

Project Report: Singapore Resale Flat Prices Prediction

Table of Contents

1. Introduction
2. Objective
3. Data Collection and Processing
4. Data Analysis
5. Model Development
6. Model Evaluation
7. Deployment
8. Conclusion
9. Future Work
10. References

1. Introduction

The purpose of this project is to develop a predictive model to estimate the resale prices of flats in Singapore. This model helps potential buyers, real estate agents, and investors make informed decisions based on data-driven insights.

2. Objective

The main objective of this project is to predict the resale prices of flats in Singapore based on various features such as floor area, lease duration, and proximity to MRT stations.

3. Data Collection and Processing

Data Collection

The dataset used in this project was collected from various real estate listings in Singapore. It includes information on flat characteristics, location data, and resale prices.

Data Cleaning and Processing

- Loading Data: Data was loaded from multiple CSV files into a pandas DataFrame.
- Handling Missing Values: Missing values were identified and handled by removing rows with missing data.
- Feature Engineering: New features such as `storey_median` were created to assist in the predictive modeling process.

Geospatial Data

Coordinates for various MRT stations and addresses were obtained via API requests to the OneMap Singapore API to calculate distances relevant to the prediction model.

4. Data Analysis

Data analysis involved exploring the relationships between different features and the target variable (resale price). Statistical analyses and visualizations were used to understand distributions and correlations within the data.

5. Model Development

- **Model Selection:** A Decision Tree Regressor was chosen for its ability to model complex non-linear relationships.
- **Training the Model:** The model was trained on a dataset prepared through feature scaling and selection of relevant features.
- **Hyperparameter Tuning:** Grid search was used to find the optimal settings for the model's hyperparameters.

6. Model Evaluation

- **Evaluation Metrics:** The model's performance was evaluated using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared.
- **Results:** The model achieved an R-squared value of 84%, indicating a good fit for the data.

7. Deployment

The model was deployed as a Streamlit web application to provide a user-friendly interface for predictions.

- **Streamlit Application:** The application allows users to input the features of a flat and receive a predicted resale price.
- **Model Loading:** The trained model and scaler were serialized and loaded into the application for making predictions.

8. Conclusion

This project successfully developed and deployed a predictive model for estimating the resale prices of flats in Singapore. The model provides valuable insights for stakeholders in the real estate market.

9. Future Work

- **Model Improvement:** Consider exploring more advanced models such as ensemble methods or deep learning for potentially better predictive performance.
- **Feature Expansion:** Include additional features such as economic indicators, and neighborhood characteristics.
- **Real-time Data Integration:** Integrate real-time data for more up-to-date predictions.

10. References

- Documentation and APIs used:
 - [OneMap Singapore API](#)

- [Streamlit Documentation](#)
- [Scikit-learn Documentation](#)