

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Teori**

Pemanfaatan teori dalam proses penelitian sangatlah penting agar landasan teori dan pedoman pelaksanaan penelitian dapat dikembangkan.

##### **2.1.1. Teori Umum**

Penting untuk dicatat bahwa poin-poin penting berikut ini merupakan uraian singkat tentang teori umum yang dibahas dalam penelitian ini.

##### **2.1.1.1. *Knowledge Discovery in Database (KDD)***

Dalam proses pengembangan suatu aplikasi data mining, terdapat beberapa tahapan yang digunakan untuk analisis data yang akan digunakan untuk aplikasi data mining tersebut, dan tahapan tersebut disebut dengan *Knowledge discovery in database*. Dalam Ilmu Basis Data, *Knowledge Discovery in Databases (KDD)* adalah Dengan mengambil, memproses, dan menganalisis data historis, suatu pola, tren, atau hubungan dapat ditemukan dalam data dalam jumlah besar dengan mengidentifikasi pola, tren, dan hubungan dalam data tersebut. Hasil penambangan data bisa dipakai untuk membantu manajer mengambil keputusan yang lebih baik dengan meningkatkan cara mereka menggunakannya untuk mengambil keputusan di masa depan (Putria, 2018).

Ada banyak teknik integrasi dan penemuan ilmiah yang terlibat dalam KDD. Dalam proses menafsirkan dan memvisualisasikan pola-pola yang ditemukan dalam kumpulan data KDD, kami mencari dan mengidentifikasi pola-pola dalam data yang valid, berguna dan dapat dipahami, dan kemudian

menggunakan pola-pola yang ada untuk mengidentifikasi dan memahami pola-pola baru. Ada beberapa tahapan dalam rangkaian proses ini, yang meliputi pembersihan data dan pengintegrasian data tersebut (cleaning and integrating). Data yang diperoleh dari berbagai database, masing-masing memiliki format atau platform berbeda, kemudian digabungkan dalam gudang data menggunakan proses pembersihan data untuk menghilangkan data yang tidak konsisten dan berisik yang sebelumnya diperoleh dari database berbeda (Zulfa *et al.*, 2020).

Ada kemungkinan untuk mengurangi jumlah data yang terdapat dalam database setelah diubah dan dipilih dengan menggunakan berbagai teknik setelah dipilih. Untuk memperoleh hasil yang berbeda, perubahan data merupakan salah satu tahapan dalam proses analisis data yang berlangsung pada tahap pertama proses untuk mengubah data setelah diolah menjadi data hasil. Berikut adalah beberapa cara di mana data dapat dimodifikasi.

1. *Clustering*,

Memanfaatkan data atribut standar yang tersedia dapat mengurangi kebutuhan penggunaan data atribut standar dan mengurangi penggunaan data atribut standar.

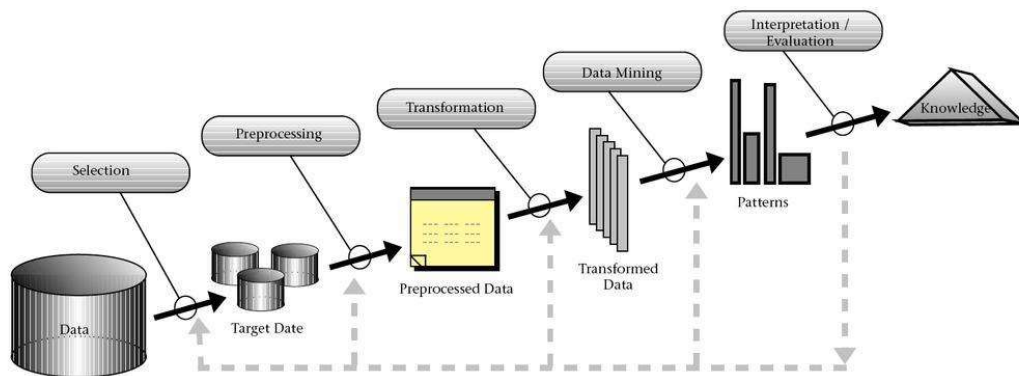
2. *Normalisasi*,

Untuk menghitung ukuran deviasi standar berdasarkan kumpulan data atribut, kumpulan data atribut dapat dianalisis.

3. *Scaling*,

Seiring bertambahnya data, ukuran penskalaan yang berbeda diterapkan hingga ukuran data tertentu tercapai, yang kemudian dianalisis lebih lanjut.

Kusrini & Lutfi dalam (Ulfa, 2018) perlu diketahui bahwa ada istilah lain yang memiliki arti sama dengan data mining, yaitu penemuan pengetahuan dalam database (KDD). Terlepas dari kenyataan bahwa istilah penambahan data dan penemuan dan pengembangan pengetahuan (KDD) umumnya digunakan secara bergantian, konsep-konsep tersebut sebenarnya mengacu pada proses menghilangkan informasi dari database besar yang telah disembunyikan. KDD hadir dalam beberapa tahapan, seperti diuraikan di bawah ini, yang dapat dilihat secara keseluruhan.



**Gambar 2.1** Tahapan dalam KDD

Keterangan dari gambar di atas dirincikan sebagai berikut:

1. Seleksi Data (*Selection*)

Untuk memulai proses penemuan informasi dalam Knowledge Discovery Database (KDD), perlu dilakukan pengumpulan data dari

kumpulan data operasional untuk memulai tahap eksplorasi informasi. Seleksi (seleksi) harus terjadi sebelum tahap eksplorasi informasi dimulai. Untuk alasan keamanan, penting untuk memisahkan database operasional dari file yang berisi data terpilih untuk ditambang sebagai bagian dari proses penambangan data.

