Deloitte.

Home in the Outback: Predicting Housing Prices in Australia

Andrew Yang

January 2023

MAKING AN
IMPACT THAT
MATTERS
since 1845

AI Academy Apprenticeship Capstone
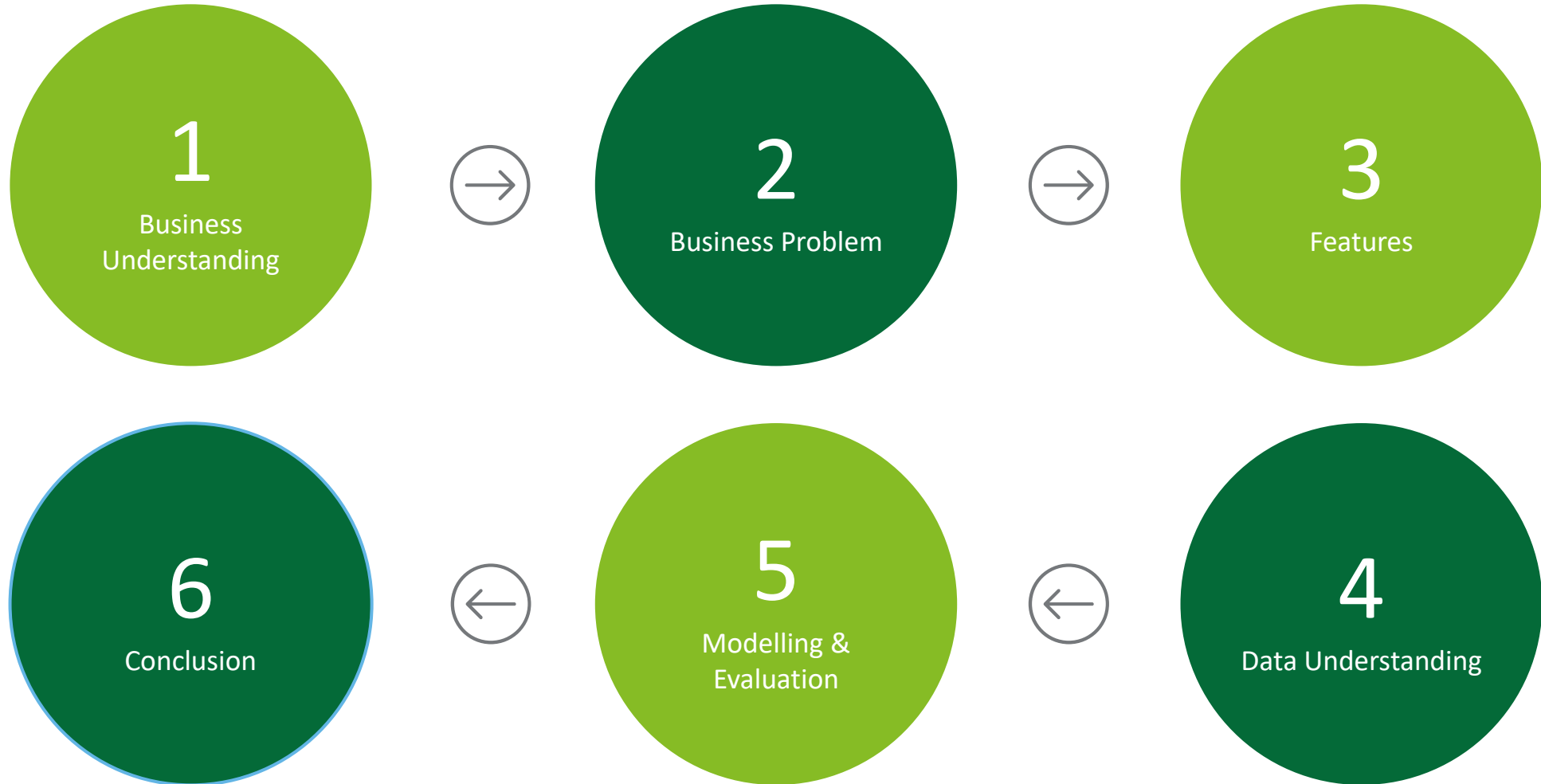
Business Understanding

REUTERS

Australia's housing crisis, largely hidden, is getting worse

Underquoting in Australian real estate industry is leaving buyers feeling betrayed

"It's misleading, it's a fraud on purchasers. And until significant penalties are handed out, regulations have been ineffective … if one does it, others have to follow."

