

# Predicting NHL Goal Scoring

---



**Capstone Sprint 3**

Taylor Gallivan

24 Nov. 2023

# Project Overview

Within the North American professional sports landscape, quality data has never been more accessible, or more in-demand:

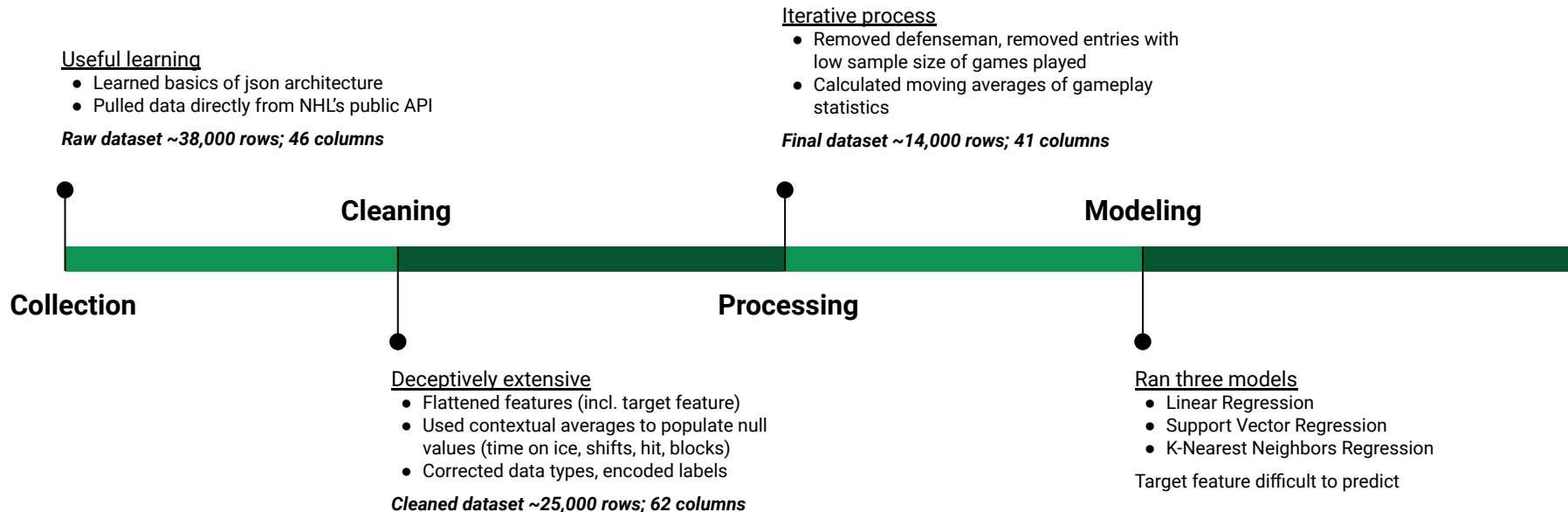
- Advances in game-tracking technology improving quality and availability of data → NHL Edge just released to public (Oct. 2023)
- US sports betting market growing ~10% / yr, growth forecasted through 2030<sup>1</sup>

## Problem Question:

Can an ML model (or models) be trained to accurately predict an NHL forward's goal output for a season, relying solely on common historical statistics?

<sup>1</sup> Source: *Grandview Research*

# Project Roadmap



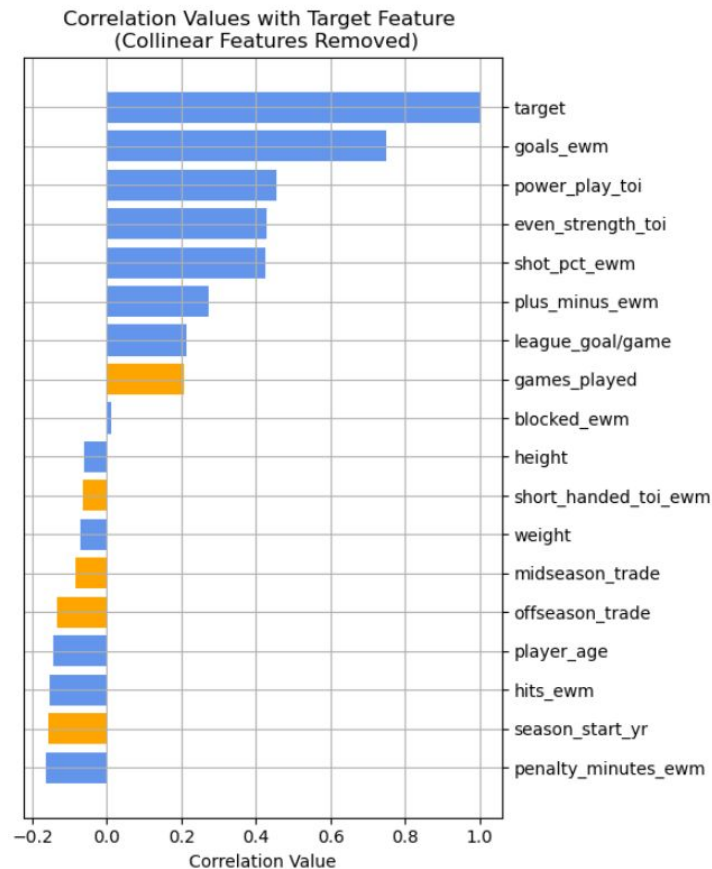
*Biggest takeaway: domain knowledge can be a blessing and a curse! Easier to identify deficiencies and dependencies; but once identified, harder to ignore*

