Guided Capstone Project Report

Problem Statement

The goal of this analysis was to build a machine learning model to find the prediction of the ticket price. Big Mountain Resort aims to optimize its ticket pricing strategy to maximize revenue while remaining competitive in the market. Additionally, the dataset only includes ticket prices and a limited set of cost factors, such as the operational cost of a new chair lift, but lacks other critical costs.

Data Wrangling

The original dataset contained 330 rows and 27 columns. However, after data cleaning and removing, the dataset reduces to 277 rows and 25 columns. Upon reviewing the data, several issues were identified, including missing values and duplicated resort entries. 82% of the resorts had complete ticket price information, 3% had one missing value (either AdultWeekday or AdultWeekend ticket price), 14% had both AdultWeekday and AdultWeekend ticket prices missing. To ensure data quality, any rows with missing price information were removed, as price was a critical target variable for this analysis. Additionally, the FastEight column was dropped due to 50.3% missing values, making it unreliable for modeling. There was also 15-16% missing data in ticket price columns, with AdultWeekday ticket prices missing slightly more values than AdultWeekend.

Analysis

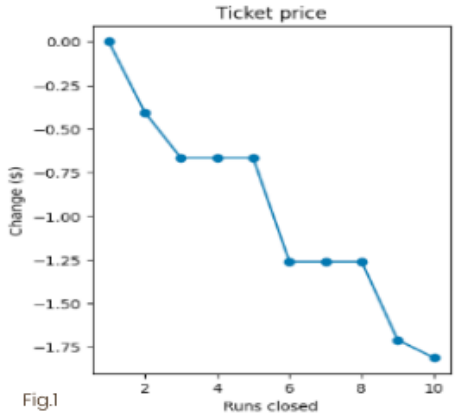

Fig.1

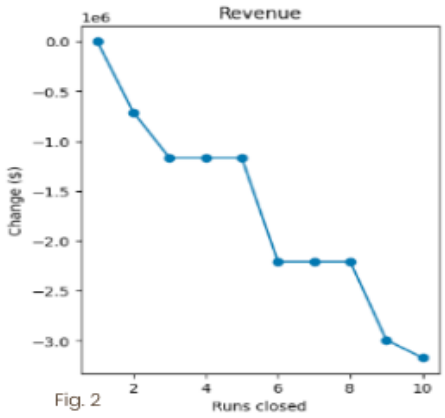

Fig. 2

The impact of closing ski runs on ticket price and revenue. Closing a single run has minimal effect on pricing and revenue, suggesting that resorts can absorb the change without significant financial consequences. However, closing two or three runs leads to a noticeable decline in both ticket price and overall revenue. When three or more runs are shut down, the impact becomes substantial, resulting in a sharp drop in revenue and potential financial losses for the resort (Fig.1 and Fig.2).

Big Mountain Resort has been evaluating various strategies to either cut costs or boost revenue, particularly by adjusting ticket prices. The following scenarios were considered:

- Closing up to 10 runs – The results show that closing a single run has no impact, but closing two or more gradually reduces support for the current ticket price, leading to a decline in revenue.
- Adding a new run – This scenario increases total seasonal revenue by $3,474,638.
- Adding two new runs – This further boosts total revenue to $6,974,638.
- Extending the longest run by 0.2 miles with 4 acres of additional snow coverage – The results indicate no significant impact on revenue.

Model Preprocessing

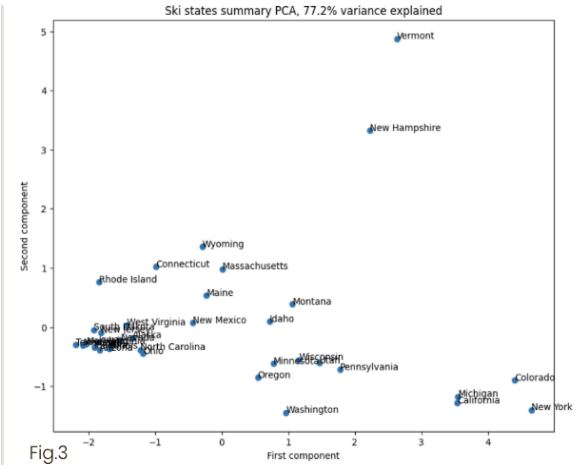

Fig.3

Principal Component Analysis (PCA) was used to find the linear combinations of the original features that are uncorrelated with one another and order them by the amount of variance. After applying PCA, the cumulative variance was plotted, revealing that 77.2% of the variance (Fig.3) was retained, allowing for dimensionality reduction while preserving critical information.
This figure visualizes relationships between the states based on features such as the total skiable terrain area.

The Random Forest Regression Model was executed to predict the ticket price for Big Mountain Resort. The dataset was divided into training and testing sets, with 70% of the data used for training the models and 30% reserved for testing their performance. The test set helps to evaluate the model's performance and selects the best model, while the training set is used to teach the model how to make predictions.

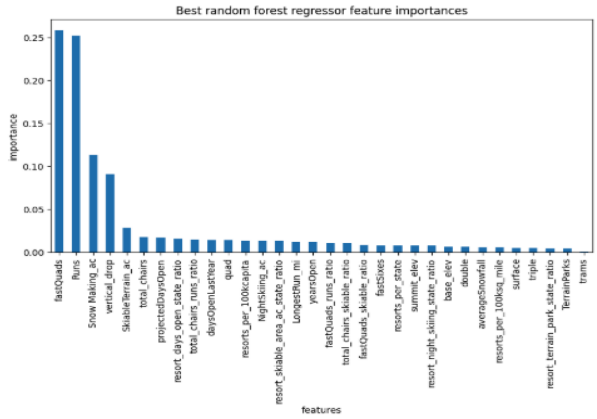

Figure 5 illustrates the Random Forest Regression Model, which predicts the weighted average while mitigating overfitting. The model was trained on numerical data, ensuring a more accurate and reliable representation of patterns within the dataset.

Linear Regression vs. Random Forest Regression

In the linear regression model there was over 80% of the variance on the train and over 70% on the test, this result indicates that there is overfitting on the model. The random forest model outperformed the linear regression model. The random forest model has a lower cross-validation mean absolute error by almost $1. It also exhibits less variability. Verifying performance on the test set produces performance consistent with the cross-validation results.

Recommendation

Based on the analysis of various scenarios, Big Mountain Resort has the opportunity to optimize its ticket pricing strategy to maximize revenue while accounting for operational costs. The analysis highlights potential revenue increases, but further refinements can enhance pricing accuracy. For instance, incorporating updated data and market trends could better inform pricing adjustments, leading to higher revenue and improved financial outcomes.

Conclusion

The analysis of Big Mountain Resort's ticket pricing strategy reveals significant opportunities for revenue optimization. By leveraging data-driven insights, the resort can strategically adjust ticket prices to maximize profitability while maintaining market competitiveness.