Predicting Rent prices in Toronto Canada

SALIM AL-HARTHI

Introduction

- People are renting houses and apartments to live in for years or months. It is important for people who rent a house or an apartment to get a fair price for the place they are renting.

- By utilizing the information of the rent of different properties we can build a model that will accurately predict a fair price for rent according to different factors such as number of bathroom and bedrooms, number of different venues nearby and more.

Data acquisition and cleaning

- Kaggel dataset here contain price, latitude, longitude, number of bedroom, and number of bathroom

- foursquare.com API provides the venues nearby the property will .

- Cleaning Kaggel dataset consist of removing ($) from **Price** column then convert it to Float

- For cleaning the API data, the **Categories icon** was reduced to main category for example.
  The following **Categories icon:**

https://ss3.4sqi.net/img/categories_v2/shops/salon_barber_

'https:', '', 'ss3.4sqi.net', 'img','categories_v2','shops','salon_barber_'

8 main categories

Category	Description
food	Restaurants, coffee shops etc.
Shops	Supermarkets etc.
Building	Library office building and others
Art entertainment	Museums
Travel	Bus stops train and airports
Nightlife	Bars and clubs
Parks outdoors	Parks
Education	Schools and universities

margining all features into one table

Column name	Description	Data type	Example
Bedroom	How many bedrooms available	Integer	2
Bathroom	How many bathrooms available	float	2.0
Den	Whether den is available or not	Boolean	1
Lat	Latitude	float	43.643051
Long	Longitude	float	-79.391643
Price	Apartment Rental price per month in CAD	Float	2450.0
food	Count of occurrence of the category near by the property	Integer	1
Shops	Count of occurrence of the category near by the property	Integer	2
Building	Count of occurrence of the category near by the property	Integer	9
Art entertainment	Count of occurrence of the category near by the property	Integer	1
Travel	Count of occurrence of the category near by the property	Integer	2
Nightlife	Count of occurrence of the category near by the property	Integer	0
Parks outdoors	Count of occurrence of the category near by the property	Integer	1
Education	Count of occurrence of the category near by the property	Integer	0

Exploratory Data Analysis

- **Relationship between features (Bathrooms, Bedrooms, Den, Latitude, and Longitude) and target (rent)**

- **Relationship between 8 main categories and target (rent)**

Relationship between features (Bathrooms, Bedrooms, Den, **Latitude**, and **Longitude**) and target (rent)
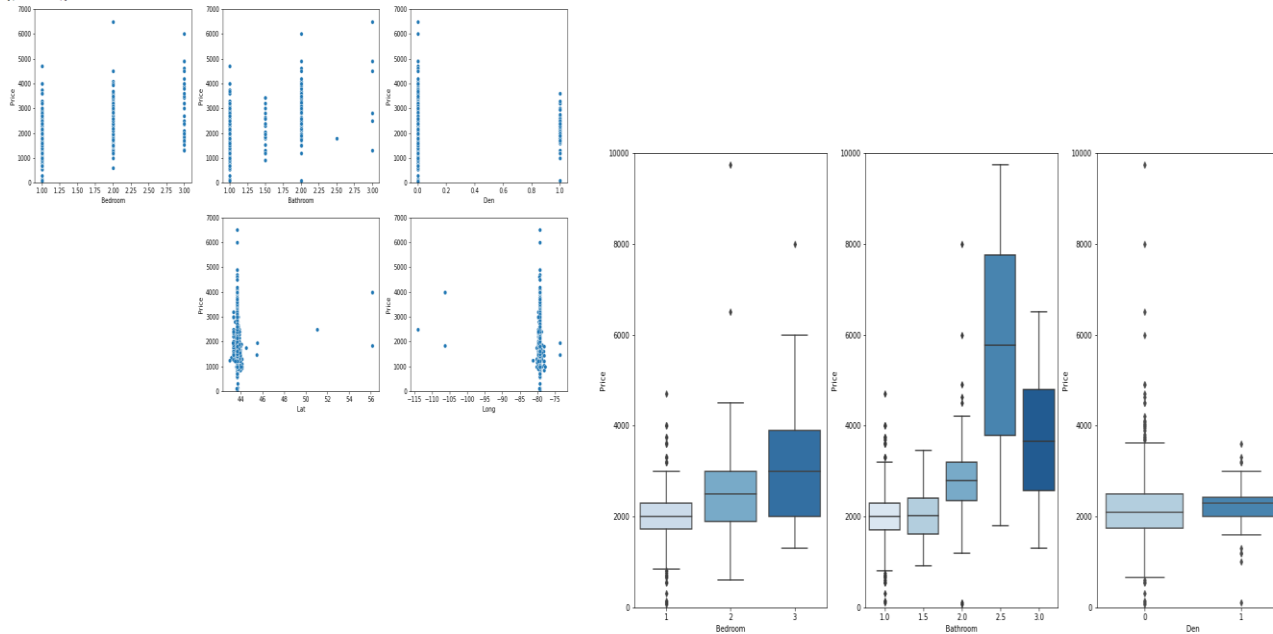
As shown in (Figure 1.1) the increase of number of bathrooms and bedrooms cause the minimum and maximum amount rent increases.

The properties with den has a higher minimum rent in compared to the properties without den as shown in (Figure 1).

As shown in (Figure 1) most of the properties are closer to (44, -80) latitude, and longitude.

