

Jeremy Anderson, Katie Shatkus, Nate Mehringer, Eric Wilens, Will Coffey, Matthew A. Lanham

Purdue University Krannert School of Management

ander650@purdue.edu; ewilens@purdue.edu; nmehring@purdue.edu; kshatkus@purdue.edu; wcoffey@purdue.edu; lanham@purdue.edu

Abstract

This study uses descriptive and predictive analytics techniques to identify the key drivers of home value, and develop an accurate prediction for the price of a home once on the market. The motivation for this investigation is to estimate a fair price for both buyers and sellers in the housing market. Taking into account the wide array of components that affect a home's value and then having to set a fair market price can be a difficult task. Purchasing a home is often one of the largest investments a person or family will make in their lifetime. Thus, listing the home at the right price with minimum error can be key in combining the right buyers and sellers, and allow the home to sell more quickly. Moreover, knowing the major factors that affect a home's value can provide decision-support to sellers in how they might want to invest in home improvement projects before they sell their home. Using 1460 homes purchased between 2006 and 2010 in Ames, Iowa that measured 81 different features about these homes, we identify the key drivers of a home's value. We show that using regularization techniques such as least absolute shrinkage and selection operator (LASSO) can help easily identify those key descriptive drivers of home value. We then show that combining multiple predictive models in an ensemble fashion via meta-modeling can lead to a better prediction than using one predictive model in isolation.

Introduction

Homeowners constantly struggle with analyzing features of their home and deciding what to upgrade prior to selling to maximize home value. Some homeowners may sell their house without performing upgrades, leaving potential cash on the table. Other homeowners may spend excessive capital on upgrades that don't translate to a worthy increase in sales price. Using our model, homeowners will be able to conveniently analyze expected ROI of renovations or attributes about their home.

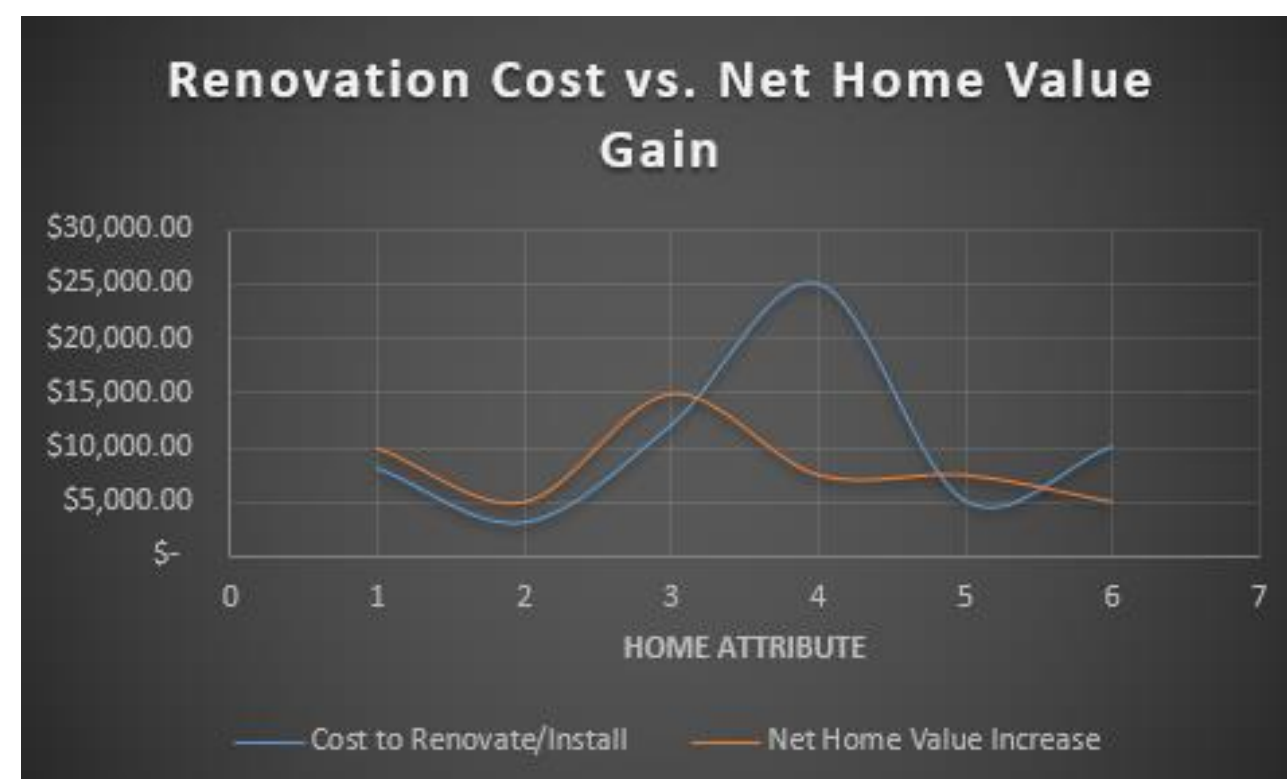


Figure 1. What should a homeowner renovate?