2. Pemilihan Data (*Preprocessing/Cleaning*)

Ada sejumlah tugas yang dilakukan sebagai bagian dari proses prapemrosesan, termasuk menghapus data duplikat, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan dan kesalahan ketik pada data. Untuk melengkapi data yang diterima KDD maka dilakukan proses pengayaan yaitu dengan menambahkan data atau informasi tambahan pada data tersebut, seperti data atau informasi eksternal, sehingga data tersebut “diperkaya”.

3. Transformasi (*Transformation*)

Data mining adalah serangkaian proses yang saling berhubungan yang melibatkan transformasi bentuk data yang memiliki data null atau tidak beroperasi menjadi bentuk data yang memiliki lubang yang dapat dengan mudah dimanfaatkan oleh proses penambangan data yang siap untuk dilakukan proses penambangan data. berjalan terus.

4. Data Mining

Selama fase ini, algoritma dan metode untuk mengambil pengetahuan diterapkan sehingga data dapat diambil.

5. Interpretasi atau Evaluasi (*Interpretation or Evaluation*)

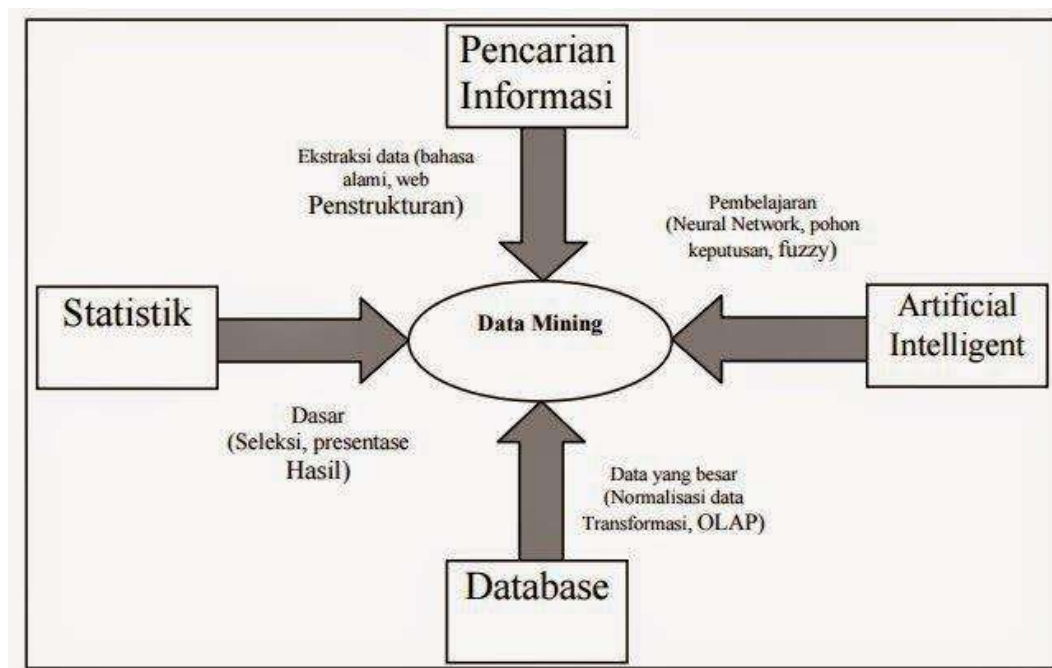
Langkah terakhir dari proses ini adalah membuat keluaran yang mudah dipahami, berdasarkan pola informasi yang diidentifikasi oleh proses Data Mining, namun dikumpulkan sebagai hasil dari proses Data Mining.

#### **2.1.1.2. Data Mining**

Penting untuk diklarifikasi bahwa ekstraksi informasi mengacu pada proses memperoleh informasi dari kumpulan data dengan menggunakan teknik algoritmik dan teknis yang dapat diterapkan pada berbagai bidang, seperti statistik, ilmu komputer, dan database umum. Ini adalah metode penambangan data yang efektif dalam mengekstraksi informasi dari kumpulan data besar yang mungkin berisi informasi penting yang tersembunyi. Ada istilah "penambangan data" yang menggambarkan proses menemukan berlian pengetahuan berharga di dasar kumpulan data yang besar setelah dikumpulkan dari waktu ke waktu (Lestari & Hafiz, 2020).

Menganalisis kumpulan data besar dalam penyimpanan menggunakan teknik pengenalan pola seperti matematika dan statistik, penambangan data bertujuan untuk menemukan hubungan, pola, dan tren yang bermakna (Ependi & Akbar, 2019). Data mining berfokus pada penemuan pola dan aturan dari kumpulan informasi melalui analisis sekumpulan variabel untuk memberikan kita pengetahuan tentang informasi tersebut. Untuk mendapatkan pola yang dapat digunakan dalam transaksi masa depan dan database relasional, digunakan data transaksi (Haidar, 2021).

