**Members:**

**AU2040011 - Varun Parekh**

**AU2040098 - Vanshit Shah**

**AU2040253 - Ushmay Patel**

**AU2040162 - Madhvendra Jhala**

**1)** Tasks Performed in the week.

**2)** Outcomes of the tasks performed.

**3)** Tasks to be performed in the upcoming week.

(You can also put Images. Graphs, etc., if required)

Firstly we worked through the dataset, and figured out the factors that we would be considering in order to predict the Heart-Diseases.

And then we finalized 14 of the parameters from the available parameters:

1. #3 (age)

2. #4 (sex)

3. #9 (cp)

{

cp: chest pain type

-- Value 1: typical angina

-- Value 2: atypical angina

-- Value 3: non-anginal pain

-- Value 4: asymptomatic

}

4. #10 (trestbps) {resting blood pressure (in mm Hg on admission to the hospital)}

5. #12 (chol) {Cholestrol}

6. #16 (fbs) {fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)}

7. #19 (restecg) {resting electrocardiographic results

-- Value 0: normal

-- Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)

-- Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria}

8. #32 (thalach) {maximum heart rate achieved}

9. #38 (exang ){exercise induced angina (1 = yes; 0 = no)}

10. #40 (oldpeak) {ST depression induced by exercise relative to rest}

11. #41 (slope) {the slope of the peak exercise ST segment

-- Value 1: upsloping

-- Value 2: flat

-- Value 3: downsloping

}

12. #44 (ca) {number of major vessels (0-3) colored by fluoroscopy }

13. #51 (thal) { 3 = normal; 6 = fixed defect; 7 = reversible defect }

14. #58 (num) (the predicted attribute) {

diagnosis of heart disease (angiographic disease status)

-- Value 0: < 50% diameter narrowing

-- Value 1: > 50% diameter narrowing }

We tried to work on the remarks and understand and analyze different classifier methods.

We also tried to figure out some classifiers that could be used for our project like SVM, Naive Bytes, Logistic Regression, Adaboost, Xgboost. For the tree based approach, we are looking forward to the Decision tree. Also, we then decided upon division of the dataset and concluded that we keep 70% training set and 30%..

The methodologies that we decided that we would be using and understood the basics of that:

1. **Linear SVM:** Linear SVM is used for linearly separable data, which means if a dataset can be classified into two classes by using a single straight line, then such data is termed as linearly separable data, and classifier is used called as Linear SVM classifier.
2. **Non-Linear SVM:** Non-Linear SVM is used for non-linearly separated data, which means if a dataset cannot be classified by using a straight line, then such data is termed as non-linear data and classifier used is called as Non-linear SVM classifier.
3. **Tree-Based Approach:** Tree-based models use a decision tree to represent how different input variables can be used to predict a target value. Machine learning uses tree-based models for both classification and regression problems, such as the type of animal or value of a home.Although it may be used to solve classification and regression problems, Decision Tree is a Supervised learning technique that is typically used for classification problems.

We will be comparing the results obtained through various methodologies by comparing the outcomes of the various techniques.