**MACHINE LEARNING DAY-2**

**HEART CSV**

STEPS TO COMPLETE THIS PROJECT USING AGILE SOFTWARE METHODS

1) READ AN UNDERSTAND THE DATA IN CSV FILE

2)EXTRACT FEATURES AND LABELS USING PANDAS

3)DIVIDE INTO X\_TEST,Y\_TEST,X\_TRAIN,Y\_TRAIN

4)BUILD MODEL USING DIFFERENT ALOGORITHM

5)BUILD MODEL USING X\_TRAIN AND Y\_TRAIN

6)CHECK SCORES AND CHECK WHICH HAS MORE ACCURACY

**IN CSV**HERE WE ANALYZE THE CSV FILE AND WE NEED TO FIND PERSON PRONE TO HEART ATTACK SO OUR OUTPUT OR LABEL SHOULD BE PERSON PRONE TO HEART ATTACK

The features[INPUTS] are age , sex , cp , trestbps , chol , fbs , restecg , thalach , exang , oldpeak, slope , ca and thal.

The label[OUTPUT] is the target.

80% of the data is used as training data.[X\_TRAIN,Y\_TRAIN]

20% of the data is used as testing data.[X\_TEST,Y\_TEST]

Th  
e different Machine Learning algorithms used are:

1.GaussianNB

2.DecisionTree

3.Random Forest

4.KNN

5.SVM

6.Logistic Regression

fit():To build a model taking input and output.

predict():To predict a output for a new value

score():Accuracy of prediction.

Accuracy of prediction

Decision Tree Accuracy: 0.9707792207792207

SVM Accuracy: 0.6753246753246753

Naive Bayes Accuracy: 0.814935064935065

KNN Accuracy: 0.7142857142857143

Random Forest Accuracy: 0.990259740259740

Logistic Regression- 0.8054474708171206

**The most accurate Machine learning algorithm for heart.csv is RANDOM FOREST**