



DAEGU PROPERTY

Daegu Apartment Price analysis
using Machine Learning

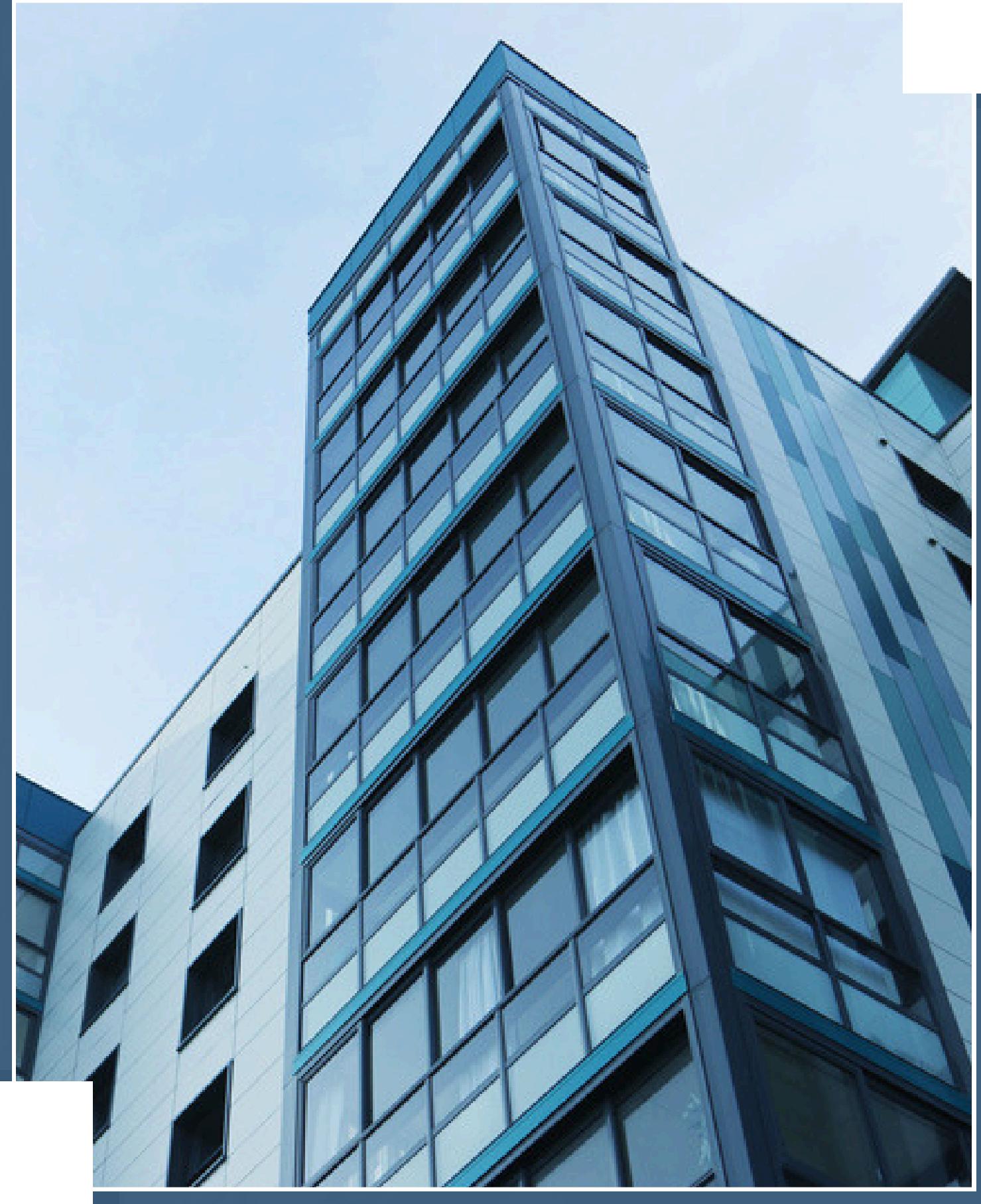


TABLE OF CONTENT

- AGENCY PROPERTY BACKGROUND
- DATA UNDERSTANDING
- EXPLORATORY DATA ANALYSIS
- CONCLUSION
- CLOSING



AGENCY PROPERTY BACKGROUND

Daegu Properti bergerak di bidang jual-beli apartemen sebagai solusi hunian modern di tengah keterbatasan lahan perkotaan.

Namun, perusahaan sering menghadapi tantangan dalam menentukan harga beli dan jual unit yang sesuai dengan kondisi pasar.

