

Description: This week you all will be working in pairs to build a model for the kaggle house price competition. Each pair will consist of an experienced programmer along with a less experienced programmer who will work through the entire ml development pipeline. At next Sunday's meeting, you all will present your models and your results.

Dataset link: <https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

Download train.csv only

Requirements:

- build at least one ANN with pytorch
 - you may use sklearn or other libraries to build generic ml models as well to see how it fares, but the ANN is the bare minimum
- practice using software style guidelines
- use mse loss to train and evaluate your models
- Make slideshow presenting model pipeline and results
 - talk about preprocessing steps
 - data visualization
 - feature extraction
 - model architecture

Results Evaluation:

- Two factors will be considered in deciding the winner
 - the test set mse loss
 - the number of features used for the model
- The score formula
 $(\text{features used} / 79) * (\text{test set mse loss})$