Reinforcement Learning Approach to Dynamic Ticket Pricing

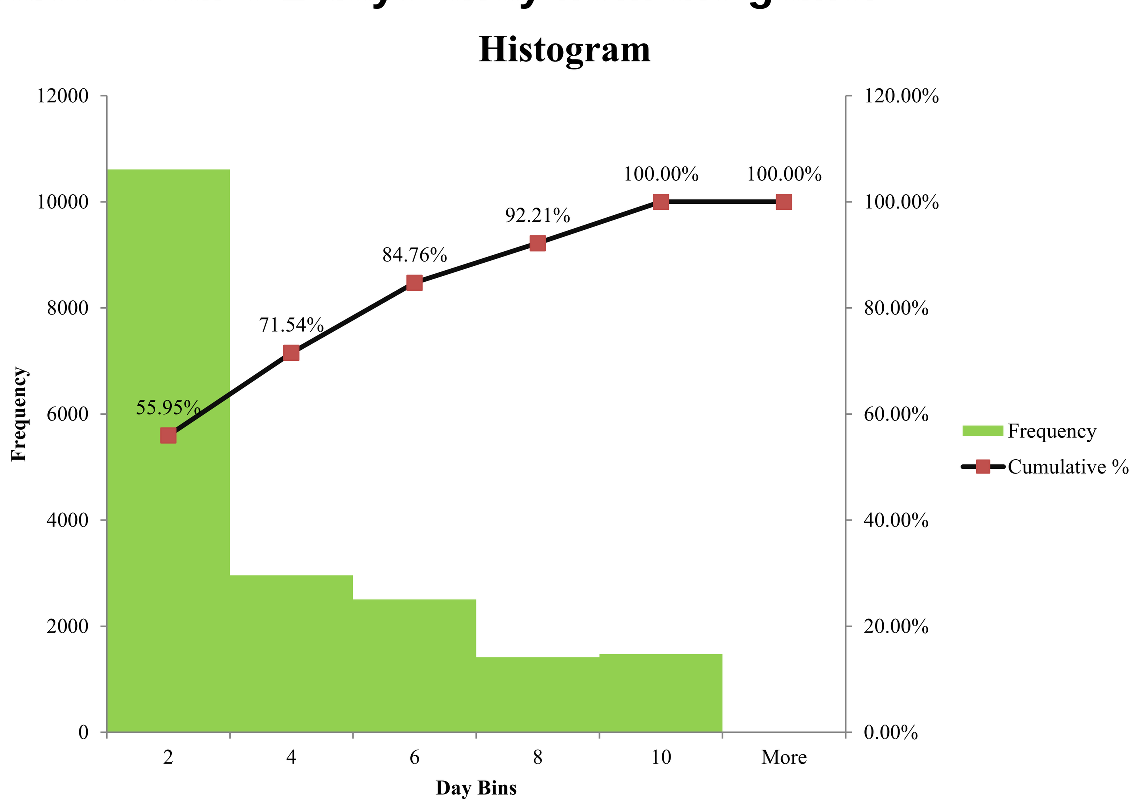
Brian T. Cain, Matthew A. Lanham
Purdue University, Krannert School of Management
cain43@purdue.edu; lanhamm@purdue.edu

ABSTRACT

We develop a Reinforcement Learning (RL) agent that serves as a basic template for making pricing decisions as a function of time. In this scenario we use sports ticket data from an NFL partner, both sold and unsold to extract the probability of a ticket being sold at a certain price and number of days prior to a game. Using these probabilities, we derive an environment containing rewards as expected revenue from selling a ticket at a certain price level on a certain day. The RL agent then iteratively takes actions/decisions within its environment based off the state it is currently in and receives an expected revenue reward. The agent continuously iterates through its environment until it has found an optimal policy of actions to take in certain states to maximize the long-run reward. RL is a relatively new field of artificial intelligence with a variety of applications that include sequential decision-making. Here we will show the power RL has on time-driven pricing decisions.

INTRODUCTION

While sporting tickets are often posted for sale a few months or weeks before an event, data from an NFL partner indicates that of tickets sold, 56% of these sales occur 0-2 days away from the game.



We focus on ticket prices and revenues as a function of days until the game.

