Springboard School of Data Science

Predictive Modeling in the National Hockey League

An Exploration of Aggregate Player Statistics and Team Performance with Machine Learning

Logan Schmitt

October 24, 2023

I. Background

The National Hockey League is one of the “big four” sports leagues in the United States and generates billions in revenue each year [1]. Hockey, however, is unique among the major sports in that it is very difficult to predict the outcome of a given game or season. The reasons for this phenomenon are twofold.

The first cause for unpredictability is randomness. While no sport would be exciting if the outcomes were completely predictable, hockey is considered the “luckiest” sport of the big four [2]. Scoring events are rarer in hockey than in the other three sports. Games often come down to a single bounce of the puck to decide a winner.

The second reason for unpredictability is that hockey is considered a “weak link” sport [3]. In a “strong link” sport, like NBA basketball, team performance is driven by the best players. Those players get to play the important minutes of the game and drive scoring. Simply put, the team with the better top-performing players tends to win. In contrast, the outcome of weak link sport is driven by the weaker players. In the NHL, teams deploy four lines of attackers on a rolling basis. The best offensive players may play just over a third of any given game, meaning there is much more time for the weaker players to impact the game. When considering the relative rarity of goals in hockey, it makes sense that the team with better depth, even at the expense of top-line talent, is in better shape to win games.

The combination of these two factors means we cannot simply project each team’s performance based on expectations for its star players. We have to take the entire roster into account. I have attempted to do so in this project, as detailed below.

II. Problem Identification

There are two main groups of stakeholders for a predictive NHL model. The first group, obviously, comprises the people who assemble NHL teams. Each team has a general manager, assistants, player scouts, player analytic departments, and coaches. These people rely on data to make decisions to improve their team’s chances of winning games. The second group comprises sports bettors and oddsmakers. Since sports betting was legalized in the United States within the past 5 years, sports betting has exploded in popularity, generating nearly $10 billion in revenue in 2022, despite not being legal in all 50 states yet [4]. Oddsmakers use predictive models to set their odds and bettors are increasingly using them to find good value bets.

Having established that the stakeholders for this project are numerous and diverse, it is important to understand the goal of the project. I set out to predict which NHL teams would qualify for the 2022-23 playoffs. In particular, I wanted my model to be applicable for roster-building decisions and evaluation. I imagined a person in a roster-building role using the model when deciding which players to trade or sign during the offseason. I also imagined sports bettors and oddsmakers using the model to evaluate those trades and signings as they are announced during the offseason, to update playoff odds and make wagers on teams. Ultimately, I tried to answer the following question: Can the aggregate statistics of each team's roster from the 2021-22 regular season reliably predict the team's performance in the 2022-23 season?

