# **The Machine Learning Project Template**

The idea is to summarize the steps followed in a basic ML project and to use all the ML algorithms including the ensemble method and them picking up the best model out of all.

There are many steps involved in this procedure from preparing the problem to finalizing the model.

Beginning from the start:

1. **Prepare Problem**

Firstly, understanding the situation and the problem which needs to be solved. And then devise how Machine Learning would solve that efficiently.

1. **Load libraries**

Importing all the libraries and the packages required.

1. **Load dataset**

Loading/importing the data and making sure that data is stored in the same working directory as the current working directory.

1. **Summarize Data**

Checking how the data looks like and what does it contains and then making sure of number of rows and columns and distinguishing the data types.

1. **Descriptive statistics**

Descriptive statistics describes the data in terms of its statistics like mean, standard deviation, quantiles etc.

1. **Data visualizations**

Visualizations using Matplotlib, Seaborn will be used to check the correlations within the features and with the target, scatter plots of data, histograms and boxplots for checking the spread and skewness and much more.

1. **Prepare Data**

Transforming data in order to make it suitable for algorithms to process and work more efficiently in order to give more accurate and precise results.

It includes 3 steps as shown below:

1. Data Cleaning

b) Feature Selection

c) Data Transforms

1. **Evaluate Algorithms**

Checking the performance of the various classification algorithms based on the type of problem.

Evaluating the algorithms will be done in these few steps:

1. Split-out validation dataset

b) Test options and evaluation metric

c) Spot Check Algorithms

d) **Compare Algorithms**

The algorithms giving consistently high scores will be the target

1. **Improve Accuracy**

After we find the best performing algorithms, their parameters and the Hyperparameters can be tuned to give maximum results.

a) Algorithm Tuning

b) Ensembles

1. **Finalize Model**

a) Predictions on validation dataset

b) Create standalone model on entire training dataset

c) Save model for later use