

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, ilmuwan, 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*)

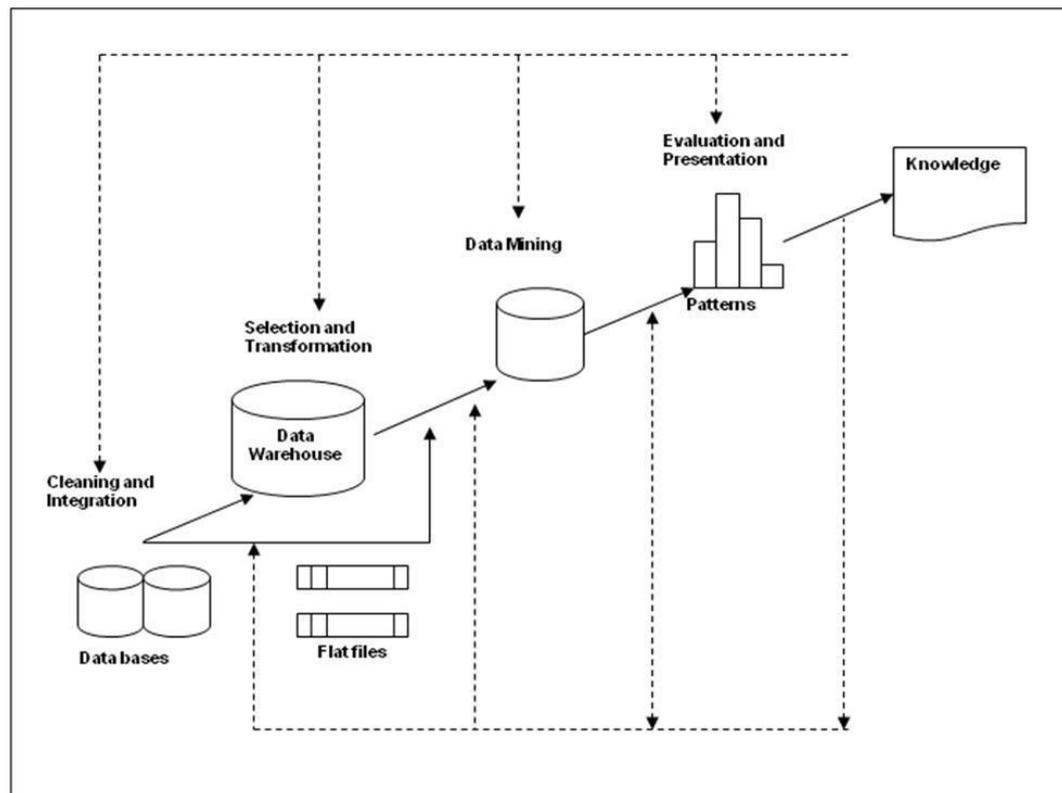
Merupakan tahapan penemuan pengetahuan dari data yang tersembunyi.

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

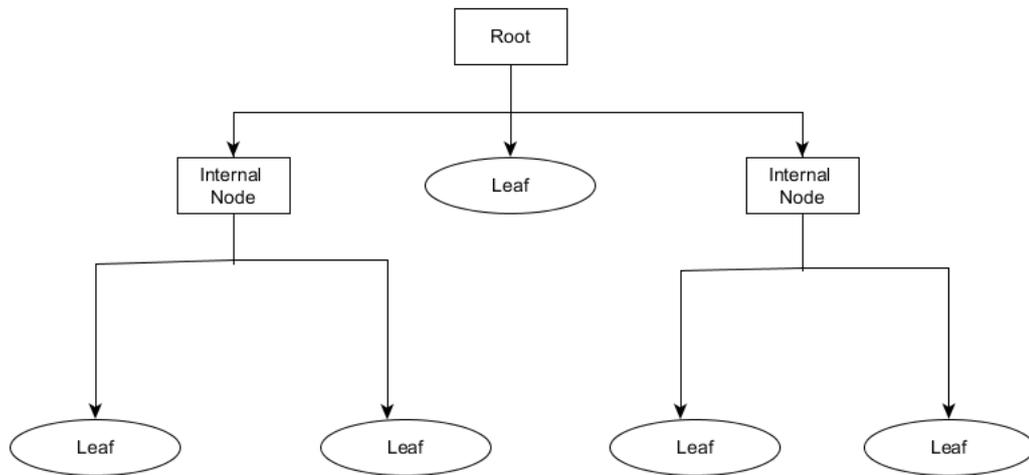
2.1.2.6 Pohon Keputusan (*Decision Tree*)