III. Data Wrangling

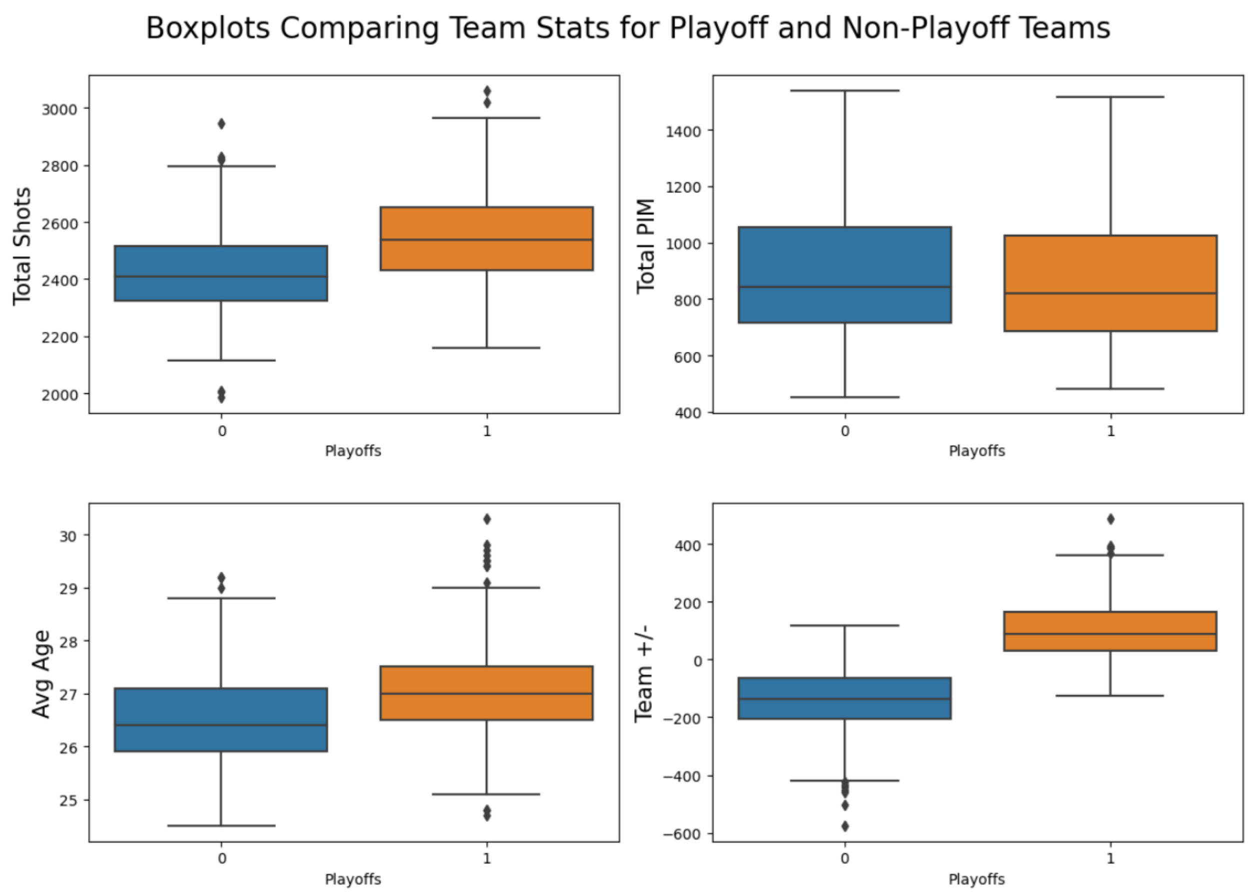
While there are a growing number of advanced metrics for NHL players, most of them are relatively new, meaning there isn’t much history for training a model. I chose instead to pick statistical measures that go back as far as the most recent lockout in the 2004-05 season. By using these statistics, I could use closer to 20 seasons’ worth of data to train the model. I also capture the entire careers of most modern NHLers. I sourced the data for this project from Hockey Reference, which keeps a complete historical archive of NHL data.

For each of the years, I pulled a table of overall standings, with each team’s final points percentage and playoff status. Then I pulled each team’s roster and each player’s respective stats from the following categories: Age, Games Played, Goals, Assists, Points, +/-, Shots, Shooting Percentage, Offensive Point Share, Defensive Point Share, and Cumulative Point Share. Once I had the statistics for each player and each team for each season, I began to process them into a usable form.

The main step I took at this point was aggregation. I wanted a statline that reflected each team’s season, with metrics like total goals, average age, highest individual goal scorer, and so on. I quickly noticed that I would have to adjust some entire seasons’ worth of data. Due to a lockout and two COVID-shortened seasons, there were a total of three seasons in which teams played, on average, fewer than 72 games. I aggregated all the statistics and then projected them out to an 82 game season. This way, I could compare those seasons fairly in my EDA.

IV. Exploratory Data Analysis

My goal with exploratory data analysis was to understand how playoff teams and non-playoff teams differ in a number of aggregate statistical measures. This entire project was based on the assumption that hockey rosters must be evaluated holistically as opposed to focusing on just a handful of key players. As such, the exploratory data analysis phase was crucial in understanding whether or not I should continue down this path. I plotted my season aggregate features as boxplots and split them by playoff status. I visualized each individual statistical measure as shown in Figure 1 and could easily see where the playoff and non-playoff team clusters were located.