Perusahaan khawatir penetapan harga yang terlalu tinggi berisiko membuat unit sulit terjual, sementara harga yang terlalu rendah dapat mengurangi potensi keuntungan. Di sisi lain, saat perusahaan ingin membeli unit apartemen juga ingin memastikan harga yang dibayarkan tidak melebihi nilai pasar agar investasi tetap menguntungkan.

Untuk mengatasi tantangan tersebut, perusahaan ingin mencoba memanfaatkan Machine Learning guna membantu memprediksi harga apartemen secara lebih objektif dan akurat, berdasarkan faktor seperti lokasi, fasilitas, usia bangunan, dan karakteristik unit lainnya.



DATA UNDERSTANDING

Penjelasan dari setiap kolom data:

- **Hallway Type** : Tipe koridor apartemen, terdapat 3 tipe koridor dalam apartemen yaitu Terraced, Mixed, Corridor, dengan paling banyak jenis apartemennya adalah Terraced.
- **TimeToSubway** : Waktu yang dibutuhkan untuk mencapai stasiun subway terdekat, terdapat 5 kelompok waktu yaitu 0-5min, 10min~15min, 15min~20min, 5min~10min, no_bus_stop_nearby, dengan paling banyak apartemen menawarkan fasilitas 0-5min untuk ke stasiun subway terdekat.
- **SubwayStation** : Nama stasiun subway terdekat, terdapat 8 stasiun subway yaitu Kyungbuk_uni_hospital, Chil-sung-market, Bangoge, Sin-nam, Banwoldang, no_subway_nearby, Myung-duk, Daegu, Lokasi apartemen paling banyak ada di sekitar stasiun subway Myung-duk.
- **N_FacilitiesNearBy (ETC)** : Jumlah fasilitas umum di sekitar apartemen, yaitu dari 0 (tidak ada) sampai jumlah fasilitas yang paling banyak ada di 5.
- **N_FacilitiesNearBy (PublicOffice)** : Jumlah fasilitas kantor pemerintahan di sekitar apartemen, yaitu dari 0 (tidak ada) sampai paling banyak fasilitas kantor pemerintahan di sekitar apartemen adalah 7.
- **N_SchoolNearBy (University)** : Jumlah universitas di sekitar apartemen, yaitu dari 0 (tidak ada) sampai paling banyak jumlah universitas di sekitar apartemen adalah 5.
- **N_Parkinglot (Basement)** : Jumlah tempat parkir di basement, yaitu dari 0 (tidak ada) sampai paling banyak ada terdapat 1321 slot parkir.
- **YearBuilt** : Tahun pembangunan apartemen, yaitu paling lama adalah bangunan yang dibangun pada tahun 1978 dan bangunan paling baru dibangun pada tahun 2015.
- **N_FacilitiesInApt** : Jumlah fasilitas yang tersedia di dalam apartemen, yaitu paling sedikit ada 1 dan paling banyak ada 10 jenis fasilitas.
- **Size (sqft)** : Luas apartemen (dalam kaki persegi), dengan apartemen yang luas paling kecil adalah 135 sqf dan paling luas adalah 2337 sqf.
- **SalePrice** : Harga jual apartemen (Won), dimana apartemen paling murah seharga 32,743 won dan paling mahal seharga 585,840 won.

EXPLORATORY DATA ANALYSIS

DATA MANIPULATION

Mengecek isi data sudah
sesuai dengan konteks kolom

Mengecek data kosong

Mengecek duplikasi data

Mengganti format data dari
float menjadi integer

DATA TRANSFORMATION

Model IA
Ordinal Encoder - TimeToSubway dan Year Built
Binary Encoder - HallwayType dan SubwayStation

Model IB
Ordinal Encoder - TimeToSubway, Year Built,
HallwayType dan SubwayStation

EXPLORATORY DATA ANALYSIS

YearBuilt
1978: 1
1980: 2
1985: 3
1986: 4
1992: 5
1993: 6
1997: 7
2003: 8
2005: 9
2006: 10
2007: 11
2008: 12
2009: 13
2013: 14
2014: 15
2015: 16

Time to subway
0~5min: 1
5min~10min: 2
10min~15min: 3
15min~20min: 4
no_bus_stop_nearby: 5

Subway Station
Bangoge: 1
Banwoldang: 2
Chil-sung-market: 3
Daegu: 4
Kyungbuk_uni_hospital: 5
Myung-duk: 6
Sin-nam: 7
no_subway_nearby: 8

