Charlie Marshall

Prof Katsaggelos

ELEC\_ENG 375

31 May 2020

Final Project Proposal

For my final project, I plan to use the Ames, Iowa Housing dataset to predict the sale prices of residential property. This data is similar to the commonly used Boston Housing dataset, but is larger and has a much greater amount of factors which are included in the dataset. One of the challenges of this project will be to use feature engineering to clean the data and to limit analysis to important variables.

Additionally, I think this project will be interesting because I will be to spend time using regression techniques such as XGBoost and random forests. I do not have any experience using either of these techniques in practice, but I know they are commonly used in industry. I think I will gain a lot from learning to tune and improve an XGBoost model. Additionally, I expect to gain some important time practicing cross-validation as well.

Lastly, I will compare the performance of my models using RMSE, the root mean square error of the predicted values from the actual sales values. I will be working on this project alone, so I think it will be a much more feasible project than my proposal. However, I think I will be able to spend some valuable time with different techniques such as XGBoost that I have not used before.