Assignment 1

Data Science and Machine Learning Bootcamp

Predicting Heart Disease

- Your task is to predict the presence of heart disease in patients using data and AI. You will be provided with a dataset containing information about various attributes of patients such as age, sex, cholesterol levels, etc. along with the presence of heart disease (0 = no disease, 1-4 = varying degrees of disease). Your goal is to build a machine learning model that can accurately predict the presence of heart disease based on these attributes.

- Data set*

https://archive.ics.uci.edu/ml/datasets/Heart+Disease

*_Read the instructions about the data set properly._

You need to perform

1. Data Exploration

Explore the dataset and perform various operations such as cleaning, feature selection, and feature engineering to prepare the data for analysis.

2. Data Analysis

Analyze the dataset and perform exploratory data analysis to identify any trends or patterns in the data.

3. Data Modeling

Do you think whether you can use logistics regression model to predict the heart disease. Explain.

4. Report

Write a report summarizing your findings, including an explanation of the steps you took to explore and analyze the data, showcase the trends and patterns you have identified.

5. Repository

Create a GitHub repository to store all of your code and report. Make sure to include a README file with instructions on how to run your code.

6. Submission

Submit a link to your GitHub repository containing your code and report.

Tools to be used

1.Python programming language

2.Jupyter Notebook

3.ChatGPT or other similar tools for generating code

4.GitHub for storing your code and report.

You have seven days to complete the assignment and submit the link to your GitHub repository containing your code and report.

Deadline is 9th May 2023.

Good luck!