**WNS Ad Click Prediction**

**Chosen Approach:**

1. **Exploratory Data Analysis**:- Merged view log and train\_data, mean imputed nan rows for user\_ids that were not present in view+log. Removed outlier for ‘item\_price’ > 99th quantile.
2. **Feature engineering**: Constructed Target Mean encoded features for all categorical features except 'user\_id', 'session\_id'. Also constructed combined features taking ‘app\_code’ and ‘product\_type’.
3. Trained a baseline Logistic Regression model, got ROC\_AUC = 0.73397
4. Trained LGBClassifier with intense hyperparameter tuning, got ROC\_AUC = 0.74308

**Experimentations:**

1. Since most of the features have high cardinality, I tried to combine all the categorical features except impression\_id as one and constructed BOW features. But results were sub-optimal as there was no extra weightage given for more important features.
2. Tried taking all the categorical features expect Id features without any feature transformation and trained a CatboostClassifier model on it and got ROC\_AUC = 0.69.
3. Tried OHE for ‘app\_code’, ‘product\_type’ and other low cardinality categorical features the results after traning LR and LGBClassifier were even worse with ROC\_AUC ~ 0.52.