UNIVERSITI TEKNOLOGI MARA

Predictive Modeling of Condominium Prices Using a Particle Swarm Optimization-Random Forest Approach

CHE WAN SUFIA BINTI CHE WAN SAMSUDIN

BACHELOR OF COMPUTER SCIENCE
(Hons.)

FEBRUARY 2025

ACKNOWLEDGEMENT

Alhamdulillah, praises and thanks to Allah because of His Almighty and His utmost blessings, I was able to finish this research within the time duration given. Firstly, my special thanks goes to my supervisor, madam Ummu Fatihah binti Mohd Bahrin.

Special appreciation also goes to my beloved parents Che Wan Samsudin bin Che Wan Sulaiman and Rodah binti Hashim.

Last but not least, I would like to give my gratitude to my dearest friends.

Abstract

The development of a predictive model for condominium prices using the Particle Swarm Optimization-Random Forest approach is the key focus of this research project. This is to establish a trustworthy analytical tool for buyers, sellers, and other stakeholders in the real estate market to gauge condominium costs with great accuracy. Essential phases of the project include data collection, data preprocessing, and the implementation of the Particle Swarm Optimization-Random Forest price prediction algorithm. Both simulated and real-world experiments are used as a basis to rigorously test and validate the predictive capability of the model. The results of the study demonstrate the effectiveness of the proposed model. Specifically, with a dataset split of 90-10, the model achieved an RMSE of 81699.40, an MAE of 136058.74, and an R-squared value of 0.8835. These findings indicate a high level of predictive accuracy, showcasing the robustness of the model in estimating condominium prices. This research will lead to the enrichment of our understanding of how various factors affect condominium prices, with the hope that users will become more informed in their real estate market decisions. The model is highly practical and easy to interpret, making it very suitable for real-world applications. The methods used in this simulation study will be summarized within a comprehensive report, along with the process used in evaluating the appropriateness of using search heuristics in tandem with parallel systems and prospective real-world applications. In general, the condominium price prediction model developed using the Particle Swarm Optimization-Random Forest approach has high value for all real estate sector stakeholders, such as legal professionals, investors, and real estate developers, as it provides accurate price forecasts and practical insights.

TABLE OF CONTENTS

CHAPTER ONE 1

1.1 Background of Study 1

1.2 Problem Statement 2

1.3 Objective 5

1.4 Project Scope 5

1.5 Project Significance 6

1.6 Overview of Research Framework 7

1.7 Conclusion 9

CHAPTER TWO 10

2.1 Predictive Modeling of Condominium Prices 10
2.1.1 Benefits of Predictive Modeling 11

2.2 Condominium Prices 12

2.3 Random Forest Algorithm 14
2.3.1 Overview of Random Forest 14
2.3.2 Features of Random Forest 14
2.3.3 Applications for Random Forest in Real Estate 15

2.4 Particle Swarm Optimization (PSO) 16
2.4.1 Overview of PSO 16
2.4.2 Integrating PSO with Random Forest 18

2.5 Implementation of Random Forest and PSO Algorithms in Various Domains 21
2.5.1 Implementation of Random Forest Algorithm 21

2.5.2 Implementation of PSO Algorithm 25

2.6 Implementation of PSO-Random Forest for Price Prediction 29

2.7 Similar Works and Comparative Analysis 34
2.7.1 Overview of Similar Works 34

2.9 CONCLUSION 41

CHAPTER THREE 43

3.1 Overview of Research Framework 43
3.1.1 Detailed of Research Framework 43

3.2 Preliminary phase 44
3.2.2.1 Data Preparation 47
3.2.2.2 Data Cleaning 51

3.3 Design & Modelling Phase 51
3.3.1 System Architecture 54
3.3.3 User Interface Design 58
3.3.4 Pseudocode of Selected Algorithm 59
3.3.5. Prototype Implementation 62

3.4 Evaluation Phase 63

3.5 Gantt chart 66

3.6 Conclusion 66

CHAPTER FOUR 68

4.1 Overview of Project System 68

4.2 Prototype Development 70
4.2.1 Data Understanding 70
4.2.2 Data Pre-Processing 70
4.2.2.1 Data Cleaning 70
4.2.2.2 Data Transformation 73
4.2.2.3 Outlier Detection and Handling 74
4.2.2.4 Raw Data and Cleaned Data 76
4.2.3 Features Selection using Particle Swarm Optimization 76
4.2.4 Data Splitting 84
4.2.5 Splitting, Training and Testing the Selected Features 85
4.2.6 Evaluation 93
4.2.6.1 Evaluation of Features Selection using PSO 93
4.2.6.2 PSO Hyperparameter Tuning Experiment 98
4.2.6.3 Evaluation Results Training and Testing Split 101
Experiments 101
4.2.7 Condominium Price Prediction Prototype 105

4.3 Functionality Testing 106

4.4 Knowledge Acquisition 106

4.5 Discussion 107

4.6 Conclusion 107