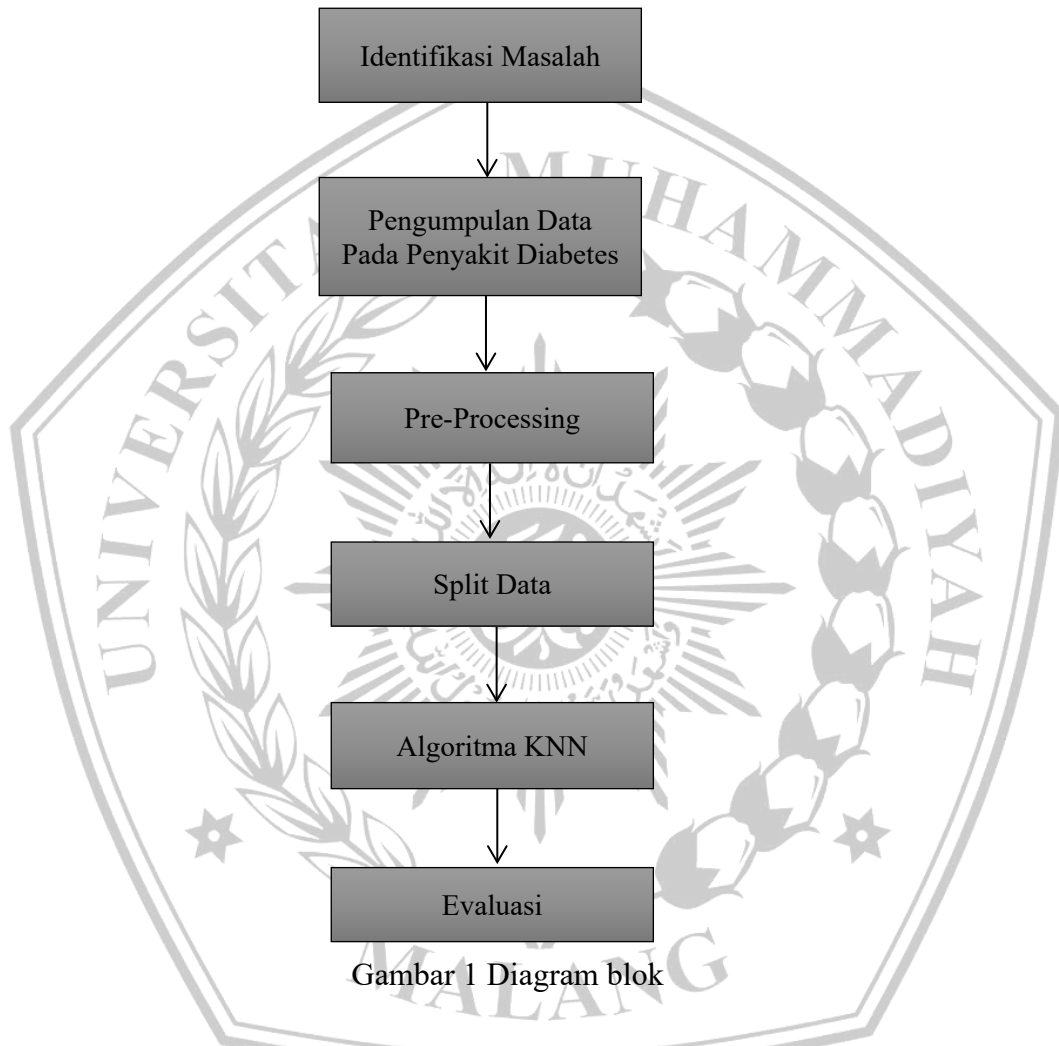


BAB III

METODE PENELITIAN

3.1 METODE

Dalam rangka menjawab permasalahan yang telah dirumuskan diatas, maka diusulkan metode yang secara umum digambarkan dengan blok diagram pada Gambar 1 berikut.



Dari Gambar 1, Dalam tahap identifikasi masalah dilakukan identifikasi masalah dengan mencari permasalahan yang ada pada objek penelitian. Tahap berikutnya adalah proses pengumpulan data berupa history dari diagnosis penyakit diabetes yang dikumpulkan dari data hasil pengukuran atau data real. Pada tahap pre-processing dilakukan pengolahan terhadap kumpulan history diagnosis penyakit diabetes dari data hasil pengukuran atau data real sebelum dilakukan klasifikasi. Proses selanjutnya adalah proses split data ini

mempunyai dua bagian yaitu training dan testing. Pada proses training sendiri ialah tahap algoritma KNN dilakukan, tahap ini juga merupakan tahap seleksi yang dimana pada tahap ini individu-individu yang memiliki kinerja terbaik dipilih untuk melanjutkan kegenerasi berikutnya, sementara individu-individu yang kurang baik mungkin ditinggalkan. Setelah dari proses split data dilakukanlah tahap evaluasi. Pada tahap evaluasi inilah yang dilihat data yang tidak memiliki nilai atau data yang hilang, sehingga diperlukan dilakukannya imputasi menggunakan metode KNN untuk mencari data atau nilai yang hilang tersebut.

