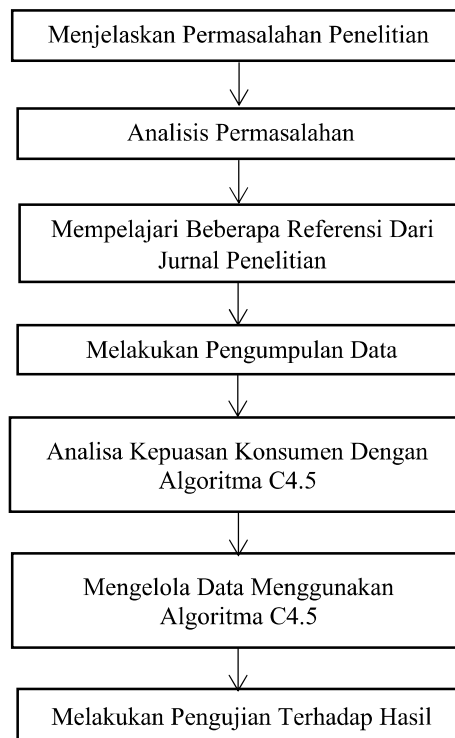


## BAB III

### METODE PENELITIAN

#### 3.1. Desain Penelitian

Dalam pelaksanaan penelitian kepuasan konsumen pada UMKM MiniMall, digunakan rencana yang berisi langkah-langkah untuk memberikan arah dalam membahas permasalahan penelitian. Gambaran mengenai desain dari penelitian yang dilakukan, sebagai berikut.



**Gambar 3.1.** Desain Penelitian  
**Sumber:** (Penulis, 2021)

Penjelasan mengenai gambar desain dari penelitian diatas, sebagai berikut:

1. Menjelaskan Permasalahan Penelitian

Pada tahap ini diuraikan mengenai permasalahan yang terdapat pada penelitian, yaitu terdapat kondisi penurunan penjualan barang dan kebutuhan akan informasi mengenai kepuasan konsumen pada UMKM MiniMall.

2. Analisis Permasalahan

Pada tahap ini merupakan kegiatan lanjutan setelah menjelaskan masalah penelitian. Setelah mendapatkan permasalahan, maka ditetapkan keputusan untuk melakukan analisis kepuasan konsumen sehingga informasi yang dihasilkan dapat menjadi pedoman dalam membuat kebijakan bagi UMKM MiniMall sehingga dapat kembali meningkatkan penjualannya.

3. Mempelajari Beberapa Referensi Dari Jurnal Penelitian

Untuk mendukung pelaksanaan kegiatan penelitian, maka dilakukan proses pencarian informasi dari beberapa referensi seperti jurnal penelitian yang membahas mengenai analisa kepuasan konsumen. Sehingga pada akhirnya, penelitian ini dilakukan menggunakan algoritma C4.5 dan beberapa indikator yang dapat memberikan pengaruh kepada kepuasan konsumen.

4. Melakukan Pengumpulan Data

Proses dalam pengumpulan data dilakukan dengan memberikan kuesioner kepada sampel penelitian, yaitu konsumen dari UMKM Minimall.

#### 5. Analisa Kepuasan Konsumen Dengan Algoritma C4.5

Data dianalisa dengan algoritma dari C4.5 mulai dari proses pembuatan struktur *decision tree* hingga menghasilkan pola berupa aturan dari kepuasan konsumen.

#### 6. Mengelola Data Menggunakan Algoritma C4.5

Selanjutnya dilakukan perhitungan nilai *gain* tertinggi dan hasilnya menjadi acuan dalam membuat *decision tree*.

#### 7. Pengujian Hasil

Pada tahapan ini, dilakukan pengujian kembali mengenai hasil pengelolaan data yang sudah dilakukan secara manual menggunakan algoritma C4.5 dengan memanfaatkan aplikasi WEKA versi 3.9.5.

