jawaban dari rumusan masalah penelitian yang masih harus dibuktikan kebenarannya. Jawaban yang diberikan baru didasarkan pada teori atau anggapan dasar, belum didasarkan pada fakta-fakta yang akurat dari hasil pengumpulan data (Sugiyono, 2014).

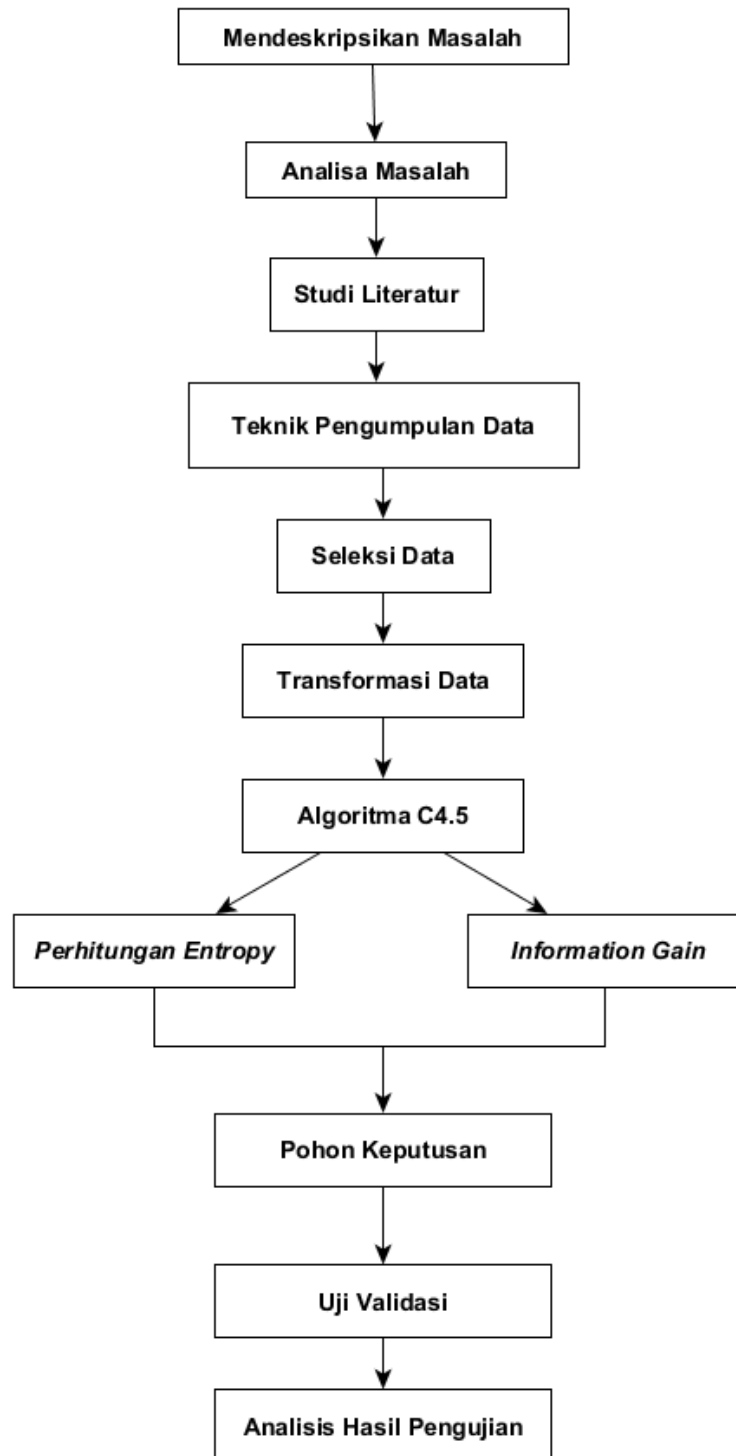
Hipotesis penelitian berlandaskan dari kerangka pemikiran ialah:

1. *Data mining* dengan metode algoritma C4.5 dapat mengetahui tingkat kepuasan konsumen dari kuesioner yang diisi oleh konsumen.
2. *Decision tree* yang terbentuk dari algoritma C4.5 dapat digunakan untuk mengukur tingkat kepuasan konsumen.

BAB III METODE PENELITIAN

3.1. Desain Penelitian

Dalam melaksanakan suatu penelitian, salah satunya tingkatan yang wajib dilakukan ialah menggambarkan desain penelitian. Dengan desain penelitian bisa dijadikan sebagai acuan untuk menyelesaikan tujuan dari penelitian. Adapun desain penelitian ini dapat dilihat pada gambar 3.1.



