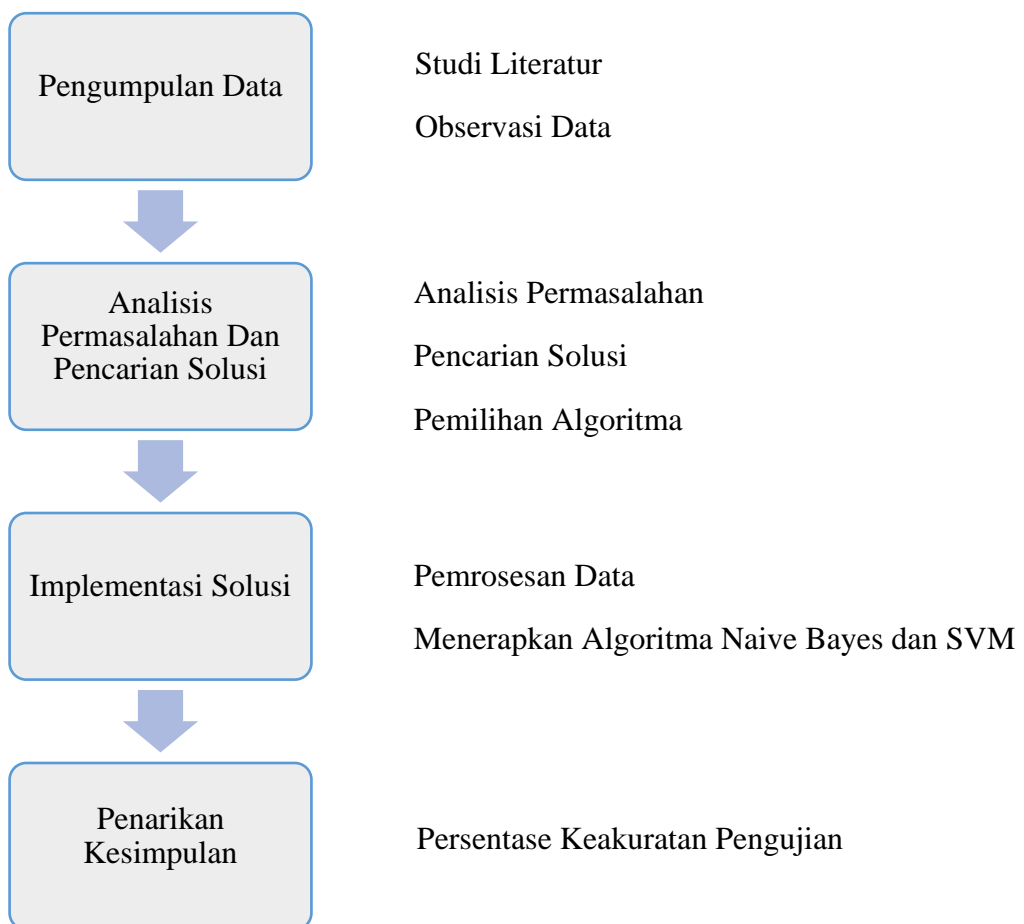


## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Tahapan Penelitian

Tahapan penelitian dimulai dari proses pengumpulan data, analisis permasalahan dan pencarian solusi, implementasi solusi sampai pada proses penarikan kesimpulan yang dijelaskan pada gambar 3.1



Gambar 3.1 Tahapan Penelitian

### 3.1.1 Pengumpulan Data

Data merupakan penunjang penelitian yang diperoleh melalui studi literatur dan observasi dengan pengamatan langsung. Studi literatur berisi uraian tentang teori, temuan dan bahan penelitian lainnya yang diperoleh dari jurnal nasional maupun jurnal internasional yang berupa *survey paper* dan *technical paper*. Observasi data dilakukan dengan mencari data yang tepat untuk melakukan penelitian.

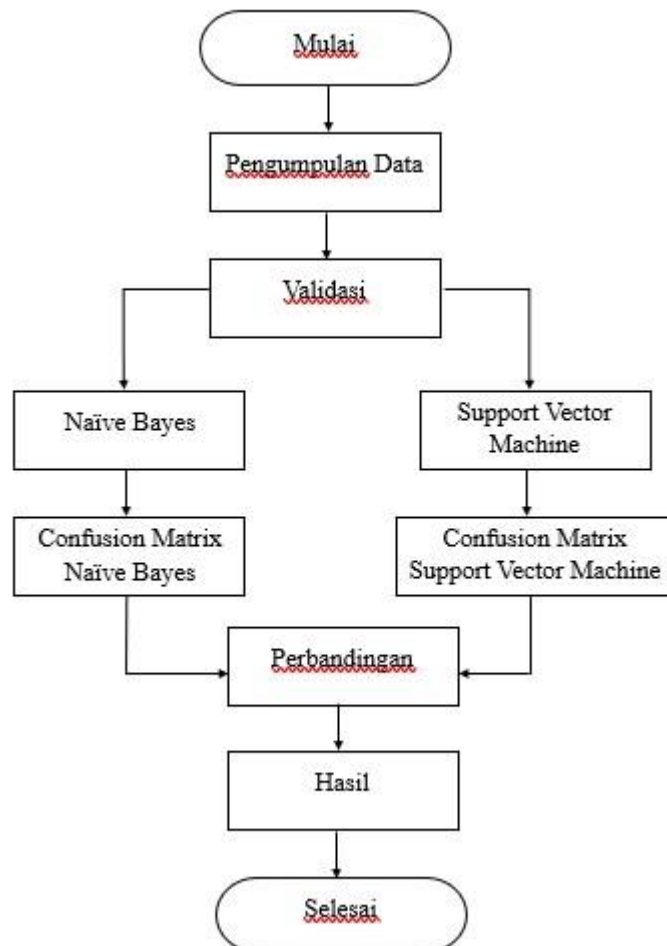
### 3.1.2 Analisis Permasalahan dan Pencarian Solusi