Relationship between 8 main categories and target (rent)

As shown in (Figure 3) the in crease in number of travel, parks_outdoors, and arts_entertainment, cause the minimum and maximum amount rent increases.
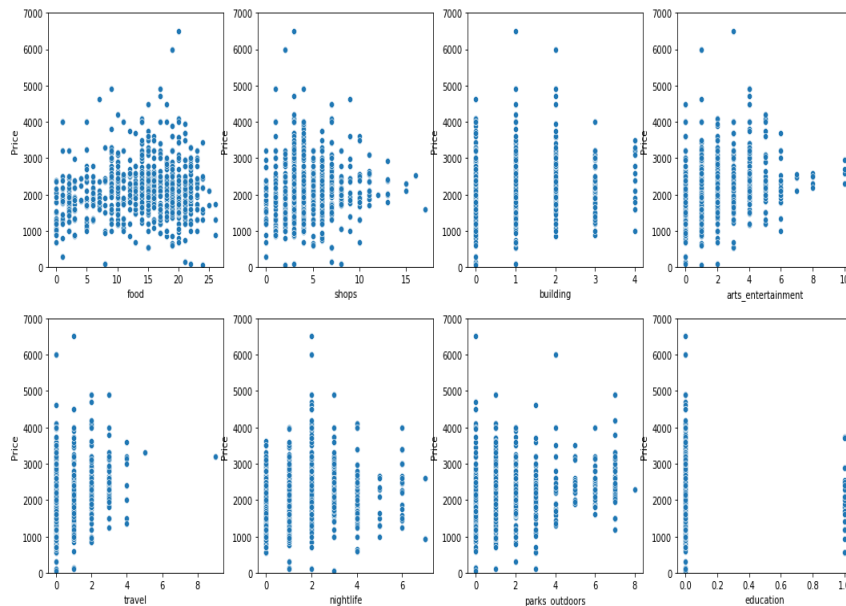As shown in (figure 3.1) the change in number of occurrences changes the range of the price.

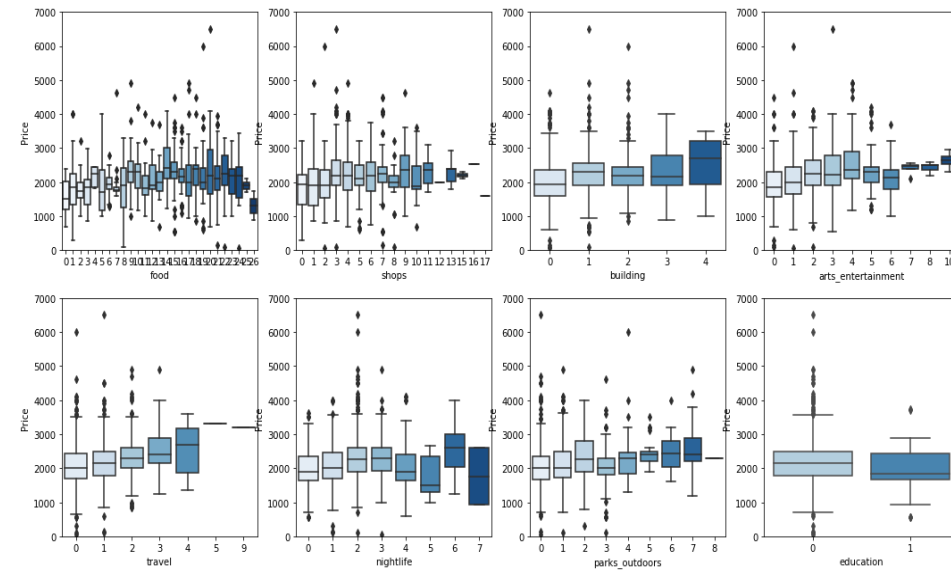*Figure 3 Relationship between 8 main categories and Price*

Figure 4.1 box plot Relationship between features (Latitude, and Longitude) and Price

Predictive Modeling

To Predict the rent of a property a regression model is the best fit. In this case 3 regression models have been used Gradient Boosting Regressor, Random Forest Regressor, and KNN Regressor

Model	Gradient Boosting Regressor	, Random Forest Regressor	KNN Regressor
Accuracy (R^2)	88.3%	77%	84.0%
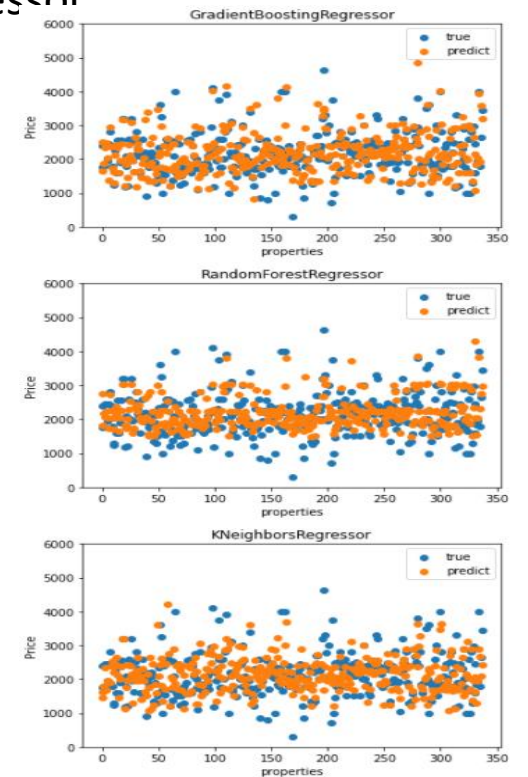
Conclusion and future directions

- In conclusion in this study, I attempted to predict the rent of some properties in Toronto. I have performed analyses on the features from Kaggel dataset as well as foursquare.com API. I used three different models Gradient Boosting Regressor, Random Forest Regressor, and KNN Regressor to predict the rent of the properties and found that KNN Regressor performed the best in term of accuracy (R^2).

- For the future of this project, I'm planning to build a simple website that will take the latitude, longitude, number of bathroom bedroom and if there is a den and output a predicted rent price. In addition, this project must be updated regularly because rent prices are affected by economical changes.