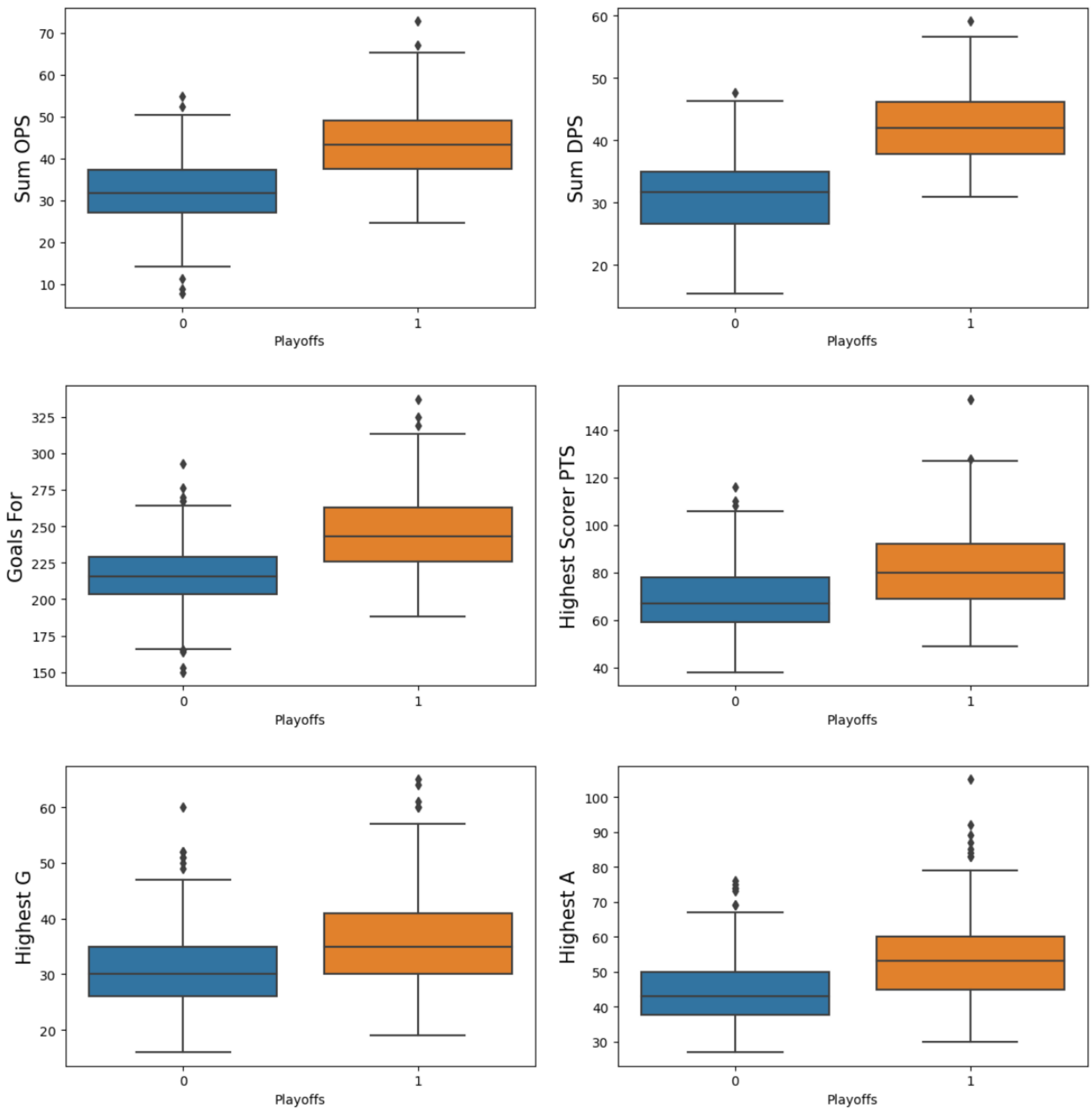


Figure 1. Boxplots showing playoff and non-playoff team statistics from 2005-2022

