ARTIFICIAl INTELLIGENCE

TOPIC:DIABETES PREDICTION SYSTEM USING AI

-PROBLEM DEFINITION:

Diabetes is a medical disorder that impacts how well our body uses food as fuel. Most food we eat daily is converted to sugar, commonly known as glucose, and then discharged into the bloodstream

Diabetes can cause blood sugar levels to rise if it is not continuously and carefully managed, which raises the chance of severe side effects likeheart attack and stroke.

DESIGN THINKING:

It helps to make prediction in medical field.This project propose an effective method tofor earlier

detection of the diabetes disease using machine learning

STEPS:

1. stalling the Libraries
2. Importing the Dataset
3. Filling the Missing Values
4. Exploratory Data Analysis
5. Feature Engineering
6. Implementing Machine Learning Models
7. Predicting Unseen Data
8. Concluding the Report

CODING:

```python
# Import libraries
import numpy as np # for linear algebra
import pandas as pd # for data processing, CSV file I/O (e.g. pd.read_csv)
import seaborn as sns # for data visualization
```

```python
5. import matplotlib.pyplot as plt # to plot data visuali
   zation charts
6. from collections import Counter

7. import os
8.
9. # Modeling Libraries
10. from sklearn.metrics import confusion_matrix, a
    ccuracy_score, precision_score
11. from sklearn.preprocessing import QuantileTran
    sformer
12. from sklearn.linear_model import LogisticRegre
    ssion
13. from sklearn.neighbors import KNeighborsClass
    ifier
14. from sklearn.tree import DecisionTreeClassifier
15. from sklearn.ensemble import RandomForestCl
    assifier, AdaBoostClassifier, GradientBoostingClassi
    fier
```

```python
16. from sklearn.model_selection import GridSearchCV, cross_val_score, StratifiedKFold, learning_curve, train_test_split
17. from sklearn.svm import SVC

17.# Importing the dataset from Kaggle
18.data = pd.read_csv("../input/pima-indians-diabetes-database/diabetes.csv")  .
19# First step is getting familiar with the structure of the dataset
20.data.info()
21# Showing the top 5 rows of the dataset
22.data.head()
23.# Exploring the missing values in the diabetes dataset
24.data.isnull().sum()
25.# Replacing 0 values of Glucose
26.data['Glucose'] = data['Glucose'].replace(0, data['Glucose'].median())
27.
28.# Filling 0 values of Blood Pressure
```

```python
29.data['BloodPressure'] = data['BloodPressure'].repl
ace(0, data['BloodPressure'].median())
30.
31.# Replacing 0 values in BMI
32.data['BMI'] = data['BMI'].replace(0, data['BMI'].me
an())
33.
34.# Replacing the missing values of Insulin and Skin
Thickness
35.data['SkinThickness'] = data['SkinThickness'].repla
ce(0, data['SkinThickness'].mean())
36.data['Insulin'] = data['Insulin'].replace(0, data['Insu
lin'].mean())
37.data.head()
38.# Reviewing the dataset statistics
39.data.describe()
```

CONCLUSION:

It is used to describe the symptoms using machine learning