The chart above shows six different attributes, their relative costs and their relative net home value gain. Attributes 1, 2, 3 and 5 show profit potential, whereas attributes 4 and 6 would be detrimental to sales profit.

Literature Review

Home price value can be predicted using various methods. Modelers have used historic recession trends to predict future housing market trends. Other modelers have used online web searching and the effect it has on increasing home prices. Another model uses characteristics and properties throughout a house to help determine and predict sales and listing prices for real estate agents.

Our model is unique because it is created for the homeowner. Our model takes unique attributes into account and allows homeowners to easily analyze the profit potential in renovating and installing certain features. Predictive model's seen online focus on predicting home prices through various methods

intended for real estate agents. Our model has a different focus; The homeowner. This study is novel because we compare and contrast all the methods used previously, but also ensemble them together.

Studies	LASSO	Random Forest	Support Vector Machine	Neural Net	Meta-Model
Wu, Brynjolfsson	✓				
Glaseser, Nathanson				✓	
Ferlan, Nadja			✓		
Imberman, Lovenheim		✓			
Our Study	✓	✓	✓	✓	✓



Methodology

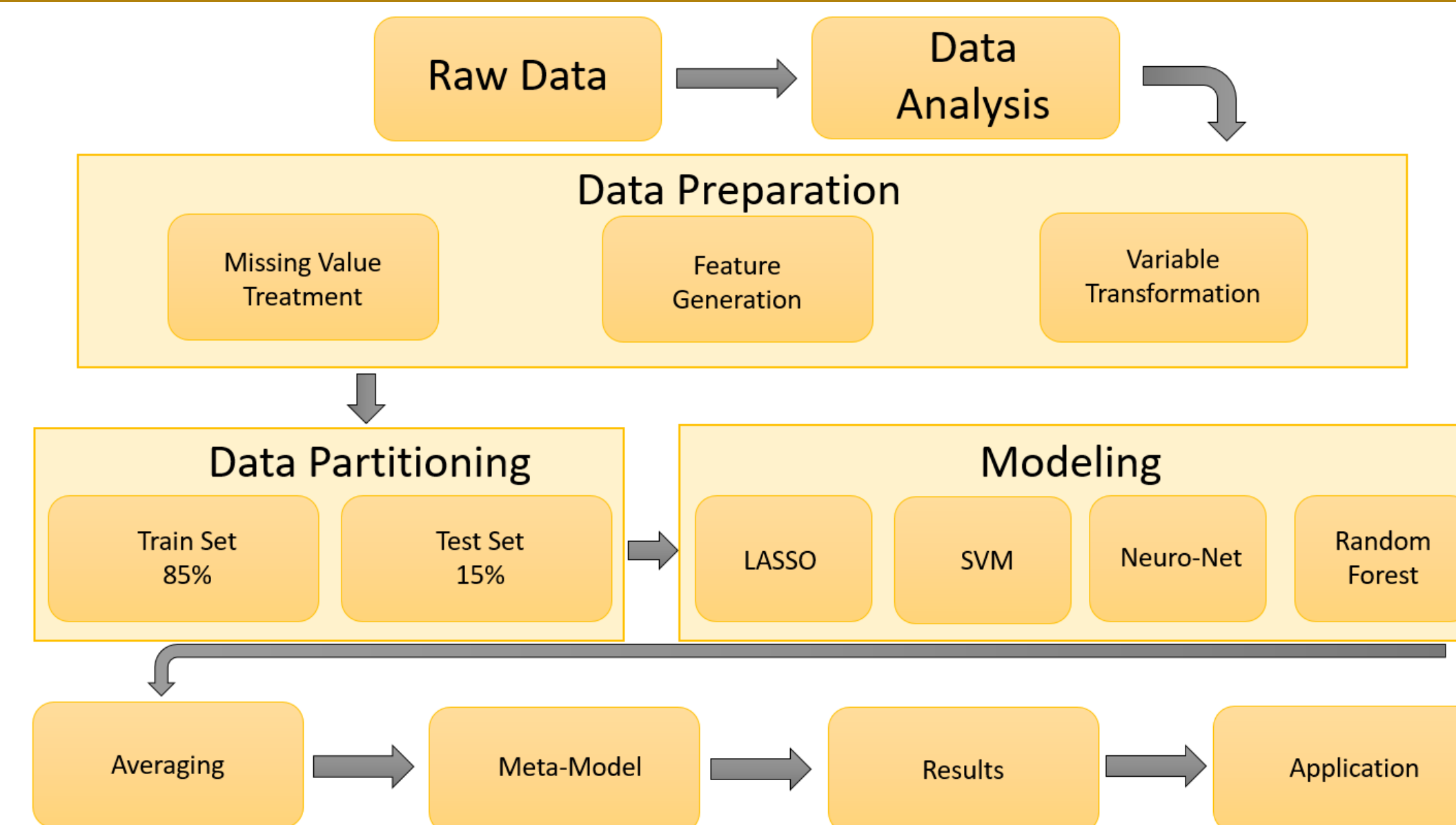


Figure 2 uses Random Forest to show the 8 factors we chose are all at least in the 75th percentile on this graph