Gambar 3. 1 Desain Penelitian

Berlandaskan dari desain penelitian yang tertera dalam gambar 3.1, setiap atributnya didetailkan seperti berikut ini:

1. Mendeskripsikan Masalah

Menjelaskan masalah di dalam penelitian dengan memastikan serta menginterpretasikan batasan dari masalah yang akan dicermati, hingga mendukung dalam menghasilkan sebuah penyelesaian dari masalah itu. Jadi, sebagai tahapan awal pada penelitian ini.

2. Analisa Masalah

Tahapan analisa ialah tahapan agar bisa menginterpretasikan masalah dengan pedoman dari batasan masalah. Dengan melakukan penganalisaan dari masalah yang sudah ditetapkan, maka dinantikan masalah tersebut mampu ditafsirkan.

3. *Studi Literature*

Untuk mencapai tujuan, maka dipelajari beberapa *literature-literature* yang diperkirakan dapat digunakan. Kemudian *literature-literature* yang dipelajari tersebut diseleksi untuk dapat ditentukan *literature-literature* mana yang akan digunakan dalam penelitian.

4. Metode Pengumpulan Data

Penelitian ini menggunakan beberapa teknik pengumpulan data guna mendukung pengambilan data pada objek penelitian. Teknik yang digunakan diantaranya:

a) Wawancara

Wawancara merupakan teknik pengumpulan data yang dilakukan melalui tatap muka dan tanya jawab langsung antara peneliti dan narasumber. Pada metode ini peneliti langsung menanyakan beberapa hal tentang kebutuhan informasi berkaitan dengan perusahaan terutama pada keluhan yang sering disampaikan oleh konsumen.

b) Kuesioner *Online*

Kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan yang telah disediakan pilihan jawaban untuk dipilih objek penelitian. Kuesioner *online* merupakan kuesioner yang disebarakan secara *digital* yaitu menggunakan salah satu aplikasi yang disediakan oleh Google yaitu *Google Form*.

5. Seleksi Data

Seleksi data dengan memilih data yang akan digunakan dalam proses algoritma klasifikasi C4.5. Tujuan dari seleksi data yaitu untuk menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada *subset* variabel atau sampel data, dimana penemuan (*discovery*) akan dilakukan.

6. Transformasi Data

Transformasi data merupakan tahap untuk menyusun dan mengembangkan *set data* menjadi lebih baik sehingga tahap ini membutuhkan proses kreatif dan sangat bergantung pada jenis atau pola informasi yang akan dicari dalam *basis data*, seperti mengkategorikan data ke dalam beberapa kategori.

7. Algoritma C4.5

Pada tahap ini akan dilakukan proses perancangan dari model sistem dengan algoritma C4.5 sehingga membentuk pohon keputusan (*Decision Tree*) dan menghasilkan suatu aturan (*rule*) pola kepuasan konsumen. Dengan menggunakan perhitungan semua atribut/varibel atau yang dikenal dengan Perhitungan *Entropy* dan *Information Gain*.

a. Perhitungan *Entropy*

Perhitungan dengan menggunakan rumus 2.1.

b. *Information Gain*

Dengan menggunakan rumus 2.2 untuk mengetahui *information gain* tertinggi yang akan dijadikan simpul akar pada pembuatan pohon keputusan.

8. Pohon Keputusan

Analisis yang dikerjakan supaya dapat menegaskan tentang hasil pengujian sesuai dengan pembahasan. Pada proses ini, *software Data mining open source WEKA* yang penulis gunakan untuk menjalankan pengujian serta hasil perancangan dari penelitian. Pengujian pola menggunakan langkah-langkah untuk mendapatkan pencarian dan permodelan melalui *data set* yang ada hingga menemukan sebuah pola yang tersembunyi dari *data set* tersebut.

9. Uji Validasi

Pada tahapan ini melakukan pengujian untuk memvalidasi atau menilai akurasi dari *data set* pada penelitian ini. Sehingga menghasilkan tingkat akurasi pada penelitian *data mining* menggunakan metode algoritma C4.5. Untuk mengetahui nilai akurasi pada data menggunakan rumus 2.3.