### **3.2. Objek Penelitian**

Pada sub bab objek penelitian, penulis menjelaskan keterangan lokasi penelitian dan penjelasan jadwal selama kegiatan penelitian, sebagai berikut.

#### **3.2.1. Lokasi Penelitian**

Penelitian ini berlokasi di cabang perusahaan UMKM MiniMall dengan alamat pada Jl. Teratai Gerbang Blok VI Baloi Ruko Karimum Pinang Jaya No. 9 Windsor, Kelurahan Batu Selicin, Kecamatan Lubuk Baja, Kota Batam, Provinsi Kepulauan Riau.

### 3.2.2. Jadwal Penelitian

Demikian penjelasan penulis mengenai rencana waktu yang akan dilakukan dalam melaksanakan kegiatan analisa mengenai kepuasan konsumen, sebagai berikut.

**Tabel 3.1.** Jadwal Penelitian

Tahap	Uraian	Bulan															
		Maret 2021	April 2021				Mei 2021				Juni 2021				Juli 2021		
		4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3
1	Pengajuan judul																
2	Bimbingan dengan dosen																
3	Studi kepustakaan																
4	Pengajuan penelitian																
5	Rancangan penelitian																
6	Pengumpulan data																
7	Pengolahan data																
8	Penyusunan Laporan																

Sumber: (Penulis, 2021)

### **3.3. Populasi dan Sampel**

Dalam sub bab ini penulis akan menjelaskan mengenai keterangan yang dimiliki oleh populasi dan sampel yang dijadikan bahan dalam pelaksanaan kegiatan penelitian sebagai berikut.

#### **3.3.1. Populasi**

Populasi penelitian ini adalah konsumen yang melakukan pembelian barang pada UMKM MiniMall. Konsumen UMKM MiniMall dari Maret 2020 hingga bulan Februari 2021 sudah mencapai 218.510 konsumen.

#### **3.3.2. Sampel**

Dalam menentukan sampel, digunakan teknik *purposive sampling*. Teknik *purposive sampling* dilakukan dengan tidak acak. Dalam menentukan sampelnya, digunakan beberapa syarat tertentu sesuai dengan tujuan penelitian. Karena jumlah konsumen UMKM MiniMall sangat banyak, maka ditetapkan beberapa syarat, yaitu: hanya konsumen yang sudah melakukan pembelian barang pada UMKM MiniMall lebih dari 20 kali transaksi pembelian dan total keseluruhan barang yang dibeli konsumen tersebut lebih dari 100 barang. Berdasarkan syarat tersebut, maka didapatkan sampel sebanyak 50 koresponden.

### **3.4. Variabel Penelitian**

Beberapa indikator yang digunakan dalam penelitian khususnya yang berkaitan dengan variabel dari kepuasan konsumen, yaitu:

1. Kualitas Produk
2. Harga
3. Promosi
4. Ketersediaan Produk
5. Ketepatan Pengiriman
6. Penanganan Keluhan

Untuk hasil akhir dari keputusannya adalah puas dan tidak puas.

### **3.5. Teknik Pengumpulan Data**

Proses pengumpulan data yang menjadi kebutuhan bahan penelitian dilakukan agar penulis dapat mencapai tujuan penelitian, yaitu:

1. Studi Literatur

Mempelajari bahan dari jurnal penelitian yang sudah dilakukan peneliti lainnya namun tetap berkaitan dengan penelitian yang dilakukan penulis.

2. Observasi

Mengamati data penjualan barang secara langsung pada UMKM MiniMall, sejak bulan Maret 2020 sampai Februari 2021.

3. Kuesioner

Proses kegiatan dengan memberikan kuesioner berisi pertanyaan menggunakan indikator dari variabel kepuasan konsumen kepada para

konsumen UMKM MiniMall menggunakan Google Form untuk mempermudah konsumen dalam mengisi kuesioner.

