BAX 452 Machine Learning

Final Project: Car Rental, STUDENTUNIVERSE.COM

Section 2, Group 5

Richard Liu, Shivank Chandra

Exploration Into the Car Rental Market in San Francisco and Oakland

—

Executive Summary

Problem Statement

In recent years, the rise of operational costs and increasing competition in the rental car market have raised the difficulty of setting an optimal price point. This project is aimed to optimize rental car utilization through predictive pricing strategies. We will analyze predictive models to determine the optimal model that generates a high profit while maintaining market competitiveness.

Proposed Solution

Approach:

We employed Linear Regression, Lasso Regularization, and Random Forest Regression to describe and analyze the data. This approach can help us determine the best model for predicting and analyzing the rental car data.

Data:

The models were trained on data sourced from the Student Universe car rental platform, created by web scrapping. It aggregates rental options from various car rental agencies across different locations. The data includes Car Name, Supplier Company, Tier, Seat Number, Storage Space, Pickup Locations, and Price. The data has various rental car companies and the pickup locations are either at San Francisco International Airport, Metropolitan Oakland International Airport, and San Francisco- Yerba Buena.

Findings:

Linear Regression achieved an out-of-sample root Mean Squared Error of 53.04 and an R-squared score of 0.53. Lasso Regularization achieved an out-of-sample root Mean Squared Error of 52.94 and an R-squared score of 0.54. Random Forest Regression achieved an out-of-sample root Mean Squared Error of 45.19 and an R-squared score of 0.68. With the comparison of the results from 3 different predictive models, we expect Random Forest Regreesion to be the most ideal model for predicting optimal price points and improving business decisions.

Value:

These insights can help the business to set its price for different kinds of vehicles, learning how other key features can affect the price. In this way, the business can improve decision-making, cut costs, and increase revenue. As a result, maintains its reputation and stays competitive in the market.

Final Thoughts and Next Step

Optimal price not only allows the business to lead the market, but it also balances the cost and revenue at a reasonable market price. With this data analysis, we expect the business to employ Random Forest Regression in their system. The next step is to improve and maintain the features that are not included in the dataset, such as customer service, insurance, and processing speed.

Background, Context, and Domain Knowledge

Background:

In the modern market of car rental, companies and customers alike must adapt persistently. This adaptation is important to follow up with the fluctuating market, to keep pace with technological progress, and to meet the changing demands of consumers. The car rental industry is all about affordability and convenience, paired with a preference for a wider array of options. Numerous variables, including the date, model of the car, location, rental agency, storage capacity, and others, play a role in adjusting the price.

Our project aimed to dig into car rental prediction using data from studentuniverse.com. The result is going to be useful for both companies and students as this website is designed for students. This can provide insights to help students make better decisions and help the company adjust to a reasonable price point.

Industry traditionally Attempts

Traditionally, car rental companies relied on market research and competitive analysis to set their pricing strategies. They often use historical data and experience to understand consumer's behavior and preferences. It is also common for companies to monitor competitors' pricing and adjust their prices to remain competitive. Lastly, companies usually use a revenue management system that analyzes past trends to forecast future demand, to generate profits.

Analyses

	Linear Regression	Lasso Regularization	Random Forest Regression	XGBoost
RMSE	53.04	52.79	45.30	64.87
R-Squared	0.53	0.54	0.68	0.34

The table displayed the performance of four different regression models: Linear Regression, Lasso Regularization, Random Forest, Regression, and XGBoost.

1. Linear Regression

   The result suggests that it explains about 53% of the variance in price variables. This is decent but only explains half of the variance.

2. Lasso Regularization

   The result suggests that it explains about 54% of the variance in price variables, a little improvement from the previous. Lasso regularization performs feature selection by shrinking the less significant feature.

3. Random Forest Regression

   The result suggests that it explains about 68% of the variance in price variables, a big leap from past models. This model has better predictive accuracy. Moreover, it has the lowest RMSE (Root Mean Squared Error) among all.

4. XGBoost

   The result suggests that it explains only about 34 % of the variance in price variables, the worst among all models. At the same time, it has the highest root mean squared error. This might indicate an overfitted model or the technique is not well suited for this dataset.