3.2. Operasional Variabel

Operasional variabel ialah melakukan perumusan variabel melalui pengamatan berlandaskan karakteristik-karakteristiknya. Variabel penelitian disebut juga sebagai objek penelitian atau sebagai aspek-aspek yang berpengaruh pada fakta atau perihal yang nantinya akan diteliti. Sedikit mengambil kesimpulan dari buku (Sugiyono, 2014) variabel penelitian ialah suatu sifat atau nilai atau atribut dari orang, obyek, juga kegiatan yang memiliki ragam tertentu yang dapat ditetapkan oleh si peneliti supaya mampu dipelajari dan diambil kesimpulannya. Variabel yang akan menjadi keputusan Kepuasan Konsumen dengan hasil keputusan PUAS atau TIDAKPUAS sebagai berikut:

Tabel 3. 1 Operasional Variabel

Variabel	Definisi Variabel	Dimensi	Indikator
Kualitas Produk	Kualitas produk berarti kualitas yang merujuk pada kecepatan sinyal akses data yang digunakan konsumen	Kecepatan sinyal yang digunakan	1. Sangat Kurang 2. Kurang 3. Cukup 4. Baik 5. Sangat Baik
Fasilitas	Fasilitas merupakan fasilitas yang diberikan kepada konsumen berupa kemudahan	Fasilitas <i>top up</i> / isi ulang paket data yang diberikan	1. Sangat Kurang 2. Kurang 3. Cukup 4. Baik

	untuk interaksi dengan karyawan, kemudahan untuk melakukan <i>top up</i> (pengisian) ulang kuota paket data internet.		5. Sangat Baik
Pelayanan	Pelayanan, yaitu kemudahan menghubungi layanan konsumen, respons karyawan saat menerima keluhan konsumen dan respons karyawan saat memperbaiki gangguan pada jaringan yang digunakan.	1. Kemudahan menghubungi layanan konsumen	1. Sangat Kurang 2. Kurang 3. Cukup 4. Baik 5. Sangat Baik
		2. Respons karyawan saat menghubungi layanan konsumen	1. Sangat Kurang 2. Kurang 3. Cukup 4. Baik 5. Sangat Baik
		3. Respons karyawan saat memperbaiki	1. Sangat Kurang 2. Kurang 3. Cukup

		jaringan	4. Baik 5. Sangat Baik
Harga	Harga berarti sejumlah uang yang dibayar oleh konsumen atas layanan yang digunakan.	Hinet Batam harganya relatif terjangkau	1. Sangat Kurang 2. Kurang 3. Cukup 4. Baik 5. Sangat Baik

3.3. Populasi dan Sampel

3.3.1. Populasi

Populasi merupakan keseluruhan objek yang akan diteliti dan wilayah generalisasi yang terdiri dari subyek yang menjadi kuantitas dan karakteristik tertentu yang ditetapkan dan kemudian ditarik kesimpulannya (Sugiyono, 2014). Populasi yang digunakan pada penelitian ini adalah data penjualan Hinet Batam dalam periode penjualan 29 Januari 2019 hingga 5 Agustus 2019 yang berjumlah 131 Konsumen.

3.3.2. Sampel

Sampel merupakan sebagian dari jumlah dan karakteristik yang dimiliki oleh populasi tersebut. Adapun metode yang digunakan dalam pengambilan

sampel pada penelitian ini yaitu metode *Sampling Sistematis*. Metode *Sampling Sistematis* merupakan teknik pengambilan sampel berdasarkan urutan dari anggota populasi yang telah diberi nomor urut. Pengambilan sampel dapat dilakukan dengan nomor ganjil saja, genap saja, atau kelipatan dari bilangan tertentu (Sugiyono, 2014).

Sampel yang digunakan dalam penelitian ini yaitu nomor urutan genap yang ada dalam data populasi dan hanya menggunakan data konsumen dengan nomor telepon yang aktif. Dari 131 konsumen yang ada dalam data populasi didapatkan sebanyak 65 konsumen yang berada dalam urutan nomor genap. Dan dari 65 konsumen yang terpilih dalam urutan genap hanya 47 konsumen dengan nomor telepon aktif dan dijadikan sebagai sampel pada penelitian ini seperti pada tabel 3.2.

Tabel 3. 2 Sampel / Data Set Penelitian

NO	TANGGAL	NAMA	TELP
1	31-01-2019	Kholil	Aktif
2	31-01-2019	Andi	Aktif
3	10-02-2019	HendriPurba	Aktif
4	10-02-2019	Rohima	Aktif
5	13-02-2019	Togi Rudinato	Aktif
6	15-02-2019	Mesda Ledia Sariana	Aktif
7	16-02-2019	Abdul Qodir	Aktif
8	18-02-2019	Anggiat	Aktif
9	18-02-2019	Asih Kurniaty	Aktif
10	25-02-2019	Cahyono Budi Wibowo	Aktif
11	26-02-2019	Nur Aini	Aktif
12	01-03-2019	Siti Maimonah	Aktif
13	05-03-2019	Nurhadi	Aktif
14	16-03-2019	Angan Dewa Harefa	Aktif
15	21-03-2019	Maizura	Aktif