### **3.6. Metode Analisis Data**

Penulis melakukan penelitian, khususnya kegiatan analisis terhadap data menggunakan proses *Knowledge Discovery In Database* (KDD) dengan *data mining*, *decision tree* dan algoritma C4.5. Tahapan awal dilakukan klasifikasi sesuai dengan kategori atau kriteria data, selanjutnya dilakukan perhitungan algoritma dari C4.5 untuk mendapatkan nilai *entropy* dan *gain* tertinggi pada saat membuat struktur dari *decision tree* dilakukan secara berulang hingga simpul pada *decision* tidak memiliki turunan lagi sampai didapatkan pola berupa aturan dasar penilaian mengenai kepuasan konsumen pada UMKM MiniMall. Pada penelitian, digunakan KDD (*Knowledge Discovery In Database*), *decision tree* dan algoritma C4.5, penjelasannya sebagai berikut.

#### **3.6.1. KDD (*Knowledge Discovery In Database*)**

Proses KDD adalah menggali data yang tersimpan berdasarkan pola data agar mudah dipahami dan menghasilkan pengetahuan. Dalam KDD terdapat enam elemen paling dasar dalam teknik pencarian pengetahuan dalam KDD, yaitu (Hadi et al., 2020):

1. Mengolah sejumlah besar data
2. Diperlukan penggunaan data yang efisien
3. Mengutamakan keakuratan

4. Membutuhkan penggunaan bahasa tingkat tinggi
5. Gunakan beberapa bentuk pembelajaran otomatis
6. Menghasilkan hasil yang unik

Proses *Knowledge Discovery in Database* (KDD) dapat diuraikan, sebagai berikut (Hadi et al., 2020):

1. *Data selection* proses pemilihan data dari suatu kumpulan data dan proses tersebut dilakukan sebelum proses penggalian informasi dalam KDD dilakukan. Hasil dari proses pemilihan data tersebut akan disimpan terpisah dari basis data operasional.
2. *Pre-processing* atau *cleaning*, merupakan proses pembersihan data yang dilakukan sebelum proses *data mining* dilakukan. Proses pembersihan tersebut, diantaranya adalah menghapus data duplikat serta memeriksa data yang tidak konsisten.
3. Selanjutnya dilakukan transformasi terhadap data yang telah dipilih agar data tersebut sesuai dengan ketentuan proses *data mining*.
4. *Data mining* kemudian dilakukan untuk mencari pola atau informasi unik pada data dengan teknik atau algoritma tertentu. Namun karena bervariasi, maka pemilihan teknik atau algoritma tersebut ditentukan sesuai dengan tujuan dan proses KDD secara keseluruhan.
5. Terakhir dilakukan interpretasi untuk mengartikan pola-pola yang dihasilkan proses *data mining* agar mudah dipahami oleh para pihak yang berkepentingan. Pada tahap ini juga dilakukan proses evaluasi untuk



memastikan pola atau informasi yang ditemukan bertentangan dengan hipotesis sebelumnya atau tidak.

### 3.6.2. *Decision Tree*

*Decision tree* menghasilkan alternatif dalam memecahkan masalah dengan hasil yang kuat dan berguna sehingga *decision tree* terkenal dan banyak digunakan oleh peneliti saat melakukan penelitian dengan teknik *data mining*. *Decision tree* sangat bermanfaat apabila digunakan untuk keperluan dalam mengubah data yang ada menjadi sebuah pola berbentuk aturan sehingga dapat dipahami dengan mudah oleh manusia. Arsitektur dari *decision tree* memang dibuat mirip dengan struktur yang terdapat pada sebuah pohon dan penjelasannya, yaitu:

1. Simpul bagian akar

Simpul bagian akar terletak pada bagian paling atas dan pertama dari *decision tree*.

2. Simpul internal

Simpul dari bagian internal ini adalah percabangan lanjut dari simpul bagian akar.