The boxplots in Fig. 1 showed a clear divide along a number of metrics between the median playoff team and the median non-playoff team. Most of these differences are intuitive. The playoff teams shoot more, take fewer penalties, and have a higher +/-, meaning players are on the ice for more goals for than goals against. I was surprised to see that playoff teams are older on average, but I think it makes more sense after some pondering. Younger teams are likely populated by high draft picks, which are acquired through years of mediocrity. Once those players learn the league and mature, they are more poised to have regular season success. The boxplots showed a clear enough distinction between playoff and non-playoff teams that I felt comfortable continuing my project.

I examined goals and shots next. I was curious to see if playoff teams score more goals and score at a higher rate than their non-playoff counterparts. Figure 2 shows the shooting percentage trend line for playoff and non-playoff teams.

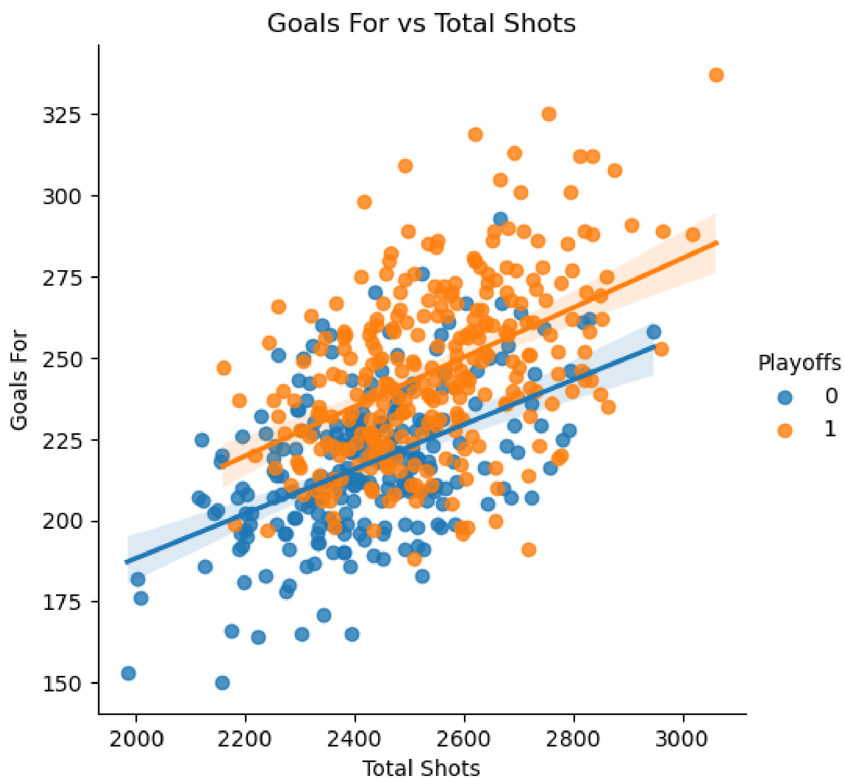


Figure 2: Scatterplot of Goals and Shots

The slopes of the trend lines are about equal, meaning that there isn't an appreciable difference in shooting percentage between playoff and non-playoff teams. The real difference in goals is just that the average playoff team generates more chances and takes more shots.

The next useful figure from my EDA was the correlation heat map, shown in Figure 3.

