The Big Mountain Pricing Strategy

LINA GAO

Introduction

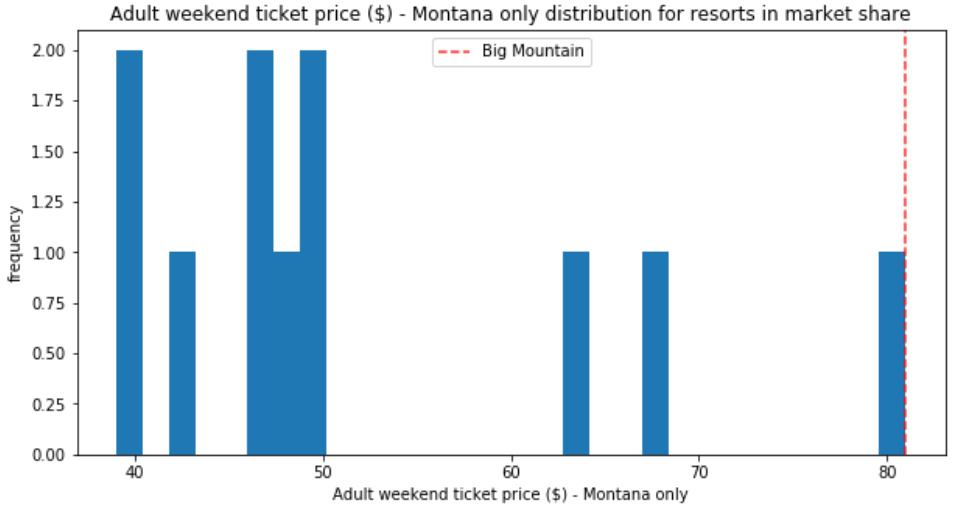
- The Big Mountain Resort, a ski resort located in Montana, offers spectacular views of Glacier National Park and Flathead National Forest. ~ 350,000 people ski or snowboard at Big Mountain each year.
- It has recently installed a chair lift, which increases the operating costs by $1,540,000 this season.
- The resort's executive team needs suggestions to develop a strategy to cover this cost, either by charging a premium fee, cutting underused facilities, or other scenarios.

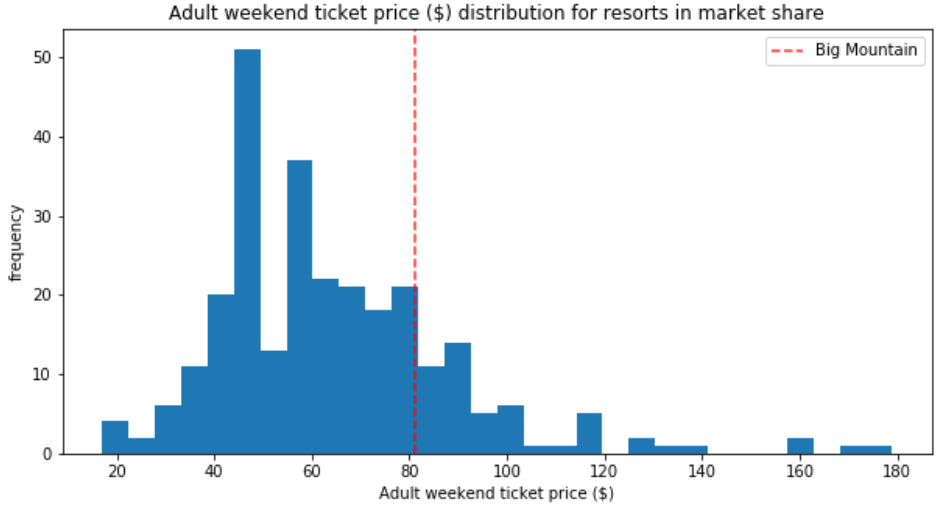
Problem identification

- Is the current ticket price of Big Mountain Resort reasonable in the ski resort market with the facilities they provided for the customers?
- If it is lower than its peers, how much can we increase the ticket price without a significant effect on the customer numbers?
- Which facility is the most valuable one to attract customers to visit this resort?
- Is any facility of the resort underused and can be eliminated?

Key findings:
Market competitiveness evaluation

- Montana is among the top five states in US with large state area, large total skiable area and high number of resorts per 100 k population.
- Big Mountain Resort is the top resort with high SnowMaking_ac, total_chairs, fastQuads, number of runs, and the most considerable amount of skiable terrain.
- The current ticket price of Big Mountain Resort is the highest in Montana. It is in the median high range in US.