Seperti namanya, ini adalah proses pengungkapan pengetahuan dari database yang disebut dengan data mining (Andriani *et al.*, 2019). Untuk mengubah data tersebut menjadi informasi yang berguna, penambahan data telah menjadi alat yang semakin penting. Ada banyak penerapan berbeda dari jenis data yang dikumpulkan, di antaranya pemasaran, pengawasan, penyelidikan ilmiah, dan penipuan penemuan (Wardah & Fitriyah, 2021). Halaman ini memberikan gambaran singkat tentang data mining, sebagaimana dimaksud dalam definisi berikut: Data mining digunakan untuk menganalisis data dan menemukan pola informasi yang sebelumnya tidak diketahui, seperti terlihat pada gambar 2.2 berikut.



**Gambar 2.2** Bidang Ilmu Data Mining

Data mining mengalami kemajuan yang luar biasa dalam beberapa tahun terakhir karena beberapa alasan, beberapa di antaranya dijelaskan dalam (Setyawati *et al.*, 2021).

1. Kumpulan data yang semakin besar dihasilkan setiap hari.
2. Basis data yang paling andal harus dipelihara dalam gudang data sehingga dapat diakses kapan saja oleh semua perusahaan, di mana pun lokasinya.
3. Seiring dengan semakin populernya akses data melalui internet dan navigasi web, penggunaan internet diperkirakan akan meningkat.
4. Persaingan yang diarahkan untuk meningkatkan kekuatan pasar menjadi semakin penting bagi dunia usaha sebagai akibat dari globalisasi ekonomi.
5. Data mining dapat dicapai dengan menggunakan teknologi perangkat lunak yang mampu mendukung proses tersebut (ketersediaan teknologi)
6. Selain peningkatan kapasitas penyimpanan media, kemampuan komputasi juga ditingkatkan secara signifikan.

### **2.1.1.3. Pengelompokan Data Mining**

Tugas data mining dapat dikategorisasikan menurut fungsinya berdasarkan kelompok berikut, yang dapat digunakan sebagai panduan: (Mulya *et al.*, 2019):

1. Klastering

Sebagai aturan praktis, objek-objek harus mirip satu sama lain dalam sebuah cluster, namun objek-objek tersebut tidak perlu berbagi kesamaan antar cluster, sebaiknya sebuah cluster hanya berisi objek-objek serupa di dalamnya, dan sebuah cluster tidak boleh berisi objek-objek yang serupa. melintasinya.

## 2. Klasifikasi

Karakteristik suatu objek dapat digunakan untuk menetapkan keanggotaannya pada kelompok yang ada. Diasumsikan bahwa data pelatihan akan diklasifikasikan ke dalam kelompok atau kelas berdasarkan label yang terkait dengannya, bukan dikelompokkan ke dalam kelompok.

## 3. Regresi atau Estimasi

Terlepas dari kesamaan antara estimasi dan regresi, ada satu perbedaan mencolok di antara keduanya: estimasi dan regresi keduanya memerlukan data pelatihan yang diberi label. Namun, keluaran klasifikasi bersifat diskrit dan bukan kontinu dalam regresi. Di sini disajikan model regresi yang menyatakan hubungan antara variabel-variabel yang mengidentifikasi prediktor serta variabel-variabel yang mengidentifikasi dependen, dan variabel-variabel tersebut juga merupakan variabel kontinu.

## 4. Asosiasi

Asosiasi digunakan untuk mengidentifikasi atribut-atribut yang terjadi bersamaan, dengan tujuan mengukur hubungan antara dua atau lebih atribut dengan menerapkan aturan-aturan yang dapat digunakan untuk mengidentifikasi atribut-atribut yang terkait.

### **2.1.1.4. Tahapan Data Mining**

Tahapan data mining pada (Syahdan & Sindar, 2018) :

#### *1. Data Cleaning*



Memastikan bahwa data yang tidak konsisten dihilangkan.

2. *Data Integration,*

Dimungkinkan untuk menggabungkan sumber data terpisah.

3. *Data selection*

Data analitik dikembalikan ke database berdasarkan tugas analisis.

4. *Data transformation*

Data berubah atau bersatu menjadi bentuk yang tepat untuk menambang dengan ringkasan performa atau operasi agresif.

5. *Data mining*

Proses esensial digunakan untuk mengekstrak pola data.

6. *Pattern evolution*

Untuk mengidentifikasi pola tindakan yang mewakili pengetahuan, beberapa pola yang sangat menarik perlu diidentifikasi.

7. *Knowledge presentation*

Agar pengguna dapat memahami pengetahuan yang ditambang, teknik visualisasi dan pengetahuan digunakan.

#### **2.1.1.5. Manfaat Data Mining**

Penambangan data digunakan dalam berbagai cara, namun berikut adalah beberapa cara yang paling umum.

1. Sesuai dengan uraian dalam definisi model, model dapat memilih sekumpulan item yang sesuai.
2. Terdapat perbedaan antara clustering dan peramalan, perbedaannya hanya pada perhitungan peramalan yang lebih cenderung memasukkan

lebih banyak variabel yang dapat bersifat prediktif dibandingkan clustering.

3. Prinsip terarah untuk menebak kelompok item yang tidak memiliki kelompok yang dapat dibentuk secara independen satu sama lain disebut pemisahan, yang memungkinkan kita menemukan pola atau mendeskripsikan dan mengidentifikasi kumpulan item yang berbeda, atau untuk mengidentifikasi pola relasional yang dapat digunakan untuk memperoleh pola.
4. Istilah cluster digunakan untuk merujuk pada sekelompok record yang mempunyai beberapa persamaan dan beberapa perbedaan dengan berbagai macam cluster record lainnya. Definisi istilah cluster mengacu pada sekelompok record yang mirip satu sama lain namun juga berbeda satu sama lain.
5. Seringkali kata "analisis" digunakan oleh para pebisnis untuk merujuk pada penggunaan simbol asosiatif dalam hal penambangan data (Zulfa *et al.*, 2020).