SubwayStation				
mapping:	SubwayStation_0	SubwayStation_1	SubwayStation_2	SubwayStation_3
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0

HallwayType		
mapping:	HallwayType_0	HallwayType_1
1	0	1
2	1	0
3	1	1

CONCLUSION

Model IA	Model IB	Model II
OLS r2 Train: 74.74 %	OLS r2 Train: 73.96 %	OLS r2 Train: 41.71 %
OLS r2 Test: 77.14 %	OLS r2 Test: 77.06 %	OLS r2 Test: 47.72 %
OLS RMSE Train: 52510.73	OLS RMSE Train: 53314.61	OLS RMSE Train: 79761.70
OLS RMSE Test : 51332.13	OLS RMSE Test : 51423.73	OLS RMSE Test : 77629.53
OLS RMSPE Train : 29.40 %	OLS RMSPE Train : 30.33 %	OLS RMSPE Train : 53.25 %
OLS RMSPE Test : 31.07 %	OLS RMSPE Test : 29.57 %	OLS RMSPE Test : 56.98 %

- Model IA menggunakan Binary Encoding untuk mengubah variabel kategorikal HallwayType dan SubwayStation menjadi numerik.
- Model IB menggunakan Ordinal Encoding untuk HallwayType dan SubwayStation, dengan pengurutan kategori berdasarkan harga apartemen dari yang paling murah hingga paling mahal.
- Model II merupakan Simple Linear Regression yang hanya menggunakan satu variabel prediktor.

CONTOH PENGGUNAAN MODEL LINEAR REGRESI

$$\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

SalePrice = - 36,730 - 1,782(TimeToSubway) + 4,765(YearBuilt) + 22,440(HallwayType_0) - 34,490(HallwayType_1) + 30,880(SubwayStation_0) + 6,620(SubwayStation_1) + 8,793(SubwayStation_2) + 26,160(SubwayStation_3) - 5,060(N_FacilitiesNearBy(ETC)) + 316(N_FacilitiesNearBy(PublicOffice)) + 1,742(N_SchoolNearBy(University)) + 32(N_Parkinglot(Basement)) + 9,332(N_FacilitiesInApt) + 134(Size(sqf))

df_2.loc[5]

HallwayType	TimeToSubway	SubwayStation	N_FacilitiesNearBy(ETC)	N_FacilitiesNearBy(Public Office)	N_SchoolNearBy(University)	N_Parkinglot(Basement)	YearBuilt	N_FacilitiesInApt	Size(sqf)	SalePrice
mixed	15min~20min	Chil-sung-market	1	7	3	56	1997	5	558	118584

SalePrice = - 36,730 - 1,782(4) + 4,765(7) + 22,440(1) - 34,490(0) + 30,880(0) + 6,620(0) + 8,793(1) + 26,160(1) - 5,060(1) + 316(7) + 1,742(3) + 32(56) + 9,332(5) + 134(558)

SalePrice = - 36,730 - 7,128 + 33,355 + 22,440 - 0 + 0 + 0 + 8,793 + 26,160 - 5,060 + 2,212 + 5,226 + 1,792 + 46,660 + 74,772
SalePrice = 172,492 Korean Won

