

Application of Machine Learning in Predicting Housing Prices in Taoyuan

Student: LIAO, SSU-MIN

Advisor: Dr. Chi-Yang

Institute of Industrial Engineering and Management Yuan-Ze University

ABSTRACT

Housing prices have different trends under different seasons and policies, which makes it more difficult to predict housing prices. With SARS in 2003, loan interest subsidy in 2004, the global financial crisis in 2008, the housing crackdown, housing hoarding, and real estate tax in 2014, and COVID-19 in 2019, it is difficult to predict the rise and fall of housing prices, which indirectly causes the marriage rate and birth rate to fall, and accelerates the widening of the wealth gap. As a result, predicting housing prices and buying real estate and pre-sell house at the right time has become a very important and meaningful issue.

According to the population information of the Department of Civil Affairs and actual price registration website, only Taoyuan is the city with positive population growth among the six municipalities, and Guishan and Guanyin District of Taoyuan City are the areas with the most population migration among the six municipalities in 2021; therefore, Taoyuan City is the core of this study. The data of actual price registration and economic indicators from 2017 to 2021 were collected, and housing prices were predicted by Random Forest (RF), Gradient Boosting Machine (GBM), and eXtreme Gradient Boosting (XGBoost), and then the predictive indicators were used to measure the model effects. The results of the study show that GBM has better results, and only the Building Shifting Total Area is more correlated with housing price, but some variables are highly correlated with each other, resulting in better prediction when multiple variables are considered simultaneously.

Keyword: Predicting Housing Prices 、 Machine Learning 、 Random Forest (RF) 、
Gradient Boosting Model (GBM) 、 eXtreme Gradient Boosting (XGBoost)