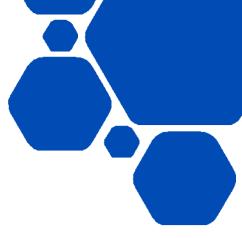


## **WEEKLY REPORT**



## **Dear** Mostafa Salama

In the below Table you will see a specified table that includes the good, the bad & the improvements that can be done on your Classification Task

Correctly Done	Can Be Improved
Encoding     Use robust model	<ul> <li>Apply Scaling</li> <li>You can apply other techniques for outlier as transformation (log), capping, fill by median (try and determine the good one)</li> <li>Handle multicollinearity using drop cols with high vif or pca (not necessary to increase performance, the integrity will be)</li> <li>Encoding using one-hot, label or manual encoder as type1:1, type2:2,</li> <li>You can use more robust models as random forest, xgbooost, adaboost,</li> <li>You can use grid-search for more hyper tunning</li> <li>You can apply cross-validation to know genera performance of model on all data</li> <li>Split data into train and test before any preprocessing</li> <li>Handle Imbalanced</li> <li>You can add more features as total_members, total_night, percent_can((p_c)/(p_c + p_not_c)</li> </ul>