- David Morrel, buyer's advocate

Where It Is Hardest to Afford a Home

Least affordable international housing markets ranked by house price to annual income ratio (2022)*

Hong Kong	23.2
Sydney	15.3
Vancouver	13.3
San Jose	12.6
Melbourne	12.1
Honolulu	12.0
San Francisco	11.8
Auckland	11.2
Los Angeles	10.7
Toronto	10.5

* median house/apartment price divided by median annual gross income
Source: Demographia International Housing Affordability Survey

statista

Business Problem

Problem

Potential buyers are having difficulties buying or renting Australian properties.

- Historical trend of rising prices

- Rampant underquoting

Objective

Develop a model for these potential homeowners that:

- Determines the true price of a property

- Aids in locating homes that fit the buyer's requirements

- Figures out which homes are sold at a discount

Features

1 Building Size	2 Land Size
3 Bedrooms	4 Bathrooms
5 Parking Spaces	6 Property Type

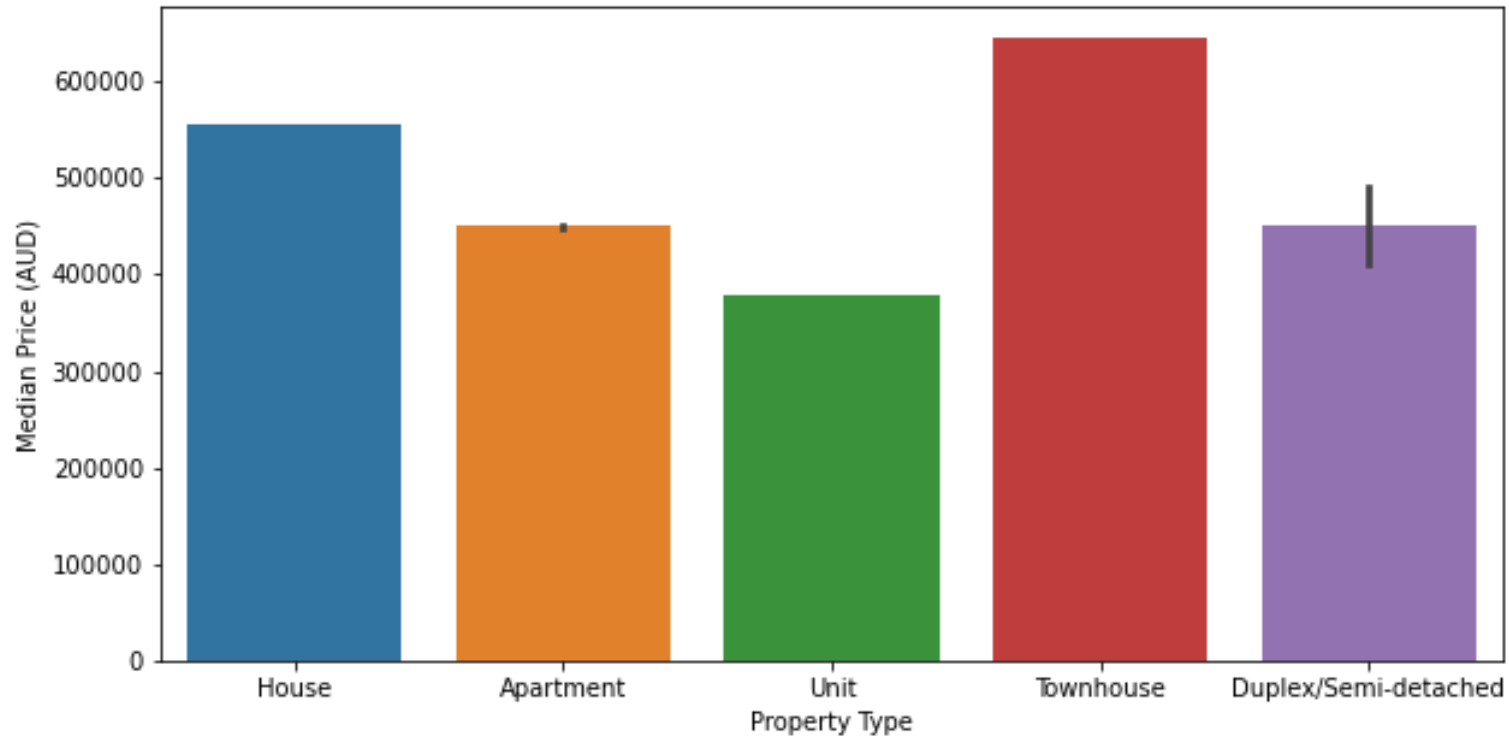7 Product Depth

Data Understanding

