Big Data Project

Supervised by: **Dr. Lydia Wahid**

Eng. Omar Samir Galal

Presented by:

Name	Sec.	B.N.
Ahmed Emad	1	7
Ahmed Mahmoud Hafez	1	11
Nour El-din Moustafa	2	33
Youssef Atef Abdo	2	41

Spring 2023

Idea:

Heart disease is the leading cause of death globally, and its prevalence is increasing rapidly. Early detection and prevention of heart disease are crucial for reducing mortality rates and improving the quality of life for patients.

In this proposal, we suggest a big data project that aims to develop a predictive model for the early detection of heart disease.

Dataset:

Link to Dataset: https://rb.gy/froyxl

Planned approach:

The dataset will be preprocessed to remove missing values and outliers. Feature engineering techniques will be used to extract useful features from the dataset. Feature selection techniques such as principal component analysis (PCA) will be used to select the most important features for the predictive model.

The predictive model will be developed using machine learning algorithms such as logistic regression, decision trees, and support vector machines (SVM) where Map-Reduce will be used. The model will be trained and tested to ensure the accuracy and reliability of the results.