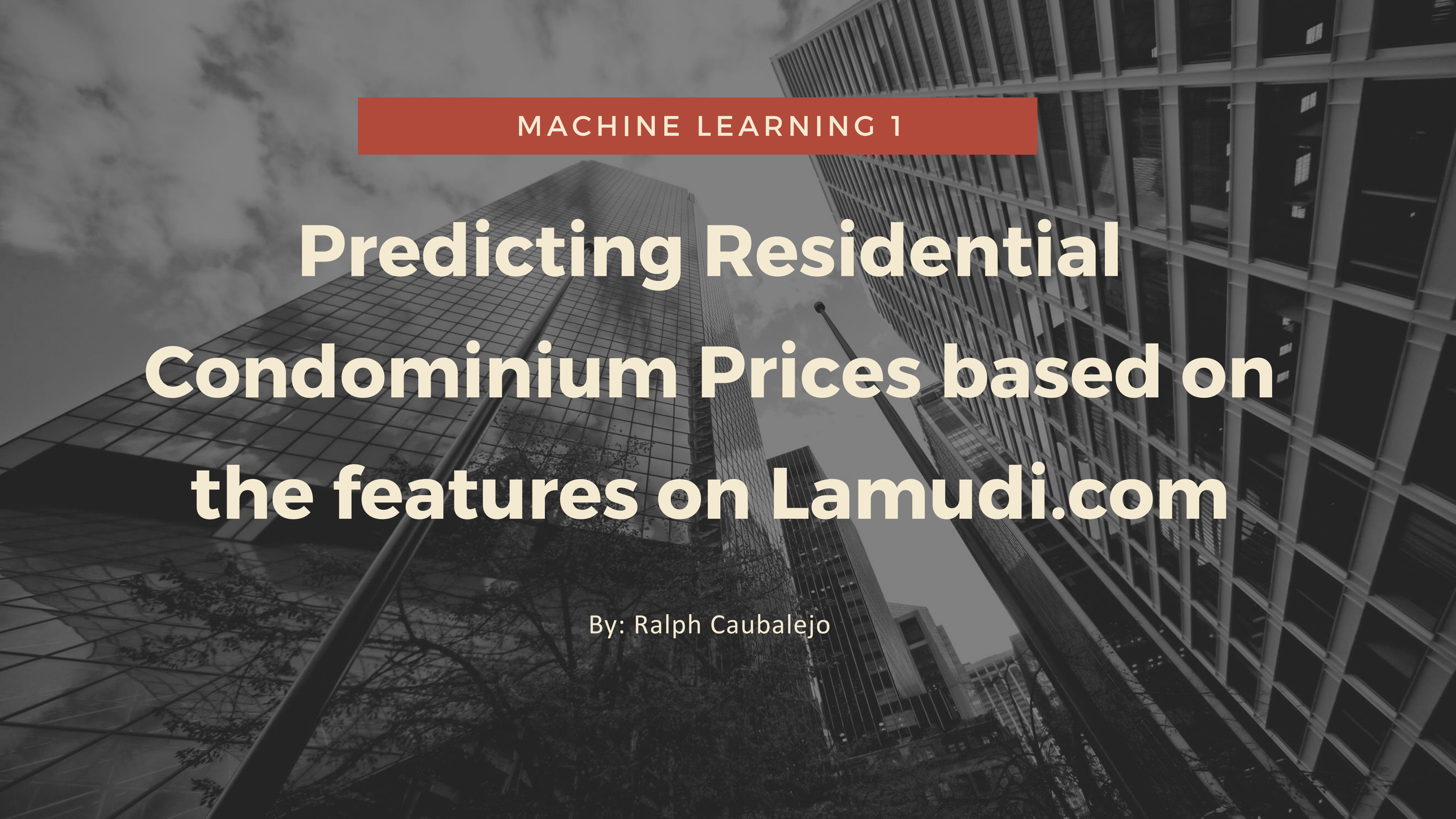
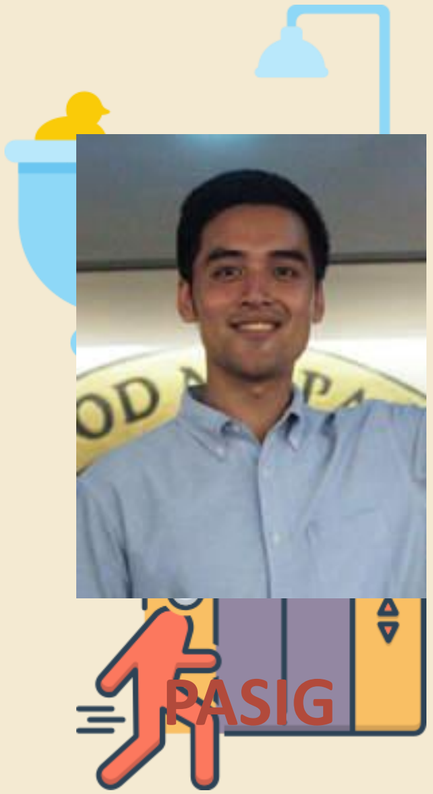