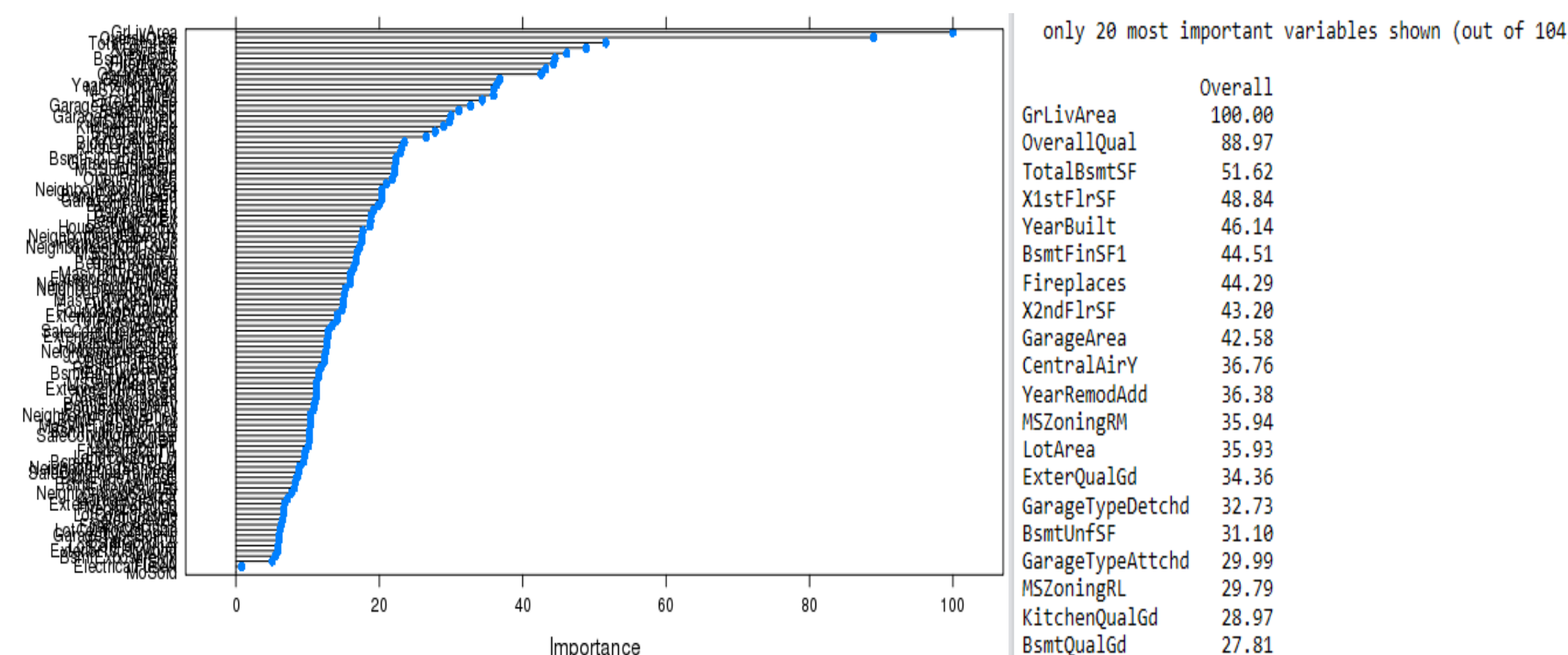