Key findings

Important features identified by the model	
Total number of runs	Total number of chairs
Vertical drop	Area covered by snow
Number of trams	Longest run
Skiable terrain area	

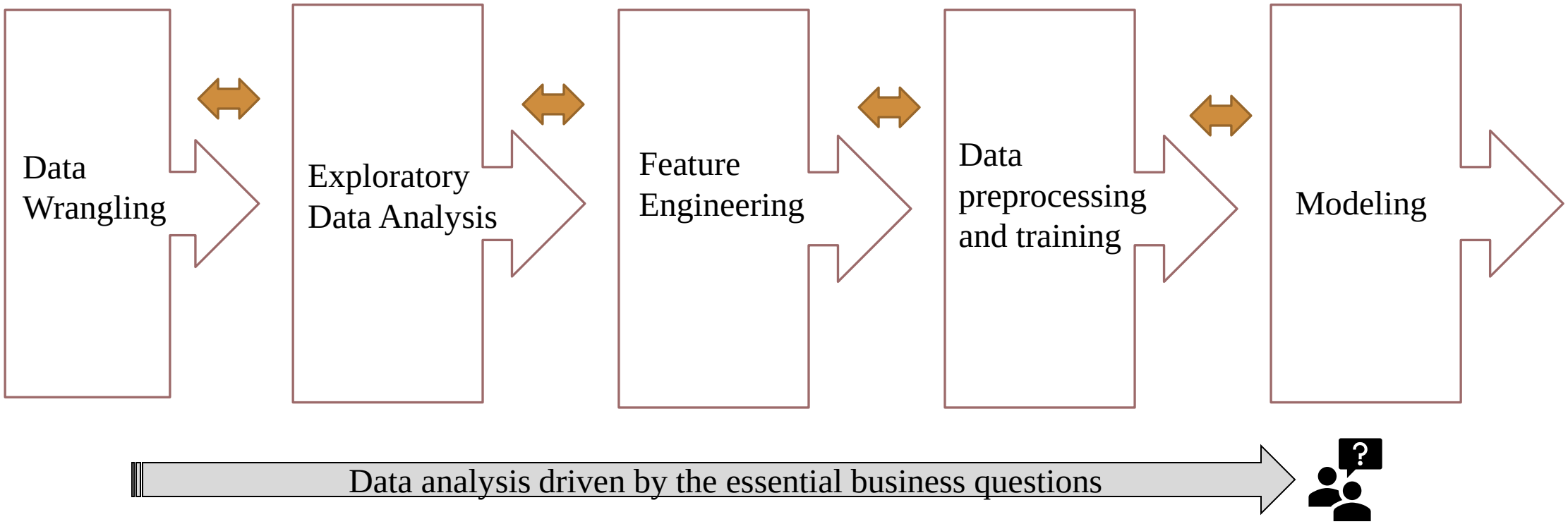
Key findings

Evaluation of the Pricing Scenarios by Machine Learning Models

S1	S2	S3	S4
Increase the ticket price from $81 to $94, the predicted ticket price by the model.	Add a new run to increase the vertical drop by 150 feet , install a chair lift	Add two acres of snow making	Close up to 4 least used runs
It will increase the revenue by 22,750,000 dollars this season. It can completely cover the cost of the new lift.	It will increase support for ticket price by $8.46. Over the season, this could be expected to amount to $14,811,594.	It will increase support for ticket price by $9.75. Over the season, this could be expected to amount to $17,068,841	It will barely have effect on the revenue

Modeling's Process and Results

Project workflow



Data wrangling & feature engineering&EDA

- 

  I. Is the dataset clean enough for exploratory data analysis
  II. Which feature is out best choice as the target variable, adultWeekend or adultWeekday?
  III. Are there enough proper features for the development of the pricing model ?