#### **2.1.1.6. Fungsi Data Mining**

Menurut Sulistiyo dalam (Marsono, 2019) bagian penting dari penambangan data adalah hal ini didasarkan pada analisis data menggunakan alat untuk mengidentifikasi pola dan aturan dalam kelompok data seperti yang ditentukan oleh namanya. Ada kemungkinan besar bahwa alat penambangan data akan mampu mendeteksi pola-pola ini dalam data tanpa banyak keterlibatan pengguna

karena alat tersebut akan mengidentifikasi aturan dan fitur dalam data, dan kemudian menampilkan pola-pola tersebut kepada pengguna.

#### **2.1.1.7. *Decision Tree***

Di bidang klasifikasi dan prediksi, pohon keputusan adalah salah satu metode yang paling efektif dan banyak digunakan untuk menarik pelanggan. Ini adalah metode pengorganisasian data dalam jumlah besar ke dalam pohon keputusan yang mewakili aturan dengan mengubah fakta yang sangat besar menjadi pohon keputusan adalah salah satu hal yang paling menantang untuk dilakukan (Sikumbang *et al.*, 2022). Sebagai alternatif dari metode klasifikasi yang dikenal dengan *Decision Trees*, yang mengandalkan representasi struktur pohon, *Decision Trees* adalah metode klasifikasi yang mengandalkan node yang mewakili atribut, cabang yang mewakili nilai atribut, dan daun yang mewakili kelas berdasarkan atribut tersebut (Fauziah *et al.*, 2020).

Pola dapat diimplementasikan menggunakan pohon keputusan dengan menerapkan berbagai tahapan dan proses yang terdapat dalam prosedur klasifikasi dan prosedur prediksi. Seiring berjalannya tahapan, pohon keputusan menganalisis data, dimulai dengan kondisi yang bervariasi dan informasi yang lengkap, dan diakhiri dengan keputusan. Penjelasan mengapa pohon keputusan berbentuk seperti pohon adalah:

1. Simpul bagian akar

Pohon keputusan memiliki simpul akar di bagian atas, yang merupakan bagian pertama dari struktur.

2. Simpul internal

Node cabang yang jauh dari node akar dapat menghasilkan maksimal dua keluaran hanya dengan satu masukan.

### 3. Simpul daun

Simpul daun adalah simpul yang berada di ujung pohon. Mereka punya masukan, tapi tidak punya keluaran (Azwanti & Putria, 2023)

Bagian selanjutnya menjelaskan kelebihan dan kekurangan pohon keputusan:

#### 1. Kelebihan dari *decision tree*

- a. Pohon keputusan digunakan ketika cakupan data yang kompleks perlu diubah menjadi data yang lebih sederhana sebelum pengambilan keputusan dapat dilakukan.
- b. Penggunaan pohon keputusan menghilangkan metode penghitungan yang tidak perlu sehingga dapat dilakukan dengan lebih efisien.
- c. Jika pohon keputusan digunakan sebagai alat untuk memilih kelengkapan atau kejelasan data, maka hasil keputusan mengenai kelengkapan atau kejelasan data akan menjadi pembeda dalam pengelompokan data berdasarkan kriteria lain jika pohon keputusan digunakan. dalam proses penyeleksian kelengkapan atau kejelasan data.
- d. Dengan bantuan pohon keputusan, suatu keputusan dapat diambil dengan memanfaatkan sejumlah kecil kategori atau kelompok,

namun pada saat yang sama tanpa mempengaruhi kualitas keputusan itu sendiri.

## 2. Kekurangan dari *decision tree*

- a. Penggunaan pohon keputusan untuk mengambil suatu keputusan menjadi sulit ketika jumlah hasil keputusan yang harus dihasilkan dan kriteria yang harus dipenuhi sangat banyak sehingga memerlukan waktu yang lama untuk mencapai suatu keputusan. Oleh karena itu, perlu digunakan sistem pendukung memori dengan kapasitas yang tinggi.
- b. Pohon keputusan berisi sejumlah kesalahan yang terjadi di setiap bagian struktur yang terdapat dalam pohon keputusan, serta jumlah agregat kesalahan yang terjadi di seluruh struktur.
- c. Jelas sulit untuk membuat struktur pohon keputusan, yang mencapai hasil maksimal, dari informasi yang diberikan.
- d. Pohon keputusan digunakan untuk menghasilkan keputusan, dan dalam hal ini kualitas atau hasil dari keputusan yang dihasilkan sangat dipengaruhi oleh struktur dan model yang digunakan untuk membuat pohon keputusan (Harryanto & Hansun, 2017).

