## Project Idea

**Intelligent Predictive Model for Early Detection and Management of Noninfectious Diseases - Cancer and Diabetes in Djibouti.**

### Problem Definition:

Djibouti faces significant challenges in the early detection and effective management of noninfectious diseases, particularly cancer and diabetes. Limited healthcare resources and infrastructure underscore the need for a tailored solution to enhance early diagnosis, optimize treatment, and improve overall health outcomes.

### Relevance to Sustainable Development Goals (SDGs):

My project directly contributes to SDG 3 by focusing on the improvement of health outcomes through early detection and management of noninfectious diseases (cancer and diabetes). Also addresses health inequalities by tailoring its approach to the unique healthcare landscape of Djibouti. By preventing the economic burden associated with late-stage diseases, the project aligns with SDG 1's goal of eradicating poverty.

### Literature Examples:
1.Application of Machine Learning Models for Early Detection and Accurate Classification of Type 2 Diabetes by Orlando Iparraguirre-Villanueva 1, Karina Espinola-Linares
2.Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment Selection: A Critical Approach by Bo Zhang, 1 Huiping Shi, 1 and Hongtao Wang

### Describe Your Data:
1.Diabetes Dataset:Diabetes_updated_Dataset
Source:kaggle
Description: Outcome: The target variable; 0 for no diabetes, 1 for diabetes.
Format: CSV
Size: 23.87 kB
Link: https://www.kaggle.com/datasets/ashishkumarjayswal/diabetes-dataset
2.Diabetes Dataset : Breast Cancer Prediction Dataset
Source:kaggle
Description:  Dataset for AI for Social Good: Women Coders' Bootcamp organized by Artificial Intelligence for Development in collaboration with UNDP Nepal.
Format: CSV
Size: 19.65 kB

### Approach (Machine Learning or Deep Learning):

Since the characteristics of my datasets are relatively small in size and related to binary classification tasks (predicting diabetes and potentially breast cancer)  , a machine learning approach is more suitable. Here's the
machine learning Models  i will use:
Logistic Regression ,Decision Trees ,Random Forest Support Vector Machines (SVM),Gradient Boosting (e.g., XGBoost)