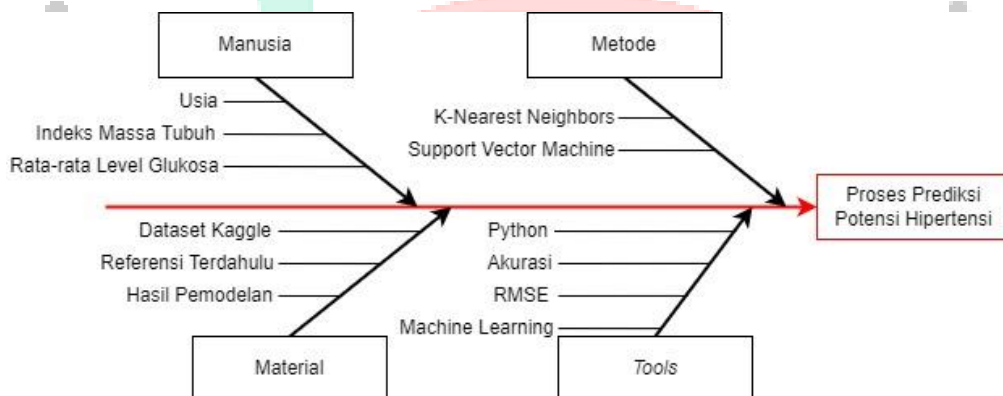


## BAB III METODE PENELITIAN

Metode penelitian diperlukan sebagai acuan untuk teknik pengembangan referensi dan realisasi pola pikir peneliti terhadap objek yang diteliti. Pada bab ini, peneliti menjelaskan secara kronologis antara latar belakang masalah, rumusan masalah, tujuan penelitian, metode untuk memecahkan masalah, proses pengumpulan data, serta hasil penelitian yang diharapkan. Selanjutnya, peneliti menjelaskan mengenai tahapan terkait pendekatan yang digunakan sebagai penelitian. Berikut ini merupakan uraian dari metode penelitian yang dipaparkan pada sub-bab di bawah ini.

### 3.1 Paradigma Penelitian



Gambar 3. 1 Diagram Tulang Ikan Paradigma Penelitian

Berdasarkan diagram di atas, terdapat beberapa faktor yang menjadi sumber tercipta sebuah akibat yaitu proses prediksi potensi hipertensi. Manusia, Metode, Material, dan Tools adalah penyebab utama yang membuat sebuah proses prediksi potensi hipertensi terjadi. Manusia dengan didasarkan pada usia, dapat menentukan riwayat hipertensi atau tidaknya. Indeks massa tubuh juga menjadi faktor yang dimiliki oleh manusia memiliki penyakit hipertensi. Selain itu, angka rata-rata level glukosa juga dapat menjadi indikasi manusia mengalami hipertensi. Sehingga, indicator-indikator yang dimiliki manusia tersebut menjadi sebab dalam proses prediksi potensi hipertensi.

Metode yang dapat digunakan dalam proses penelitian ini adalah *K-Nearest Neighbors* dan *Support Vector Machine*. Algoritma klasifikasi tersebut sering digunakan sebagai metode dalam prediksi sebuah model data. Material atau bahan yang didapat untuk proses prediksi data didapat berdasarkan beberapa pengumpulan data. Pengumpulan data dilakukan dalam 3 hal, studi literasi penelitian terdahulu, kuantitasi dataset dari Kaggle.com, dan hasil pemodelan data klasifikasi. *Tools* atau perangkat-perangkat yang digunakan untuk menghasilkan proses prediksi potensi hipertensi adalah dengan menggunakan algoritma pemrograman. Perangkat-perangkat tersebut antara lain bahasa pemrograman *Python*, akurasi, *RMSE*, dan *machine learning*.

Sistem yang sedang peneliti uji bertujuan untuk membuktikan bahwa algoritma KNN dan SVM dapat menyelesaikan masalah demikian. Maka dari itu, penerapan algoritma *K-Nearest Neighbors* dan *Support Vector Machine* sebagai sistem untuk memprediksi potensi hipertensi menggunakan bahasa Python ini diharapkan dapat menjadi motivasi perancangan sebuah sistem fasilitas medis untuk mencegah munculnya penyakit yang lebih berat diderita oleh pasien.

### **3.2 Metode Penelitian**

Metodologi yang digunakan dalam penelitian ini merupakan metode kualitatif dengan acuan pendekatan menggunakan studi literasi. Metode kualitatif merupakan pendekatan yang dilakukan dalam penelitian dengan berorientasi pada fenomena alami di mana pendekatan dilakukan melalui studi observasi dan mengutamakan masalah proses dan persepsi (Batubara, 2017, p. 97). Peneliti melakukan observasi dari beberapa literasi untuk memenuhi segala pengetahuan selama masa penelitian. Literasi-literasi ini digunakan untuk memperkuat fakta dan data terkait masalah dalam penelitian. Fakta dan data ini didasarkan pada proses dan persepsi yang peneliti observasi.