Proses penentuan nilai gain dengan hasil tertinggi ini dilakukan pada saat pohon keputusan berada pada tahap awal, yaitu pada saat memilih kelengkapan data atau informasi, dengan cara memeriksa nilai gain dengan hasil tertinggi dari total nilai kelengkapan atau informasi yang ada. Rumus berikut dapat digunakan untuk mencari nilai gain tertinggi:

$$Gain(A) = Entropi(S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi(S_i)$$

### Rumus 2.1 Perhitungan Gain

Keterangan:

$S$  = himpunan kasus

$A$  = atribut

$n$  = jumlah partisi atribut  $A$

$|S_i|$  = jumlah kasus dalam partisi yang ke- $i$

$|S|$  = jumlah kasus dalam  $S$  (Prasetyaningrum & Susanti, 2023)

#### 2.1.1.8. Algoritma C4.5

Terdapat sebuah algoritma bernama C4.5 yang dapat digunakan untuk mengambil keputusan pohon berdasarkan bentuk data yang tersedia saat ini, dan dapat digunakan untuk mengklasifikasikan data yang sangat besar ke dalam kelas-kelas tertentu serta bergantung pada bentuk datanya. data yang ada. Model pembelajaran berbasis pohon keputusan yang disebut algoritma C4.5 dapat digunakan untuk mempelajari data dalam jumlah besar dan digunakan dalam penambangan data dan pembelajaran mesin. Ini dapat digunakan untuk mempelajari berbagai kumpulan data (Jannah *et al.*, 2023). (Prasetyaningrum & Susanti, 2023) algoritma C4.5 adalah algoritma yang sangat efektif yang mengubah sejumlah besar fakta menjadi pohon keputusan yang mewakili aturan dalam bentuk pohon keputusan.

Kegunaan algoritma dari C4.5 ini banyak sekali, antara lain penggunaan angka untuk mengklasifikasikan data, atau menggunakan kategori untuk mengklasifikasikan data. Ada beberapa pola aturan yang diturunkan dari proses

klasifikasi, dan aturan tersebut dapat digunakan untuk melakukan berbagai tahapan proses prediksi dengan nilai prediksi yang berasal dari pencatatan data lama hingga terbaru setelah tahapan proses klasifikasi tersebut lengkap (Azwanti & Putria, 2023). Prosedur berikut digunakan dalam menghitung nilai entropi.

1. Menggunakan rumus yang menghitung nilai Entropi,
2. Dengan menggunakan nilai Gain Ratio setiap atribut, hitung nilai Gain Ratio keseluruhan untuk setiap atribut
3. Atribut yang memiliki nilai Gain Ratio tertinggi akan dipilih sebagai akar pohon, sedangkan atribut yang memiliki nilai Gain Ratio terendah akan dipilih masing-masing sebagai cabang.
4. Pada proses tahap sebelumnya, menghitung nilai Gain Ratio untuk setiap atribut dengan mengecualikan atribut yang dipilih sebagai akar analisis dari perhitungan.
5. Suatu cabang dipilih berdasarkan atribut dengan Gain Ratio tertinggi yang telah dipilih.
6. Hingga nilai Gain yang dihasilkan untuk semua atribut yang tersisa adalah 0 melalui penggunaan langkah 4 dan 5, harus mengulangi langkah 4 dan 5.

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

**Rumus 2.2** Perhitungan Entropy

Keterangan:

$S$  = himpunan kasus

$A$  = fitur

$n$  = jumlah partisi  $S$

$p_i$  = proporsi dari  $S_i$  kepada  $S$  (Prasetyaningrum & Susanti, 2023).

### **2.1.2. Teori Khusus**

Berdasarkan teori spesifik berikut, penelitian ini dilakukan.

#### **2.1.2.1. Kepuasan Pelanggan**

Pelanggan seringkali merasakan puas terhadap suatu produk atau jasa karena prediksi atau harapan yang mereka miliki, dan rasa percaya diri yang mereka miliki sesaat sebelum mereka membeli atau menggunakannya. Harapan masyarakat terhadap kualitas produk atau jenis layanan yang akan mereka terima setelah membeli produk atau menggunakan bentuk layanan biasanya berasal dari kualitas produk atau perkiraan layanan yang akan mereka terima. Begitu seorang pelanggan membeli atau menggunakan sebuah produk atau jasa, jika dia tidak puas pada kualitas produk atau jasa tersebut, maka pelanggan tersebut akan merasa tidak puas terhadapnya. Selain itu, hal sebaliknya juga terjadi, dan jika pelanggan merasa puas dengan kualitas produk maupun layanan yang dirasakan, kemungkinan pelanggan juga pasti puas dengan kualitas tersebut (Azwanti & Putria, 2023).

Pada dasarnya kepuasan konsumen diukur sebagai tingkat kepuasan yang dimiliki seorang pelanggan terhadap produk atau layanan yang diterimanya. Pelanggan bisa saja merasa puas bila harapannya terpenuhi melalui layanan yang diterimanya. Untuk memperoleh pelayanan yang terbaik, konsumen berusaha mencapai tingkat kualitas yang setinggi-tingginya. Untuk menjaga dan



meningkatkan hubungan baik antara penjual dan pembeli, fungsi pelayanan sangatlah penting (Dewi *et al.*, 2023).

Sebagai ukuran kepuasan konsumen, kita dapat mengartikannya sebagai ukuran kepuasan konsumen terhadap pembelian atau konsumsi suatu barang, serta faktor-faktor tambahan apa pun yang mereka anggap penting bagi mereka. Tidak ada konsensus umum tentang apa yang dimaksud dengan kepuasan konsumen, dan hal ini bervariasi dari orang ke orang. Dalam banyak kasus, konsumen akan menerima begitu banyak promosi dan penawaran kemudahan, namun tidak akan memuaskan kebutuhan mereka, dalam kasus lain, konsumen akan sangat senang dengan kondisi dan keadaan sederhana yang ditawarkan, dan kebutuhan serta keinginan mereka akan terpenuhi (Tarigan *et al.*, 2017).

