A Dynamic Pricing Model for a Professional Sports Team

Brian Cain, Nikolai Saporoschetz, Theo Ginting, Matthew A. Lanham

Purdue University Krannert School of Management

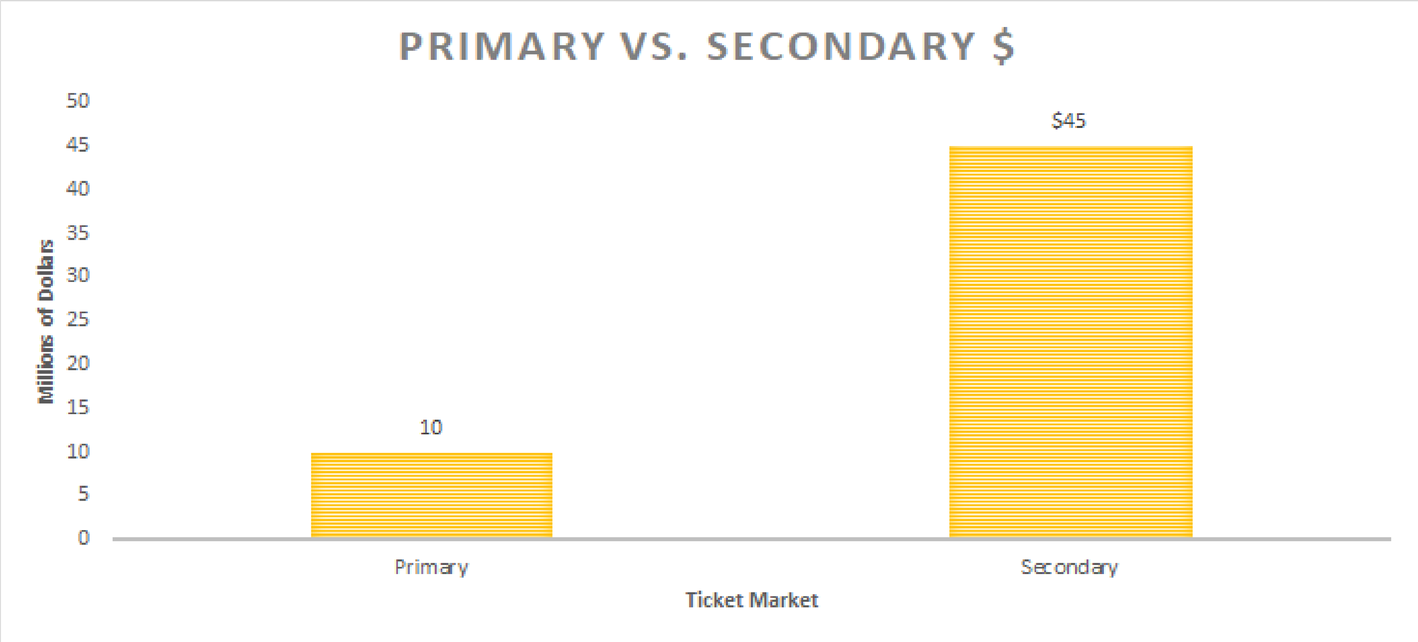
cain43@purdue.edu; nsaporos@purdue.edu; tginting@purdue.edu; lanhamm@purdue.edu

Abstract

In partnership with an NFL franchise we are developing a dynamic ticket pricing model that will optimize tickets sold while maximizing total ticket revenue. In 2015 the NFL authorized the use of dynamic pricing to recapture lost ticket sales, which has doubled some pro sports organizations single game revenues. By utilizing this model, organizations will be able to recapture lost sales and increase revenue.

Introduction

As technology has advanced and data analytics has proliferated, it has become essential for professional sports organizations to rethink their ticket pricing strategies.



The Solution: Dynamic Pricing using an optimization model

- Optimize the organization's revenue
- Fill out stadiums and compete with the secondary ticket market

Through our research, we will attempt to answer the following questions:

Research Question 1: Can an accurate and interpretable predictive model be developed to determine how likely ticket will be sold on the primary market?

Research Question 2: Can an optimization model be developed that maximizes the expected revenue of an event given market conditions, while minimizing unsold tickets?

