Project 1: Predict the Housing Prices in Ames

CS598: Practical Statistical Learning

Naomi Bhagat - nbhagat3, Michael Miller - msmille3, Joe May - jemay3

17 September 2023

Assignment Data

Program: MCS-DS Assignment post: campuswire

Team contributions:

Person	Contribution
Naomi Bhagat	
Michael Miller	
Joe May	

Section 1: Technical Details

Discuss details such as data pre-processing and other non-trivial implementation aspects of your models. Do NOT paste your code in the report. Instead, explain the technical steps in clear English. Your description should be comprehensive enough for your fellow PSL classmates to replicate your results.

For instance, when documenting your pre-processing steps, provide specifics such as: - Which variables did you exclude from the analysis? - Identify the variables treated as categorical. How were these variables encoded, were any levels merged, etc? - For numerical variables, were there any transformations applied? - You're not required to justify these pre-processing decisions; just state what was done. When documenting implementation, general statements like "We use lasso to fit a sparse regression model" are insufficient. Instead, aim for detailed descriptions, such as: "We utilized lasso for regression modeling. Specifically, we employed the glmnet function in R with the data standardized and with lambda set to lambda.min."

Section 2: Performance Metrics

Report the accuracy of your models on the test data (refer to the provided evaluation metric below), the execution time of your code, and details of the computer system you used (e.g., Macbook Pro, 2.53 GHz, 4GB memory or AWS t2.large) for all 10 training/test splits.