#### **2.1.2.2. Kualitas Produk**

Produk ialah apa saja yang ada di pasaran dapat dimanfaatkan untuk memperoleh perhatian, permintaan, konsumsi, atau pemanfaatan guna memenuhi keinginan atau kebutuhan konsumen yang dapat tersedia di pasaran. (Izzuddin & Muhsin, 2020). Penting untuk dipahami bahwa kualitas produk ditentukan oleh sejumlah faktor, antara lain daya tahan, keandalan serta kemajuan, kekuatan, kemudahan pengemasan juga perbaikan suatu produk (Yanuar *et al.*, 2017).

Kualitas Produk merupakan kemampuan suatu produk untuk memuaskan keinginan dan kebutuhan konsumen dengan menyediakan seluruh komponen baik fisik maupun non fisik yang tersedia. (Budiono, 2021). Tingkat kepuasan yang dialami seorang konsumen bisa berbeda-beda, yaitu jika ia tidak puas terhadap suatu produk setelah ia mengkonsumsinya maka ia akan merasa tidak puas, jika

produk tersebut tidak sesuai dengan harapannya. Namun apabila yang terjadi justru sebaliknya, yaitu produk tersebut sesuai dengan apa yang diharapkan pelanggan, maka saat sudah sampai pelanggan merasa puas, dan pasti bersedia untuk membeli kembali produk tersebut dalam rentang waktu dekat (Tarigan *et al.*, 2017).

### **2.1.2.3. Kualitas Pelayanan**

Pelayanan adalah suatu proses atau tindakan yang memberikan bantuan, dukungan, atau kepuasan kepada individu atau pelanggan guna memuaskan kebutuhan, keinginan, atau harapannya melalui pemenuhan kebutuhan, keinginan, atau harapannya. Pemberian layanan merupakan interaksi antara penyedia layanan dan penerima layanan, yang tujuannya untuk menciptakan nilai, meningkatkan kepuasan pelanggan, dan menanggapi kebutuhan pelanggan secara efisien dan efektif (Dewi *et al.*, 2023).

Untuk memenuhi kebutuhan konsumen, kualitas pelayanan didasarkan pada penetapan ekspektasi kualitas dan pengendalian variasi untuk mencapai kualitas tersebut. Pelanggan akan merasakan puas pada suatu kualitas pelayanan yang melebihi harapannya juga melebihi harapannya. (Putri & Arnomo, 2020). Menurut Tjiptono dalam (Yanuar *et al.*, 2017) bahwa sebagian besar skema pemasaran yang berupaya menarik konsumen mencapai keberhasilan jika dapat memberikan layanan dengan kualitas terbaik yang memenuhi kebutuhan dan harapan mereka sekaligus akurat dalam menyampaikan kebutuhan dan keinginan tersebut.

#### **2.1.2.4. Lokasi**

Dalam bisnis, lokasi berkaitan dengan tempat di mana bisnis itu dijalankan. Lokasi usaha yang baik merupakan hal yang penting baik bagi produsen maupun konsumen, karena merupakan lokasi yang disukai baik oleh produsen maupun konsumen sebagai tempat usaha dapat beroperasi. Tiga kunci sukses bisnis menurut Straub dan Attner adalah lokasi, lokasi, lokasi. Artinya, di masa depan, lokasi akan memainkan peran penting dalam menentukan mampu atau tidaknya suatu bisnis jasa bertahan. Penting juga bagi konsumen untuk mengetahui lokasi penyedia layanan mereka agar dapat mengambil keputusan yang tepat. Faktanya, lokasi merupakan aspek terpenting dari penyedia layanan kepada pelanggan, karena penyedia layanan berlokasi dekat dengan rumah atau kantor pelanggan dan oleh karena itu dekat dengan tempat tinggal atau tempat kerja pelanggan (Izzuddin & Muhsin, 2020).

Sekalipun memiliki produk yang bagus, namun jika tidak tersedia lokasi yang tepat untuk memasarkan produk tersebut ke konsumen, maka produk tersebut tidak akan laku. Pemilihan lokasi yang tepat untuk suatu usaha dipengaruhi oleh beberapa faktor, antara lain lokasi yang mudah dijangkau dan fasilitas pendukungnya. Dalam dunia bisnis, tempat dimana Anda menentukan posisi perusahaan Anda disebut dengan lokasi. Pengertian lokasi mengacu pada suatu benda dalam ruang atau suatu fenomena dalam ruang yang dapat memberikan informasi lebih lanjut mengenai objek atau fenomena geografis yang menjadi pokok persoalan (Efnita, 2017). Konsep akses lokasi mengacu pada kemampuan atau izin untuk mendapatkan informasi tentang lokasi suatu tempat dan cara

menuju ke sana. Penting untuk mempertimbangkan apakah lokasi tersebut mudah diakses atau sulit dijangkau (Dewi *et al.*, 2023)

#### **2.1.2.5. Harga**

Sebagai faktor terpenting kedua yang mempengaruhi kepuasan pelanggan, harga juga menjadi faktor yang sangat penting. Pada dasarnya, harga mengacu pada jumlah uang yang pembeli bersedia bayarkan dan penjual bersedia menerimanya sebagai pembayaran (Yanuar *et al.*, 2017). Untuk kepentingan penerimaan suatu jasa atau barang tertentu, harga adalah sejumlah uang atau nilai tertentu yang ditentukan atau diperlukan sebagai pembayaran atas jasa atau barang tersebut. Pertimbangan harga suatu produk atau jasa dari segi nilai keekonomiannya dapat dijadikan sebagai salah satu faktor dalam mengambil keputusan untuk membelinya atau tidak (Dewi *et al.*, 2023).