Fig 1. Histogram of Ticket Purchases by Days until Game

Key Research Questions:

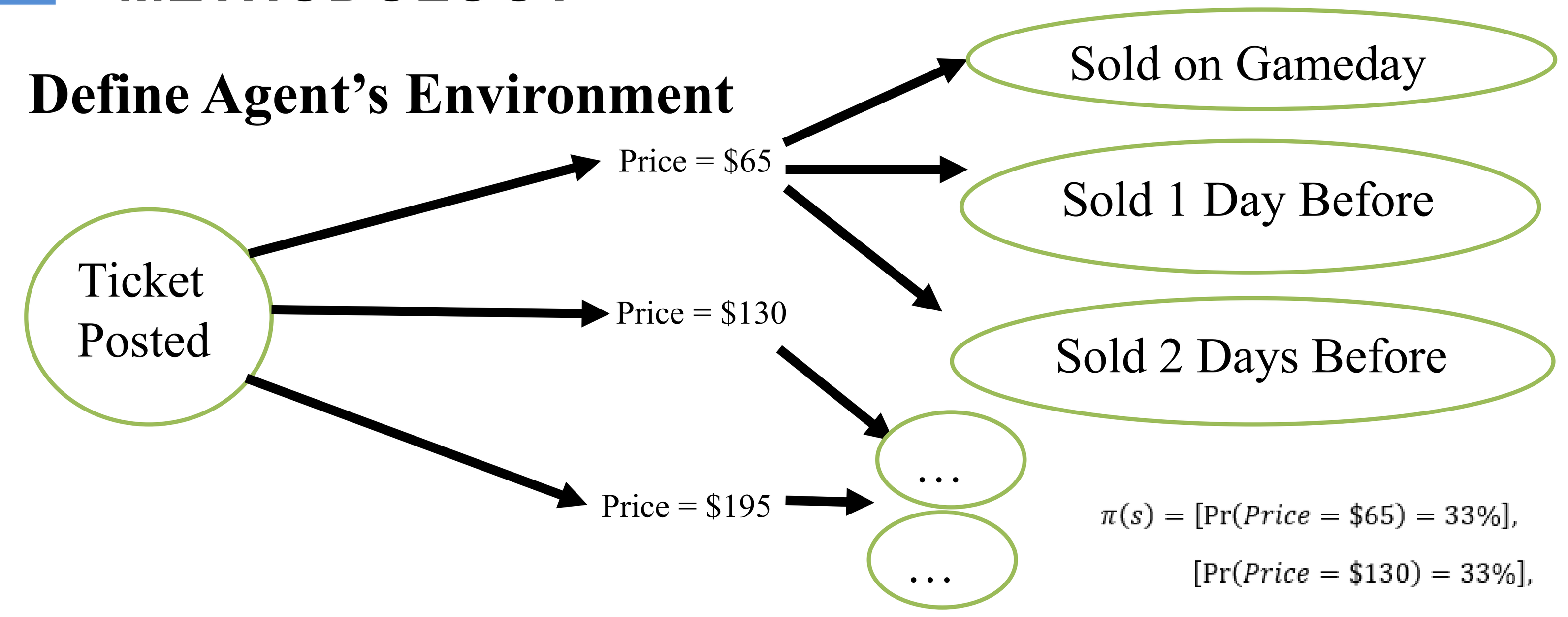
- Can a Reinforcement Learning agent learn an optimal pricing policy to maximize long-run ticket revenue?
- Can an optimal policy be obtained by using policy simulation rather than iteration?

LITERATURE REVIEW

1. Vanktesh Pandey, 2020, Deep Reinforcement Learning Algorithm for Dynamic Pricing of Express Lanes with Multiple Access Locations
   - Priced toll lanes based on current traffic in order to best reduce travel time
2. Cain, Ginting, Saporoschetz, 2020, A Dynamic Pricing Model for Professional Sports Teams
   - Used a two-tiered predictive model that fed logistic regression parameters into a linear optimization to maximize the expected revenue for a single ticket
3. Bła ̇zej Osi ́nski, 2020, Simulation Based Reinforcement Learning for Real-World Autonomous Driving
   - RL agents have been trained in simulated environments but rarely cited that the agent's policy is determined by mass simulation.
   - We examine Monte Carlo simulation in the policy segment of Reinforcement Learning in our study

