**Readme**

**Environment-**

The project uses Python3 and Jupyter notebook.

**Packages-**

1) sklearn - For models and metrics

2) warnings - For preventing warnings

3) numpy - For basic matrix handling

4) matplotlib - For figure plotting

5) pandas - For creating dataframes

6) seaborn - For figure plotting

7) timeit - For tracking times

8) os - for setting work directory

9) random - For creating random seeds

10) csv - For saving csv files

11) json - For creating json files

12) itertools - For creating iterators for efficient looping

13) pprint - For pretty printing data structures

14) pydash - for doing “stuff” in a functional way (utility library).

15) gc - Garbage collector for deleting data

16) re - Raw string notation for regular expression patterns

17) featuretools - Automated feature engineering

18) xgboost - XGBoost model

19) lightgbm - LightGBM model

20) hyperopt - Bayesian hyperparameter optimization

*Note- The required packages can be installed by uncommenting the first cell of Jupyter notebook.*

**Dataset-**

The complete dataset is available here- https://www.kaggle.com/c/home-credit-default-risk/data