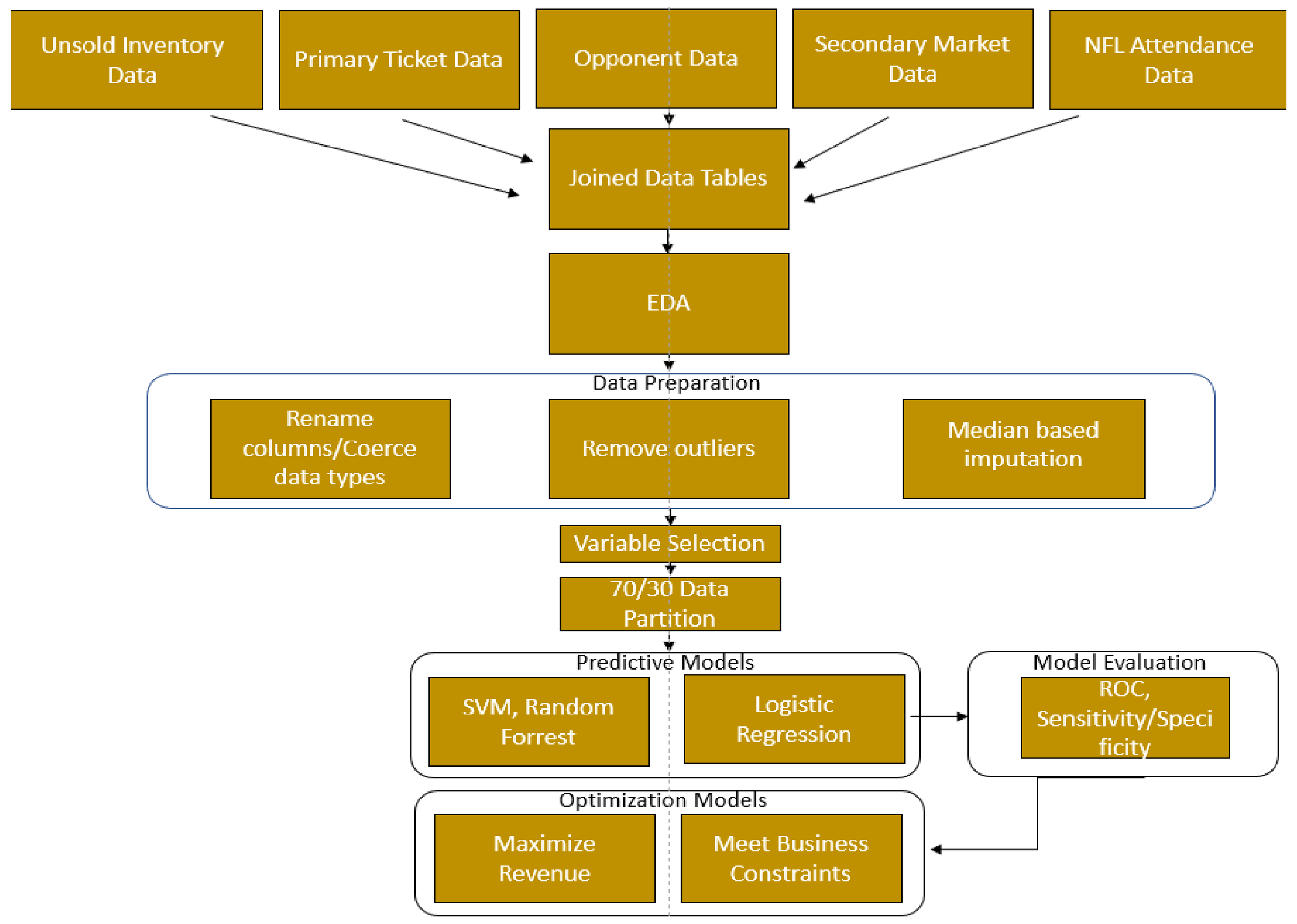
Literature Review

Published research papers on dynamic pricing in the Sports Industry are sparse. This is evident by the fact that in 2015, only 25% of all NFL adopted dynamic pricing as their pricing **(Kaplan, 2015)**. Contrary to other major Sports League in the United States, NFL has been reluctant to implement dynamic pricing. However, this is not to say that research on dynamic pricing is also scarce, it is not. In fact, research on dynamic pricing is abound, but in unrelated verticals.

