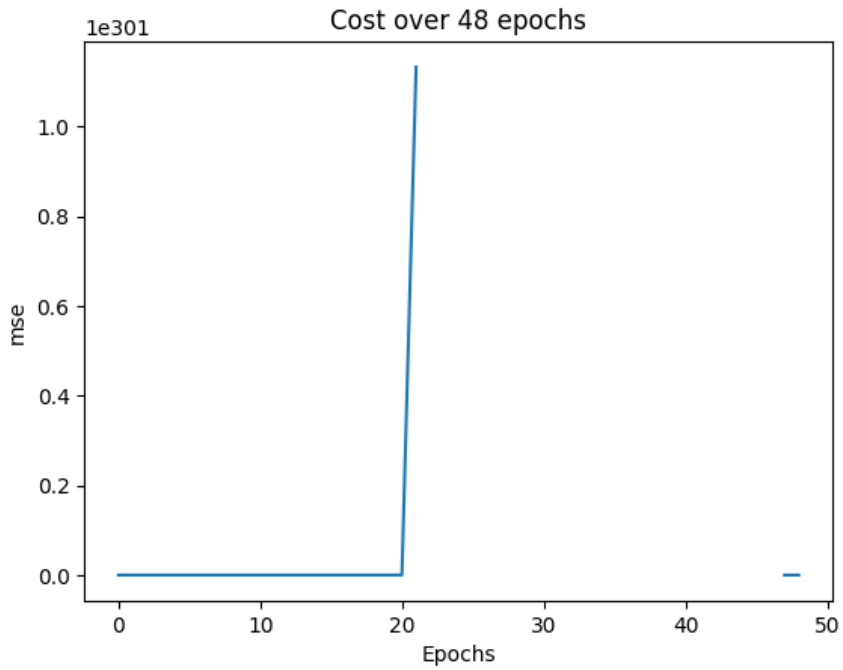
Question 1- Training without normalization

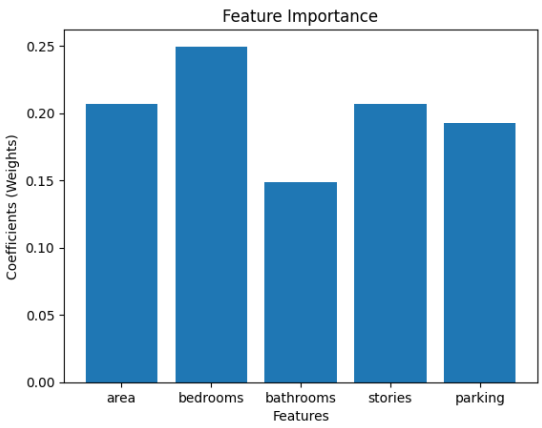
Attempting to train a model on the housing dataset as-is is futile. Every column of data is on a different scale, and the high magnitude of the values causes all sorts of errors. In this case, area will automatically have the most influence on the trained model because of how large the area values are compared to bedrooms, for example.

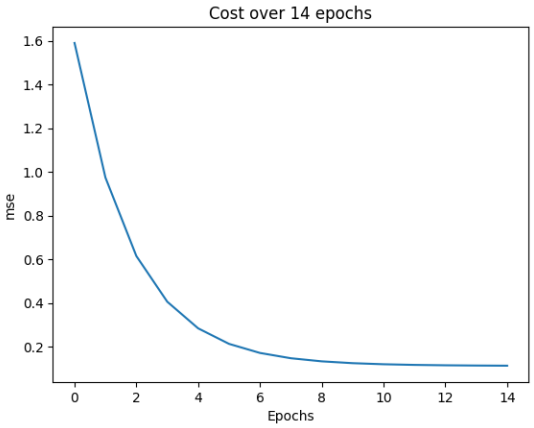


Because of the high numbers, python just reports an overflow error on each operation and the model isn't able to properly train.

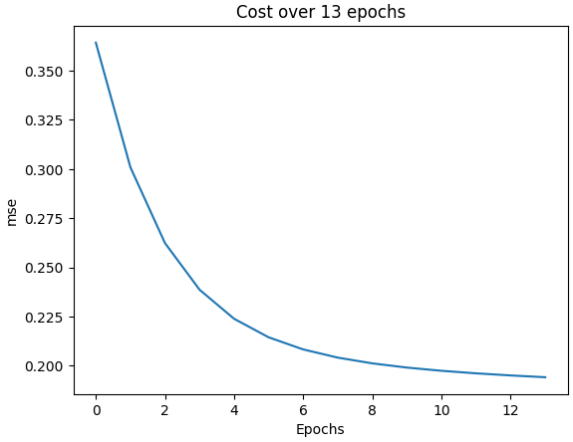
Question 2- Training with normalization and standardization

