

# Lab 02 Report – Regression

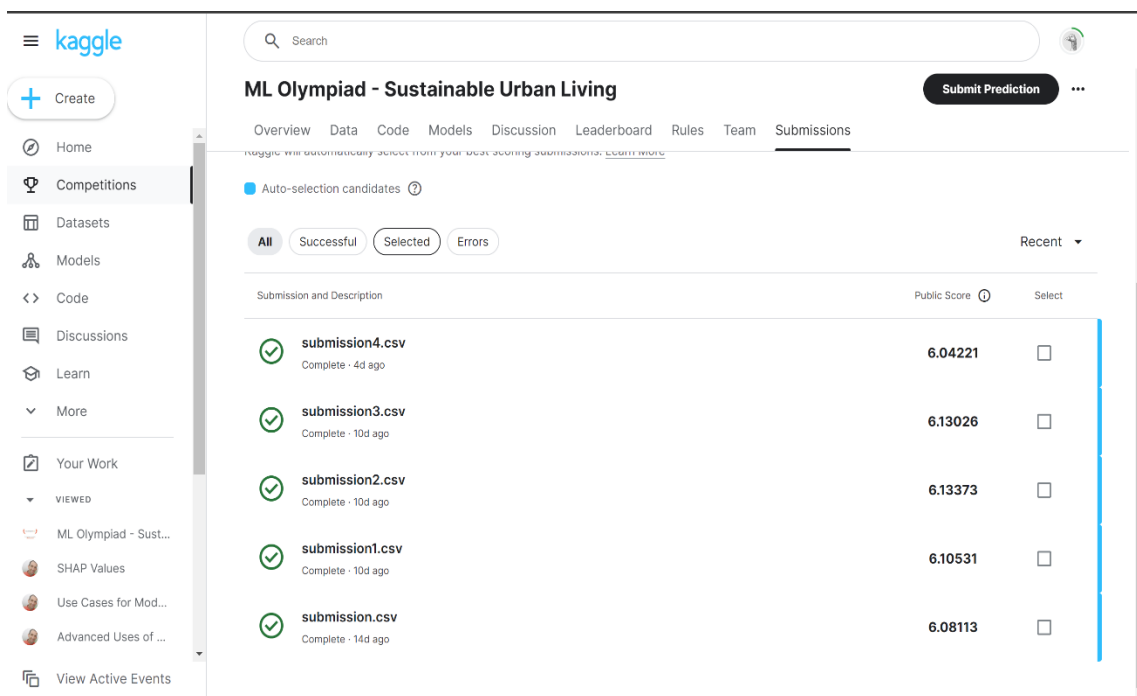
W.L.K.D.D.A.Wickramasinghe

210711T

## Introduction

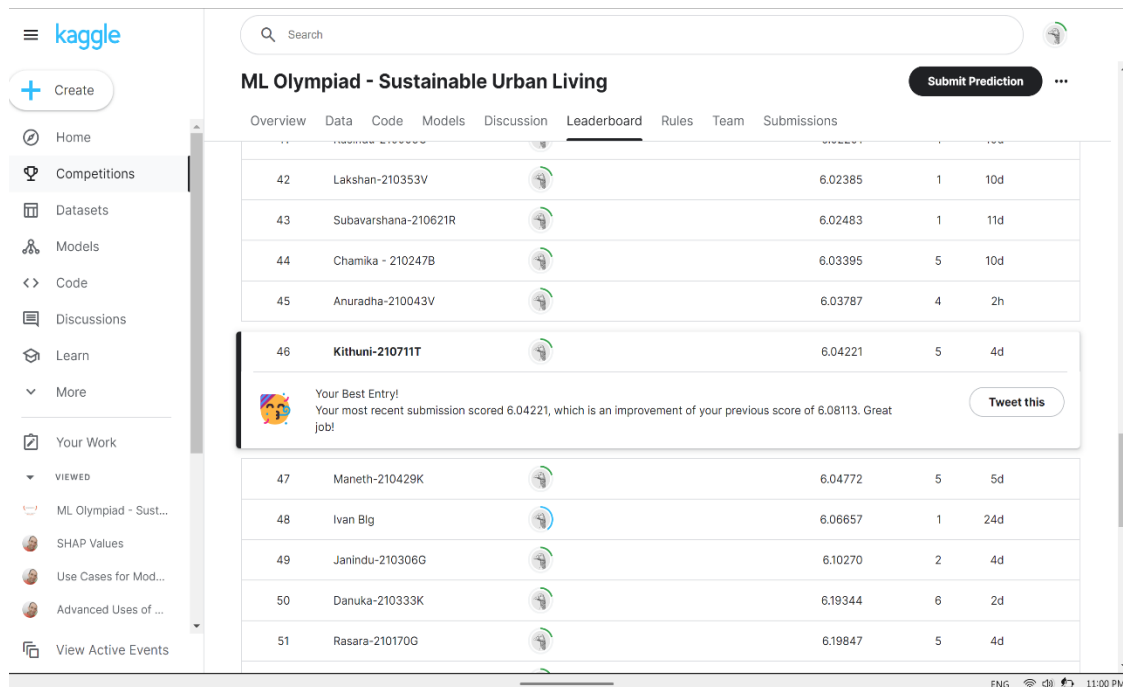
This report aims to compare the performance of four different regression model approaches applied to the dataset. The evaluation of these approaches is based on various regression evaluation metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared ( $R^2$ ) score, and Adjusted R-squared.