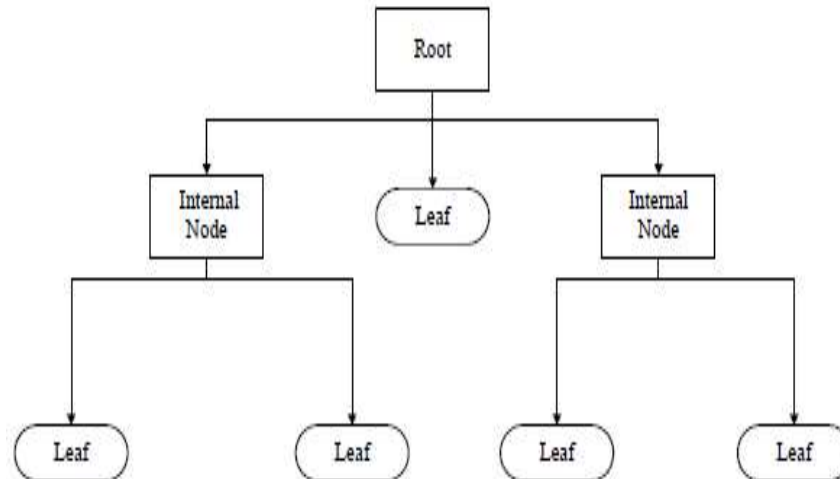
3. Simpul daun

Simpul daun merupakan simpul yang berada pada ujung bagian dari pohon.

Selanjutnya, demikian penjelasan dari kelebihan dan kekurangan yang dimiliki oleh *decision tree*, yaitu:

1. Kelebihan dari *decision tree*
  - a. Dalam penggunaan *decision tree* apabila terdapat cakupan data yang rumit maka dirubah menjadi lebih sederhana untuk kemudian dijadikan dasar pengambilan keputusan.
  - b. Kelebihan dari penggunaan *decision tree* yaitu dapat menghapus cara perhitungan yang tidak perlu sehingga dalam melakukan proses dari pengujian hanya menggunakan kriteria sesuai dengan keperluan untuk pengujian saja.
  - c. *Decision tree* bertindak untuk menghindari permasalahan lain dengan memanfaatkan menggunakan jenis kelompok atau kategori dengan jumlah sedikit tanpa mengurangi kualitas hasil keputusan akhir.
2. Kekurangan dari *decision tree*
  - a. Dalam menggunakan *decision tree* apabila jumlah dari hasil keputusan yang harus dihasilkan dan kriteria yang digunakan memiliki jumlah yang sangat banyak, maka waktu yang dibutuhkan dalam menghasilkan keputusan akan semakin bertambah serta dibutuhkan memori sistem pendukung yang tinggi.
  - b. Dalam *decision tree* terdapat akumulasi dari jumlah *error* yang terjadi dalam setiap bagian dari struktur yang terdapat pada *decision tree*.
  - c. Membuat struktur dari *decision tree* dengan hasil yang maksimal ternyata sulit untuk dilakukan.

- d. Apabila menggunakan *decision tree*, maka hasil atau kualitas dari keputusan yang dihasilkan sangat ditentukan dari model atau struktur *decision tree* yang dibuat. (Harryanto & Hansun, 2017)



**Gambar 3.2.** Model *Decision Tree*  
Sumber: (Takalapeta, 2018)

### 3.6.3. Algoritma C4.5

Pada dasarnya, pohon keputusan yang berasal dari fakta dengan jumlah yang besar dan memiliki fungsi untuk merepresentasikan sebuah aturan bisa dihasilkan melalui penggunaan algoritma yang disebut dengan algoritma C4.5. Agar masalah menjadi mudah untuk diselesaikan merupakan tujuan dibentuknya pohon keputusan tersebut. Pada saat menerapkan algoritma ini, pada dasarnya terdapat beberapa tahap. Dibentuknya model pohon yang berasal dari data yang sebelumnya sudah diubah ke dalam bentuk tabel merupakan tahap awal dari algoritma ini. Selanjutnya, pada tahap kedua dibuat aturan yang dilakukan melalui

proses pengubahan dari model pohon yang dihasilkan pada tahap pertama. Pada tahap akhir dilakukan proses membuat aturan yang dihasilkan pada tahap kedua menjadi lebih sederhana (Romli et al., 2020).

