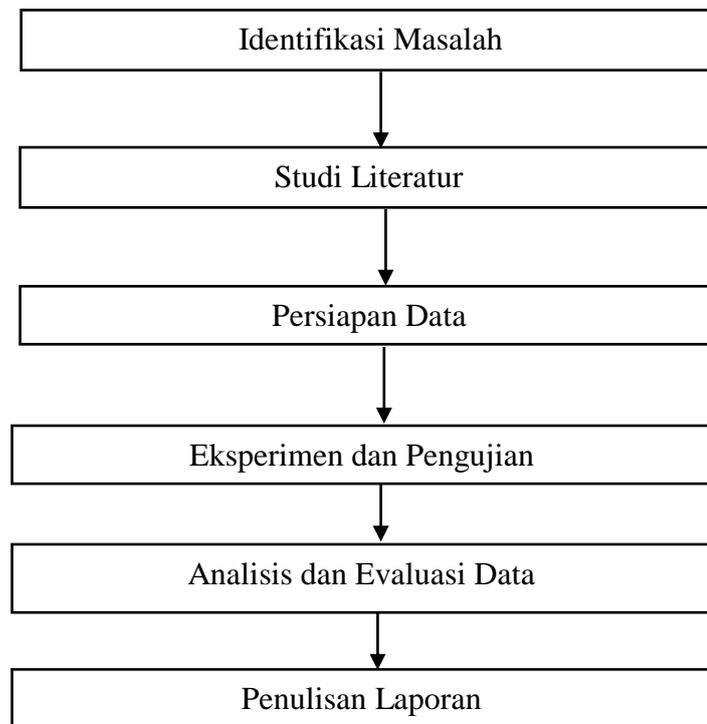


## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 KERANGKA KERJA PENELITIAN**

Metode kerangka kerja penelitian merupakan tahapan yang dilakukan selama mengerjakan penelitian. Kerangka dalam penelitian merupakan kumpulan konsep yang tersusun secara sistematis agar tujuan penelitian yang dilakukan dapat dilaksanakan dengan baik. Adapun kerangka kerja penelitian yang digunakan dapat dilihat pada gambar 3.1.



**Gambar 3.1 Kerangka Kerja Penelitian**

Berdasarkan kerangka kerja yang telah digambarkan diatas, maka dapat diuraikan pembahasan dari masing-masing tahapan penelitian yaitu:

1. Identifikasi Masalah

Pada tahap ini peneliti melakukan pengidentifikasian dari masalah yang terjadi yaitu membandingkan performa klasifikasi algoritma *K-NN* dan *C4.5* dalam mendeteksi penyakit gagal jantung dan mencari metode yang memiliki performa terbaik dalam pengklasifikasian data penyakit gagal jantung.

2. Studi Literatur

Pada tahapan ini dilakukan kajian pustaka yaitu mempelajari buku-buku, artikel, hasil penelitian sejenis yang relevan dengan permasalahan yang sedang diteliti sebagai sumber referensi dan landasan teoritis terhadap permasalahan yang diteliti.

3. Persiapan Data

Pada tahapan ini dilakukan tahap untuk mempersiapkan data yang telah diperoleh yang akan digunakan pada tahapan selanjutnya.

4. Eksperimen dan Pengujian

Tahapan ini akan membahas tahapan eksperimen dan teknik pengujian yang akan digunakan.

5. Analisis dan Evaluasi

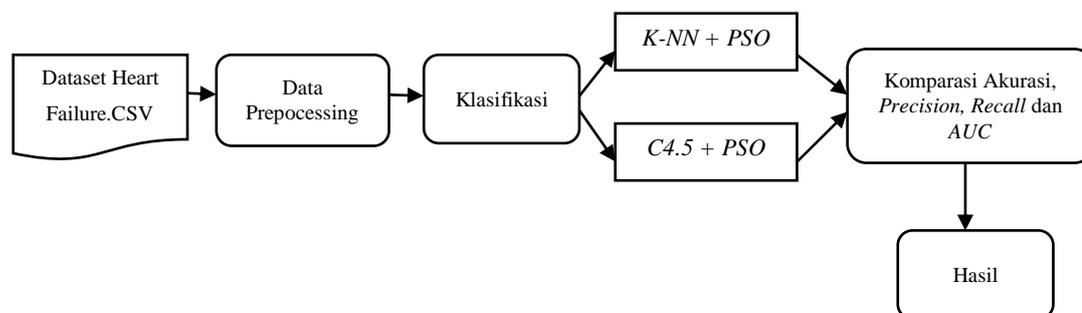
Setelah dilakukan tahapan eksperimen, maka selanjutnya pada tahapan ini dilakukan evaluasi terhadap hasil eksperimen yang telah dilakukan.

6. Penulisan Laporan

Setelah semua tahapan penelitian dilakukan maka peneliti akan membuat laporan sebagai bentuk dokumentasi penelitian agar dapat dimanfaatkan pada waktu akan datang, baik oleh peneliti sendiri maupun peneliti lainnya.

### 3.2 RANCANGAN EKSPERIMEN

Adapun alur proses pengklasifikasian atau eksperimen yang akan dilakukan pada penelitian ini dapat dilihat pada gambar 3.2.



**Gambar 3.2 Rancangan Eksperimen**

Berdasarkan pada gambar 3.2 berikut merupakan penjelasan mengenai tahapan-tahapan alur proses klasifikasi yaitu:

#### 1. Pengambilan Data

Data yang digunakan dalam eksperimen diperoleh dari data publik *kaggle* yang dapat diakses pada laman:

<https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction>.