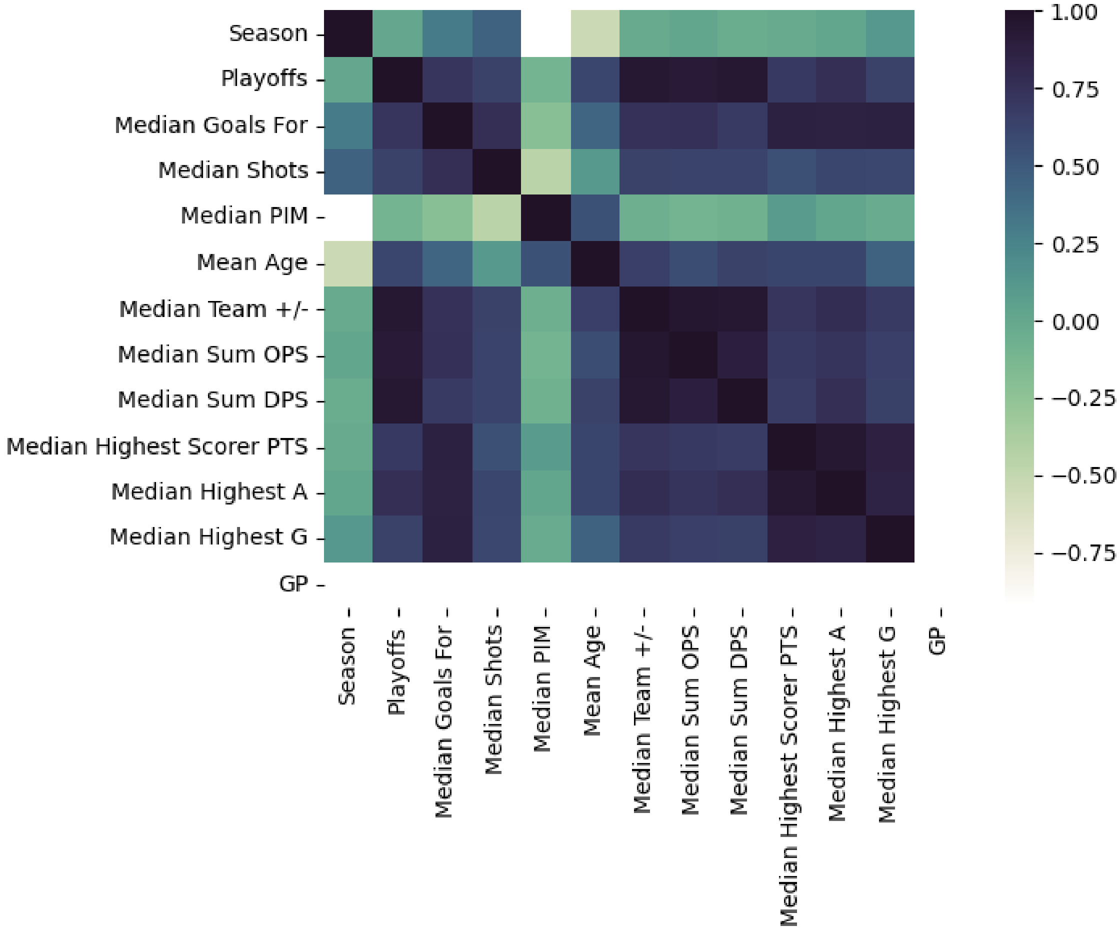


Figure 3: Correlation heat map of aggregate season features

This heat map showed me some predictable correlations but I did not see any concerning collinearity. It made sense that more goals and more offensive point shares correlated with making the playoffs.

The last exercise of my EDA process was to check out Principal Component Analysis (PCA). The explanatory strength of the principal components wasn't great, as I could only explain about 60% of the variance with the first two principal components. Nevertheless, I plotted the data, colored by playoff status, along the principal component axes in Figure 4.

