Abstract on Applied Data Science and Machine Learning in Sports Statistics

Sports is a multi-billion-dollar revenue global industry with data ranging from historical aggregate data to modern day play possession times in seconds. This report aims to apply Data Science and Machine Learning tools to sports using the National Hockey League (NHL) data. The data is sourced from online resources but is all original NHL owned data.

The project is split into two hypothetical data agents within the sports industry. The first is the sports bettor whose goal is to raise margins or stack the odds as much in their favor. The second agent is the team analyst designed to use data to improve the team’s overall performance within the organization (think Moneyball).

The goal is to provide insight into real-world problem solving and to explore the limits of data science and machine learning.