

PROYEK MODELING PREDIKSI HARGA JUAL APARTEMEN DI KOTA DAEGU, KOREA SELATAN

By Evan Yonathan Tejasukmana





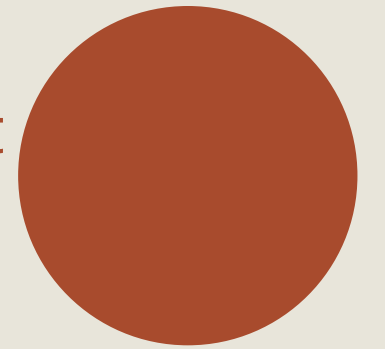
Contents

01	-----	BUSINESS PROBLEM UNDERSTANDING
02	-----	DATA UNDERSTANDING
03	-----	DATA PREPROCESSING
04	-----	MODELING
05	-----	CONCLUSION
06	-----	RECOMMENDATION

BUSINESS PROBLEM UNDERSTANDING

Context

Apartemen adalah salah satu solusi untuk kebutuhan perumahan masyarakat modern akibat terbatasnya lahan hunian dan padatnya aktivitas bisnis di daerah perkotaan.



Problem Statement

Tantangan utama adalah menentukan harga yang sesuai dengan pasar.

Goals

Membuat model prediktif untuk memprediksi harga jual apartemen (SalePrice) berdasarkan fitur-fitur yang tersedia

Analytic Approach

kita akan membangun suatu model regresi

Metric Evaluation

Evaluasi metrik yang akan digunakan adalah RMSE, MAE, dan MAPE.

DATA UNDERSTANDING

Dataset training = dataset Daegu Apartment

Attribute	Data Type	Description
Hallway Type	Object	Jenis lorong atau tipe apartemen.
TimeToSubway	Object	Waktu yang dibutuhkan untuk mencapai stasiun subway terdekat.
SubwayStation	Object	Nama stasiun subway terdekat.
N_FacilitiesNearBy(ETC)	Float	Jumlah fasilitas umum di sekitar.
N_FacilitiesNearBy(PublicOffice)	Float	Jumlah fasilitas kantor layanan publik di sekitar.
N_SchoolNearBy(University)	Float	Jumlah universitas di sekitar.
N_Parkinglot(Basement)	Float	Jumlah tempat parkir (di basement).
YearBuilt:	Integer	Tahun dibangunnya apartemen.
N_FacilitiesInApt	Integer	Jumlah fasilitas yang ada di dalam apartemen.
Size(sqft)	Integer	Ukuran apartemen (dalam kaki persegi).
SalePrice	Integer	Harga apartemen (dalam Won).

DATA PREPROCESSING

MISSING VALUE

DATA DUPLICATION

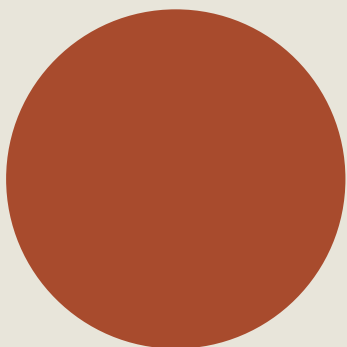
DATA CORRELATION

OUTLIERS DETECTION

MODELING

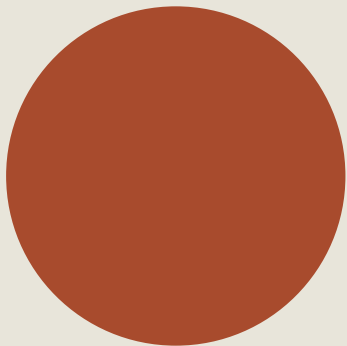
- ENCODING : One Hot Encoding
- SCALING : Standart Scaller
- MODEL : XGBoost Regressor (TheBestModel)
- HYPERPARAMETER TUNING : RandomizedSearchCV

