

## **BAB II**

### **METODE PENELITIAN**

#### **2.1 Obyek Penelitian**

Obyek penelitian ini adalah masyarakat yang menjadi penerima manfaat program PAMSIMAS di Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara. Fokus utama adalah mengumpulkan data dari para responden mengenai berbagai aspek yang berkaitan dengan kepuasan mereka terhadap program ini.

##### **2.1.1 Lokasi dan Waktu Penelitian**

Penelitian ini dilaksanakan di Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara. Pengumpulan data dilakukan pada tanggal 20 Mei s/d 31 Mei 2024 dengan menyebarkan kuesioner kepada masyarakat di wilayah tersebut. Lokasi dan rentang waktu pengumpulan data ini dipilih dengan memperhatikan aksesibilitas responden serta kebutuhan untuk memperoleh representasi yang baik dari populasi yang diteliti.

##### **2.1.2 Jumlah Data**

Jumlah data yang digunakan dalam penelitian ini adalah 100 responden. Data yang dikumpulkan melalui kuesioner mengukur berbagai aspek kepuasan terhadap program PAMSIMAS.

##### **2.1.3 Atribut (Fitur)**

Atribut dalam dataset ini meliputi beberapa faktor yang mempengaruhi kepuasan pelanggan terhadap program PAMSIMAS. Berikut adalah beberapa atribut yang digunakan yaitu (i) Jenis Kelamin (ii) Pekerjaan (iii) Pendidikan (iv) Kualitas Air: Kejernihan, rasa dan bau air (v) Kuantitas: ketersediaan air di pelanggan (vi) Kontinuitas: pengaliran air selama 24

jam (vii) Pelayanan: kecepatan, keramahan, dan ketepatan waktu pelayanan (viii) Penanganan: kecepatan, ketepatan dan penyelesaian keluhan.

#### **2.1.4 Kelas (Label)**

Kelas atau label dalam penelitian ini adalah tingkat kepuasan responden terhadap program PAMSIMAS, dibagi menjadi dua kategori yaitu (i) Puas: Responden merasa puas dengan program PAMSIMAS (ii) Tidak Puas: Responden merasa tidak puas dengan program PAMSIMAS.

#### **2.1.5 Sampel Data**

Sampel data penelitian yang didapatkan dari hasil penyebaran kuesioner dapat dilihat pada Lampiran 2.

Nilai numerik pada kuesioner mewakili keterangan pada Tabel 2.1.

**Tabel 2.1** Skala dan Bobot

<b>Keterangan</b>	<b>Bobot</b>
Sangat Baik	5
Baik	4
Cukup	3
Kurang Baik	2
Tidak Baik	1

## **2.2 Alat dan Bahan**

Alat dan bahan yang digunakan dalam menganalisis data dalam penelitian ini adalah:

### **2.2.1 Kuesioner**

Kuesioner digunakan sebagai alat utama untuk pengumpulan data melalui google formulir, kuesioner berisi pertanyaan terkait dengan kepuasan pelanggan terhadap program PAMSIMAS.

### **2.2.2 Microsoft Excel**

Microsoft excel digunakan untuk mengatur dan membersihkan data kepuasan pelanggan secara efisien, termasuk pemformatan data, penghapusan duplikat, dan pengelompokan data berdasarkan kriteria tertentu.

### 2.2.3 Google Colab

Google Colab digunakan sebagai alat pengujian data untuk mencari keputusan dalam mengukur tingkat kepuasan pelanggan terhadap program PAMSIMAS (Amanda et al., 2022).

### 2.2.4 Laptop

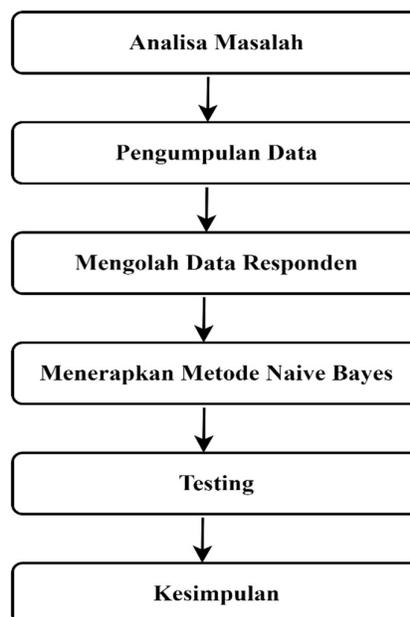
Laptop digunakan untuk merancang kuesioner, menganalisis data dan menyusun laporan.

### 2.2.5 Literatur dan Referensi

Literatur dan Referensi digunakan untuk mendukung perancangan penelitian, analisis data, dan interpretasi hasil. Ini bisa berupa buku, jurnal ilmiah, artikel, serta dokumen-dokumen terkait program PAMSIMAS.

## 2.3 Prosedur Penelitian

Untuk memastikan penelitian berjalan sesuai dengan harapan peneliti, prosedur penelitian harus jelas, dimulai dari identifikasi masalah hingga kesimpulan. Desain penelitian yang dilakukan dapat dilihat pada Gambar 2.1.