Data tersebut berisikan 918 *record* data dengan 11 atribut dan 1 atribut *class*, Atribut *class* dijadikan penentu atau target pengklasifikasian.

#### 2. Data Preprocessing

Pada tahap ini dilakukan pembersihan data serta mempersiapkan data sebelum diproses ke tahap pengklasifikasian (*data mining*), biasa mencakup memeriksa data yang tidak konsisten serta memperbaiki kesalahan data.

### 3. Klasifikasi

Tahap klasifikasi dilakukan dengan alat bantu (*tools*) yaitu RapidMiner, data yang telah siap untuk dilakukan proses *data mining* kemudian diuji menggunakan metode pengklasifikasian yang akan digunakan. Metode yang dipakai pada penelitian ini yaitu *K-NN* dan *C4.5* berbasis *PSO* (*Particle Swarm Optimization*), atau pengujian yang akan dilakukan yaitu pengklasifikasian dengan algoritma *K-NN* dan *C4.5* tanpa *PSO* dan pengklasifikasian dengan algoritma *K-NN* dan *C4.5* menggunakan *PSO*.

### 4. Komparasi (Pengujian)

Setelah didapat hasil dari eksperimen yang telah dilakukan, pengujian dilakukan untuk menentukan tingkat akurasi dari suatu algoritma. Evaluasi dan validasi untuk mengukur kinerja algoritma dapat menggunakan *confusion matrix*. *Confusion matrix* adalah salah satu alat ukur berbentuk matrik 2x2 yang digunakan untuk mendapatkan jumlah ketepatan klasifikasi *dataset*. Evaluasi model klasifikasi didasarkan pada pengujian untuk memperkirakan objek yang benar dan salah, urutan pengujian ditabulasikan dalam *confusion matrix* [20].

**Tabel 3.1 Confusion Matrix**

Classification	Predicted Class		
	Class = Yes	Class = No	
Aktual Class	Class = Yes	A	B
		True Positive (TP)	False Negative (FN)
	Class = No	C	D
		False Positive (FP)	True Negative (TN)

Keterangan :

- a. TP (*True Positive*) : Hasil diprediksi dan data positif benar.
- b. FP (*False Positive*) : Hasil diprediksi positif namun data negatif.
- c. FN (*False Negative*) : Hasil diprediksi negatif namun data positif.
- d. TN (*True Negative*) : Hasil diprediksi dan data negatif benar.

Untuk pengukuran performa algoritma, pada penelitian ini menggunakan Akurasi, *Precision*, *Recall* dan *AUC* (*Area Under Curve*) dihitung untuk mengukur perbedaan performansi [30].

a. Akurasi

Pengujian akurasi adalah salah satu ukuran seberapa dekat hasil pengukuran terhadap angka sebenarnya. Akurasi dapat diperoleh dari persentase kebenaran yaitu perbandingan data benar dengan keseluruhan data [31]. Rumus menghitung akurasi dapat dilihat pada persamaan 3.1 berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} * 100\% \dots \dots \dots (3.1)$$

b. *Precision*

*Precision* atau presisi adalah salah satu metode pengujian suatu algoritma dengan melakukan suatu perbandingan data benar yang diperoleh sistem dengan jumlah seluruh data yang diambil oleh sistem yang benar maupun salah. Rumus menghitung *precision* dapat dilihat pada persamaan 3.2 berikut:

$$Precision = \frac{TP}{TP+FP} * 100\% \dots \dots \dots (3.2)$$

c. *Recall*

*Recall* adalah salah satu metode pengujian suatu algoritma dengan melakukan perbandingan jumlah data benar yang diperoleh sistem dengan jumlah seluruh data benar yang diambil atau tidak diambil oleh sistem. Rumus menghitung *recall* dapat dilihat pada persamaan 3.3 berikut:

$$\text{Recall} = \frac{TP}{TP+FN} * 100\% \dots \dots \dots (3.3)$$

d. Kurva *ROC / AUC*

Kurva *ROC (Receiver Operating Characteristics)* menunjukkan akurasi dan membandingkan klasifikasi secara visual dengan *false positive rate* sebagai garis horizontal dan *true positive rate* sebagai garis vertikal. Untuk klasifikasi *data mining* nilai *AUC* digunakan untuk menentukan model terbaik dalam memprediksi *class*. Nilai *AUC (Area Under Curva)* dapat dibagi menjadi beberapa kelompok yaitu [25]:

- a. 0.90-1.00 = Klasifikasi Sangat Baik (*Excellent Classification*)
- b. 0.80-0.90 = Klasifikasi Baik (*Good Classification*)
- c. 0.70-0.80 = Klasifikasi Cukup (*Fair Classification*)
- d. 0.60-0.70 = Klasifikasi Buruk (*Poor Classification*)
- e. 0.50-0.60 = Klasifikasi Salah (*Failure*)

### 3.3 ALAT BANTU PENELITIAN

Dalam melakukan penelitian peneliti juga memerlukan alat bantu. Alat bantu yang digunakan dalam melakukan penelitian ini yaitu sebagai berikut:

1. Perangkat Keras (*Hardware*)

Satu unit laptop Lenovo dengan spesifikasi:

Processor : AMD

RAM : 8 GB

SSD : 500 GB

2. Perangkat Lunak (*Software*)

- a. Sistem Operasi *Windows 11*

- b. *Microsoft Word 2021*

- c. *Microsoft Excel 2021*

- d. *Mendeley*

- e. *RapidMiner*

- f. *Notepad*

- g. *Adobe Reader*