***ML  
1. Data calling.***

***2. EDA.***

***3. Data Preprocessing: i. imputation, ii. Data Scale, iii. Duplicate entries.***

***4. Data Clean (preprocessing nd cleaning can be done together).***

***5. Model Building.***

***6. Validation: KFold technique.***

***7. Grid search CV or Random Search CV} HyperParameter Tuning (finding optimum parameter, model will perform excellently, getting higher model accuracy).***

***8. Deployment.***

***loggers.py- create logs***

***\_\_init\_\_.py also works as package builder,***

***Filemode = ‘w’- write***

***‘b’ – binary***

***‘r’- read***

***‘rb’-read binary***

***SD = dist from mean to data points***

***If mean ! = Median***

***-🡪 Non Normal Distribution***

***1. Tree Based Algo***

***🡪 Random Forest***

* ***Decision Tree***

***2. Boosting Algo***

***🡪 XGBoost***

* ***GBoost***
* ***AdaBoost***

***If mean = Median***

***-🡪 Normally distributed***

***1. Classification / Recommendation***

***🡪 Logistic Regression X***

***Regression/ Prediction***

* ***Linear Regression X***