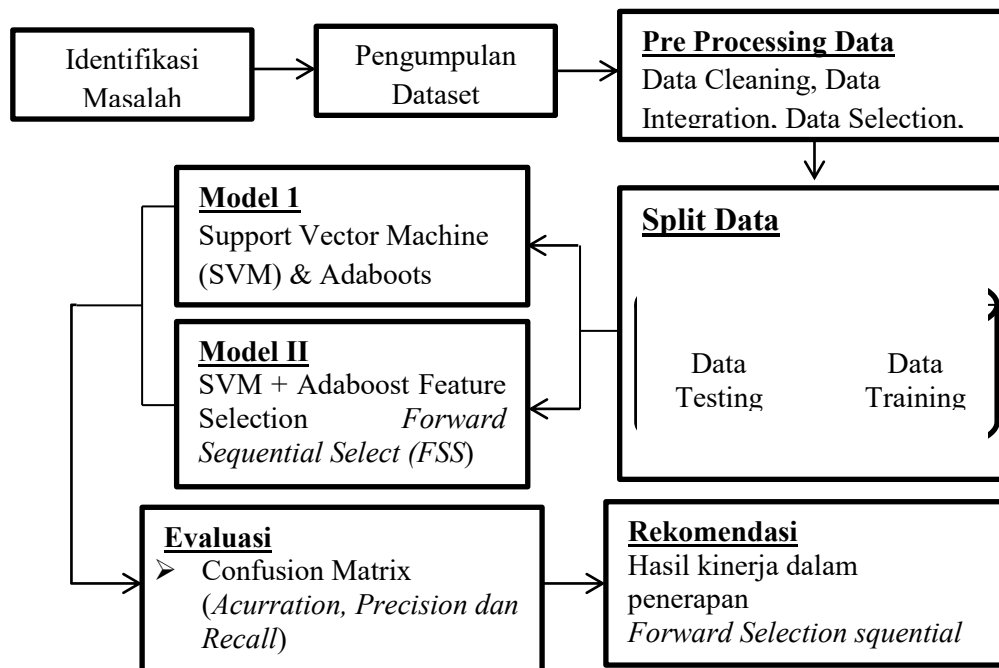


## BAB III METODOLOGI PENELITIAN

### 1.1. Metode Penelitian

Pada bab ini peneliti akan membahas terkait dengan langkah-langkah dari proses penelitian yang akan dilaksanakan. Dalam menganalisa dan mencari pola data untuk dijadikan sebuah database yang dapat memudahkan dalam penelitian sehingga dapat berjalan sesuai dengan sistematis dan memenuhi tujuan yang diinginkan maka dibuat alur dalam tahapan penelitian yang akan dilakukan sebagai berikut:



Gambar 2. Alur Dalam Tahap Penelitian

Berikut merupakan penjelasan terkait dengan gambar 2. yang merupakan alur dari penelitian ini :

1. Pertama, hal yang akan dilakukan adalah membuat suatu identifikasi masalah penyakit tiroid dan melakukan studi pustaka, sebagai penunjang dalam penelitian ini yang berkaitan dengan teknik data mining dengan metode algoritma *Support Vector Machine (SVM)*, Adaboost dengan feature selection

*Forward Sequential Select (FSS)* serta untuk menguatkan penelitian ini yang berlandaskan berdasarkan dengan teori yang akan digunakan.

2. Kedua pengumpulan dataset, data yang digunakan yang berasal dari dataset publik yang diperoleh dari [kaggle.com](https://www.kaggle.com/datasets/shubham8983/thyroid-dataset) dengan alamat <https://www.kaggle.com/datasets/shubham8983/thyroid-dataset>, memiliki 3773 data, 22 atribut dan 1 kelas yang memiliki nilai 1 dan 0. 22 atribut tersebut adalah *age, sex, on\_thyroxine, query\_on\_thyroxine, on\_antithyroid\_medication, sick, pregnant, thyroid\_surgery, I131\_treatment, query\_hypothyroid, query\_hyperthyroid, lithium, goitre, tumor, hypopituitary, psych, TSH, T3,TT4, T4U, FTI, referral\_source*.
3. Ketiga preprocessing data, Preprocessing data merujuk pada serangkaian langkah atau teknik yang dilakukan pada dataset sebelum dilakukan analisis atau pemodelan data. Tujuan dari preprocessing data adalah untuk membersihkan, mengubah format, atau mengorganisasi data agar sesuai dengan persyaratan analisis yang akan dilakukan. Preprocessing data juga bertujuan untuk meningkatkan kualitas data, mengurangi noise, atau menghilangkan outliers yang dapat mempengaruhi hasil analisis. Dalam preprocessing data terdapat empat tahapan yaitu :
  1. *Data clening* yang merupakan pendeteksian dan penanganan nilai yang hilang atau yang tidak valid dalam dataset.
  2. *Data integration* merupakan penggabungan data dari berbagai sumber yang menjadi satu dataset yang terpadu.
  3. *Data selection* merupakan pemilihan subset atau bagian dari dataset yang digunakan untuk analisis lebih lanjut.
  4. *Data transformation* merupakan perubahan format atau skala data agar sesuai dengan kebutuhan analisis.
4. Keempat, splid data mengacu pada proses membagi dataset menjadi dua atau lebih subset yang berbeda. Tujuan dari split data adalah untuk memiliki subset pelatihan dan subset pengujian (atau validasi) yang dapat digunakan dalam proses pembelajaran mesin atau evaluasi model.
5. Kelima penerapan algoritma, *support vector machine (SVM)* dan *Adaboost* untuk mengetahui nilai akurasi dari dua metode prediksi tersebut. Selanjutnya

penerapan *support vector machine* (SVM), Adaboost dengan feature selection *forward sequential select* (FSS).

6. Keenam, melakukan evaluasi dengan menggunakan confusion matrix (*accuracy, precision dan recall*) dan menggunakan kurva ROC
7. Ketujuh, rekomendasi hasil kinerja forward sequential select (FSS).

## **1.2. Alat dan Bahan**

### **1.2.1. Alat**

Pada penelitian ini alat yang digunakan sebagai berikut :

1. Laptop Acer
2. Prosesor : AMD E1-7010 APU with AMD Radeon R2 Graphics 1.5 GHz
3. Memory : 4,00 GB
4. System : 64-bit

### **1.2.2. Bahan**

Bahan yang digunakan dalam penelitian ini adalah perangkat lunak Microsoft Excel dan RapidMiner Studio versi 8.0 sebagai pengolah data.