Studi literasi merupakan salah satu jenis pendekatan menggunakan metode kualitatif. Pendekatan ini melakukan pengumpulan data pustaka melalui artikel, jurnal, buku, dan lain-lain untuk dibaca dan dicatat lalu diolah menjadi bahan penelitian (Zed, 2014, p. 3). Pada penelitian ini, peneliti menggunakan beberapa

literasi terkait penggunaan algoritma KNN dalam penerapannya di bidang kesehatan.

### **3.3 Metode Pengumpulan Data**

Pengumpulan data dilakukan melalui tiga tahapan penelitian, pertama, kedua, dan ketiga. Tahapan tersebut dipaparkan dalam sub-bab di bawah ini.

#### **3.3.1 Pengumpulan Data Pertama : Studi Literatur**

Pada tahap ini, dilakukan penelitian menggunakan metode kualitatif yaitu studi literasi. Digunakan studi literasi guna memperkuat pemikiran dan pendapat peneliti terkait topik yang sedang dibahas. Peneliti memberikan fakta dan data dari hasil penelusuran pada platform Google Scholar. Studi literasi yang didapatkan berupa jurnal, artikel, dan buku. Pada pengumpulan data pertama ini, data yang didapatkan merupakan data primer karena merupakan hasil ringkasan pokok penelitian terdahulu yang dilakukan oleh peneliti.

#### **3.3.2 Pengumpulan Data Kedua : Pengumpulan *Dataset***

Pada tahapan pengumpulan data kedua, peneliti melakukan studi menggunakan metode kuantitatif di mana data yang peneliti dapatkan merupakan data sekunder karena didapatkan dari *dataset* Kaggle.com yang diunggah oleh Fedesoriano. Dataset ini berupa file berekstensi *comma sparated values (.csv)*. Dataset ini berisi 5110 baris data yang berisi kolom pasien dengan 12 fitur, antara lain : id pasien, jenis kelamin, umur, hipertensi, penyakit jantung, sudah menikah, jenis pekerjaan, jenis tempat tinggal, rata2 level glukosa, *BMI*, merokok, dan stroke. Pada pengumpulan data tahap ini, hanya dibutuhkan pelabelan data berupa level glukosa, *BMI/IMT*, dan usia. Asal pelabelan ini tidak berdasarkan pada konsultasi melalui dokter, akan tetapi merupakan hasil studi literasi yang peneliti lakukan, dan telah dipaparkan pada sub-sub-bab 2.2.1.

Kekurangan pada dataset ini adalah bahwa data tersebut tidak dapat dipastikan kecocokan antara data tersebut dengan data pasien sebenarnya di

lapangan. Namun demikian, dataset tersebut sudah digunakan dalam beberapa penelitian terkait prediksi penyakit stroke dan sudah dipublikasikan secara internasional pada penerbit yang bereputasi (Rana, et al., 2021). Dengan demikian, dataset tersebut cukup merepresentasikan kasus hipertensi di bidang kedokteran.

### **3.3.3 Pengumpulan Data Ketiga : Pembangunan dan Analisa Model**

Dalam pengumpulan data ketiga dilakukan pembangunan machine learning yang ditekankan pada model *K-Nearest Neighbors* dan *Support Vector Machine*. Kedua algoritma tersebut dibangun dalam bahasa pemrograman Python untuk otomatisasi dan memudahkan dalam pengumpulan data. Machine learning dibangun untuk menganalisis data dengan menampilkan perbandingan hasil prediksi algoritma mana yang paling akurat dari kedua model klasifikasi terhadap pembagian data latih dan data uji. Pada tahapan ini, data bersifat primer karena luaran yang dihasilkan merupakan perhitungan mandiri peneliti dengan referensi ambang batas minimum hasil prediksi yang didasarkan pada referensi penelitian terdahulu.

### **3.4 Metode Analisis Data**

Metode analisis data yang digunakan adalah hasil sampel file berekstensi *comma sparated values (.csv)* berupa data pasien hipertensi sebanyak 5110 baris data. Data tersebut diolah ke dalam Bahasa pemrograman *Python* dengan menerapkan *Library-library* yang diperlukan untuk pengujian. Setelah itu, digunakan metode analisis algoritma *machine learning* yakni KNN dan SVM. Hasil uji prediksi *machine learning* yang akan dilakukan didasarkan pada pengujian RMSE agar meminimalisasi error yang dihasilkan serta meningkatkan hasil akurasi yang tinggi.

### **3.5 Metode Pengujian**

Dilakukan beberapa metode dalam menghasilkan akurasi terbaik untuk pemodelan data pasien hipertensi. Akurasi yang tinggi mempengaruhi hasil ketepatan prediksi. Dengan demikian, dilakukan pengujian terhadap beberapa pembagian data latih dan data uji serta fitur-fitur yang tersedia untuk dicari mana

yang memiliki tingkat error terendah menggunakan RMSE. Pembagian data terendah dengan fitur terbaik akan digunakan dalam pemodelan data pasien hipertensi tersebut, sehingga dapat menghasilkan akurasi prediksi terbaik.

