Project Timeline



Import the datasets into data frames.



Perform EDA to visualise distributions/o utliers/correlat ions/excess zeroes.



Perform a train-validation-test split of the dataset.



Review the training dataset's columns and handle them accordingly.



Review and consider any other sources of data leakage.

STAGE 01

STAGE 02

STAGE 03

STAGE 04

Project Timeline Template



Perform feature engineering (using training data).



Perform feature selection (using validation data).



Define the classifiers/mod els used.



Perform hyperparamete r optimisation via grid-search methods (using validation data).



Train each model (using training data).

STAGE 06

STAGE 07

STAGE 08

STAGE 09

Project Timeline



Evaluate the performance of each model (using validation data)



Assess & determine the best performing model, based on the metrics discussed above



Generate predictions/pro bability estimates using the best model (using test data).



Evaluate the performance of the final model (using test data).



Review model performance in the context of "feature importance" and explainability.

STAGE 11

STAGE 12

STAGE 13

STAGE 14

Project Timeline



Review for areas of improvement.