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.

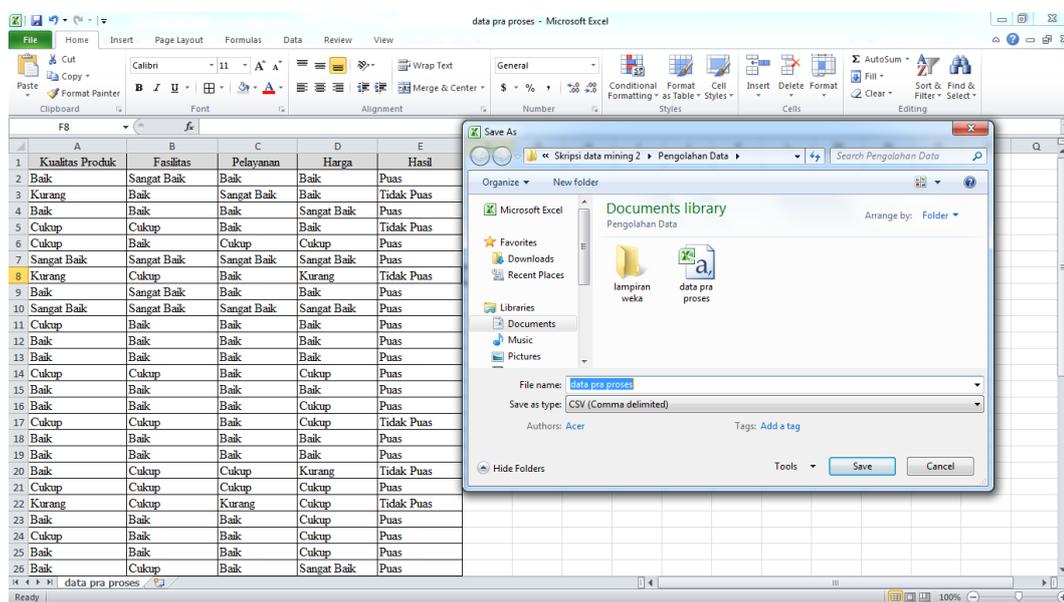
Waikato Environment for Knowledge Analysis (WEKA) ialah suatu perangkat lunak yang mengimplementasikan algoritma *machine learning* untuk dilakukannya berbagai proses yang bersangkutan dengan sistem temu pada informasi atau *data mining*. *Software WEKA* menyediakan teknik klasifikasi

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*.