Gambar 2.1 Alur Penelitian

Dari Gambar 2.1 dapat dilihat bahwa alur penelitian dimulai dari :

### 2.3.1 Analisa Masalah

Menganalisis masalah yang terkait dengan tingkat kepuasan pelanggan terhadap program PAMSIMAS Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara.

### **2.3.2 Pengumpulan Data**

Data dikumpulkan dengan memberikan kuesioner kepada masyarakat di wilayah tersebut selama 12 hari dari tanggal 20 Mei s/d 31 Mei 2024.

### **2.3.3 Mengolah Data Responden**

Mengelola data dengan *Data Transformation Service* (DTS) melibatkan proses transformation dari format kuesioner ke format Microsoft Excel (.xlsx). Untuk menghitung nilai evaluasi dalam penelitian ini, melalui beberapa langkah evaluasi standar dalam machine learning. Berikut adalah langkah-langkah Evaluasi yang digunakan (i) Memisahkan Data: Sebelum melakukan evaluasi, data harus dipisahkan menjadi data latih (*training set*) dan data uji (*testing set*) (ii) Melatih Model: Model *Naive Bayes* dilatih menggunakan data latih. Algoritma akan mempelajari hubungan antara atribut-atribut dan label kepuasan (puas atau tidak puas) (iii) Menguji Model: Model yang telah dilatih diuji menggunakan data uji untuk memprediksi label kepuasan. Prediksi ini akan dibandingkan dengan label sebenarnya untuk menghitung metrik evaluasi.

### **2.3.4 Menerapkan Metode**

Dalam penelitian ini, penulis menggunakan Teknik data mining klasifikasi dengan algoritma *Naive Bayes* untuk mengukur tingkat kepuasan pelanggan terhadap program PAMSIMAS Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara.

### **2.3.5 Testing**

Pengujian data dilakukan dengan menggunakan aplikasi Google Collab sebagai sistem untuk menghasilkan keputusan dalam mengukur tingkat kepuasan pelanggan terhadap program PAMSIMAS Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara.

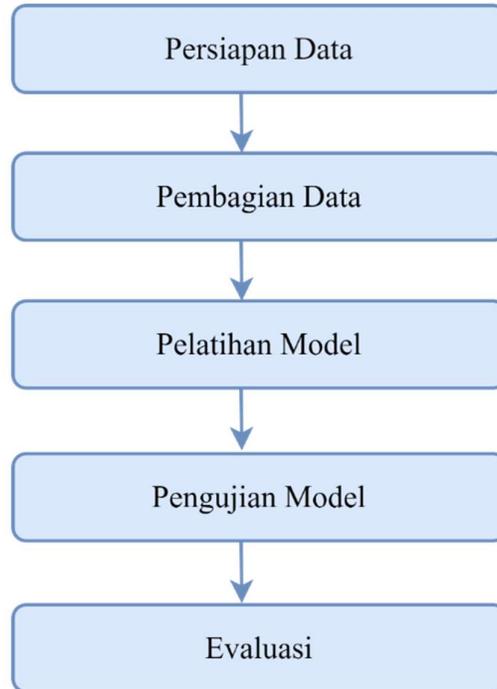
### 2.3.6 Kesimpulan

Kesimpulan dari klasifikasi tingkat kepuasan pelanggan terhadap program PAMSIMAS dapat memberikan wawasan yang berharga mengenai efektivitas program tersebut dalam memenuhi kebutuhan dan harapan masyarakat terkait penyediaan air bersih.

## 2.4 Algoritma *Naïve Bayes*

Penelitian ini menggunakan metode algoritma *Naïve Bayes* untuk menganalisis data yang telah dikumpulkan. Algoritma *Naive Bayes* digunakan untuk memprediksi kepuasan responden berdasarkan atribut-atribut yang ada. *Naïve Bayes* merupakan pengklasifikasian dengan bentuk model probabilistik dan statistik yang disederhanakan berdasarkan pada Teorema Bayes dengan asumsi bahwa setiap atribut bersifat *Independence* (bebas).

Dalam penelitian ini, kita akan melatih model *Naïve Bayes* menggunakan data yang ada dan kemudian menguji model untuk memprediksi apakah pelanggan merasa puas atau tidak puas dengan program PAMSIMAS berdasarkan atribut-atribut yang diberikan. Adapun langkah-langkah analisis yang digunakan dapat dilihat pada Gambar 2.2.



**Gambar 2.2** Langkah Analisis

Dari Gambar 2.2, dijelaskan bahwa untuk menentukan tingkat kepuasan pelanggan terhadap program PAMSIMAS, metode *Naïve Bayes* digunakan untuk melakukan beberapa penyelesaian berikut (i) Persiapan Data: Mengumpulkan dan membersihkan data, termasuk mengubah data kualitatif menjadi kuantitatif (ii) Pembagian Data: Memisahkan data menjadi dua bagian, yaitu data latih (*training data*) dan data uji (*testing data*) (iii) Pelatihan Model: Menggunakan data latih untuk melatih model Naive Bayes (iv) Pengujian Model: Menggunakan data uji untuk mengevaluasi performa model (v) Evaluasi: Menilai performa model dengan metrik seperti *accuracy*, *presisi*, *recall*, dan *F1-score* untuk memastikan model dapat memprediksi tingkat kepuasan dengan baik.

