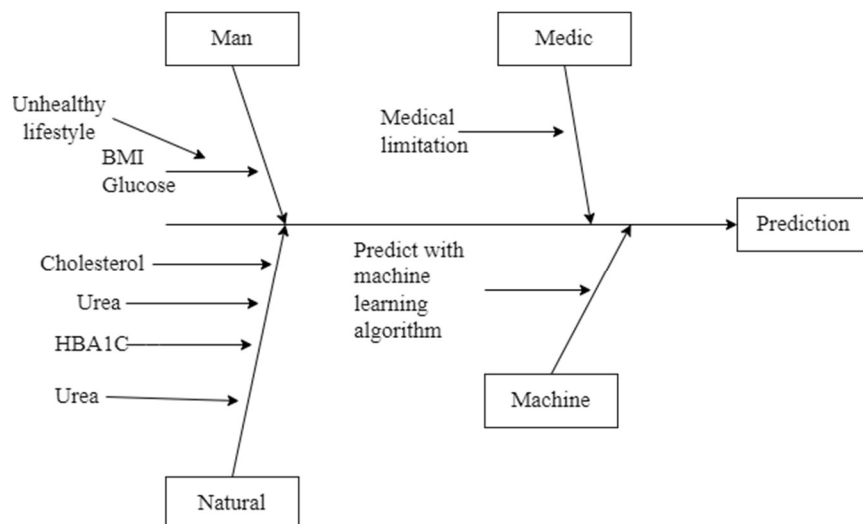


## BAB III METODE PENELITIAN

Metode penelitian diperlukan sebagai acuan untuk teknik pengembangan referensi dan realisasi pola pikir terhadap objek yang diteliti. Pada bab ini, dijelaskan secara kronologis antara latar belakang masalah, rumusan masalah, tujuan penelitian, metode untuk memecahkan masalah, proses pengumpulan data, serta hasil penelitian yang diharapkan. Selanjutnya, dijelaskan mengenai tahapan terkait pendekatan yang digunakan sebagai penelitian.

### 3.1 Paradigma Penelitian

Paradigma penelitian merupakan sebuah pola pikir atau cara pandang terhadap keseluruhan proses serta hasil dari penelitian yang dilakukan (Muslim, 2015). Penelitian ini menggunakan *fishbone diagram* dalam mendapatkan sebuah paradigma penelitian. Diagram tulang ikan (*fishbone diagram*) merupakan sebuah metode dalam analisis faktor-faktor yang berpengaruh dalam keberhasilan terhadap penelitian. Masalah yang terjadi dilambangkan sebagai kepala ikan dan penyebab masalah dilambangkan sebagai tulang-tulang ikan yang kemudian dihubungkan dengan kepala ikan untuk mendapatkan ringkasan kemudahan dalam pembahasan (Jayusman, 2018).



*Gambar 3.1* Diagram Tulang Ikan Paradigma Penelitian

Berdasarkan diagram tulang di atas, terdapat faktor-faktor yang menjadi sumber terjadinya permasalahan yang diteliti. Faktor manusia (*man*), medis (*medic*), faktor alami (*natural*), dan mesin menjadi penyebab yang terhadap penelitian yang dilakukan dalam mendapatkan hasil dari prediksi penyakit diabetes. Faktor manusia dimana kurang menjalani pola hidup sehat sehingga meningkatkan risiko terjangkit penyakit diabetes. Sementara itu, faktor alami yang menjadi penyebab utama penyakit diabetes beberapa diantaranya yaitu, tekanan darah, gula darah, juga umur yang semakin tua semakin rentan menderita penyakit diabetes. Kemudian faktor medis dimana masih dilakukan secara manual dengan pergi ke rumah sakit sehingga manusia tanpa sadar sudah terjangkit sebelumnya. Teknologi juga menjadi faktor penyebab terhadap masalah dimana bantuan dari penggunaan teknologi sangat diperlukan untuk meningkatkan kesadaran dalam menjaga pola hidup sehat.