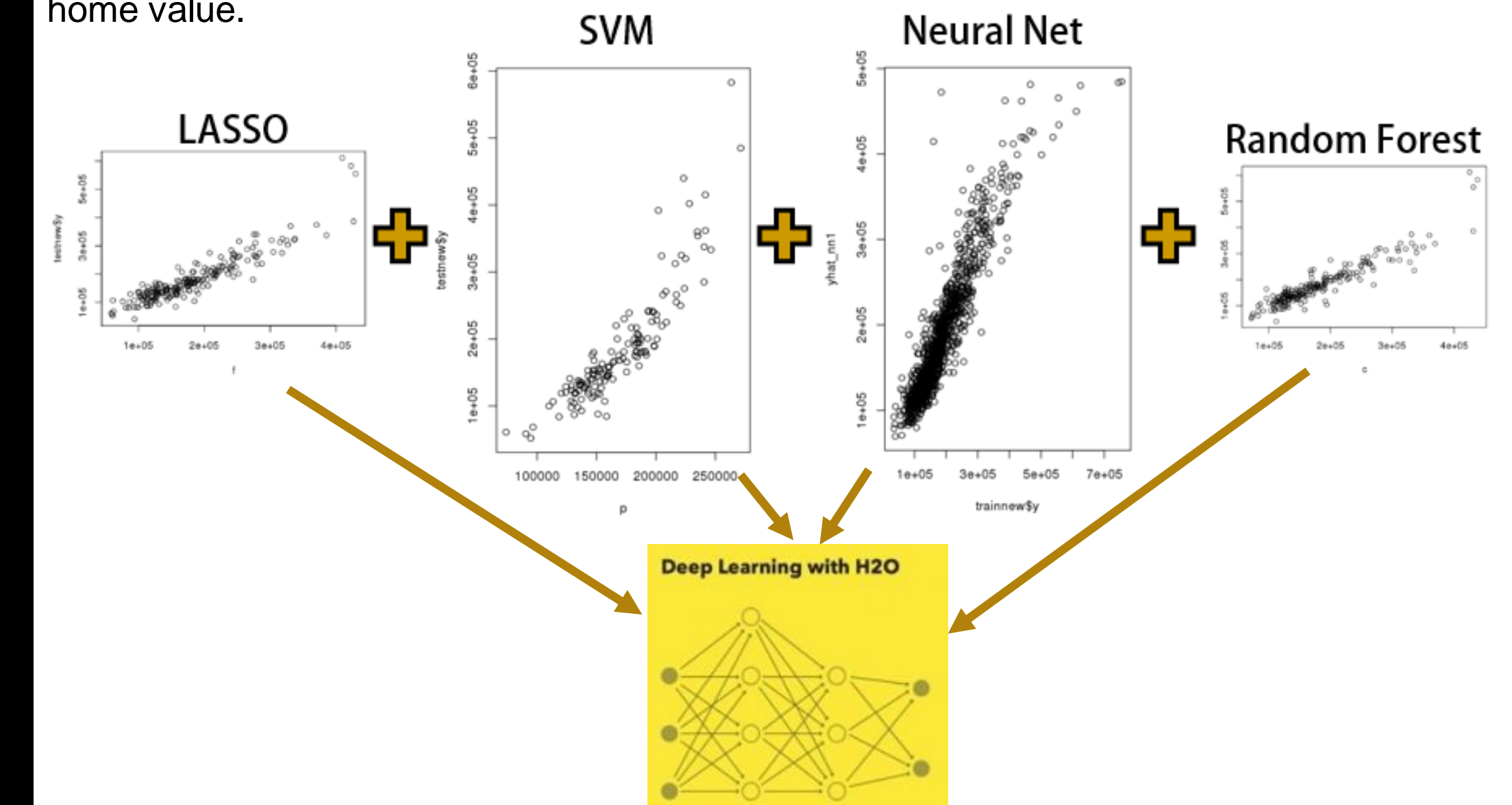