MODELS EVALUATION



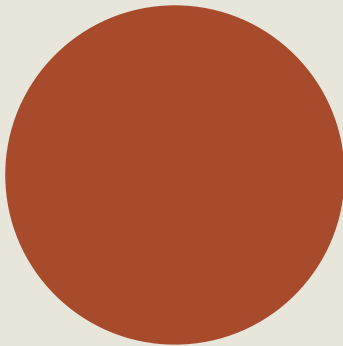
	Model	Mean_RMSE	Std_RMSE	Mean_MAE	Std_MAE	Mean_MAPE	Std_MAPE
0	Linear Regression	-52437.463406	2002.906996	-42119.180060	1560.070630	-0.213133	0.015114
1	KNN Regressor	-50190.395549	809.284301	-39471.735387	848.008392	-0.198765	0.006249
2	DecisionTree Regressor	-45712.082497	382.214777	-36525.602117	280.772195	-0.185379	0.006654
3	RandomForest Regressor	-45798.441325	309.573928	-36569.598230	312.106837	-0.186141	0.008251
4	XGBoost Regressor	-45731.065410	347.912909	-36530.409185	310.683979	-0.185700	0.007512

MODELS EVALUATION



	RMSE	MAE	MAPE
XGB	44318.596679	35616.073711	0.190959
RandomForest	44350.019375	35646.422424	0.191390
DecisionTree	44319.306591	35615.777209	0.190948

RESULT EVALUATION



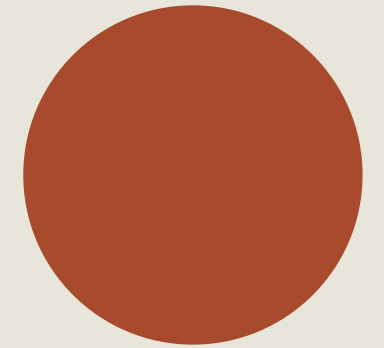
BEFORE TUNING

	RMSE	MAE	MAPE
XGB	44318.596679	35616.073711	0.190959

AFTER TUNING

	RMSE	MAE	MAPE
XGB	43967.939412	35399.184356	0.190352

PREDICTION RESULT

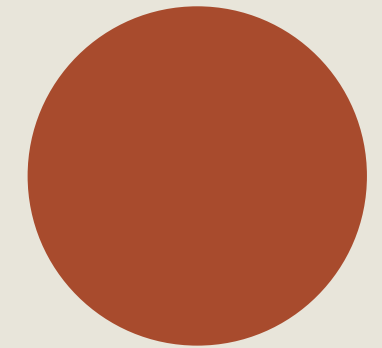


berikut merupakan dataset yang kita Punya:

- 'HallwayType': mixed
- 'TimeToSubway': 10min~15min
- 'SubwayStation': Daegu
- 'N_FacilitiesNearBy(ETC)': 10
- 'N_FacilitiesNearBy(PublicOffice)': 5
- 'N_SchoolNearBy(University)': 3
- 'N_Parkinglot(Basement)': 2
- 'YearBuilt': 2005
- 'N_FacilitiesInApt': 15
- 'Size(sqf)': 1200 sqf

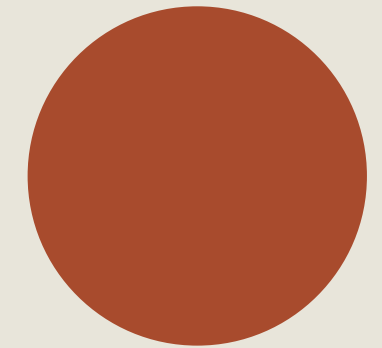
Hasil prediksi harga jualnya sebesar 219,896.62 Won

CONCLUSION



Model sudah berhasil memprediksi harga apartemen menggunakan metrik yang cukup akurat, dengan perkiraan harga rata-ratanya akan meleset kurang lebih sebesar 19% dari harga seharusnya.

RECOMMENDATION



- Menambahkan fitur yang lebih korelatif dengan target ('SalePrice'), seperti harga tanah dan lain-lain.
- Menggunakan model yang lebih kompleks untuk penambahan data yang banyak, seperti recursive neural networks (RNN).
- Model yang sudah dibangun ini bisa dimanfaatkan untuk pengembangan pembuatan model lainnya.

CLOSING

“You can have data without information,
but you cannot have information without data.”

- Daniel Keys Moran