Ada dua aspek dalam istilah harga - jumlah uang (dalam satuan moneter) dalam suatu pertukaran dan/atau aspek non-moneter dari harga yang mengandung kegunaan atau kegunaan tertentu yang diperlukan untuk memperoleh produk. Sebagai bagian dari strategi pemasaran utama yang berorientasi pada permintaan suatu perusahaan, tujuan penetapan harga dapat mendukung strategi pemasaran yang berorientasi pada permintaan jika perusahaan yakin bahwa harga yang lebih rendah akan menyebabkan peningkatan jumlah pengguna, tingkat penggunaan, atau pembelian berulang atas produk tertentu. dalam kategori atau bentuk produk tertentu (Efnita, 2017).

#### **2.1.2.6. Aplikasi WEKA**

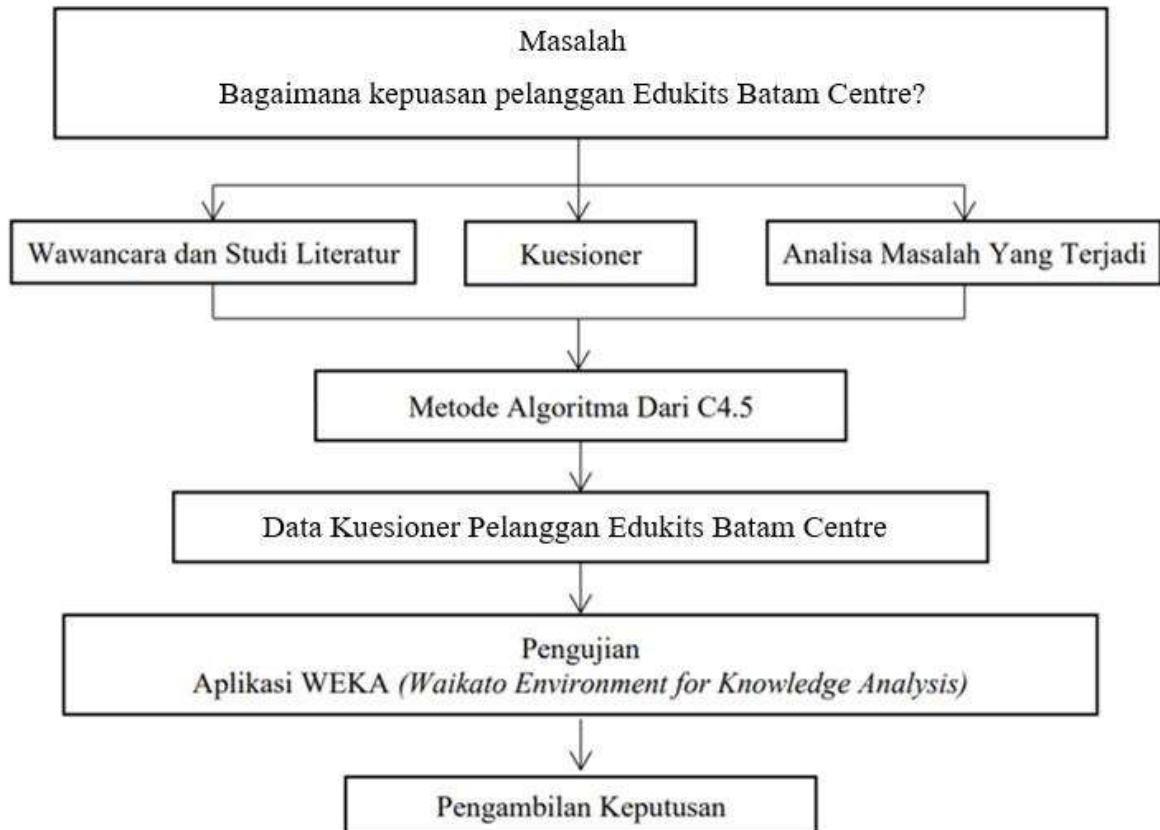
Awalnya dikembangkan oleh University of Waikato di Selandia Baru sebelum menjadi bagian dari Pentaho, aplikasi ini pertama kali dikembangkan oleh University of Waikato. Sebagai algoritma pembelajaran mesin, Weka adalah kelas algoritma yang dapat digunakan untuk menggeneralisasi dan merumuskan kumpulan data pengambilan sampel secara keseluruhan. Walaupun kekuatan Weka terletak pada algoritmanya yang lebih komprehensif dan canggih, namun kekuatan data mining tetap terletak pada pengetahuan komunitas yang mengimplementasikannya. Untuk memastikan keberhasilan formulasi, penting untuk mengumpulkan data berkualitas tinggi dan memiliki pengetahuan tentang pemodelan dan cara memanfaatkan algoritma dengan cara yang tepat.

WEKA ( Waikato Environment for Knowledge Analysis) merupakan Proses penambangan data yang didukung oleh perangkat lunak sumber terbuka ini fleksibel dan

menyediakan banyak jenis metode penambangan data yang siap pakai. Framework Java yang menjadi dasar Weka memungkinkan pengguna menggunakannya dalam berbagai situasi untuk menganalisis data (Bilqisth & Amalia, 2022).

### **2.2. Kerangka Pemikiran**

Di bawah ini adalah flowchart yang menggambarkan gambaran singkat kerangka berpikir seperti terlihat pada gambar di bawah ini.