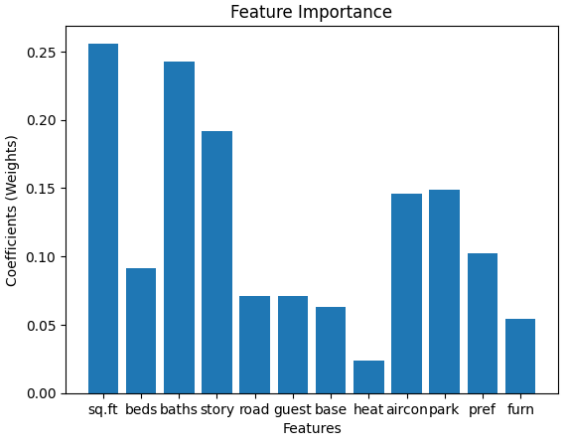
With normalization:





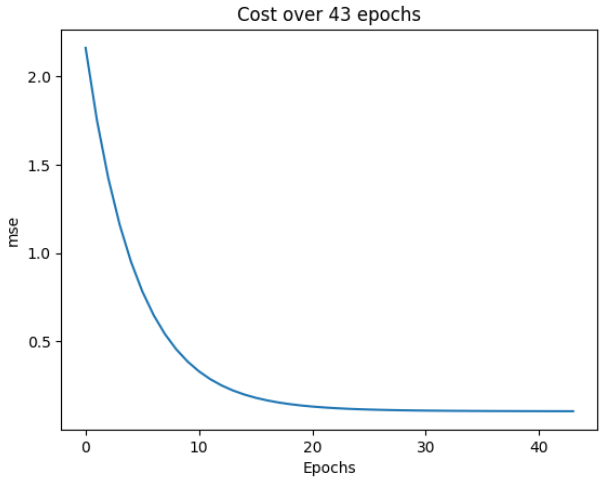
With standardization:

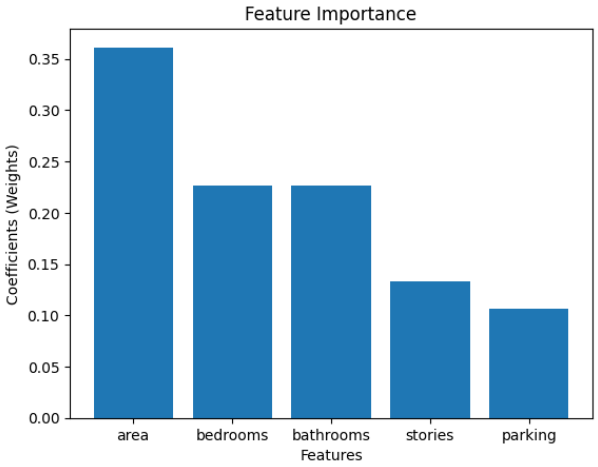




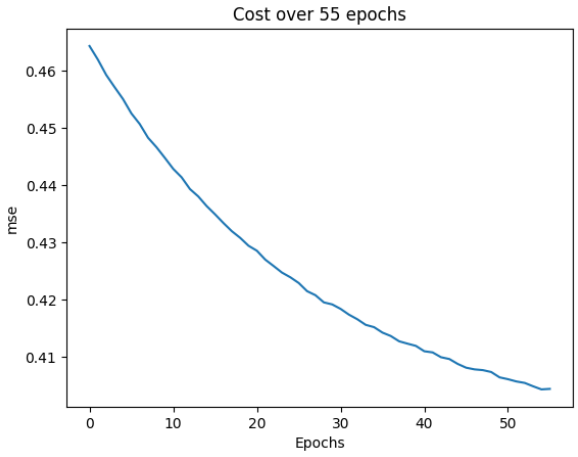
Question 3- Training with Lasso regulation (parameter penalty)

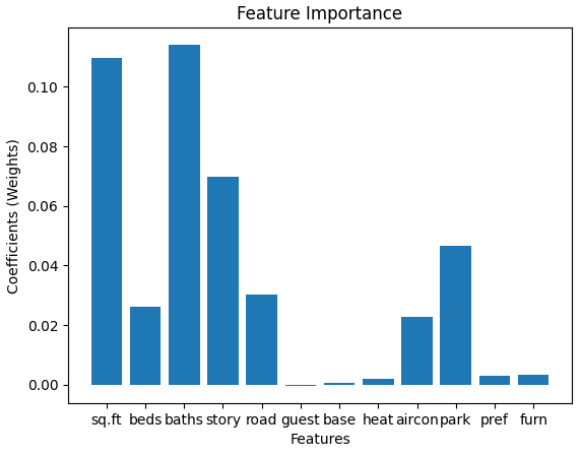
With normalization:





With standardization:





[Anu78/intro-to-ml-hw: homework for ECGR-4105 @ uncc (github.com)]