- 

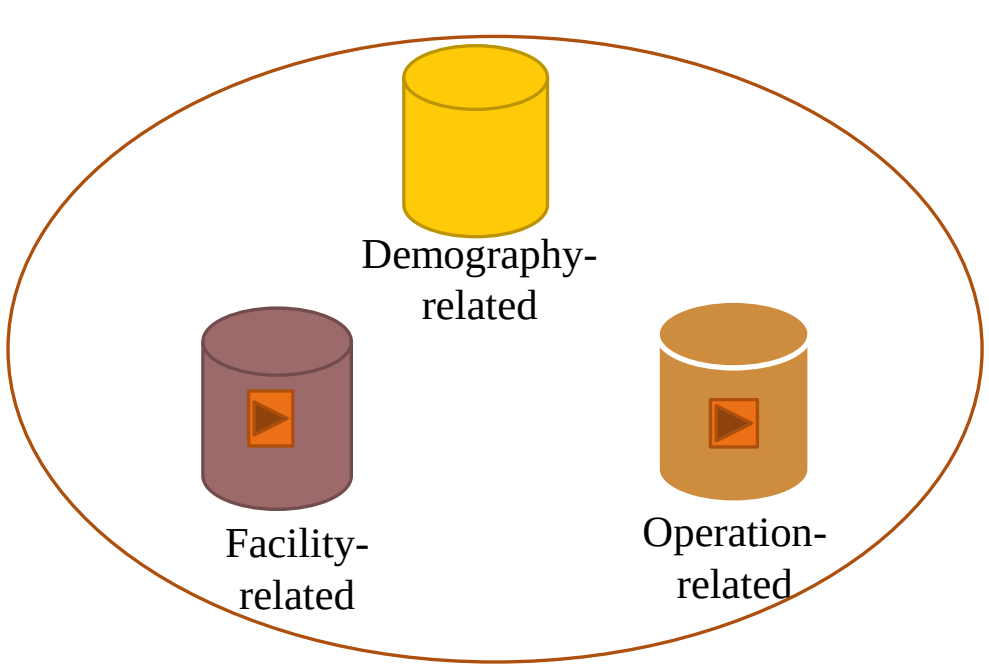
  Pandas, numpy, matplotlib, seaborn

- 

  I. Cleaned the missing data, deleted negligible features, corrected suspicious data
  II. AdultWeekend is chosen as the proper target variable for the following analysis
  III. Added State summary for this project and finalized the data for EDA analysis

