Problem Definition:

The Diabetes Health Indicators Dataset contains healthcare statistics and lifestyle survey information about people in general along with their diagnosis of diabetes. The target variable for classification is whether a patient has diabetes or is healthy (0 or 1) .

Features:

1. **ID : Patient ID.**
2. **HighBP** : 0 = no high BP 1 = high BP.
3. **HighChol :** 0 = no high cholesterol 1 = high cholesterol.
4. **CholCheck** : cholesterol check in 5 years – 0= no1 = yes.
5. **BMI**:Body Mass Index
6. **Smoker** :Have you smoked at least 100 cigarettes in your entire life? [Note: 5 packs = 100 cigarettes] 0 = no 1 = yes.
7. **Stroke** : (Ever told) you had a stroke. 0 = no 1 = yes.
8. **HeartDiseaseorAttack** : coronary heart disease (CHD) or myocardial infarction (MI) 0 = no 1 = yes.
9. **PhysActivity** : physical activity in past 30 days - not including job 0 = no 1 = yes.
10. **Fruits** : Consume Fruit 1 or more times per day 0 = no 1 = yes.
11. **Veggies** :Eat Vegetables 1 or more times per day 0 = no 1 = yes.
12. **HvyAlcoholConsump** : Heavy drinkers (adult men having more than 14 drinks per week and adult women having more than 7 drinks per week) 0 = no 1 = yes.
13. **AnyHealthcare:**Have any kind of health care coverage, including health insurance, prepaid plans such as HMO, etc. 0 = no 1 = yes.
14. **NoDocbcCost** : Was there a time in the past 12 months when you needed to see a doctor but could not because of cost? 0 = no 1 = yes.
15. **GenHlth**:Would you say that in general your health is: scale 1-5 1 = excellent 2 = very good 3 = good 4 = fair 5 = poor.
16. **MentHlth**:Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good? scale 1-30 days.
17. **PhysHlth**:Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good? scale 1-30 days.
18. **DiffWalk**:Do you have serious difficulty walking or climbing stairs? 0 = no 1 = yes.
19. **Sex** :   Sex   0 = female 1 = male.
20. **Age**   :13-level age category (_AGEG5YR see codebook) 1 = 18-24 9 = 60-64 13 = 80 or older.
21. **Education**:Education Level   Education level (EDUCA see codebook) scale 1-6 1 = Never attended school or only kindergarten 2 = Grades 1 through 8 (Elementary) 3 = Grades 9 through 11 (Some high school) 4 = Grade 12 or GED (High school graduate) 5 = College 1

year to 3 years (Some college or technical school) 6 = College 4 years or more (College graduate).

22. **Income**:Income scale (INCOME2 see codebook) scale 1-8 1 = less than $10,000 5 = less than $35,000 8 = $75,000 or more.

Target:

Diabetes_binary   :   0 = no diabetes 1 = prediabetes or diabetes.