**Gambar 2.3** Kerangka Pemikiran

**Sumber:** (Penulis, 2023)

### 2.3. Hipotesis Penelitian

Hipotesis penelitian dirumuskan berdasarkan kerangka di atas, dan diringkas sebagai berikut:

1. Kepuasan pelanggan Edukits Batam Center dapat dianalisis dengan algoritma C4.5 yang digunakan dalam data mining.
2. Dengan menggunakan proses data mining dan perhitungan algoritma C4.5, hasil penelitian dapat membantu memecahkan permasalahan

dalam pengambilan keputusan mengenai kepuasan pelanggan Edukits Batam Center.

#### **2.4. Penelitian Terdahulu**

Mengingat masalah dan judul penelitian, maka perlu untuk menetapkan apa yang telah terungkap dari penelitian sebelumnya dari sudut pandang yang berbeda untuk mendapatkan pemahaman yang lebih baik tentang fenomena serupa dari sudut pandang yang berbeda. Tulisan ini akan mencoba melakukan hal tersebut. Untuk mendukung penelitian yang diusulkan, berikut adalah ringkasan penelitian terdahulu yang mendukung gagasan tersebut:

1. Dari penelitian “Index Kepuasan Pelanggan Informa dengan Menggunakan Algoritma C.45” yang dilakukan oleh (Febriyani *et al.*, 2021) didapatkan informasi bahwa kepuasan pelanggan Informa dengan penggunaan Algoritma C.45 menunjukkan hasil akurasi yang didapat yaitu sebesar 84,16 % dengan rincian sebagai berikut Hasil Prediksi Puas.
2. Dari penelitian “Analisa Kepuasan Konsumen pada Laundry Menggunakan Algoritma C.45” yang dilakukan oleh (Dewi *et al.*, 2023) didapatkan informasi bahwa Penelitian ini menerapkan algoritma C4.5 dengan akurasi sebesar 88.00 %. Dapat disimpulkan bahwa ya dihasilkan karena perasaan puas dari konsumen terhadap Laundry Bunda.
3. Dari penelitian “Implementasi Algoritma C.45 terhadap Kepuasan Pelanggan” yang dilakukan oleh (Tarigan *et al.*, 2017) didapatkan

informasi bahwa Hasil penelitian terdapat tingkat akurasi sebesar 91%, dengan nilai presisi pada prediksi puas sebesar 92.21% dan nilai presisi pada prediksi tidak puas sebesar 90,91%.

4. Dari penelitian “Analisa Tingkat Kepuasan Pelanggan pada Percetakan CV. Mega Media Menggunakan Algoritma C.45” yang dilakukan oleh (Prasetyaningrum & Susanti, 2023) didapatkan informasi bahwa kepuasan pelanggan yang di prediksi dengan mengaplikasikan algoritma C.45 sehingga mendapatkan hasil akurasi dengan nilai 93.00% pada K-Fold validation 2 dan termasuk kategori good classification.
5. Dari penelitian “Penerapan Algoritma C.45 untuk Menentukan Tingkat Kepuasan Pelanggan Kartu Telkomsel Prabayar” yang dilakukan oleh (Sikumbang *et al.*, 2022) berdasarkan hasil pengujian menggunakan algoritma C4.5 maka dapat disimpulkan bahwa nilai rata-rata setiap variabel yang menjadi akar dalam pohon keputusan adalah variabel Kualitas Layanan (X4) dengan nilai gain tertinggi sebesar 0.266396957.
6. Dari penelitian “Penerapan Data Mining Klasifikasi Kepuasan Pelanggan Transportasi Online Menggunakan Algoritma C.45” yang dilakukan oleh (Jannah *et al.*, 2023) berdasarkan hasil pengujian menggunakan algoritma C4.5 didapatkan hasil klasifikasi kepuasan pelanggan pengguna transportasi online, jumlah pelanggan pengguna transportasi online dengan klasifikasi “Tidak Puas” sebanyak 91 orang pelanggan, sedangkan yang “Tidak Puas” sebanyak 9 orang pelanggan.



7. Dari penelitian “Analisa Klasifikasi Data Mining pada Tingkat Kepuasan Pengunjung Taman Hewan Pematang Siantar dengan Algoritma C.45” yang dilakukan oleh (Batubara & Windarto, 2019) berdasarkan hasil pengujian menggunakan algoritma C4.5 didapatkan data *performance* yang ditunjukkan terhadap kesesuaian metode C4.5 dengan kepuasan pengunjung taman hewan adalah 95%.
8. Dari penelitian “Analysis of sales levels of pharmaceutical products by using data mining algorithm C45” yang dilakukan oleh (Sovia *et al.*, 2021) hasil penelitian menunjukkan C45 menghasilkan rules yang dapat mengidentifikasi pewarisan suatu jenis produk obat sebesar 55% dari kategori objektif yang ada.
9. Penelitian berikutnya “Prediksi Duraksi Kuliah Dengan Algoritma *Classifier*.” Diteliti oleh (Harman & Rosiska, 2020). Hasil yang didapatkan ialah diketahui hampir 99% mahasiswa yang kuliah di Universitas Terbuka Batam (POKJAR) sudah bekerja, dimana 80% mampu menyelesaikan studi tepat waktu, dan 20% lainnya tidak. Hal ini berdasarkan data Kelompok Belajar Universitas Terbuka Batam (POKJAR).