Detecting the price of real estate property in Mexico city?

Introduction

In Mexico city it's hard to find a nice place to live because of its large population. Detecting this problem the project main objective is to determine the price for a selling property it can be a house or department. The variables in which the price of the property depends are location, number of beth, number of bathrooms, number garage, the distance from the location to the metro and more.

Methodology

Extract Data: This project used data from scraping web pages do tectect price of the properties and the other variable mentioned.
The only satisfied web page where the data was stisfite extracted , inmuebles 24 and lab mundy. Other web pages had missing values for example longitude and latitude from https://www.metroscubicos.com. The web page https://www.vivanuncios.com.mx After getting the data there were a lot of missing values and cleaning it was extremely difficult.

Web Scraping

1. Price
2. longitude and latitude
3. Address
4. Size of property
5. #bathrooms
6. #rooms
7. #garage
8. Agency name
9. Description

Cleaning Data:Cleaning these data was extremely difficult because of the web scraping it had a lot of strange values in the price. After scraping the data it was stored in data frames. There were many occurrences in the data. It was hard to convert the price into floats because it had many values and symbols that made it hard to convert for these reasons web pages like viva anuncios did not work to analyse data. At the end to understand. Clean at finally with deleting the

Data Architecture:This assignment has a flask that connects the data from a H5 file which has the regression model.This file transports data to the flask which is also connected to the html file. This file shows the graphs and the map of Mexico city.

Machine Learning: For this project there were applied many models like linear regression, multiple regression and the best model that worked was random forest. This model got the highest score in r2. This model measure in y axis the price of the property and in the x axis thats has colonies,size of property, #bathrooms, #rooms, #garage.After assigning these values it makes the prediction based on the x axis.

Data Visualisation: In this final part of the project the visualisations were made in jyson to have a more interesting conted. There is a map of the city of Mexico showing the property that complained with the information selected. The graphs below are scatter plots showing baths, rooms ,and garages .

Next-Steps

To expand the model to different cities in Mexico in order to help most of the population to buy their dream house

Applying this model will help real estate investors to reduce risk and increase profits in new real estate developments

Include an interactive map feature to be able to choose a specific point in the city to evaluate.