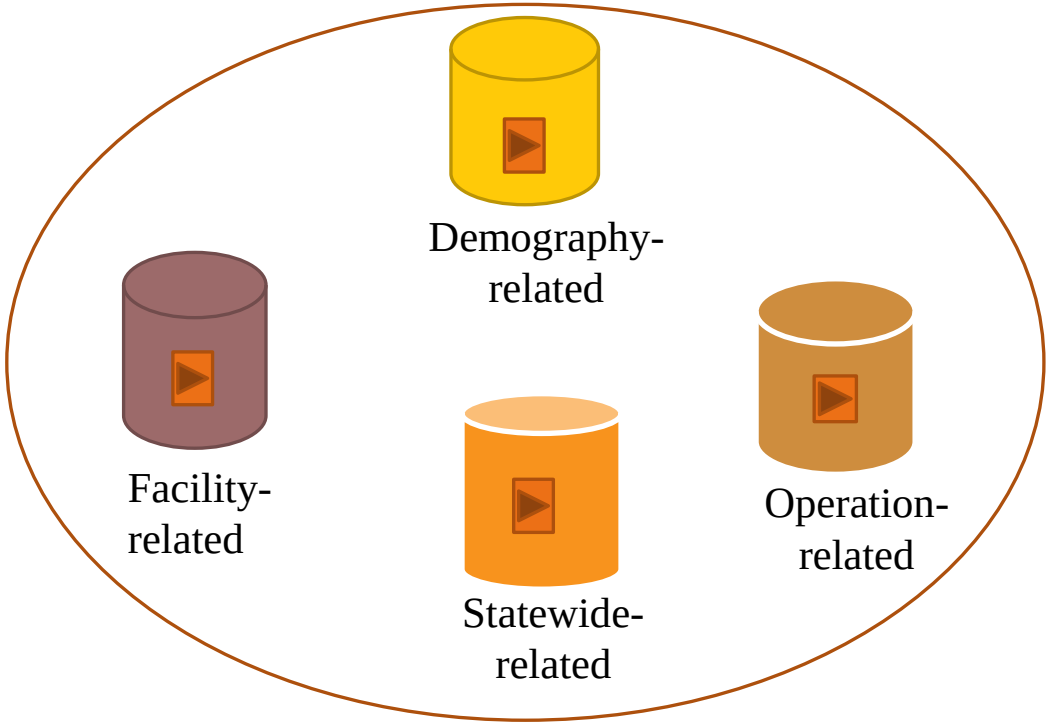
Dataset Overview

Before vs. After



ski_data

330 rows ×27 cols



ski_data

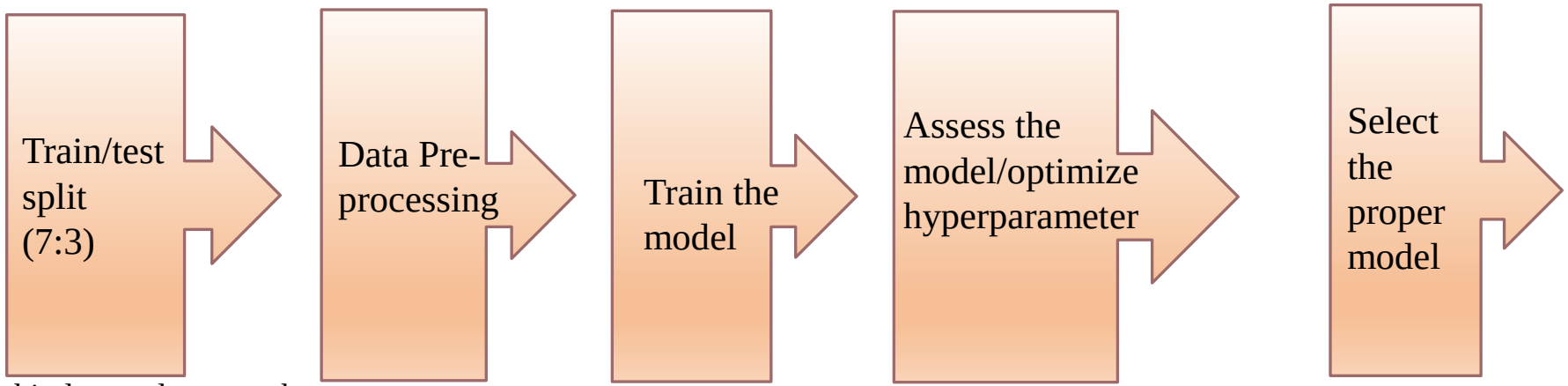
277 rows ×36 cols

Data Pre-processing and Modeling

I. Primary question: How can we develop a good predictive pricing model?

II. Minor questions:

1. Which imputation technique should we use for filling in the missing data, mean or median?
2. How do we train the model?
3. How can we find the best hyperparameters?
4. How can we evaluate the model? ….



X: ski_data-y-demography variables
y: 'AdultWeekend'

Impute missing values
Scale the data

Optimization technique:
RandomSearchCV, GridSearchCV

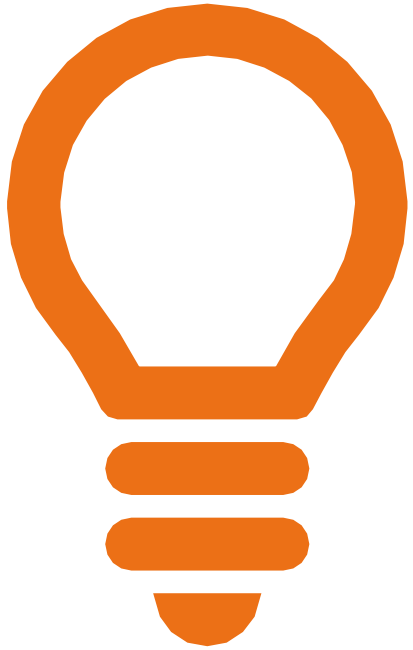
Regression model metrics:
R-squared, Mean absolute error, Mean squared error

Data Pre-processing and Training
Random Forest vs. Linear Regression

MAE: 9.50

MAE: 11.79

highly **accurate** and robust	slow in generating predictions
not suffer from the overfitting problem	difficult to interpret compared with linear regression
Tolerate missing values	
Display feature importance	
Simple to implement, easy to interpret	Assume normal distribution in variables
Less complex compared with random forest	Assume the linear relationship exists
Susceptible to overfitting	Outlier can have a huge effect on the model



I. Combine customer information in the study
II. Cluster analysis and market segmentation
III. Collect social media data to improve marketing accuracy

suggestions