# Baseline Modeling Results

Model	Linear Regression	Support Vector Regression	K-Nearest Neighbors Regression
$R^2$ - train	61.0%	61.4%	69.4%
$R^2$ - test	60.8%	60.7%	54.3%
MAE	5.5	5.5	6.0
MSE	54.0	54.1	63.0
Accuracy*	36.4%	37.9%	32.0%

SVR model was the most accurate,  $R^2$  scores & errors among lowest  
All data scaled using MinMax Scaler

\* Accuracy compares predictions vs. actual values of target, accepted margin of error of +/-3 goals



# Optimized SVR Model

Model	Support Vector Regression*
$R^2$ - train	64.4%
$R^2$ - test	62.9%
MAE	5.3
MSE	49.1
Accuracy	38.3%

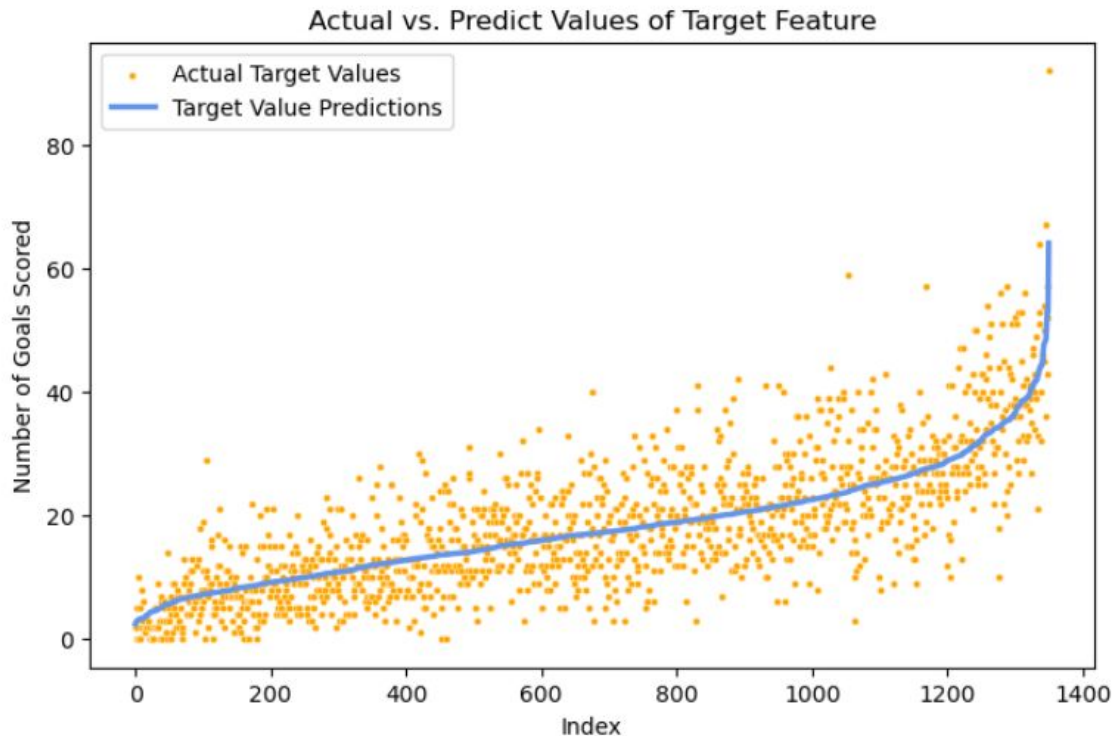
All metrics improved following optimization

\* Model parameters:

Kernel = rbf

C = 10

$\epsilon$  = 5



# Next Steps

1. Assess model effectiveness:
  - a. Collect historical betting odds for “player goals - season total”
  - b. Re-run best model w/ odds in test set
  - c. Compare oddsmakers’ predictions against my model’s
2. Prepare web demo of goal predictor
3. Try running a neural network on the full cleaned dataset

