Draft Technical Paper Working Title

Elizabeth Gabel
Indiana University
eligabel@iu.edu

Alexander Mervar
Indiana University
amervar@iu.edu

Aidan Rosberg
Indiana University
arosberg@iu.edu

Abstract

This paper presents a backpropagation neural network method for predicting the average home cost of a given neighborhood. It builds from previous work in the area, and refines neural parameters and feature space to increase accuracy, and compares performance to a linear regression model and support vector model. It serves to demonstrate the usefulness of neural networks in the problem space.

1 Introduction

The Californian housing market is impacted not only by the features of a given house, but by the unique features of the surrounding landscape. This can mean that buyers do not only look at the qualities of a given home of interest, but also the qualities of the neighborhood the homes reside in that may be unique to the coastal state. While previous papers have focused on individual home costs and features, this paper seeks to demonstrate the efficacy of a backpropagation neural network in predicting the average cost of homes in a neighborhood given the neighborhood's proximity to the ocean, median income, population, and households, along with the average age and home features of the homes themselves.

Previous work examines housing market prediction with both pre-neural and neural methods. In pre-neural methods, regression models such as linear regression, support vector models (SVMs), kohonen neural networks (KNNs), and random forest models have all been used to various degrees of success. Pow et al. (**?**) demonstrated that KNNs and random forest models outperform baseline linear regression and SVMs, and speculate this is likely due to ability to consider a higher vector space and draw connections beyond a linear plane. Later studies, such as Ćetković et al (**?**), examined the efficacy of neural network methods for market prediction, and found the results reasonable enough to continue refinement of parameters and further development of neural network methods in this field.

We seek to continue investigation into how best to harness the processing abilities of neural networks to solve the problem. This paper demonstrates that our neural network outperforms linear regression models, and shows how to refine the parameters of the neural network for best results.

2 Dataset

The dataset that we have selected is a selection of house data used in the second chapter of Aurélien Géron's book 'Hands-On Machine learning with Scikit-Learn and TensorFlow' (**?**). The data contains information from the 1990 California census. Although this not necessarily mean that our model can predict current housing prices, due to the volatility of the current economy due to the COVID-19 pandemic and the nature of this project, it was decided that applying this methodology to current datasets would be outside the scope of this paper.

2.1 Preprocessing Methods

3 Methods

3.1 Linear Regression Model

3.2 SVMs????

3.2.1 Backpropagation Neural Network

4 Results

4.1 Linear Regression Results

4.2 SVM????? Results

4.3 Neural Results

5 Discussion

6 Conclusion

Acknowledgements