Pada umumnya algoritma dari C4.5 digunakan dalam proses pembangunan struktur dari *decision tree*, dan demikian penjelasan proses atau tahapannya yaitu:

1. Pertama adalah memilih kelengkapan atau keterangan data sebagai simpul akar dari *decision tree*.
2. Kemudian dilanjutkan dengan membuat struktur cabang dari *decision tree* yang sudah berisi nilai-nilai.
3. Pada tahap ketiga, dilakukan pembagian atau pengelompokan dari permasalahan yang terdapat pada struktur cabang dari *decision tree*.
4. Dan tahap akhir adalah melakukan proses pengulangan yang ada pada tahapan pertama hingga ketiga sampai seluruh cabang pada struktur *decision tree* memiliki kelas dan kategori yang sama.

Pada saat mencari nilai *gain* tertinggi dapat dilakukan dengan rumus berikut:

$$\mathbf{Gain (A) = Entropi (S) - \sum_i^n \frac{|S_i|}{|S|} \times Entropi (S_i)} \quad \mathbf{Rumus\ 3.1.}\ \text{Perhitungan Gain}$$

Keterangan dari simbol yang terdapat pada rumus, yaitu:

$S$  adalah himpunan

$A$  adalah keterangan atau kelengkapan yang digunakan.

$n$  adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian  $A$

$|S_i|$  adalah jumlah dari kasus yang terdapat dalam partisi yang ke- $i$

$|S|$  adalah jumlah dari kasus yang terdapat dalam  $S$

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

**Rumus 3.2.** Perhitungan *Entropy*

Keterangan dari simbol yang terdapat pada rumus, yaitu:

$S$  adalah himpunan

$A$  adalah keterangan atau kelengkapan yang digunakan.

$n$  adalah jumlah partisi yang terdapat pada keterangan atau kelengkapan bagian  $S$

$p_i$  adalah proporsi dari  $S_i$  kepada  $S$  (Elisa, 2017).

#### 3.6.4. *Cross Validation*

Penilaian terhadap hasil statistik analisis akan menggeneralisasi kumpulan data independen dapat dilakukan dengan penggunaan model teknik validasi yang disebut dengan *cross validation*. Tujuan dari penggunaan teknik tersebut, yakni untuk prediksi model dan prediksi tingkat keakuratan sebuah model prediktif ketika digunakan. Bias yang terdapat pada data diketahui dapat dihilangkan melalui penggunaan K-fold *cross validation* (Tempola et al., 2018).

#### 3.6.5. *Confusion Matrix*

Pada dasarnya *recall*, akurasi, *error rate* serta *precision* merupakan hasil perhitungan yang dapat dihasilkan melalui penggunaan *confusion Matrix*. Diketahui bahwa, pada sistem terdapat sebuah kemampuan dan kemampuan yang berfungsi untuk menemukan peringkat yang paling relevan dapat dievaluasi melalui penggunaan hasil perhitungan dari *precision*. Selain itu, kemampuan pada sistem yang berfungsi untuk menemukan semua item yang relevan dari koleksi dokumen dapat dievaluasi melalui penggunaan hasil perhitungan dari *recall*. Perbandingan

antara jumlah keseluruhan kasus dengan kasus yang diidentifikasi benar dapat diketahui melalui hasil perhitungan dari *accuracy* serta perbandingan antara jumlah keseluruhan kasus dengan kasus yang diidentifikasi salah dapat diketahui melalui hasil perhitungan dari *error rate* (Arini et al., 2020).