

Puck Predictions: Unraveling the NHL Game Forecasting Riddle

Jason Vasquez Dylan Skinner Jeff Hansen
Benjamin McMullin

March 26, 2024

Abstract

The goal of this project is simple: predict the outcomes of NHL games from any given state. As simple as the problem statement is, however, the solution is not so straightforward. To solve this problem, we will use a variety of machine learning techniques, including logistic regression, XGBoost, and ARIMA models. Additionally, we utilize a form of MCMC to simulate the outcomes of games from any given state. Our hypothesis is that we will be able to successfully predict the outcomes of NHL games with a high degree of accuracy using these tools.

1 Problem Statement and Motivation

2 Data

Our data came from the hockeyR Github repository[1]. This repository contains an abundance of data about every NHL game that has occurred since the 2010-11 season. This data includes information about the events that transpire in a game (hits, shots, goals, etc.), which teams are playing, who is on the ice, and the final score of the game.

3 Methods

4 Results

5 Analysis

6 Ethical Considerations

7 Conclusions

References

- [1] Dan Morse. hockeyR-data: A collection of hockey datasets for use with the hockeyR package. <https://github.com/danmorse314/hockeyR-data>, 2024. Accessed: March 2024.