**LARGE GROUP PROJECT**

**CARDIO- HEART DISEASE PREDICTION WITH FRAMINGHAM HEART STUDY DATA**

**TEAMMATES:**

1. LIKITHA KANAKA LAKSHMI KHAMBHAMPATI

2. NAGA SASI SRAVYA TEEGALA

3. SRILEKHA NANABALA

**Link** : <https://www.kaggle.com/code/muzammilrizvi1/heart-disease-random-forest/notebook>

GitHub : <https://github.com/srilekhananabala/CARDIO-HEART-DISEASE-PREDICITION.git>

This is the dataset.

Columns - 16

Rows – 4240

Target variable – 10year CHD

**Preprocessing:**

We got missing values after running preprocessing.

**Hyperparameter Tuning:**

We have used Random Forest and did the hyperparameter tuning.

# **Best Parameters:**

# {'max\_depth': 10, 'min\_samples\_leaf': 4, 'min\_samples\_split': 5, 'n\_estimators': 100}

**Model Building:**

By using these best parameters, we built the random forest model.

**Model Evaluation:**

|  |  |  |
| --- | --- | --- |
| **Model** | **Training Accuracy** | **Testing Accuracy** |
| SVM | 0.90 | 0.84 |
| XGB Classifier | 0.86 | 0.84 |
| Random Forest | 0.87 | 0.83 |