



Student Name: Tharusha Pasindul Jayadeera	Matriculation Number: 2121988
Supervisor: Dr. Alroy Mascenghe	Second Marker:
Project Title: Machine Learning-based House Price Prediction System for Sri Lanka	
Submission Type: Research Project Report	Start Date: 27 January 2023
	Submission Date: 23 September 2023

CONSENT

I agree
I do not agree

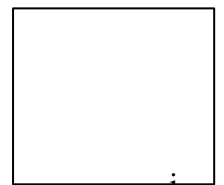
That the University shall be entitled to use any results, materials or other outcomes arising from my project work for the purposes of non-commercial teaching and research, including collaboration.

DECLARATION

I confirm:

- **That the work contained in this document has been composed solely by myself and that I have not made use of any unauthorized assistance.**
- **That the work has not been accepted in any previous application for a degree.**
- **All sources of information have been specifically acknowledged and all verbatim extracts are distinguished by quotation marks.**

Student Signature: 	Date Signed: 23/09/2023
--------------------------------------------------------------------------------------------------------	----------------------------



Abstract

This study undertakes the task of developing a comprehensive house price prediction system for Sri Lanka. The system leverages various machine learning regression algorithms to predict house prices based on a range of features such as land size, house size, no. of bedrooms, no. of bathrooms, city, and district. In-depth data exploration, pre-processing, and feature selection precede the application of regression models, ensuring an optimal dataset for analysis.

The model performance has been evaluated using metrics like Mean Absolute Error and R^2 score, and the best performing models - Gradient Boosting, Random Forest, Extra Trees and K-Nearest Neighbors were integrated into the final system. This system was developed in Django framework, owing to its ease of Python script integration and robustness, offering a user-friendly interface to input the house features and display the average predicted price along with a graphical representation of the prediction results from each model.

Web scraping was employed to automate the data extraction process from online house sale websites, providing an up-to-date dataset for the system. This study contributes to the domain of real estate price prediction by offering a reliable, robust, and user-friendly system to forecast house prices, aiding both sellers and buyers in the market. However, the study acknowledges limitations, including the need for re-training bots for different data sources, negotiation influences on actual prices, and exclusion of environmental variables. Future work is envisioned to overcome these limitations and enhance the model's accuracy and comprehensiveness.