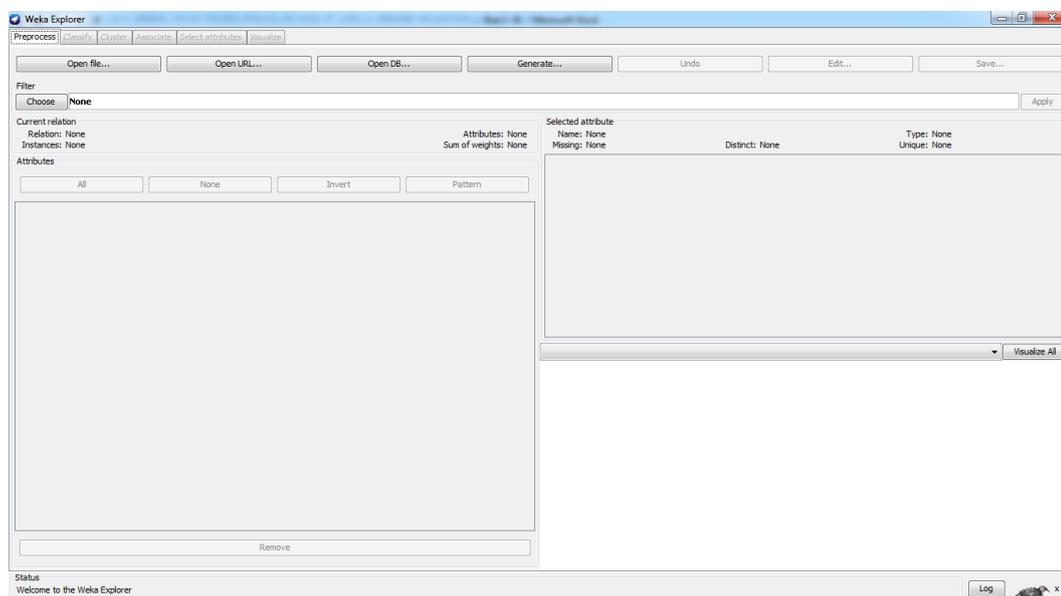
Gambar 2.3 Data Pra Proses

2. Menghubungkan Data Dengan Aplikasi WEKA



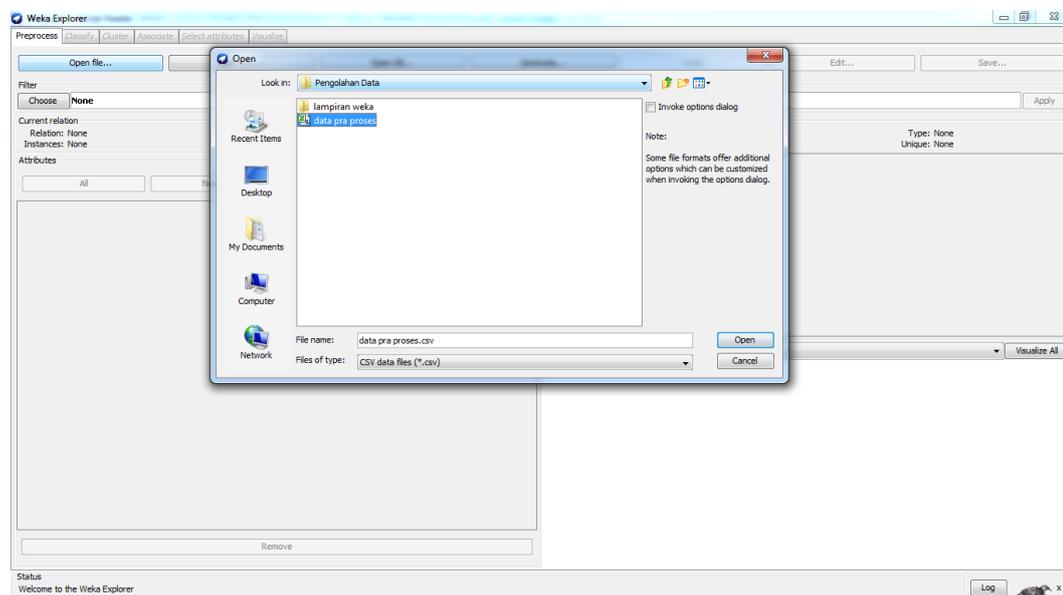
Gambar 2. 4 Halaman Welcome WEKA

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



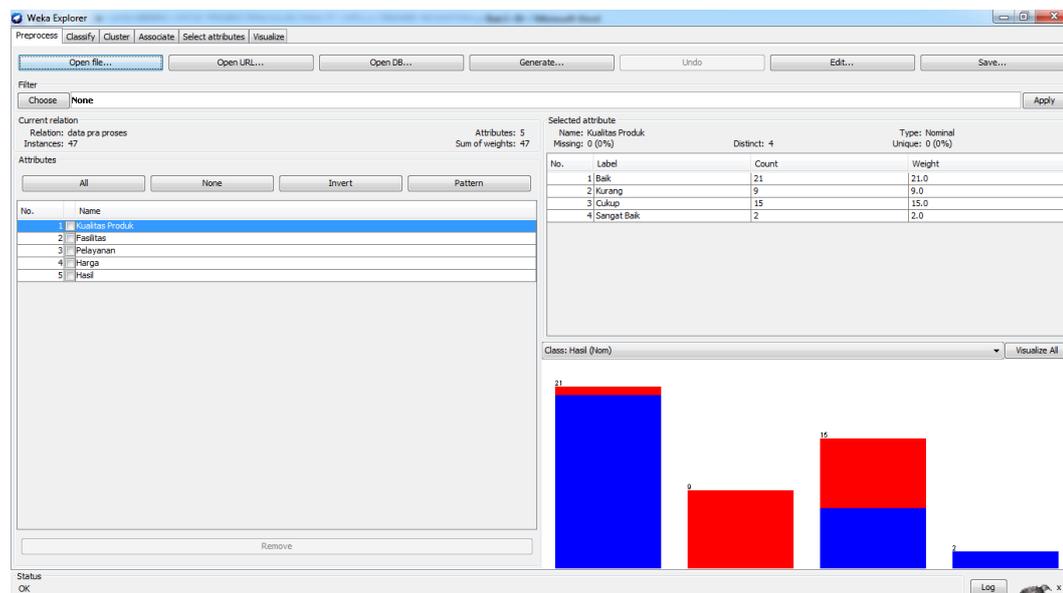
Gambar 2. 5 Design Preprocess WEKA