Setiap individu dalam populasi dinilai berdasarkan seberapa baik mereka memenuhi kriteria tertentu terkait dengan penyakit diabetes, misalnya, tingkat gula darah, tekanan darah, atau indeks massa tubuh. Metode klasifikasi dilakukan menggunakan algoritma KNN. Pembobotan untuk KNN dilakukan dengan konfigurasi nilai K. Proses perhitungan klasifikasi.

3.2 SUMBER DATA (METODE PENGUMPULAN DATA)

Sumber data dalam penelitian ini menggunakan Continuous Glucose Monitoring (CGM) sebagai mesin pembaca dan menyediakan data kadar glukosa darah. Sedangkan metode pengumpulan data adalah dari histori data (CGM) yang melakukan pengecekan kadar glukosa darah setiap 10 menit menghasilkan 1 data dan selama 1 jam memperoleh data 6. Kemudian dalam waktu 1 hari menghasilkan 144 data. Dalam waktu 3 hari akan menghasilkan data 432 data dan yang digunakan sebagai populasi dan sampel data yang dianalisis dalam penelitian ini.

Penelitian dimulai dari penelitian awal berupa pengumpulan data sampai kepada pasien dan menggunakan KNN (Google Colab) untuk menganalisis data sehingga ditemukan kecocokan antara hasil akhir terhadap masing-masing perhitungan. Output dari penelitian ini adalah Classification terhadap hasil pengecek glukosa darah pada pasien pengidap penyakit diabetes. Pada penelitian ini, penulis menemukan beberapa keadaan yang membuat peneliti tertarik untuk melakukan suatu penganalisaan terhadap kemungkinan – kemungkinan yang akan membuat seseorang terkena penyakit diabetes dan peningkatan gula darah pada pasien. Seperti yang peneliti ketahui, penyakit diabetes tidak hanya dapat dijumpai

pada orang yang memiliki keturunan penyakit diabetes, namun juga berpotensi pada seseorang yang memiliki gaya hidup tidak sehat serta beberapa faktor-faktor pemicu lainnya.

3.3 PROCESSING DATA, ALGORITMA KNN DAN EVALUASI DATA

```
[ ] import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report
from sklearn.preprocessing import MinMaxScaler
import seaborn as sns

[ ] diabetes_data = pd.read_excel ('/content/IIn-lowokwaru.xls')

[ ] diabetes_data.head()

[ ] diabetes_data.describe()

[ ] diabetes_data.info(verbose=True)

[ ] diabetes_data.columns.values

▶ diabetes_data.isnull().sum()

[ ] diabetes_data.shape

▶ import missingno as msno
p=msno.bar(diabetes_data)

[ ] plt.figure(figsize = (10, 10))
plotnumber = 1

for column in diabetes_data:
    if plotnumber <= 20:
        ax = plt.subplot(3, 3, plotnumber)
        sns.distplot(diabetes_data[column])
        plt.xlabel(column, fontsize = 10)

        plotnumber += 1
    plt.show()

[ ] diabetes_data.hist(figsize = (10,10))

▶ diabetes_data.plot(kind='box', subplots=True, layout=(3,3), sharex=False, sharey=False, figsize=(10,10))
plt.show()

[ ] plt.plot([0,1],[0,1], 'k--')

▶ plt.plot([0,1],[0,1], 'k--')
plt.plot(label='knn')
plt.xlabel('Unnamed:')
plt.ylabel('Unnamed:')
plt.title('Knn(n_neighbors=11) ROC curve')
plt.show()

▶ def closest(lst, K):

    return lst[min(range(len(lst)), key = lambda i: abs(lst[i]-K))]

# Driver code
lst = [3., 5., 7., 9., 15.,]
K = 3.
print(closest(lst, K))

3.0
```

Gambar 3.2 Processing Data

Library yang digunakan program ini adalah numpy, pandas dan sklearn. pd.Read_excel untuk penginputan sebuah data berformat excel lalu sebagai pembaca head() dataframe value berfungsi menjadikan model data. model

digunakan ialah insull() sum () sebagai klasifikasi KNN imputer akurasi dan mengisi missing value pada data.

Tahapan ini sangat perlu dilakukan untuk mengetahui langka-langkah dalam menganalisis data. Adapun tahapannya sebagai berikut:

- a. Pengklasifikasian data – data pasien penderita diabetes. Setelah data–data dikumpulkan, maka dilanjutkan dengan proses mengelompokkan data ke dalam MS.Excel.
- b. Algoritma ialah metode non-parametrik, yang dapat digunakan untuk melaksanakan klasifikasi pada suatu objek, prinsip dari KNN adalah jika memiliki sekumpulan data sebagai training. Metode ini mudah dipahami bila dibandingkan dengan metode lain mengklasifikasikan berdasarkan jarak terdekat dengan objek yang lain.
- c. Analisis dan melakukan pengolahan data dengan menggunakan Google Colab.
- d. Setelah data diklasifikasikan ke dalam Google Colab, maka data tersebut kemudian akan evaluasi. Dengan menggunakan google colab tersebut maka data akan dapat dianalisa untuk mendapatkan sebuah keputusan.