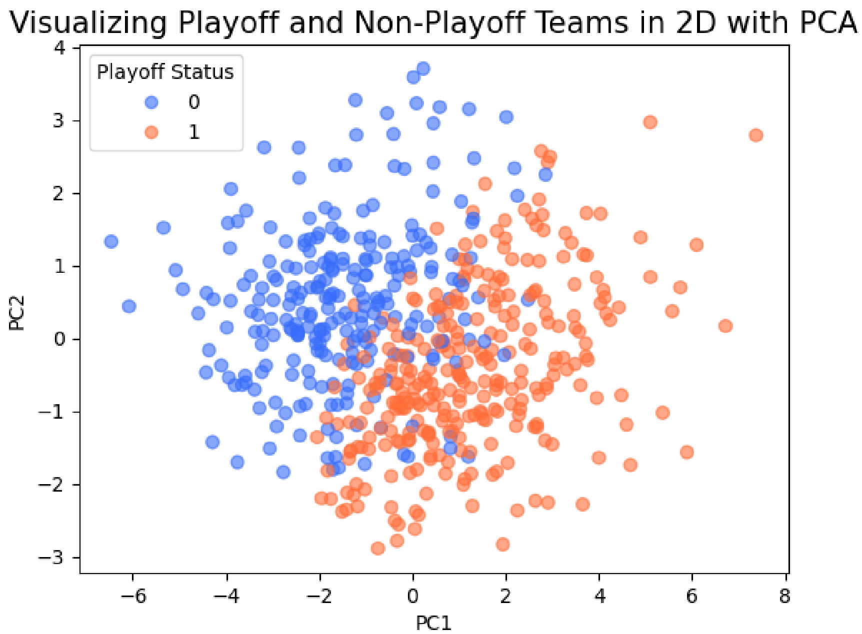


Figure 4: Playoff teams plotted along Principal Component axes

There is a region with plenty of mixing in the center of the graph, but the playoff and non-playoff clusters are pretty well-defined. This view made me wonder about using a Support Vector Machine in my modeling process to find a hyperplane that separates the classes.

The result of my EDA was an increased confidence in my approach. I saw clear divides between playoff and non-playoff teams and was ready to continue processing and feature engineering.

V. Preprocessing

My goal for processing was simpler to achieve than to describe. Since my model needed to predict a team's playoff status at the beginning of the season, I couldn't use the players' stats from that season, as it hadn't happened yet. I thought about making a model to predict each player's stats, but it was quickly apparent that this was not within the realm of possibility for a number of reasons. Was I going to create one model and treat goalies the same as centers? If I created a model for positions, should I also create a different model for top-line players and depth players? Would I create a model for each of the statistics I hoped to use in my final model? I realized I could skip this layer of modeling and simplify my process. Instead, I just made a roster for each team and used the statistics for all those players in the preceding season. There are certainly some flaws with this approach — some players will improve, some will decline, some didn't play the previous season and therefore will have no statistics. By and large, however, it is rare to see a player's per-game numbers change dramatically from one season to the next. With this approach, I effectively created a null-hypothesis model and predicted each player's performance would not change from last season to this season.

In order to generate statistics for each season, first I had to aggregate each player's statistics from each season. Some players had played for multiple teams in a given season as a result of trades, so I had to combine those statistics into a single statline. Now I had a DataFrame with each player's aggregate statistics for each season. Next, I created a new DataFrame with just player, team, and season information. Then, I simply incremented the season by one year in my aggregate DataFrame and merged it with the new DataFrame on the player and season columns. The effect was a shift of each player's statistics forward by a year and an association of that statline with the player's team in that season. This operation handled all the offseason moves like trades and signings. With this transformation complete, I could wrap up my feature engineering.

The final feature engineering steps were aggregating and scaling. I aggregated many of the same statistics as I examined in my EDA, only with the team and season transformation applied. I added features for each team's top three goal scorers and top three assist earners. Then I applied a standard scaler to all of the relevant columns. The scaler was especially important because league-wide scoring has fluctuated so dramatically over the past 20 seasons. There were several seasons in the mid-2010's where no player exceeded 100 points. In the 2021-22 season, 8 players hit or exceeded that mark. In order to account for these fluctuations, I fit and applied the standard scaler across all teams for one season at a time. This way, a team with a 100 point scorer in 2014 was treated as just as much of an outlier as a team with a 120 point scorer in 2022. Since teams are only being compared to other teams within a given season, it made the most sense to do this. Finally, I added my target column, points percentage, and I was ready for modeling.