METHODOLOGY

Define Agent's Environment



$\pi(s) = [\Pr(Price = \$65) = 33\%],$
$[\Pr(Price = \$130) = 33\%],$
$[\Pr(Price = \$195) = 33\%]$

Agent Creation/Execution

Obtain Raw Data

Calculate Transitional Probabilities

$\Pr(S', R | S, A)$

$S' = Successor\ State$

$S = Initial\ State$

Define Reward/Value Function

	EV[Ticket Sale]
Gameday	154.0183075
1 Day Before	63.7005709
2 Days Before	74.37019036

Policy Evaluation w/ Bellman Updating

$$v(s) = \sum_a \pi(a|s) \sum_{s',r} p(s',r|s,a)[r + \gamma v(s')]$$

Monte Carlo Policy Iteration

10,000 Simulations: $\pi(s) = [\Pr(Price = \$65) = Random\ \%],$
$[\Pr(Price = \$130) = Random\ \%],$
$[\Pr(Price = \$195) = Random\ \%]$

Select Optimal Policy

STATISTICAL RESULTS

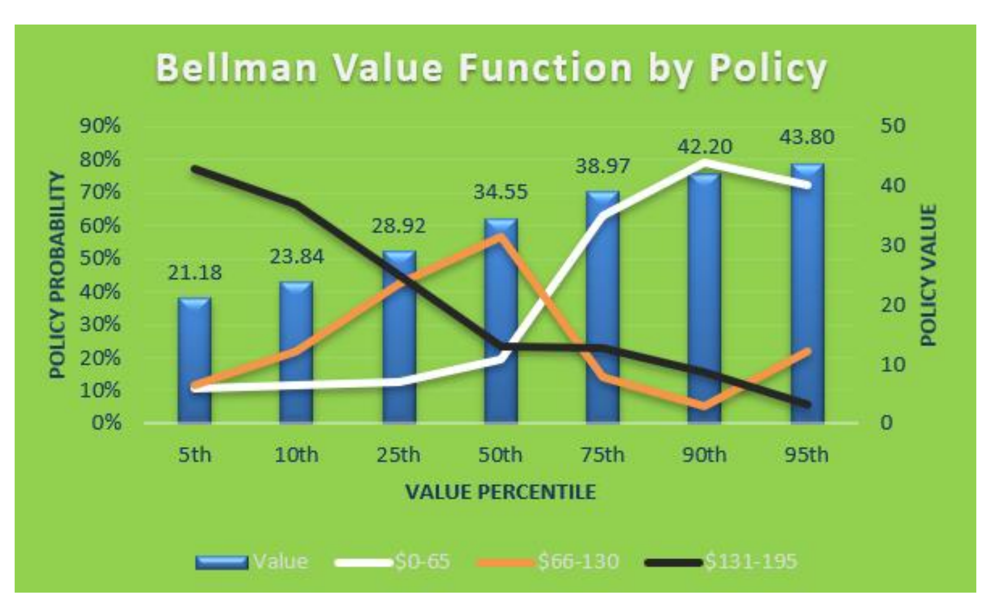


Fig 3. Policy Simulation Results

Percentile	$0-65	$66-130	$131-195
5th	11%	12%	77%
10th	11%	22%	66%
25th	13%	43%	45%
50th	20%	57%	24%
75th	63%	14%	23%
90th	79%	5%	16%
95th	72%	22%	6%

Fig 4. Policy Value Percentile

Figure 3 and 4 display our Monte Carlo policy iteration simulation results. After randomly simulating 10,000 different decision policies we analyze the RL agent's ability to maximize the value of selling tickets at specific prices.

