# ADS Assignment 8.

This is a continuation of the analysis done from assignment 7. Base on the feature columns identified, use them in your prediction analysis. You are advised to confirm back to the noes as well as make good use of the Sklearn and Keras documentation.

1. Separate df_all to two dataframes df_regression and df_classificaion each having all the features but only the respective target column.
2. Model the data effectively for the ML.
3. Identify the best algorithm for each use case from the list of possible ML methods under consideration for each task. Your decision should be based on training time, prediction time, training and testing score.

| **Classification** | **Regression** |
|---|---|
| Logistic Regression | Linear Regression |
| Support vector machine | Decision Tree |
| K nearest Neighbours | Bagging |
| Random Forest | Adaboost |
| MLP ANN | |

4. When the best model for each task has been identified, use GridSearch on each of the models to identify the best hyper parameters.
5. Set up an new instance of both the regression and classification model with the identified parameters, train the models and report both the training and testing cross validation mean score and standard deviation over 5 cross folds.
6. Using the following metrics, report the testing accuracy of each of the models using the following:
    a. Classification:
        i. Classification report.
        ii. Heatmap confusion matrix.
    b. Regression:
        i. Mean Absolute Error
        ii. Mean Squared Error
        iii. R2-Score
7. Save both models as a pickle file bearing the respective names:
    a. Regression – ‘regression_model’
    b. Classification – ‘classification_model’