

Predicting Airbnb Listing Outcomes in Chicago, IL

Introduction

Airbnb has transformed the way people travel and experience new locations by offering a platform for short-term rental accommodations. However, hosts face challenges in optimally pricing their listings. This project leverages machine learning to predict Airbnb rental prices in Chicago, aiming to provide hosts with data-driven insights to make informed pricing decisions. I will focus on determining the factors influencing rental prices and predict rental prices based on listings features.

Data Overview

- **Dataset:** Chicago Airbnb Listings

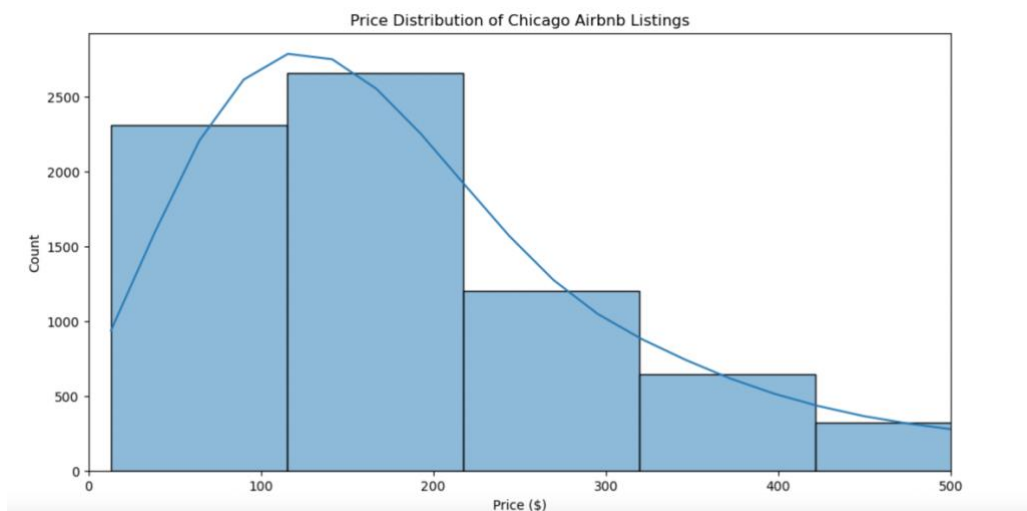
- **Target Variable:** price

- **Features Analyzed**

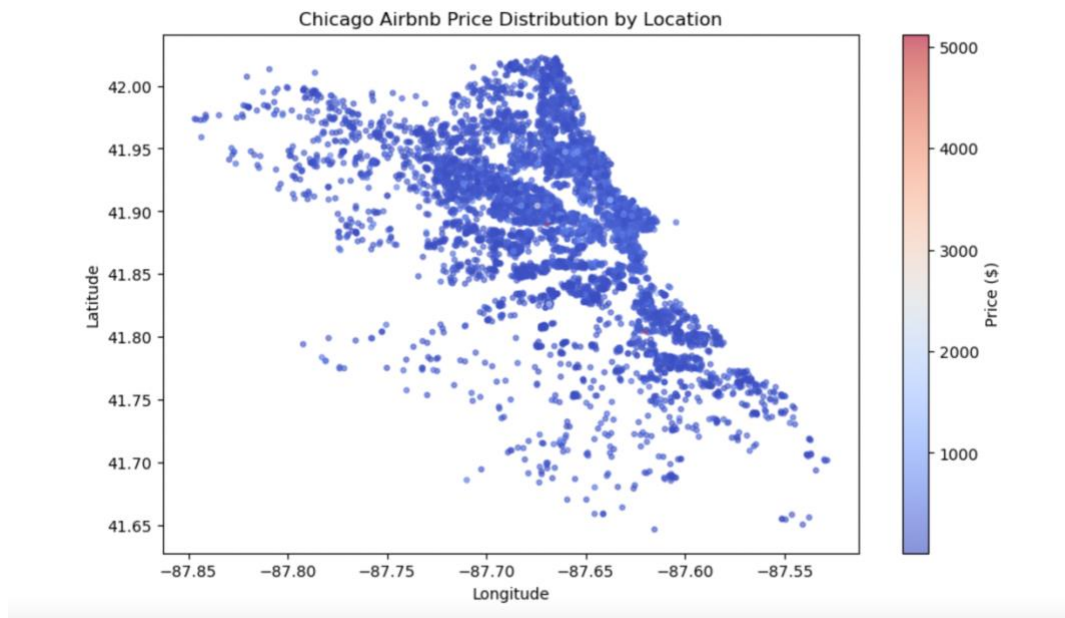
- Property Characteristics: accommodates, bedrooms, beds, bathrooms
- Location Data: latitude, longitude
- Host Attributes: host_is_superhost, instant_bookable, host_response_rate, host_acceptance_rate
- Guest Reviews: review_scores_rating, number_of_reviews
- Availability: availability_365

Exploratory Data Analysis (EDA)

Before building a model, I analyze the data distribution and identify potential patterns.



Geographic Price Distribution



- Most listings have prices below \$200, with a few high-end outliers.
- Higher-priced listings tend to cluster in specific regions.

Feature Engineering

To improve model performance, I cleaned and transformed various data points:

- Converted categorical variables (host_is_superhost, instant_bookable) to numerical (1/0).
- Filled missing values in review_scores_rating, beds, and bathrooms using median imputation.
- Normalized numerical features using StandardScaler.

Model Training & Evaluation

I tested two machine learning models to predict Airbnb prices:

Model	RMSE	R ² Score
Random Forest	156.09	0.66
Tuned Random Forest	151.47	0.68
XGBoost	176.29	0.57

Key Findings

Random Forest performed best after hyperparameter tuning, while XGBoost had a lower R^2 , meaning it struggled with price variations.

Using feature importance analysis, we identified the most influential factors:

Feature	Impact on Price
bedrooms	High
bathrooms	High
latitude & longitude	Medium
host_is_superhost	Medium
review_scores_rating	Low

Larger properties (more bedrooms & bathrooms) significantly increase price.
Location plays a role, but less than expected.
Being a "Superhost" has a moderate impact on pricing.

Price Optimization Strategy

Based on the analysis, we recommend strategies for Airbnb hosts to optimize pricing:

- Airbnb pricing fluctuates based on seasonality & demand.
- A/B Testing for Pricing Optimization

Future Work

Improve time-series forecasting to track seasonal trends.
Experiment with deep learning models (e.g., LSTMs for price trends).