MACHINE LEARNING 1

Predicting Residential Condominium Prices based on the features on Lamudi.com

By: Ralph Caubalejo

Why Condo?

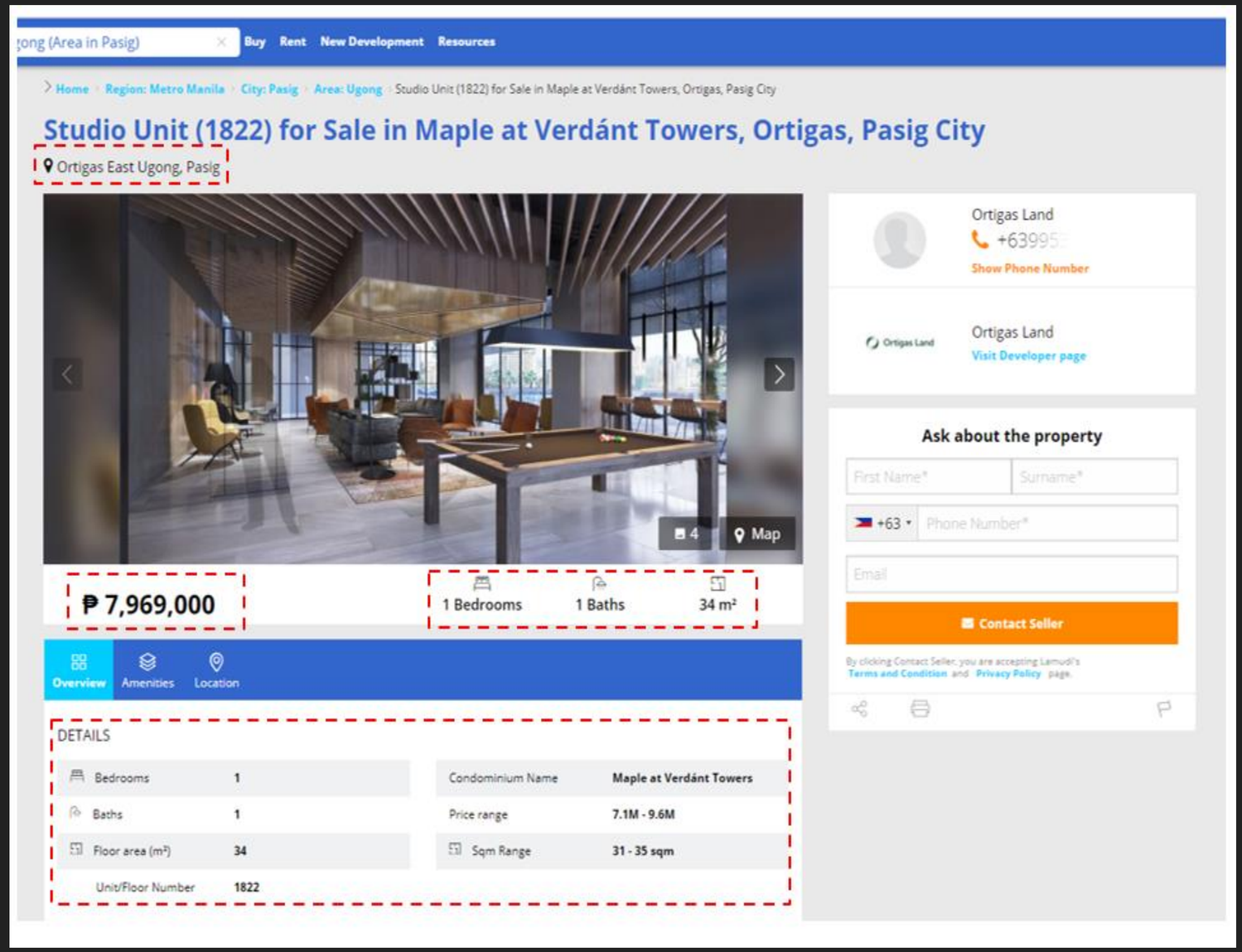


Can we predict condominium prices based on its features/location?