Error/residu = 172,492 - 118,584 = **53,908 Korean Won**

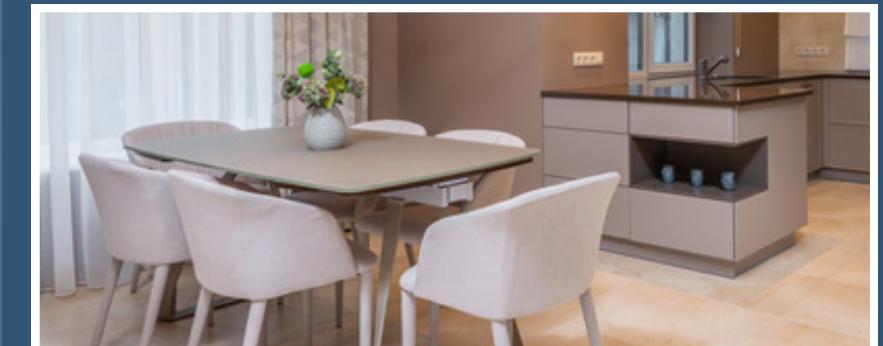
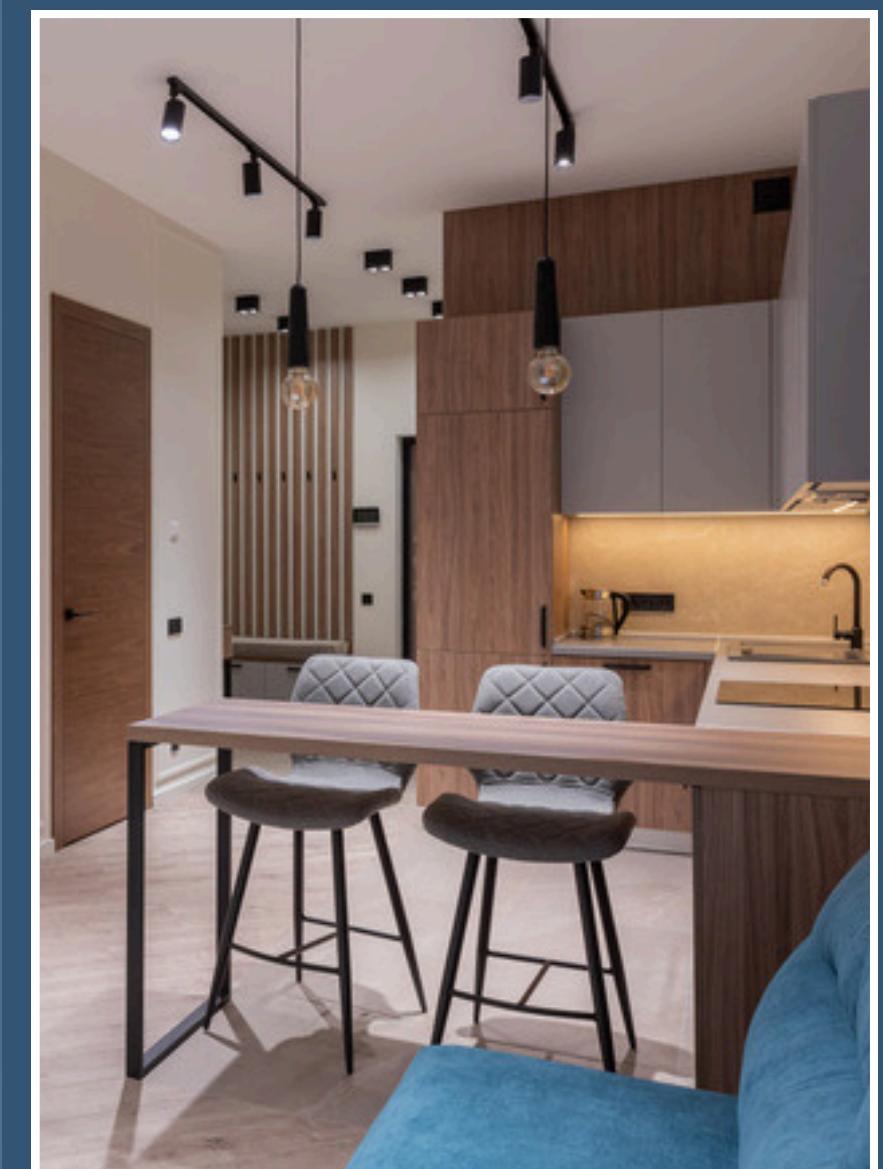
CLOSING - KETERBATASAN MODEL & FAKTOR EKSTERNAL

Model Machine Learning dalam penelitian ini mampu menjelaskan sekitar 77,14% variasi harga apartemen. Artinya, masih terdapat sekitar 22,86% variasi harga yang dipengaruhi oleh faktor lain di luar data, terutama kondisi makroekonomi.

Sektor properti sangat dipengaruhi oleh kebijakan keuangan, khususnya suku bunga pinjaman. Sebagian besar pembelian apartemen dilakukan melalui skema cicilan bank, sehingga perubahan suku bunga berdampak langsung pada daya beli masyarakat.

- Saat suku bunga rendah, permintaan properti meningkat dan harga cenderung naik.
- Saat suku bunga tinggi, daya beli menurun dan harga apartemen dapat tertekan.

Oleh karena itu, hasil prediksi dari model tidak dapat dijadikan satu-satunya dasar dalam menentukan harga apartemen. Prediksi perlu dikombinasikan dengan analisis kondisi ekonomi dan tren pasar properti, agar keputusan yang diambil lebih akurat dan relevan dengan kondisi nyata.





Daegu
Property

Discover Your Dream Residence