**Predictive Model for Heart Disease Diagnosis**

**Introduction**

In the global effort to achieve Sustainable Development Goal (SDG) 3: Good Health and Well-being, cardiovascular diseases remain a leading cause of morbidity and mortality. Heart disease, in particular, poses significant health challenges worldwide. Early detection and intervention can dramatically improve patient outcomes and reduce healthcare costs. This project aims to develop a machine learning model that predicts the presence of heart disease in patients using historical patient data.

**Problem Statement**

Heart disease diagnosis often relies on invasive procedures and subjective assessments. Many patients suffer from undiagnosed heart conditions due to the lack of efficient and accessible diagnostic tools. The problem lies in the need for a reliable, non-invasive predictive model that can assist healthcare providers in identifying patients at risk of heart disease early.

**Objective**

The primary objective of this project is to develop a machine learning model that accurately predicts the presence of heart disease in patients. This model will:

1. Enhance early detection of heart disease.

2. Provide a non-invasive diagnostic tool for healthcare providers.

3. Improve patient outcomes and reduce healthcare costs.

**Why This Problem?**

Addressing heart disease through early detection aligns with SDG 3 by improving health and well-being. Early intervention can prevent complications, reduce hospital admissions, and save lives. By leveraging machine learning, we can create a scalable solution that enhances the diagnostic process and supports healthcare providers in making informed decisions.

**Solution: Developing the Heart Disease Predictor**

**Overview**

The Heart Disease Predictor is a machine learning model trained on historical patient data to predict the presence of heart disease. This model utilizes various patient attributes, including age, gender, blood pressure, cholesterol levels, and more, to make accurate predictions.

**Features**

**- Data-Driven Insights:** The model analyzes patient data to identify patterns associated with heart disease.

**- Non-Invasive:** Utilizes readily available patient data, reducing the need for invasive diagnostic procedures.

**- Scalability:** Can be integrated into healthcare systems to support large-scale screening.

**Dataset Description**

The dataset used for this project originates from four databases: Cleveland, Hungary, Switzerland, and Long Beach V. It contains 76 attributes, with a subset of 14 key attributes used for model training and prediction. These include:

- Age

- Sex

- Chest Pain Type (cp)

- Resting Blood Pressure (trestbps)

- Serum Cholesterol (chol)

- Fasting Blood Sugar (fbs)

- Resting Electrocardiographic Results (restecg)

- Maximum Heart Rate Achieved (thalach)

- Exercise Induced Angina (exang)

- ST Depression Induced by Exercise (oldpeak)

- Slope of the Peak Exercise ST Segment (slope)

- Number of Major Vessels Colored by Fluoroscopy (ca)

- Thalassemia (thal)

- Target (heart disease presence)

**Technical Implementation**

* **The IBM Cloud platform:** It includes a console, identity management, extensive catalog, resource search, and robust account and billing system for a consistent cloud experience.
* **IBM Watson Studio:** IBM Watson Studio streamlines AI development with tools for data preparation, model building, deployment, and collaboration, supporting scalability, governance, and security across hybrid environments.
* **IBM Watson AutoAI:** It automates building predictive models by preparing data, selecting features, tuning hyperparameters, and ranking model pipelines for accuracy and precision.

**Conclusion**

The Heart Disease Predictor model demonstrates the potential to enhance early detection of heart disease, aligning with SDG 3 by improving patient outcomes and reducing healthcare costs. This model can serve as a non-invasive, scalable tool for healthcare providers, supporting better diagnostic decisions and timely interventions.

**Project Link**

<https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/0c6874fa-a946-42ab-83e4-bc65072a309d/view?access_token=dc56ab2978d13fd05519ae1f042fbc692c88d9722bdc15cefa375a084118c90b&context=cpdaas>

**References**

- Kaggle Datasets: Heart Disease Dataset (https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset)

- IBM Cloud: IBM Watson Studio, AutoAI

- IBM SkillsBuild: Getting Started with Artificial Intelligence, Artificial Intelligence Fundamentals