

# Report: Comparing Random Forest and Decision Tree Models for Housing Price Prediction

## Introduction

In this report, we compare the performance of Random Forest and Decision Tree models for predicting housing prices using the provided dataset. The dataset contains various attributes of houses such as transaction date, house age, latitude, longitude, and house price of unit area.

## Methodology

- Data Preparation: The dataset was split into features (X) and the target variable (y), followed by splitting the data into training and testing sets with a ratio of 80:20.
- Model Training:  
Decision Tree: A Decision Tree model was trained using the training data.  
Random Forest: A Random Forest model was trained using the training data.
- Model Evaluation:  
Both models were evaluated on the testing set using Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) metrics.
- Comparison:  
The performance of Random Forest and Decision Tree models was compared based on the evaluation metrics.
- Results  
Decision Tree RMSE: 8.153  
Random Forest RMSE: 5.693

## Conclusion

Based on the evaluation results, the Random Forest model outperforms the Decision Tree model in terms of RMSE, with a lower value indicating better predictive performance.

The Random Forest model is preferred for analyzing the data due to its lower RMSE value, suggesting better accuracy in predicting housing prices.

## Future Considerations

- Further optimization of the Random Forest model through hyperparameter tuning.
- Exploration of additional features or data preprocessing techniques to improve model performance.

Github link for the code: <https://github.com/alyakbar/Assignment2ML>