Capstone Project Submission

Instructions:

i) Please fill in all the required information.

ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:
Name - Rohan Jagadale **Email id - rjagadale202@gmail.com** **Contribution - Whole project**
Please paste the GitHub Repo link.
Github Link:- https://github.com/Rohan20202/Cardiovascular-Risk-Prediction-Project
Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)
The dataset is from an ongoing cardiovascular study on residents of the town of Framingham, massachusetts. The classification goal is to predict whether the patient has a 10-year risk of future coronary heart disease (CHD). Machine learning is now an emerging field due to the increasing amount of data. Machine learning makes it possible to acquire knowledge from a massive amount of data, which is very heavy for man and sometimes impossible. The objective of this paper is to prioritize the diagnostic test and to see some of the health habits that contribute to CVD. In this project, manually classified data is used. Manual classification is healthy or unhealthy. Based on a machine learning technique called classification, 80% of the data is supervised or trained and 20% is tested as part of this project. Used variance inflation factor to check the collinearity among the variables. Standard scalar was used to scale the data. I have used the Decision Tree Classifier,Logistic Regression, Random Forest Classifier, K Near Neighbor, Naive Bayes Classifier, SVM Classifier, Gradient Boosting Classifier to train the model. Logistic Regression , SVM classifier, Gradient Boosting Classifier(Tuning) these three models give good accuracy on test dataset with percentage of 86%, 85%, 85% respectively. If we want to choose only the best one model it is better to train the model with Logistic Regression which has 85.32% accuracy on test data.

Google drive link :
https://drive.google.com/drive/folders/1j0jqMLL7wOVILCj8Qy7_TrQfHiGok9KU