

Research Project Proposal

Tianyu Wang

10/01/2017

Predicting Melbourne Housing Sales Price

Some experts say that Melbourne is experiencing a spectacular housing bubble – where house prices are over-inflated compared to a benchmark – and when it might burst. Many people may ask, which suburbs are the best to buy in? Which houses are value for money? Where should I buy a 2-bedroom unit? And more importantly what will the growing trend of housing price be like? It may be helpful to analyze data about Melbourne's housing sales of these years and find a trend of the house prices and determinants of housing sales.

Background

Melbourne house prices will soon crunch into reverse then slide for three years, according to a report by a leading global ratings agency. But contrary to popular belief, the city's apartment market will defy gravity and keep climbing, the report says. Commonwealth Bank economist Kristina Clifton said the decline was likely to be temporary, with strongly rising house prices in several capital cities suggesting strong momentum in new lending for the next few months of data. But she said new Australian Prudential Regulation Authority rules limiting interest-only loans to 30 per cent of new mortgage lending should cause a slowdown in loans from April onwards.

There is also a study supports firmly that Melbourne's on the cusp of a market downturn. Property price falls are forecast due to massive car manufacturing job losses and rampant overbuilding of houses and apartments.

Data Source

I will use the dataset on Kaggle which was scraped from publicly available results posted every week from Domain.com.au. I have cleaned this dataset as best I can, such as dropping all rows containing NaN values and outliers. The dataset includes Suburb, Address, Number of Rooms, Type of Real estate, Method of Selling, Price, Real Estate Agent, Number of Carspots, Date of Sale and Distance from C.B.D, Land Size, Location of House, etc. This dataset contains 22 columns and 5181 rows.

Algorithms

- 1) Linear Regression: Linear regression is a model that predicts a relationship of direct proportionality between dependent variables and predictor variables. It is very suitable for analyzing house prices based on reasonable variables. It will be discussed in greater details during the modeling process.
- 2) Decision Tree: A decision tree is built through a so-called recursive partitioning process, providing a powerful tool for the description, classification, regression and prediction of data. Through an appropriate decision tree, we can help people who want to buy houses to make better choices.
- 3) Support Vector Machine: By solving a quadratic programming problem, SVM can find the global optimal solution. But when the samples are a lot, it leads to a higher complexity. For solving the problems of less number of samples, sample data is nonlinear and samples have high dimensions; SVM has great advantages. SVM not only has a strong learning ability but also has a strong generalization ability.

References:

[1] Aaron Ng, Machine Learning for a London Housing Price Prediction Mobile Application, June 2015

[2] Boston Home Prices Prediction and Evaluation, link: <http://www.ritchieng.com/machine-learning-project-boston-home-prices/>

[3] Ilya Bershadskiy, Predicting House Price Using Regression Algorithm for Machine Learning, link: <https://yalantis.com/blog/predictive-algorithm-for-house-price/>

[4] Jingyi Mu, Fang Wu, and Aihua Zhang, Housing Value Forecasting Based on Machine Learning Methods, Abstract and Applied Analysis Volume 2014 (2014), Article ID 648047