NHL Player Valuation System: Using Performance Metrics to Understand and Predict Player Salary

Michael Kuby, Sangmun Kim, Yuyang Chen, and Haichen Sun

Table of Contents

[Motivation 2](#_Toc163336419)

[Problem Statement 2](#_Toc163336420)

[Limitations of quantitative analysis 3](#_Toc163336421)

[Data Science Pipeline 3](#_Toc163336422)

[Methodology 3](#_Toc163336423)

[Evaluation 3](#_Toc163336424)

[Data Product 3](#_Toc163336425)

[Lessons Learned 3](#_Toc163336426)

[Summary 3](#_Toc163336427)

# Motivation

The NHL is commonly referred to as a results-based business. Players get paid to perform at the highest levels, and some players can transcend barriers and do the unimaginable on the ice, year after year. Those who find a way into the league and succeed are rewarded handsomely come contract time. However, for those involved in the contract negotiations, including general managers, player agents, players themselves, and those on the outside looking in with curiosity, the methods and models for explaining a player’s value are somewhat opaque. This motivates the project herein, which attempts to peel back the curtain.

## Existing solutions: Player Cards from The Athletic

Predicting a player’s value is not a new phenomenon. It is an integral element of all professional sports and is done regularly by general managers, player agents, fans and players alike.

One such tool that attempts to value players in a similar vein to the project herein is The Athletic Player Cards (<https://theathletic.com/5015509/2024/03/30/nhl-player-cards-pacific-division/>). As stated on their website, the “value is according to Net Rating, our all-in-one player value stat that’s based on each player’s Offensive and Defensive Rating. The cards showcase all the states that determine how strong each player’s ratings are on and off the puck, a weighted combination of their production (goals, assists, expected goals, blocks, penalty differential) and their play-driving (on-ice expected and actual goal stats).”

While interesting from a fan perspective, it is rather complicated to determine the utility of the Player Cards due to their seeming lack of reasonableness. For example, The Athletic Player Card for Connor McDavid values him at $19.0 M, representing 22.75% of the salary cap, which is hard to reconcile with the fact that no player in the salary cap era has ever been paid above 16% of the salary cap, nor could they be since the most a player can be paid under the collective bargaining agreement is 20%. Hence, our model tries to reconcile the valuation of a player based on their performance in a realistic, practical, and understandable manner.

# Problem Statement

Given our motivation for the project, we try to answer the following questions:

* What is player A worth? Why?
* Why does player A get paid X and player B get paid Y?
* What do general managers find valuable in a player?
* Who is overperforming relative to their salary?
* Who is underperforming relative to their salary?
* What explains the valuation of a Forward (Center, Left-wing, Right-wing)?
* What explains the valuation of a Defenceman?

## Limitations of quantitative analysis

Given the problem statements we wish to answer, it is vital to recognize the limitations of quantitative analysis. Many elements lead to success in team sports that numbers cannot capture. Leadership, mentorship, spirit, camaraderie, sportsmanship, humour, supportiveness and kinship are all necessary, desirable, valuable, and sought-after characteristics that cannot be captured by performance metrics alone. As such, any quantitative model that uses statistics, machine learning, or otherwise inherently fails to capture aspects of a player’s value. As such, we recognize the limitations of our model in the sense that it strictly uses quantitative analysis to drive value and should be used as a starting point, not an endpoint.

# Data Science Pipeline

# Methodology

# Evaluation

# Data Product

# Lessons Learned

# Summary