Linear Regression For Rental Prices

By: Asteway Kebede

Abstract

To build a predictive linear regression model that can predict the rental price of a property in the Sacramento, CA. The model is trained on over 2,500 listings posted on Craigslist with over ten features. A few of the critical features are soft, number of bedrooms and bathrooms. Before selecting the best model, other linear regression models such as Ridge, Lasso and Polynomials were attempted.

Design

Over the past year, the rental price has increased in the region by 16% which is significant. Therefore, building this model will give us a better picture of the inventory available but with the focus of building a rental price predictive model. The data will be scraped from Craigslist and the necessary exploratory data analysis will be performed before running model selection. Additional work such as regularization and feature selection will be performed to optimize the model.

Data

Data is scraped from Sacramento Craigslist. It has over 2, 700 listings over the past 60 days. The features included are: SQFT, BedRooms, BathRooms, Laundry, Parking, Smoking, Location, Pets

---

Tools:

- Python, Requests, BeautifulSoup

Algorithms

- Ridge
- Lasso
- Elastic Net
- Polynomials

Tools:

- Sklearn, statsmodels

Other Tools

- NumPy and Pandas for data manipulation
- Matplotlib and Seaborn for visualization

Communication

The results will be communicated via data visualization in a presentation. All work performed is on Github https://github.com/kasteway/NY-MTA-Turnstile-EDA.