Adapun rumus *Naïve Bayes* yang digunakan dalam penelitian adalah :

Menghitung probabilitas prior dari setiap kelas dalam data latih:

$$P(C_k) = \frac{\text{jumlah sampel dalam kelas } C_k}{\text{total jumlah sampel}}$$

Keterangan:

$P(C_k)$  : Probabilitas prior dari kelas  $C_k$

Menghitung *likelihood*, yaitu menghitung probabilitas kondisi dari setiap fitur diberikan kelas tertentu.

$$P = (X_i|C_k)$$

Menghitung probabilitas posterior, menggunakan *teorema bayes* untuk menghitung probabilitas posterior untuk setiap kelas diberikan *instance* dari fitur.

$$P(C_k|X) = \frac{P(C_k) \cdot \prod_{i=1}^n P(X_i|C_k)}{p(X)}$$

Keterangan:

$P(C_k|X)$  : Probabilitas posterior dari kelas  $C_k$  diberikan fitur  $X$

$P(C_k)$  : Probabilitas prior dari kelas  $C_k$

$P(X_i|C_k)$  : Likelihood dari fitur  $X_i$  diberikan kelas  $C_k$

$P(X)$  : Probabilitas total dari fitur  $X$

Untuk menghitung nilai akhir kelas menggunakan rumus:

$$C_{MAP} = \operatorname{argmax}_{c \in C} P(X|C)$$

Keterangan:

$C_{MAP}$  : Hipotesa nilai tertinggi

$\operatorname{argmax}_{c \in C}$  : Nilai rata-rata dari setiap kelas

*Bayesian Classification* terbukti memiliki tingkat akurasi dan kecepatan yang tinggi. Berikut merupakan *Teorema Bayesian* dalam bentuk umum:

$$p(A|B) = \frac{(p(B|A) \cdot p(A))}{p(B)}$$

Keterangan:

$B$  : Data dengan kelas yang belum diketahui

$A$  : Hipotesa data  $B$  merupakan suatu kelas spesifik

$P(A)$  : Probabilitas hipotesa  $A$  (*prior probability*/probabilitas awal)

$P(B)$  : Probabilitas  $B$

$P(B|A)$  : Probabilitas hipotesa  $B$  berdasarkan kondisi  $A$

$P(A|B)$  : Probabilitas hipotesa  $A$  berdasarkan kondisi  $B$  (*Posterior probability*/probabilitas akhir)

#### 2.4.1 Rumus Pencarian Tingkat Akurasi

Akurasi adalah seberapa jauh nilai sebenarnya dan prediksi berbeda. Untuk mengukur akurasi model, digunakan *matrix confusion* yang menitik beratkan pada kelasnya. Sebuah *array* yang digunakan untuk mencatat hasil kerja klasifikasi disebut *confusion matrix*. Pada langkah ini, *matrix confusion* menguji data dan menggunakan model untuk menemukan tingkat akurasi terbaik. *Confusion Matrix* mengacu pada Tabel 2.2.

**Tabel 2.2** *Confusion Matrix*

Klasifikasi <i>Naïve Bayes</i>	Puas (+)	Tidak Puas (-)
Puas (+)	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
Tidak Puas (-)	<i>False Positive (FP)</i>	<i>True Positive (TP)</i>

Rumus *confusion matrix* yang digunakan adalah :

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan :

- Accuracy* adalah jumlah prediksi yang benar yang menyatakan puas.
- $TP$  (*True positive*) adalah jumlah pelanggan yang puas dan diprediksi puas oleh model
- $TN$  (*True negative*) adalah jumlah pelanggan yang tidak puas tetapi diprediksi puas oleh model
- $FP$  (*False positive*) adalah jumlah pelanggan yang puas tetapi diprediksi tidak puas oleh model

- e. FN (*False negative*) adalah jumlah pelanggan yang tidak puas dan diprediksi tidak puas oleh model

## 2.5 Jadwal Penelitian

Penelitian ini dilakukan di Desa Batuah Kecamatan Loa Janan Kabupaten Kutai Kartanegara selama lima bulan, dari Maret hingga Juli 2024. Jadwal penelitian mengacu pada Tabel 2.3.

**Tabel 2.3** Jadwal Penelitian

No	Kegiatan	Bulan (Tahun 2024)				
		Maret	April	Mei	Juni	Juli
Persiapan Penelitian						
1	1. Penyusunan dan Pengajuan Judul	█				
	2. Pembuatan Proposal		█			
Tahap Pelaksanaan						
2	1. Pengumpulan Data			█		
	2. Pengolahan Data				█	
	3. Penyusunan Laporan Skripsi				█	
	4. Seminar Hasil					█