Key Findings:

- **Value increases as agent sells more $0-65 tickets**
- **Agent learns value tradeoff between selling tickets $66-130 and $131-195, ultimately learning $66-130 contributes more to expected value**

Expected Business Impact

Current Pricing

$\pi(s) = [\Pr(Price = \$65) = 38\%],$
$[\Pr(Price = \$130) = 43\%],$
$[\Pr(Price = \$195) = 18\%]$ → **Agent Found Value = 37.58**

$\pi(s) = [\Pr(Price = \$65) = 72\%],$
$[\Pr(Price = \$130) = 22\%],$
$[\Pr(Price = \$195) = 06\%]$ → **Agent Found Value = 43.80**

Using Monte Carlo Optimal Policy: → $\%\ Value\ Differnce = +16.55\%$

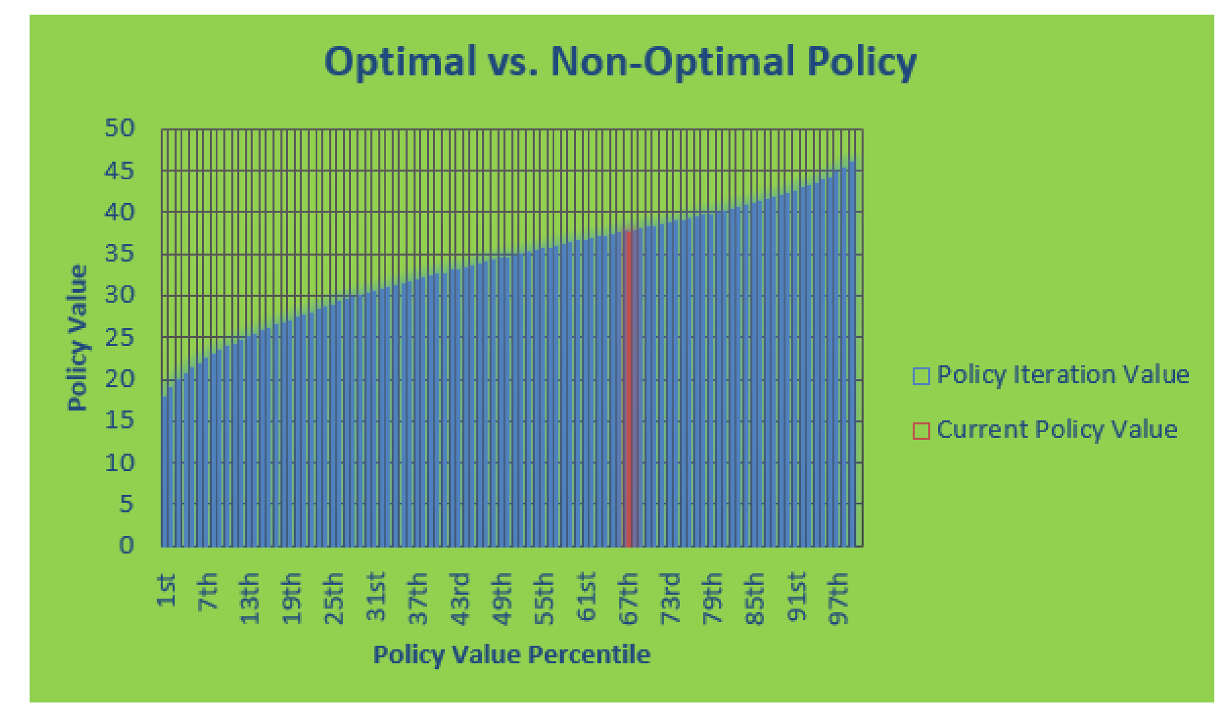


Figure 5 displays the location of the NFL team's current pricing policy 2 days before a game. As seen above this policy value is around 37.58 which lies around the 67th percentile of all policies evaluated during the Monte Carlo policy simulation.

By using this less-than-optimal policy we are losing a value of 16.55% reward in the long-run.

Fig. 5 Percentile Bar Chart Displaying Location of Current State Policy

CONCLUSIONS

In this study we show that a Reinforcement Learning agent can learn an optimal ticket pricing policy. By setting up an environment of actions composed of ticket pricing decisions and states composed of sale date we were able to model ticket price as a function of time in the eyes of the Reinforcement Learning agent.

After developing a policy evaluation algorithm in which the agent performs trial-and-error by pricing certain tickets at certain values with different probabilities of being sold we were able to find an optimal policy to maximize reward by simulating 10,000 random policies for the agent to act by.

Secondly, by performing policy iteration using a randomized simulation of different policies, we were able to get a vast distribution of many policies. From these policies we were able to select one in the 95th percentile that makes us confident our agent will select a pricing scheme that maximizes long-run revenues.

ACKNOWLEDGEMENTS

I would like to thank Professor Matthew Lanham for his mentorship.