16	26-03-2019	Berliana Natalia Turnip	Aktif
17	27-03-2019	Clerisa Sirait	Aktif
18	30-03-2019	Karim	Aktif
19	31-03-2019	Isnalindra	Aktif
20	05-04-2019	Qindiyani	Aktif
21	06-04-2019	Pungun Martina Marpaung	Aktif
22	07-04-2019	Ibu Tri	Aktif
23	07-04-2019	Isriandi	Aktif
24	09-04-2019	Yusuf Rambe	Aktif
25	11-04-2019	Firmansyah	Aktif
26	12-04-2019	Ramses	Aktif
27	13-04-2019	Ida Ayu Sihombing	Aktif
28	14-04-2019	Jonpaif Pandiangan	Aktif
29	16-04-2019	Lilis	Aktif
30	20-04-2019	Inayah	Aktif
31	25-04-2019	Suhaimi	Aktif
32	26-04-2019	Yesma	Aktif
33	06-05-2019	Liana	Aktif
34	15-06-2019	Wiwit	Aktif
35	22-06-2019	Ponco	Aktif
36	25-06-2019	Slamet Subiyanto	Aktif
37	04-07-2019	Fitri Yeni	Aktif
38	06-07-2019	Yuliah	Aktif
39	13-07-2019	Erikson Dalimunte	Aktif
40	16-07-2019	Suparmo	Aktif
41	17-07-2019	Tengku Hariani Riska Putri	Aktif
42	21-07-2019	Sujono	Aktif
43	24-07-2019	Siti Aisyah	Aktif
44	31-07-2019	Winarto Hadi Purnomo	Aktif
45	03-08-2019	Wahono	Aktif
46	03-08-2019	Parriman	Aktif
47	05-08-2019	Yulizar	Aktif

3.4. Teknik Pengumpulan Data

Kualitas pengumpulan data menjadi aspek yang berpengaruh terhadap kualitas data dari hasil penelitian. Melakukan pengumpulan data bisa dikerjakan menggunakan berbagai sumber, *setting* dan cara. Dalam penelitian ini aspek atau metode yang digunakan untuk mengumpulkan data yaitu:

1. Wawancara, penulis menanyakan langsung perihal masalah yang sering dikeluhkan oleh konsumen kepada perusahaan.
2. Kuesioner *online*, penulis menyebarkan kuesioner *online* dengan menggunakan bantuan dari aplikasi yang ada di dalam *Google* yaitu *Google Form*. Kuesioner *online* disebarkan kepada konsumen yang menjadi sampel dari penelitian. Dengan memberikan seperangkat pertanyaan yang telah disediakan pilihan jawaban untuk dipilih objek penelitian.
3. Studi Pustaka, ialah metode untuk mendapatkan acuan dari beberapa referensi dan literatur sebagai penunjang penelitian seperti jurnal atau buku yang berhubungan dengan obyek permasalahan.

3.5. Metode Analisa Data

Metode analisa data yang digunakan dalam penelitian ini ialah dengan menggunakan langkah-langkah dari *Knowledge Discovery in Database (KDD)*. Langkah-langkah yang digunakan sebagai berikut:

1. *Data cleaning* (Pembersihan data)

2. *Data integration* (Integrasi data)
3. *Data selection* (Seleksi data)
4. *Data transformation* (Transformasi data)
5. Tahapan *mining* menggunakan algoritma C4.5 dan klasifikasi pohon keputusan (*decision tree*).
6. *Pattern evaluation* (Evaluasi pola)

3.6. Lokasi dan Jadwal Penelitian

3.6.1. Lokasi Penelitian

Lokasi penelitian yaitu Hinet Batam yang merupakan perusahaan penyedia jasa layanan akses data internet. Hinet Batam beralamat di Ruko Anggara Niaga No. 2 Jl. Kavling Lama RT. 001 RW. 012. Kec Sagulung Kota Batam 29439.



Gambar 3. 2 Kantor Hinet Batam

3.6.2. Jadwal Penelitian

Tabel 3. 3 Jadwal Penelitian

No	Kegiatan	2019-2020											
		November				Desember				Januari			
		1	2	3	4	1	2	3	4	1	2	3	4
1	Pengurusan Perizinin Perusahaan untuk Penelitian	■											
2	Proposal Penelitian		■	■									
3	Pengumpulan Data (Kuesioner)				■	■	■	■					
4	Rekap Data								■				
5	Pengolahan Data (Algoritma C4.5 dan <i>Decision Tree</i>)									■	■	■	
6	Pengujian <i>software</i>											■	
7	Kesimpulan												■