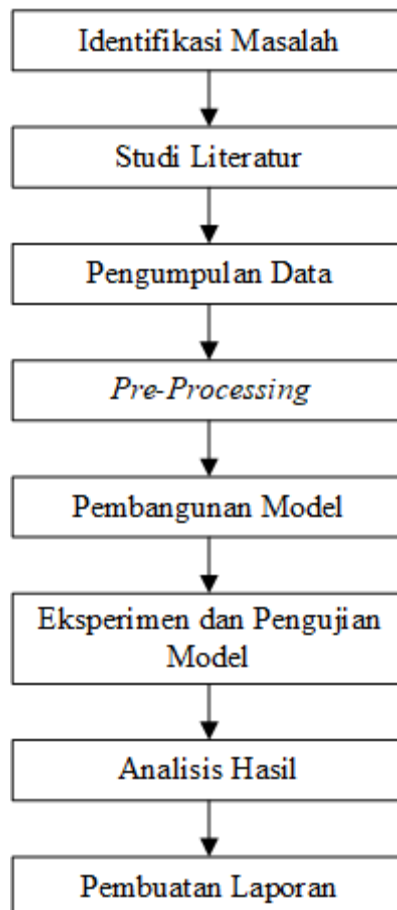


BAB III

METODOLOGI PENELITIAN

3.1 KERANGKA KERJA PENELITIAN

Kerangka kerja penelitian merupakan tahapan atau proses yang akan dilakukan peneliti dalam menyelesaikan masalah penelitian. Kerangka kerja penelitian dibuat agar penelitian dapat diselesaikan sesuai dengan waktu yang telah ditentukan dan dapat berjalan sesuai dengan yang direncanakan. Adapun kerangka kerja penelitian yang digunakan dapat dilihat pada gambar 3.1 berikut ini :



Gambar 3. 1 Kerangka Kerja Penelitian

3.1.1 Identifikasi Masalah

Dalam tahap ini peneliti melakukan identifikasi permasalahan yaitu melihat data penderita penyakit stroke. Sehingga peneliti dapat menentukan rencana kerja serta menentukan data apa saja yang dibutuhkan dalam penelitian ini.

3.1.2 Studi Literatur

Pada tahapan ini peneliti melakukan kajian pustaka, yaitu mempelajari buku-buku referensi, artikel-artikel, dan hasil penelitian sejenis yang relevan dengan permasalahan yang sedang diteliti. Studi literatur ini bertujuan untuk mendapatkan landasan teoritis mengenai permasalahan yang akan diteliti seperti teknik data mining, proses klasifikasi menggunakan algoritma naive bayes, seleksi fitur Greedy Forward Selection, tool data mining yang akan digunakan, data penderita penyakit stroke. Hal ini bertujuan agar peneliti dapat memahami permasalahan yang diteliti dengan benar dan sesuai dengan pembahasan yang dilakukan.

3.1.3 Pengumpulan Data

Pada penelitian ini, peneliti menggunakan data sekunder. Pengertian data sekunder adalah sumber data penelitian yang diperoleh melalui media perantara atau secara tidak langsung yang berupa buku, catatan, bukti yang telah ada atau arsip yang telah dipublikasikan secara umum.

Kelebihan dari data sekunder adalah waktu dan biaya yang dibutuhkan untuk mengklasifikasi permasalahan dengan mengevaluasi data, relative lebih sedikit dibandingkan dengan pengumpulan data primer, sedangkan kekurangan dari data sekunder adalah jika sumber data terjadi kesalahan, kadaluarsa atau sudah tidak relevan dapat mempengaruhi hasil penelitian.

Pada tahap ini penulis melakukan pengumpulan data yang dibutuhkan dan sesuai dengan topik penelitian, agar data tersebut dapat dianalisis. Data didapatkan dengan mengunduh dataset penyakit stroke pada website Kaggle, dataset ini dapat diunduh di <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

3.1.4 Pre-Processing

Pada tahap ini peneliti melakukan proses persiapan dataset final atau dataset yang nantinya akan dijadikan inputan kedalam modelling data mining.

a. *Data Selection*

Pemilihan (seleksi) data baru sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam data mining dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

b. *Cleaning Data*

Proses menghilangkan data noise (data yang tidak relevan/berhubungan langsung dengan tujuan akhir proses data mining) dan tidak konsisten. Data yang tidak berkualitas akan menghasilkan data mining yang tidak berkualitas. Keputusan yang berkualitas harus didasarkan pada data yang berkualitas (data ganda atau data yang telah hilang dapat menyebabkan ketidak benaran).

c. *Transformation*

Transformation merupakan proses mentransformasikan data ke dalam bentuk yang lebih sesuai untuk di mining. Kualitas dari hasil data mining dapat ditentukan dari transformasi dan pemilihan data. Algoritma Naive Bayes Classifier mampu melakukan proses data yang bernilai nominal,

ordinal, atau kontinu. Sehingga nilai dari tiap-tiap atribut yang terdapat pada dataset tidak perlu ditransformasikan.

d. *Reduction*

Data *reduction* merupakan proses yang dilakukan dengan mereduksi suatu data set yaitu mengurangi jumlah atribut dan record yang ada sehingga lebih sedikit tetapi masih tetap informatif. Proses ini dilakukan untuk mendapatkan hasil representasi pada volume data yang sebelumnya sudah berkurang namun tetap menghasilkan analisis yang sama.

3.1.5 Pembangunan Model

Pembangunan model yang diusulkan dalam penelitian ini yaitu dengan menggunakan algoritma klasifikasi Naïve Bayes dan seleksi fitur Greedy Forward Selection, seleksi fitur digunakan untuk menyeleksi fitur yang relevan sehingga dapat meningkatkan kinerja dari model yang digunakan. Data yang digunakan sudah melalui proses cleaning data, untuk melakukan pengukuran dalam penelitian ini menggunakan tool *RapidMiner*.

