Template ML

Problem Definition

- Understand the problem
- Ask relevant questions

Preprocessing

- Grab the Data
- Separate the Data into Dependent and Independent variables.
- Deal with Missing values (Impute / Drop)
- Deal with Outliers (Impute / Drop)
- Check Co-Relation between Data and Multi-Co-linearity

Preprocessing 2

- Encode Categorical Data and create dummy variables
- Eliminate the Non Influencing factors (Independent Variables)
- Feature Scaling
- Apply Dimensionality Reduction
- Handle Data Imbalance

Data Visualization

- Make the charts based on type of the data you have and charts should be easily interpretable.

Modeling

- Split data into training and testing data
- Build the Model
- Run it on Training set
- Validate using Testing set and capture the performance score of the model using various evaluation matrix.
- Check Bias and Variance

Modeling 2

- Run it on K-Fold Test
- Recheck Bias and Variance
- Tune the Hyper-parameters

Interpret the result of your model