

1. You are provided the titanic dataset. Load the dataset and perform splitting into training and test sets with 70:30 ratio randomly using test train split.
2. Use the Logistic regression created from scratch (from the prev question) in this question as well.
3. Data cleaning plays a major role in this question. Report all the methods used by you in the ipynb.

-->

- i. Check for missing values
- ii. Drop Columns & Handle missing values
- iii. Create dummies for categorical features

you are free to perform other data cleaning to improve your results.

4. Report accuracy score, Confusion matrix, heat map, classification report and any other metrics you feel useful.

dataset link : https://iiitaphyd-my.sharepoint.com/:f:/g/personal/apurva_jadhav_students_iiit_ac_in/Eicctt5_qmoxNqezgQQiMWelBph4sxlA6jWAJNPnV2SF9Q?e=mQmYN0

(titanic.csv)

