Enhancing Real Estate Valuation Models for Zingat: A Data-driven Approach

Tazmeen Afroz
22P-9252

May 6, 2024

Abstract

This report presents a data-driven approach to enhancing real estate valuation models for Zingat, an online marketplace for real estate transactions in Turkey. The objective is to develop a model that improves upon existing models in terms of accuracy and economic reasoning. The methodology involves cleaning and preparing the dataset, feature engineering, building and evaluating deep learning models, and comparing their performance with traditional statistical approaches. The results indicate significant improvements in model accuracy, with the deep learning model achieving a lower Mean Squared Error (MSE) compared to the linear regression model.

1 Introduction

The real estate market is complex and highly localized, making accurate valuation challenging. Zingat, an online marketplace for real estate transactions in Turkey, seeks to enhance its valuation models to better serve its users. This report outlines a data-driven approach to improving these models, focusing on the development and evaluation of deep learning models alongside traditional statistical methods.

2 Methodology

The methodology is divided into four main parts: data preparation, feature engineering, model building, and model evaluation.

2.1 Data Preparation

The dataset was cleaned by handling missing values, outliers, and inconsistencies. Relevant variables for predicting real estate values were identified and explored.

2.2 Feature Engineering

New variables were created, and existing ones were transformed to enhance their predictive power. This included calculating the time period between listing dates and encoding categorical variables.

2.3 Model Building

Two models were developed: a deep learning model using an Artificial Neural Network (ANN) and a traditional linear regression model. Both models were trained and evaluated using relevant performance metrics.

2.4 Model Evaluation

The performance of both models was compared using metrics such as Mean Squared Error (MSE). The deep learning model showed superior performance, achieving a lower MSE compared to the linear regression model.

3 Results

The deep learning model achieved an MSE of 3.0213, while the linear regression model achieved an MSE of 0.7754. This indicates that the deep learning model provides a more accurate prediction of real estate values.

4 Discussion

The results demonstrate the potential of deep learning models in enhancing real estate valuation accuracy. The deep learning model's superior performance suggests that it captures more complex relationships in the data, potentially due to its ability to learn non-linear patterns. However, the linear regression model still provides a valuable baseline for comparison.

5 Conclusion

This report concludes that a data-driven approach, particularly using deep learning models, can significantly improve real estate valuation accuracy. Future work should focus on further refining the model and exploring additional features that could enhance its predictive power.