Lamudi.com

An Online Real Estate Marketplace

- Expanded in Philippines in 2014
- 6.6 Million Monthly Pageviews
- 140,000 Monthly Leads
- Currently Operates in 34 Countries



References: https://www.lamudi.com.ph/about-us/

Dataset

6,521 Unique Residential Condominium Listings

Preprocessing

a) Drop Null Values

b) Drop Anomalous Rows

c) No Scaling

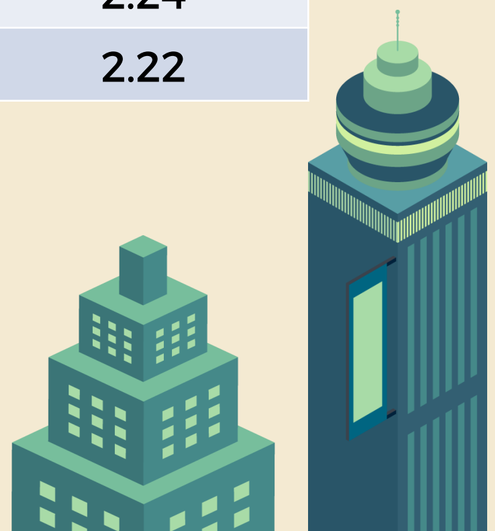
FEATURES

1. City
2. No. of Bedrooms
3. No. of Bathrooms
4. Floor Area
5. No. of Amenities
6. Latitude Coordinate
7. Longitude Coordinate
8. Floor Number
9. Price/Amount - Target

ML Models

Fitting the Data to the applicable models

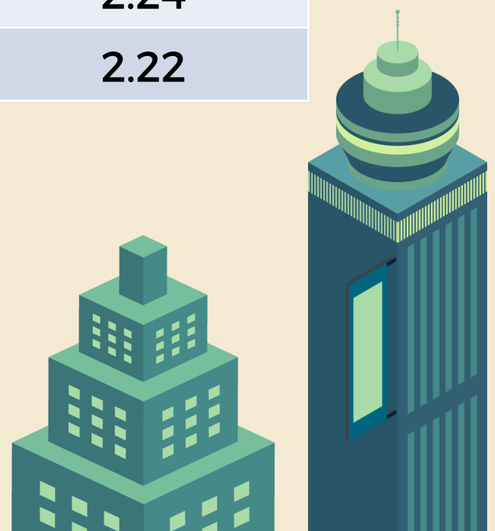
ML MODEL W/ Optimized Parameters	Train Accuracy (R2 Score)	Test Accuracy (R2 Score)	MAE (in millions)	RMSE (in millions)
Linear	62.25 %	63.05 %	2.45	3.55
KNN	69.40 %	64.54 %	2.32	3.52
Decision Tree	83.52 %	75.58 %	1.78	2.87
Random Forest	89.99 %	83.34 %	1.53	2.20
Gradient Boosting	93.73 %	85.31 %	1.42	2.17
Extra Trees	90.44 %	83.14 %	1.58	2.21
XGB	90.56 %	84.45 %	1.48	2.24
Lightboost	90.79 %	84.93 %	1.44	2.24
Catboost	89.02%	84.00 %	1.51	2.22