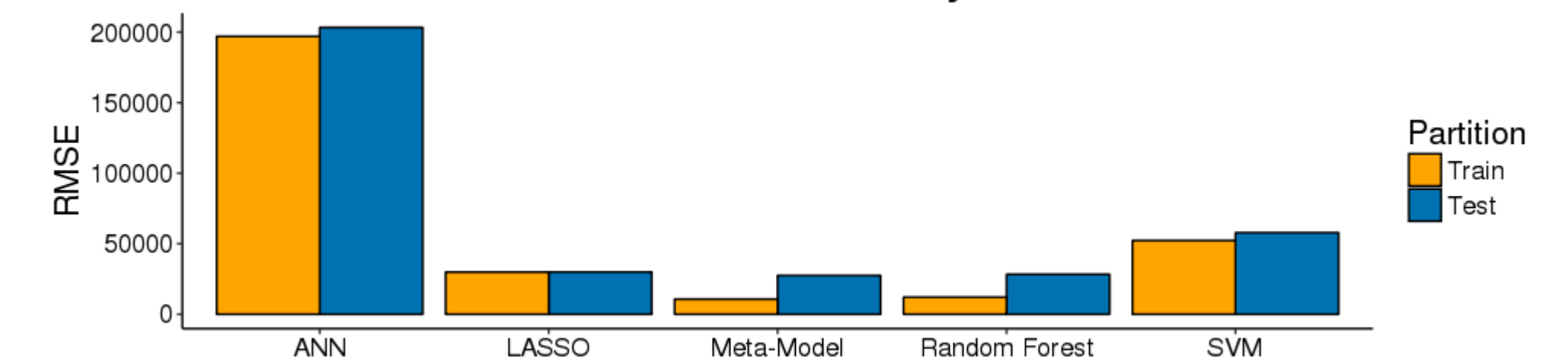
Figure 2: Random Forest, feature importance

Results

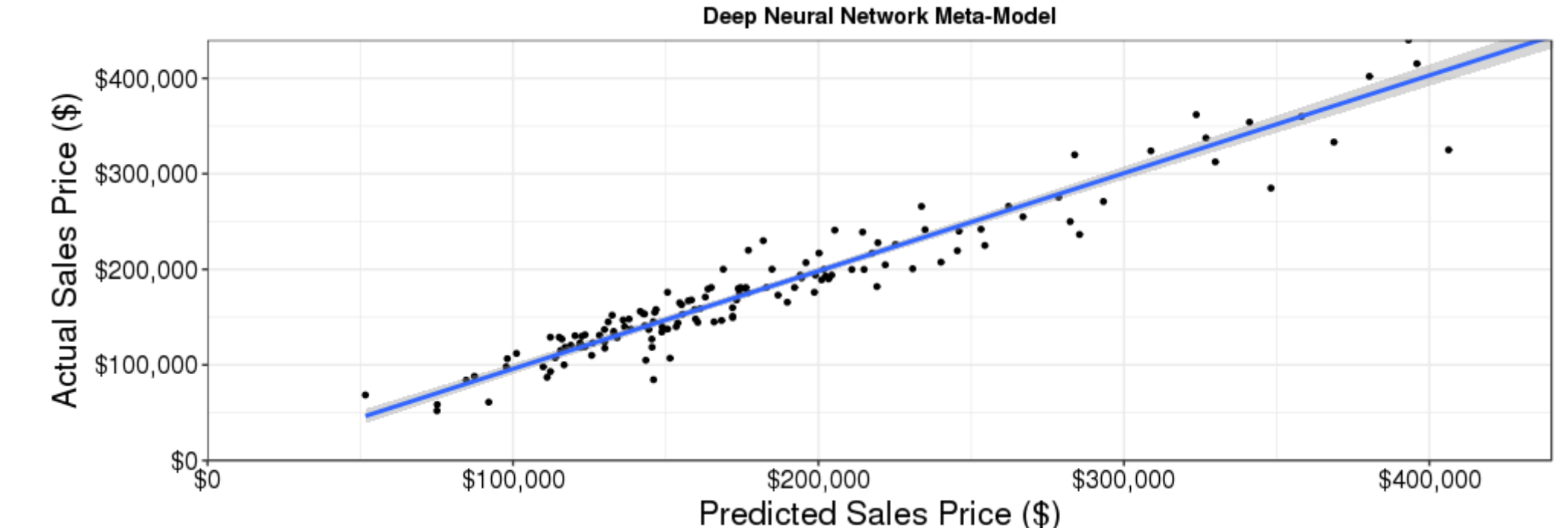
The figures below display the prediction of four different models compared to the actual home value.



Statistical Performance by Model



Best Model Prediction



Conclusions

An accurate prediction of a home price is difficult to come by. Knowing which specific features will increase your home value will always be sought after by home owners.

1. Our model provides an understanding of important features using a Random Forest and LASSO (a regularization technique) and found 8 features in common.
2. We were able to obtain a highly accurate prediction of home price by combining previously studied modeling approaches using a meta-model approach, which had not been investigated previously. We found that the deep learning architecture from h2o to this problem was the best model to learn from previous models.

Acknowledgements

We thank Professor Matthew Lanham for constant guidance on this project.