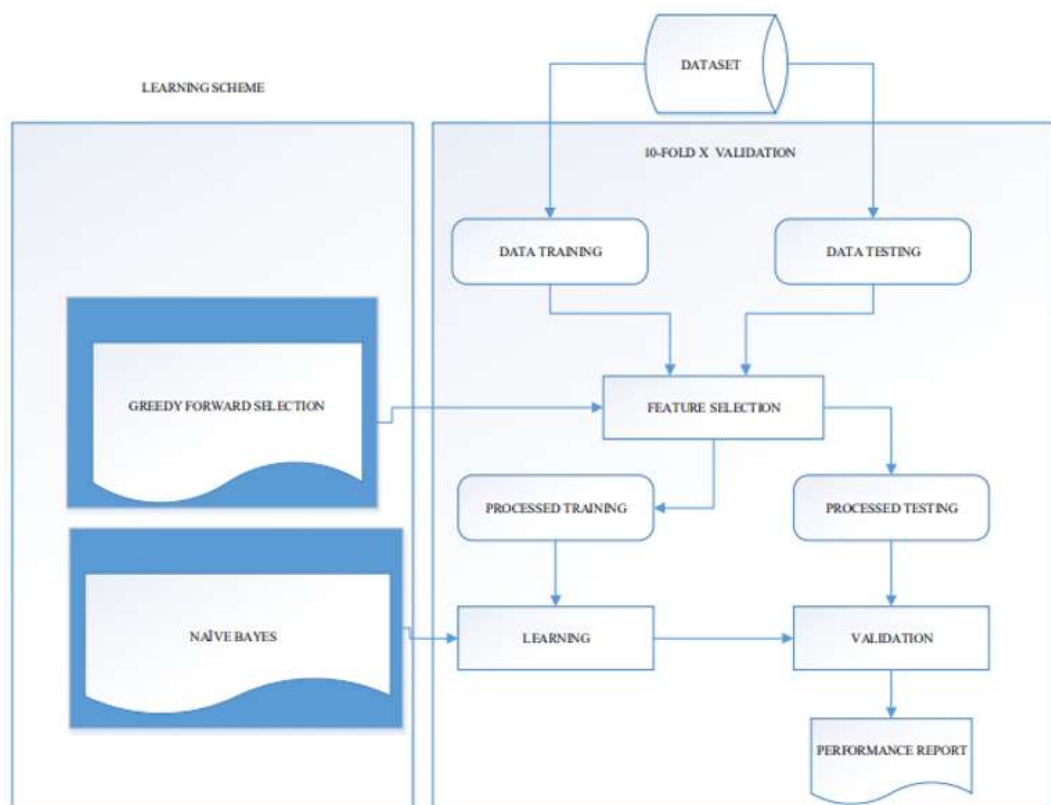
3.1.6 Eksperimen dan Pengujian Model

Untuk mempermudah penelitian, maka diperlukan alur eksperimen yang berisi tahapan – tahapan pengerjaan penelitian. Alur eksperimen pada penelitian ini dapat dilihat pada gambar 3.2

Eksperimen dan pengujian dilakukan dengan beberapa tahapan sebagai berikut:

1. Menyiapkan dataset untuk eksperimen.
2. Mendesain arsitektur Naïve Bayes.

3. Melakukan training dan testing terhadap model Naïve Bayes dan mencatat hasil kinerja dari model.
4. Mendesain arsitektur Naïve Bayes dan Greedy Forward Selection.
5. Melakukan training dan testing terhadap model Naïve Bayes dan Greedy Forward Selection, kemudian mencatat hasil kinerja dari model.



Gambar 3. 2 Desain Eksperimen

3.1.7 Analisa Hasil

Pada saat eksperimen dan pengujian selesai akan keluar hasil dari kinerja model yang diusulkan yaitu nilai akurasi sebagai ukuran seberapa bagus model yang digunakan.

Pada tahap ini, penulis melakukan analisa terhadap hasil yang telah didapatkan. Setelah melakukan analisis maka didapatkan hasil yang berupa solusi dari permasalahan yang diteliti. Hasil dari penelitian tersebut akan dibahas pada tahap ini. Penulis akan mengambil kesimpulan sebagai hasil akhir dari penelitian ini.

3.1.8 Pembuatan Laporan

Tahap ini merupakan pembuatan laporan yang disusun berdasarkan hasil penelitian. Setelah semua tahapan penelitian dilakukan, maka akan dibuat laporan sebagai dokumentasi penelitian agar dapat dimanfaatkan pada waktu yang akan datang. Baik oleh peneliti sendiri maupun peneliti lainnya.

3.2 ALAT BANTU PENELITIAN

Dengan melakukan penelitian, penulis juga menggunakan alat bantu. Alat bantu yang digunakan, yaitu:

1. Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan peneliti sebagai pengembang adalah sebagai berikut :

- a. Laptop : VivoBook_ASUSLaptop
- b. *Processor* : AMD Ryzen 3 3250U with Radeon Graphics,
2.60 Ghz
- c. RAM : 8 GB
- d. *Harddisk* : 1 TB HDD
- e. *Flashdisk* : SanDisk 8 GB
- f. *Mouse* : FANTECH CRYPTO VX7

2. Perangkat Lunak (*Software*)

Perangkat lunak yang digunakan oleh pengembang dalam penelitian yaitu :

- a. *Windows 11 Home Single Language*
- b. *Microsoft Office 2016*
- c. *Microsoft Visio*
- d. *Google Chrome*
- e. *Tools RapidMiner*
- f. *Mendeley Dekstop*
- g. *Notepad*