Recommendations and Business Value Provided

Recommendation:

For future recommendations, we recommend prioritizing the Random Forest Regression model as it outperformed the others.

Business Value:

With better prediction capabilities, the company can improve in various approaches. First, the company can adjust its price according to market conditions. With the more accurate pricing decisions provided by the Random Forest model, it can increase revenue and customer satisfaction. With the low prediction error or the root mean square error, it indicates that the company can expect a lower possibility of overpricing and underpricing a vehicle. As a result, the integration of a better predictive model can affect the strategic planning of the entire company.

Summary and Conclusions

Our machine-learning project delved into the car rental market around San Francisco and Oakland Airport, mainly focusing on listings relevant to students. By employing Linear Regression, Lasso Regularization, Random Forest Regression, and XGBoost. We successfully discovered the most ideal model to predict the price on studentuniverse.com, a popular platform catering to students. The dataset includes car models, prices, capacities, suppliers, and rental conditions. A better prediction model can help both the students and the company analyze the optimal price point to rent cars.

Through our project, we demonstrated the effectiveness of different prediction models on our specific dataset, which is collected by web scrapping the real-time data. The chosen model, Random Forest Regression, outperformed others with the lowest RMSE (Root Mean Squared Error) and the highest percentage of R^2(R Squared). This indicated its higher predictive accuracy and lower error rate.

The project illustrates the importance of applying a machine learning model in data-driven strategies for businesses to understand the demand of the market and consumer preferences. It can ultimately contribute to improving business decisions, cost efficiencies, and market competitiveness.

	Car_Name	Supplier	Tier	Seats	Storage	Pickup	Price
0	Ford Focus	Thrifty	Compact	5	2	San Francisco International Airport	84.96
1	Ford Fusion	Thrifty	Full-Size	5	3	San Francisco International Airport	101.12
2	Ford Escape	Thrifty	Intermediate	5	3	San Francisco International Airport	101.50
3	Toyota Yaris	Fox	Economy	4	1	Metropolitan Oakland International Airport	57
4	Toyota Yaris	Fox Rent A Car	Economy	4	1	Metropolitan Oakland International Airport	71
5	Wild Card	Fox	Compact	5	2	Metropolitan Oakland International Airport	57
6	Nissan Versa	Fox	Compact	5	2	Metropolitan Oakland International Airport	58
7	Wild Card	Fox Rent A Car	Special	5	2	Metropolitan Oakland International Airport	72
8	Nissan Versa	Fox Rent A Car	Compact	5	2	Metropolitan Oakland International Airport	74
9	Volkswagen Jetta	Thrifty	Standard	5	2	Metropolitan Oakland International Airport	76.65
10	Ford Fiesta	Thrifty	Economy	4	1	Metropolitan Oakland International Airport	77.81
11	Ford Focus	Thrifty	Compact	5	2	Metropolitan Oakland International Airport	77.83
12	Ford Fusion	Thrifty	Full-Size	5	3	Metropolitan Oakland International Airport	79.82
13	Volkswagen Jetta	Thrifty Car Rental	Standard	5	2	Metropolitan Oakland International Airport	81
14	Ford Fiesta	Thrifty Car Rental	Economy	4	1	Metropolitan Oakland International Airport	82
15	Ford Focus	Thrifty Car Rental	Compact	5	2	Metropolitan Oakland International Airport	82
16	Buick Encore	Thrifty	Compact	5	3	Metropolitan Oakland International Airport	82.44
17	Ford Fusion	Thrifty Car Rental	Full-Size	5	3	Metropolitan Oakland International Airport	84
18	Ford Fiesta	Thrifty	Economy	4	1	San Francisco International Airport	84.90
19	Ford Fiesta	Thrifty Car Rental	Economy	4	1	San Francisco International Airport	84
20	Ford Focus	Thrifty Car Rental	Compact	5	2	San Francisco International Airport	84
21	Hyundai Accent	Payless	Economy	5	1	Metropolitan Oakland International Airport	85