Sistem yang diuji bertujuan untuk membuktikan bahwa algoritma KNN dapat menyelesaikan masalah demikian. Maka dari itu, penerapan algoritma *K-Nearest Neighbors* sebagai sistem untuk memprediksi penyakit 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**

Metode pada penelitian ini menggunakan metode kualitatif dengan pendekatan menggunakan studi literasi. Metode kualitatif merupakan suatu proses penelitian dan pemahaman berdasarkan pada metodologi untuk menyelidiki suatu fenomena sosial dan masalah manusia Pada penelitian ini dilakukan observasi dari beberapa sumber literasi sebagai pemenuh pengetahuan selama penelitian dilakukan. Sumber-sumber literasi kemudian peneliti gunakan untuk memperkuat fakta dan data terkait masalah dalam penelitian.

Studi literatur merupakan cara yang dipakai dalam menghimpun data-data atau sumber-sumber yang berkaitan dengan topik penelitian (Habsy, 2017). Pada penelitian ini mengambil data dari berbagai sumber terkait dengan algoritma KNN dan *Random Forest* serta pembahasan terkait penyakit diabetes melitus.

### 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

Tahap pertama dalam pengumpulan data dengan menggunakan metode kualitatif yaitu studi literasi. Data yang dikumpulkan didapatkan dari berbagai data yang berdasarkan terhadap hasil penelusuran baik di platform Google Scholar maupun halaman *website* lainnya berupa jurnal, artikel, dan buku. Pengumpulan data tahap pertama termasuk kepada data sekunder, dimana data yang didapatkan berasal dari sumber-sumber yang sudah ada.

#### 3.3.2 Pengumpulan Data Kedua

Tahap kedua dalam pengumpulan data yaitu didapatkan data berupa *dataset* dari data.mendeley.com dengan nama *dataset* yaitu “*Diabetes Dataset*” yang diterbitkan oleh Ahlam Rashid berdasarkan data asli yang berasal dari laboratorium *Medical City Hospital and the Specializes Center for Endocrinology and Diabetes-AI-Kindy Teaching Hospital* dengan format file *comma sparated values (.csv)*. *Dataset* ini berisi 1000 *instances* dengan beberapa atribut kelas antara lain yaitu “*No. of Patient, Sugar Level Blood, Age, Gender, Creatinine ratio(Cr), Body Mass Index (BMI), Urea, Cholesterol (Chol), HBA1C, Class*”. *Dataset* ini digunakan pada penelitian terhadap sistem diagnosa diabetes menggunakan *machine learning* dengan metode *big data analytic* pada tahun 2021 oleh Usman Ahmed dan Chunxiao Li di *School of Information Engineering, Yangzhou University, China* (Ahmed & Li, 2021).

Pada penelitian ini, digunakan beberapa atribut pada *dataset* sebagai fitur yang digunakan dalam pemodelan untuk mendapatkan target prediksi berupa *Class* yang berupa apakah pasien *diabetic, non-diabetic*, atau *predict-diabetic* serta tingkat akurasi dari algoritma yang digunakan.

#### 3.3.3 Pengumpulan Data Ketiga

Tahapan ketiga dalam pengumpulan data setelah dilakukannya pengumpulan data pertama dan kedua, dilakukan pengujian untuk menganalisa algoritma *K-Nearest Neighbors* dan *Random Forest*. Kedua algoritma tersebut nantinya akan

menghasilkan tingkat akurasi antara data *training* juga data *testing* dan dibandingkan satu sama lain. Pada tahap ini bersifat data primer karena luaran yang didapatkan menggunakan parameter evaluasi akurasi serta *confusion matrix* dengan beberapa pengukuran setelah dilakukan pemodelan *machine learning*.

### **3.4 Metode Pengujian**

Pengujian dilakukan dengan cara melakukan evaluasi terhadap skor dari masing-masing algoritma menggunakan *confusion matrix*. *Confusion matrix* digunakan dalam memberikan informasi terhadap perbandingan dari hasil klasifikasi yang dilakukan model dengan hasil klasifikasi yang sebenarnya. Terdapat tiga pengukuran yaitu berdasarkan akurasi, presisi, dan *recall*. Selain itu dilakukan pengujian terhadap beberapa data uji yang didapatkan setelah melakukan *splitting data*.