**HW3 Instructions**

We will use the same data set as in HW1 and HW2. You will add more models to your HW2 notebooks. We will use r2\_score as scoring method for regression task and accuracy for classification task. If any model in HW2 has used different scoring method, re-run those models with these metrics only.

**Regression Task:**

* In HW2 you were required to run following models: KNN regressor, linear regression, linear regression with SGD, Ridge, Lasso, ElasticNet, Polynomial regression, SVM both simple and with kernels (try rbf, poly, and sigmoid kernel), Decision Tree regression.
* If any of the model in HW2 was missing in your HW2 notebook, please add those models.
* You will loose points if any model is missing in your notebooks.
* **In HW3 you will now run the following models:** 
  + Two models with Bagging (best model from HW2 and Decision Tree)
  + Two models with Pasting (best model from HW2 and Decision Tree)
  + Random Forest
  + Ada Boost (with decision tree and SVM)
  + Gradient Boost
  + XGBoost (you will need to install the library xgboost using pip install xgboost).
* Predict your outcome on test data for all the models
* Create a correlation matrix for prediction from all the models
* Use Voting Regressor to combine results of top 5 models (all the models including models you ran for HW2)
* Use Voting Regressor to combine results of 5 models with least correlation.
* If your model has a scaling parameter(s) use Grid Search to find the best scaling parameter using cross validation score.
* Choose the final model - The final model should also be chosen based on Cross-Validation Score.
* After selecting the best model, report the test and train error based on the chosen best model. Provide this information in the first cell of your notebook.

**Classification Task:**

* In HW2 you were required to run following models:: KNN classification, Logistic Regression, Linear Support Vector Machine, Kernelized Support Vector Machine (try rbf, polynomial and sigmoid kernels), Decision Tree classification.
* If any of the model in HW2 was missing in your HW2 notebook, please add those models
* **In HW3 you will now run the following models:** 
  + Two models with Bagging (best model from HW2 and Decision Tree)
  + Two models with Pasting (best model from HW2 and Decision Tree)
  + Random Forest
  + Ada Boost (with decision tree and SVM)
  + Gradient Boost
  + XGBoost (you will need to install the library xgboost using pip install xgboost).
* Predict your outcome on test data for all the models
* Create a correlation matrix using prediction from all the models
* Use Voting Classifier to combine results of top 5 models (all the models including models you ran for HW2.
* Use Voting Classifier to combine results of 5 models with least correlation.
* If your model has a scaling parameter(s) use Grid Search to find the best scaling parameter using cross validation score.
* Choose the final model - The final model should also be chosen based on Cross-Validation Score.
* After selecting the best model, report the test and train error based on the chosen best model. Provide this information in the first cell of your notebook.

**Deliverables:**

* Use "download as" in the "file" menu to convert your ipython file to a .pdf file
* Submit two files each for Regression and Classification task: .ipynb, and .pdf files to the eLearning