ML Models

Fitting the Data to the applicable models

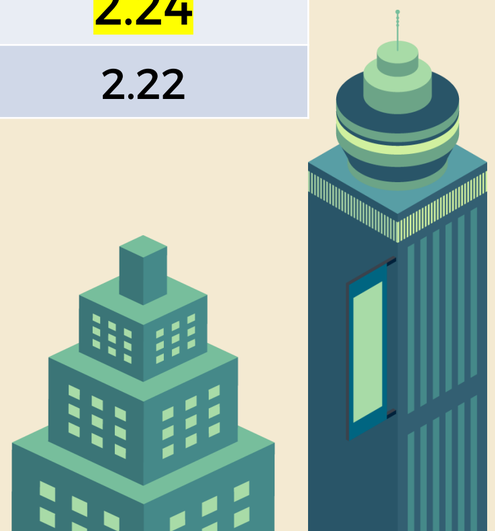
ML MODEL W/ Optimized Parameters	Train Accuracy (R2 Score)	Test Accuracy (R2 Score)	MAE (in millions)	RMSE (in millions)
Linear	62.25 %	63.05 %	2.45	3.55
KNN	69.40 %	64.54 %	2.32	3.52
Decision Tree	83.52 %	75.58 %	1.78	2.87
Random Forest	89.99 %	83.34 %	1.53	2.20
Gradient Boosting	93.73 %	85.31 %	1.42	2.17
Extra Trees	90.44 %	83.14 %	1.58	2.21
XGB	90.56 %	84.45 %	1.48	2.24
Lightboost	90.79 %	84.93 %	1.44	2.24
Catboost	89.02%	84.00 %	1.51	2.22



ML Models

Fitting the Data to the applicable models

ML MODEL W/ Optimized Parameters	Train Accuracy (R2 Score)	Test Accuracy (R2 Score)	MAE (in millions)	RMSE (in millions)
Linear	62.25 %	63.05 %	2.45	3.55
KNN	69.40 %	64.54 %	2.32	3.52
Decision Tree	83.52 %	75.58 %	1.78	2.87
Random Forest	89.99 %	83.34 %	1.53	2.20
Gradient Boosting	93.73 %	85.31 %	1.42	2.17
Extra Trees	90.44 %	83.14 %	1.58	2.21
XGB	90.56 %	84.45 %	1.48	2.24
Lightboost	90.79 %	84.93 %	1.44	2.24
Catboost	89.02%	84.00 %	1.51	2.22



ML Models

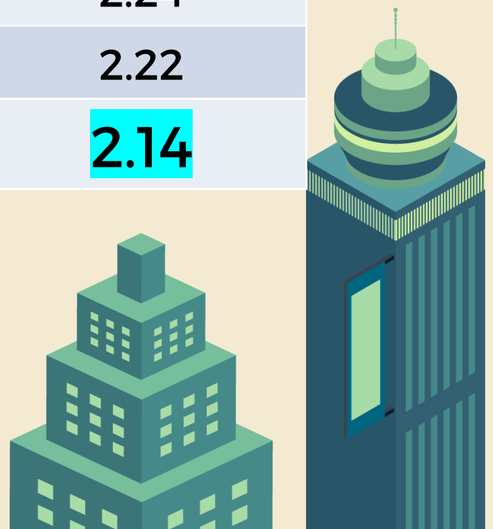
Fitting the Data to the applicable models

ML MODEL W/ Optimized Parameters	Train Accuracy (R2 Score)	Test Accuracy (R2 Score)	MAE (in millions)	RMSE (in millions)
Linear	62.25 %	63.05 %	2.45	3.55
KNN	69.40 %	64.54 %	2.32	3.52
Decision Tree	83.52 %	75.58 %	1.78	2.87
Random Forest	89.99 %	83.34 %	1.53	2.20
Gradient Boosting	93.73 %	85.31 %	1.42	2.17
Extra Trees	90.44 %	83.14 %	1.58	2.21
XGB	90.56 %	84.45 %	1.48	2.24
Lightboost	90.79 %	84.93 %	1.44	2.24
Catboost	89.02%	84.00 %	1.51	2.22
Voting	92.58 %	85.46 %	1.41	2.14

85.46%

Voting Regressor Method

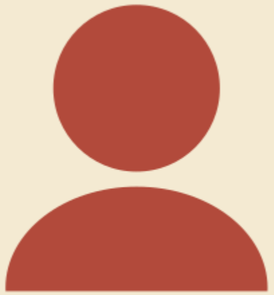
Using GBM and LB as the Base Estimators



Who can benefit from this model?

Real Estate Companies

Real Estate Agents



Buyers

Thank you

“Love is Patient, Love is Kind,
It always protects, trusts, hopes and perseveres”
1 Corinthians 13: 4-8a,NIV