## Proofs of the Submissions



Submission and Description		Public Score	Select
✓	submission4.csv Complete · 4d ago	6.04221	<input type="checkbox"/>
✓	submission3.csv Complete · 10d ago	6.13026	<input type="checkbox"/>
✓	submission2.csv Complete · 10d ago	6.13373	<input type="checkbox"/>
✓	submission1.csv Complete · 10d ago	6.10531	<input type="checkbox"/>
✓	submission.csv Complete · 14d ago	6.08113	<input type="checkbox"/>

## Score and the rank.



	Overview	Data	Code	Models	Discussion	Leaderboard	Rules	Team	Submissions
42	Lakshan-210353V					6.02385	1	10d	
43	Subavarshana-210621R					6.02483	1	11d	
44	Chamika - 210247B					6.03395	5	10d	
45	Anuradha-210043V					6.03787	4	2h	
46	<b>Kithuni-210711T</b>					6.04221	5	4d	
<div> Your Best Entry! Your most recent submission scored 6.04221, which is an improvement of your previous score of 6.08113. Great job!</div> <div>Tweet this</div>									
47	Maneth-210429K					6.04772	5	5d	
48	Ivan Blg					6.06657	1	24d	
49	Janindu-210306G					6.10270	2	4d	
50	Danuka-210333K					6.19344	6	2d	
51	Rasara-210170G					6.19847	5	4d	

## Evaluation metrics

### 1. Random Forest Regressor:

- Mean Squared Error (MSE): 36.29
- Root Mean Squared Error (RMSE): 6.02
- Mean Absolute Error (MAE): 4.64
- R-squared ( $R^2$ ) Score: 0.82
- Adjusted R-squared: 0.82

### 2. XGBoost Regressor:

- Mean Squared Error (MSE): 36.35
- Root Mean Squared Error (RMSE): 6.03
- Mean Absolute Error (MAE): 4.67
- R-squared ( $R^2$ ) Score: 0.8
- Adjusted R-squared: 0.82

### 3. LightGBM Regressor:

- Mean Squared Error (MSE): 37.48
- Root Mean Squared Error (RMSE): 6.12
- Mean Absolute Error (MAE): 4.81
- R-squared ( $R^2$ ) Score: 0.81
- Adjusted R-squared: 0.81

#### 4. Linear Regressor:

- Mean Squared Error (MSE): 86.46
- Root Mean Squared Error (RMSE): 9.30
- Mean Absolute Error (MAE): 7.45
- R-squared ( $R^2$ ) Score: 0.57
- Adjusted R-squared: 0.57

### Comparison of results

- Random Forest, XGBoost, and LightGBM regressors demonstrate similar performance in terms of MSE, RMSE, MAE,  $R^2$  score, and Adjusted  $R^2$  score.
- These tree-based ensemble methods outperform the Linear Regressor across all evaluation metrics.
- Random Forest and XGBoost regressors achieve the lowest MSE, RMSE, and MAE, indicating superior predictive accuracy and smaller prediction errors compared to LightGBM and Linear regression models.
- Linear Regressor exhibits the highest MSE, RMSE, and MAE values, signifying poorer predictive performance and larger prediction errors compared to tree-based models.
- The  $R^2$  scores and Adjusted  $R^2$  scores are highest for Random Forest and XGBoost regressors, indicating better model fit and higher explanatory power compared to LightGBM and Linear regression models.

### Conclusion

In conclusion, tree-based ensemble methods such as Random Forest and XGBoost regressors outperform traditional linear regression and LightGBM regressors in terms of prediction accuracy and model fit. These findings suggest that ensemble methods are more effective for modeling complex relationships in the dataset, resulting in improved predictive performance and better explanatory power.