Author	Sports-Industry	Dynamic Pricing	Predictive Model	Optimization Model
Drayer, Rascher (2013)	✔			✔
Paul and Weinbach (2013)	✔	✔	✔	
Fisher, Gallino, et al. (2018)		✔	✔	
Bertsimas, Vayanos, et al. (2014)				✔
Ferreira, Lee, et al (2017)			✔	✔
Liu, Chou, et al. (2019)		✔	✔	✔
Ye, Qian, et al. (2018)		✔		
Our Study	✔	✔	✔	✔

Methodology

Figure 2 outlines our study design, starting with joining and cleaning our data tables, then exploratory data analysis, to reducing outliers, preprocessing our data, and training our models.



Data

- Primary
- Secondary
- Unsold data
- Opponent data

Data Cleaning & Pre-Processing

- Removed missing values or used median based imputation
- Data was unbalanced so up sampling was utilized

Feature Selection

- An initial logistic regression model was trained with 61 variables
- Predictors with the six highest betas were selected for final model

Model Design

The data used for the model was for a single stadium section, consisting of 2,000 records. A tenfold cross validation was performed for a 90/10 train/test split.

Methodology (Approach) Selection

Our solution consisted of a two-tiered approach with a probabilistic classification model feeding into an optimization model. The probabilistic classifier provides a function to estimate the probability that a ticket is sold on the primary ticket market. The optimization model then maximizes the expected revenue for each ticket by setting the price of a ticket to a certain level.

$$Expected\ Revenue = \sum Price * \Pr\left(Ticket\ is\ Sold\ on\ Primary\ Market\right)$$

Model Evaluation & Business Performance Measures

- Sensitivity and specificity
- ROC
- Expected revenue

Results

We created several features to help predict the probability that a ticket is sold on the primary market. The 6 most powerful predictors were used in the final model.

Predictor	Beta
Home to Opposing Team Win Index	4.12397
Ticket Price Divided by Row Number	-7.09938
Opponent Success and Popularity Index	-2.91929
Opponent Teams Fan Spending	5.78219
Home Team Playoff Eligibility	15.68487
Final Ticket Price	7.23588

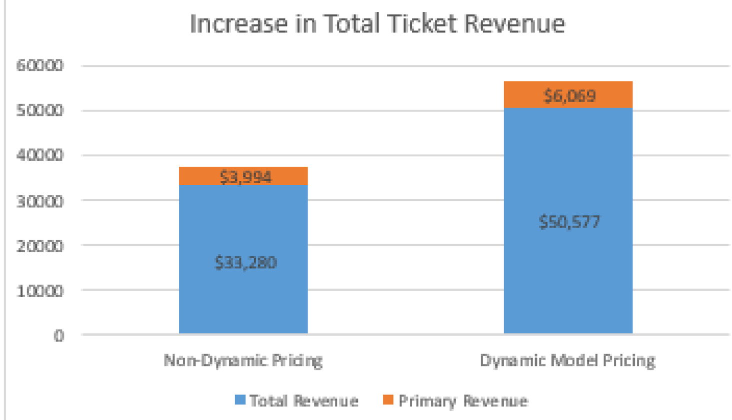
Using these predictors, the model proved accurate in predicting probability moving from the train to test set for factors like sensitivity, specificity and as well ROC.

Model Type	ROC
Logistic Regression	.68
Boosted Classification Tree	.97
Neural Network	.90

The results using this two-tiered model are positive. Total Ticket Revenue for Section 642 has a projected increase of 52%, leading to an approximately $20,000 increase in non-primary market revenue, and a $2,000 increase in Primary ticket revenue. Both revenue source increases will lead to a significant amount more money going back into the Sports Team Organization.

Main Impacts

- Logistic regression leading to a *52%* increase of captured revenue
- Potential to capture > *80%* of expected revenue through ADA, ANN optimization on the Primary market



Conclusions

Every season NFL franchises lose $45 million dollars to the secondary ticket market. We developed a two-tiered solution to help NFL teams recapture this lost revenue.

Through our research we were able to build an accurate and interpretable predictive model to determine how likely a ticket will be sold on the primary market. We also successfully built an optimization model to maximize expected revenue.

Acknowledgements

We thank Professor Matthew Lanham for constant guidance on this project.