**Documentary reports on the project**

Title ; Titanic machine learning from disasters.

Aims and Objectives : to build a predictive model that can classify whether a passenger survived or not based on the available features contained in the dataset.

Application used: Anaconda (jupyter notebook)

Procedures:

1, I loaded the necessary library needed to carry out the task.

2. I loaded the dataset into the jupyter notebook,

3, I carried out exploratory data analysis by trying to derive some information from the data set (column, shape, type, null, outliers)

4.During EDA I discovered some element in the datasets are null so I had to fill them using the pd.fillna() with mean, forward fill or backward fill.

5. I encoded a column in the data set to make it accessible for modelling (SEX. COLUMN IN THE DATASET) the encoding was carried out using labelencoder from sklearn module.

6.I split the data set into dependent and independent form to enable the ML model understand what I was trying to do.

7.Before modelling another important procedure is data preprocessing I carried out the preprocessing using standard scaler module from sklearn. I only carry out the preprocessing on what I assumed to be my input.

8.On the verge of data modelling by splitting the data into training and testing size, I used the train test spilt module from sklearn to carry out that.

9.Now, my data is ready for training, I imported the RANDOMFOREST model from the ensemble library then I passed in my training splits in order to train the data.

10.After Training, I also passed in my testing splits to the model.

11.After modelling I went ahead to check for accuracy\_score , precision score and f1\_score.

Observation : through model tuning I observed that there is a slight increase in the number of survivals after modelling.

**ENGIMEDICS**