

PREDICTING RENTAL VALUE OF RESIDENTIAL PROPERTIES IN COLOMBO CITY USING MACHINE LEARNING TECHNIQUES

Uhane Gedara Vinusha Dewshan Kumarathunga

A dissertation submitted in partial fulfilment of the requirement for
Master of Science degree in Business Analytics.

**Department of Business
Informatics Institute of Technology, Sri Lanka
in collaboration with
Robert Gordon University, Aberdeen, Scotland, UK**

2022

Abstract

Real estate is a well stabilized and an actively engaging sector throughout the globe with standardized set of guidelines that are initiated towards the benefit of both landlords and the tenants. This is not the case to be observed in Sri Lanka and there is a doubt whether there is a proper framework to determine and value a price for properties. Sri Lanka being a third world developing country, people find it hard to construct their own house or property which makes them vulnerable towards renting a property. The most common aspect is renting out residential properties and majorly renting out in the capital city due to many reasons. With the increase in demands for jobs within the Colombo city regions, people tend to find residential properties for rent within this area which is more convenient for them in terms of accessibility and the factor of heavy traffic. Also, with the urbanization and the development of resources in the Colombo city limits, temporary migration is a common factor which has been influencing the determination of rental values of properties.

One important aspect of this study is that there is no literature to be found where the residential area has been taken into consideration where the number of housing units per square kilometer obtained and ranked in the descending order. Also, the use of machine learning techniques when it comes to residential rental property value prediction is completely new in the local context.

Hence, in this study major cities within the Colombo city limit were selected to ensure that there is no major deviations are observed in terms of the rental value. Various common parameters were chosen to determine its relationship with the property rental value. Both statistical and Machine Learning approaches are used in this study to construct models and see how precisely it can predict the rental value. Machine Learning techniques like Linear Regression, Ridge Regression, Lasso Regression, Partial Least Square Regression, K-Nearest Neighbor Regression, Decision Tree Regression, Random Forest Regression and Gradient Boost Regression was used. Each of these models were trained and tested and the evaluation matrices were recorded. Evaluation of the models were done in terms of R², RMSE and Cross Validation CV. Hyper parameter tuning was used to determine the best combination of parameters in order to construct the model.

Keywords: Machine Learning, Regression, Feature Engineering, Feature Selection