Analisis merupakan proses pemecahan persoalan yang dimulai dengan kegiatan memilah, mengurai, membedakan sesuatu untuk digolongkan dan dikelompokkan sesuai kriteria yang diinginkan. Proses yang dilakukannya yaitu membandingkan data menggunakan algoritma *Naive Bayes* terlebih dahulu kemudian dilanjutkan oleh algoritma *Support Vector Machine (SVM)*. Sehingga dikelompokkan sesuai persentase nilai dari algoritma *Naive Bayes* dan *Support Vector Machine (SVM)*.

### 3.1.3 Implementasi Solusi

Langkah yang dilakukan pada tahap implementasi solusi adalah mencari dataset yang diperoleh setelah melakukan observasi data. Dataset awal yang diperoleh sebanyak 100.000. Dataset yang telah didapatkan diolah menggunakan *tools* WEKA. Kemudian data yang sebelumnya bernilai numerik diubah terlebih dahulu menjadi nominal pada aplikasi WEKA, agar *class* datanya dapat terbaca oleh aplikasi. Data diproses dengan *Explorer* dan *Classify Rule*, kemudian proses selanjutnya adalah menerapkan algoritma *Naive Bayes* dan *Support Vector*

*Machine (SVM)* sebagai *classifier*. Hasil dari proses pengujian berupa item *summary*, *detailed accuracy by class* dan *confusion matrix*. Berikut adalah alur implementasi solusi dari penelitian ini.



Gambar 3.2 Implementasi Solusi

Gambar 3.2 merupakan proses menemukan solusi dimana pada tahapan tersebut meliputi pengumpulan data yang kemudian dilakukan validasi. Setelah dilakukan validasi diterapkan algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* dan dilakukan perhitungannya menggunakan *confusion matrix*. Selesai dilakukan perhitungan maka proses selanjutnya yaitu dilakukan perbandingan antara algoritma *Naïve Bayes* dan *Support Vector Machine (SVM)* yang kemudian

membuahkan hasil dari pengujian yang dilakukan. Tabel *summary* yang menggambarkan hasil dari proses pengujian dataset secara garis besar dijelaskan pada tabel 3.1.

Tabel 3.1 *Overall Summary*

<b>No.</b>	<b>Kategori</b>
1.	<i>Correctly Classified Instances</i>
2.	<i>Incorrectly Classified Instances</i>
3.	<i>Kappa Statistic</i>
4.	<i>Mean Absolute Error</i>
5.	<i>Root Mean Squared Error</i>
6.	<i>Relative Absolute Error</i>
7.	<i>Root Relative Squared Error</i>
8.	<i>Total Number Of Instances</i>

Tabel tingkat keakuratan pemrosesan data berdasarkan kelas yang digunakan pada tabel 3.2

Tabel 3.2 *Detailed Accuracy By Class*

<b>No.</b>	<b>Kategori</b>
1.	<i>TP Rate</i>
2.	<i>Fp Rate</i>
3.	<i>Precision</i>
4.	<i>Recall</i>
5.	<i>F-Measure</i>
6.	<i>MCC</i>
7.	<i>ROC Area</i>
8.	<i>PRC Area</i>

Tabel parameter pengujian merupakan proses data perbandingan berdasarkan parameter yang digunakan pada tabel 3.3

Tabel 3.3 Parameter Pengujian

No.	Parameter
1.	<i>Accuracy</i>
2.	<i>Precision</i>
3.	<i>Recall</i>

Tabel 3.4 merupakan contoh bentuk *Confusion Matrix* dari hasil pemrosesan data.

Tabel 3.4 *Confusion Matrix*

A	B	< - - <i>Classified as</i>
9	0 /	<i>a = positif</i>
1	4 /	<i>b = negative</i>

### 3.1.4 Penarikan Kesimpulan

Kesimpulan adalah proses akhir dari analisis. Pengelompokkan objek yang didapat dari hasil analisis merupakan hasil akhir dari penelitian. Hasil yang diperoleh adalah nilai keakuratan dari hasil uji coba penerapan algoritma *Naive Bayes* dan *Support Vector Machine (SVM)* terhadap *diabetes prediction dataset*. Yang kemudian disimpan sesuai kriteria masing-masing analisis.