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



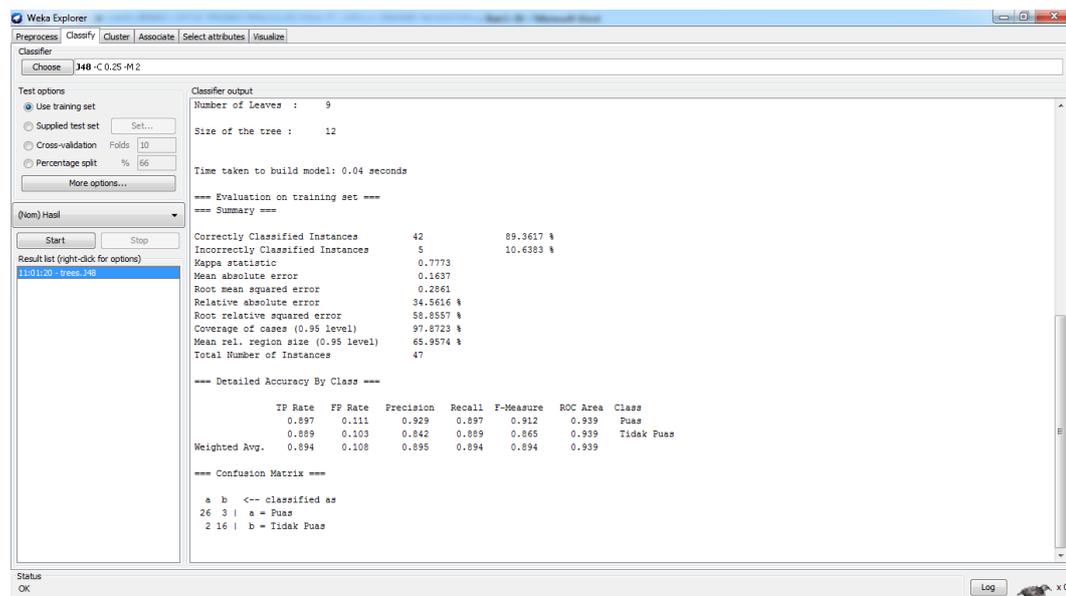
Gambar 2. 6 Open File

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



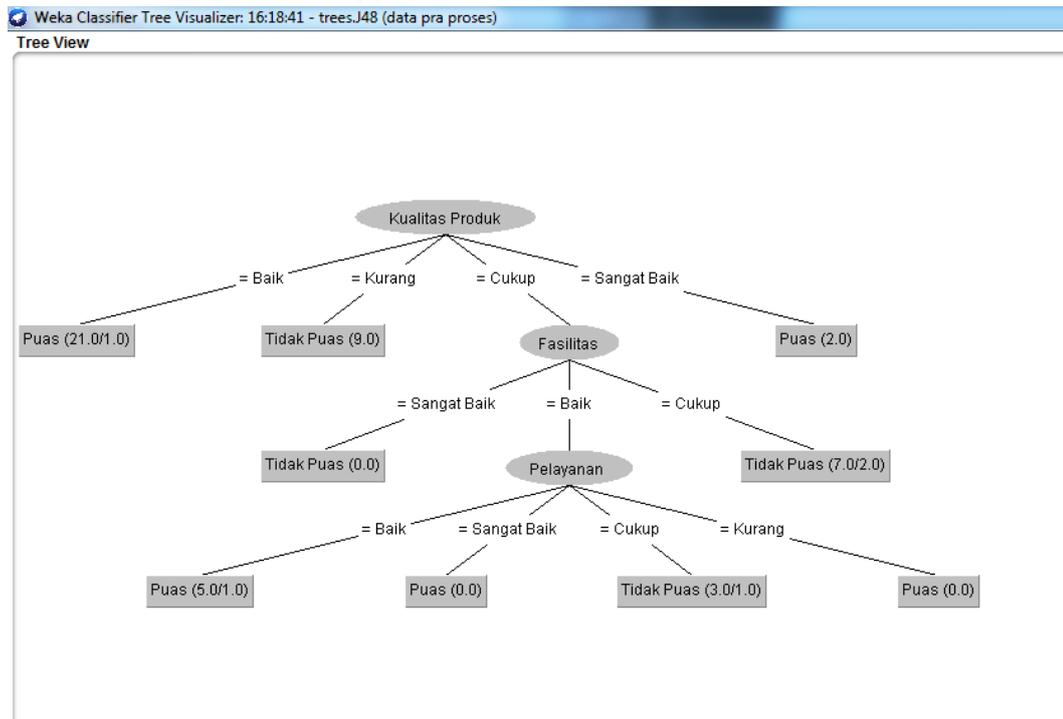
Gambar 2. 7 Tampilan Setelah Terhubung

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>
--	--	--	---

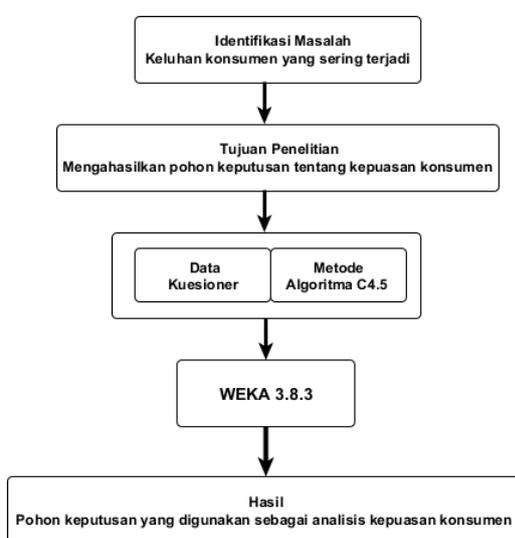
				proyek yang akan datang.
4	Uji Klasifikasi dan Validasi Data Pada Penggunaan Metode Membership <i>Function</i> Dan Algoritma C4.5 Dalam Penilaian Penerima Beasiswa	(Budiman et al., 2018)	2252-2983	Dari hasil uji validasi silang (<i>cross validation</i>) diperoleh nilai error terendah adalah pada metode 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.