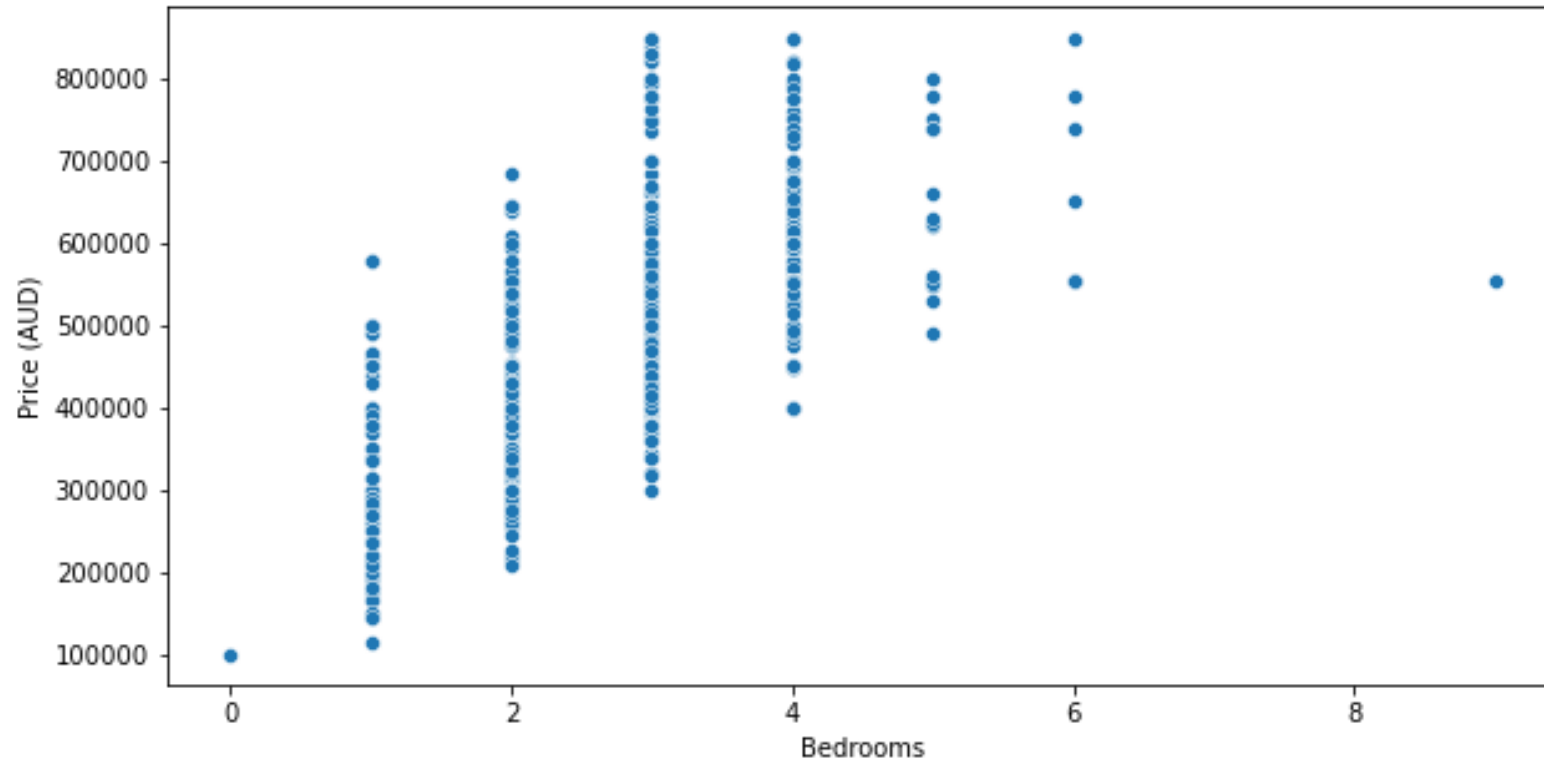Figure 1: Median housing prices by property type.

Data Understanding

Figure 2: Housing prices over number of bedrooms.

Data Understanding

	Median	IQR
Building Size (m^2)	147	103
Land Size (m^2)	294.5	650
Bedrooms	3	1
Bathrooms	2	1
Parking Spaces	2	1
Price (AUD)	485,000	177,500

Table 1: Descriptive statistics of features and target.

Modeling & Evaluation

Model	RSME (AUD)
Baseline	124,670.31
Decision Tree	94,997.47
Gradient Boosting	87,969.37
Gradient Boosting (subsampled)	85,967.02

Table 2: Comparison of model test performance.

Conclusion

On average, the best model predicts the true value of Australian housing within 85,967.02 AUD. This is an approximately 31% improvement in performance compared to the baseline.

Limitations

- Listing data is subject to underquoting.
- Small public dataset (1000 listings)
- Inability to implement geographic information.

Future Improvements

- Use sold housing data for true property value.
- Implement neural networks on a larger dataset.
- Investigate the effects listing agencies have on pricing.