

ML Summer Training 2024

Final Project

Objective

You are required to create a code from scratch that solves this problem and make a suitable interface for it (AI & Software).

Project Description

Cardiovascular diseases (CVDs) are the leading cause of death globally, responsible for approximately 17.9 million deaths each year, which accounts for 31% of all deaths worldwide. The majority of CVD-related deaths result from heart attacks and strokes, with a significant portion occurring prematurely in individuals under the age of 70. This dataset focuses on predicting heart disease, a common consequence of CVDs, using 11 key features such as age, sex, chest pain type, resting blood pressure, and cholesterol levels. Individuals at high cardiovascular risk, due to factors like hypertension, diabetes, or hyperlipidemia, require early detection and management. Machine learning models can significantly aid in this process.

Dataset

Check the dataset [here](#).

List of Requirements

- **Algorithms to Implement:**
 1. KNN
 2. Machine Learning approach
 3. Deep Learning approach
- **Interface:**
 1. User interface to input individual details and get a detection result
- **Deployment:**
 1. Upload it on Cloud (AWS or Azure)

- **Documentation:**

- Include a document covering the following:
 - a. Preprocessing
 - b. Model selection
 - c. Evaluation criteria
 - d. Explanation of your work
 - e. Comparison between approaches

GuideLines

1. Understand the Data:
 - a. Try to conduct a thorough Exploratory Data Analysis (EDA) to understand the dataset.
 - b. Hint: Google "What is Exploratory Data Analysis" and explore Kaggle notebooks to discover useful functions and techniques for EDA.
2. Choose Your Framework:
 - a. You are free to choose any machine learning or deep learning framework that you prefer for this project 😊.
3. Code Readability:
 - a. Ensure that your code is well-documented and includes comments for clarity.
 - b. The comments should explain the purpose of each significant block of code to make it understandable to others.
4. Originality of Code:
 - a. Do not copy and paste code from any source. You must write original code and fully understand every line.
 - b. Important: **Avoid using code auto-completion tools like GitHub Copilot.** You will be required to explain your code during the project discussion.
5. Dataset Usage:
 - a. Only use the dataset provided for this project. **Using any other dataset will result in a zero grade.**
6. Academic Integrity:
 - a. **Any sign of cheating, plagiarism, or unauthorized collaboration will not be tolerated and will result in a zero grade for the project.**

Deliverables

1. Your Code with results shown (You can either use notebooks or python scripts)
2. Document for the documentation requirements listed above.
3. You need to submit Code and documentation before **Monday 9/9/2024**.