VI. Model Evaluation

The target of my model was to predict which teams would make the playoffs, which makes the problem seem like one of binary classification. Unfortunately, it wasn't quite so simple. The NHL's playoff criteria have changed over time, but they currently dictate that 16 teams make the playoffs. Those teams consist of the top three teams from each of four divisions and two wildcards from each conference. There was no way to ensure that a binary classification model would understand those parameters. Instead, I approached the problem from a regression angle and attempted to predict each team's points percentage.

Points percentage is simply the number of points a team collects divided by the number of points available in a season. Like so many sports statistic "percentages," it is not multiplied by 100 and is instead represented to three decimal places. Typically, the number of points would make just as much sense for this problem, but because COVID interrupted the 2020 season, points percentage is a better choice for the data. A team gets two points for any type of win, zero for a regulation loss, and one point for an overtime or shootout loss. In a normal, 82-game season, each team can earn up to 164 points, with 100 points being a safe bet to make the playoffs. That number corresponds with a points percentage of .610.

I wrote a function to convert points percentages into playoff seedings depending on the team and the season. I also created a function to compare those playoff seeding predictions against an answer key of true playoff teams and return a confusion matrix and classification report.

Next, I had to determine how to split the data into training and testing sets. Due to the nature of the data, I took a time-series split approach. I knew the final test year had to be 2021-22, since 2022-23 was my holdout prediction season. I wrote a function which iterated through the years in the dataset, incrementally adding more years to the training set and testing on the following year. This allowed me to perform a grid search-style cross validation on the hyperparameters of each model. The function returned performance graphs for three models with the highest mean weighted accuracy. I put more weight on more recent test years because they were supported with more training data and they were also closer in time proximity to my holdout test point, 2023. I chose to output the performance of the three best models so I could see which hyperparameters were consistent, which changed, and what effects those changes had on the accuracy over time graphs. With these functions created, I began exploring models.

VII. Model Selection

Linear Regression

I wanted to try a variety of regressors, but I started with an ordinary least squares linear regression. With no hyperparameters to tune, this model was very simple to implement. The performance was surprisingly good throughout the cross-validation process as shown below in Figure 5.

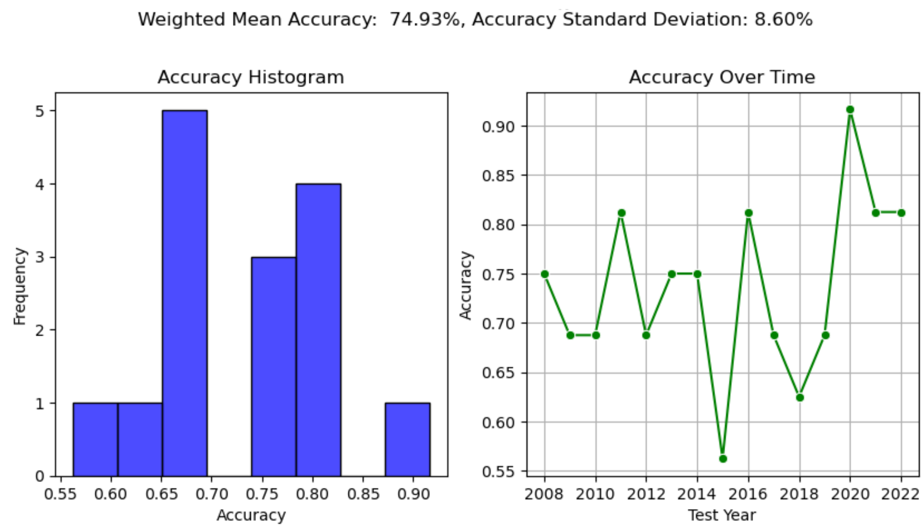


Figure 5: OLS Linear Regression Performance Graphs

The accuracy over time graph demonstrates that the high accuracy in recent years really pulled up the weighted average. The model excelled at predicting playoff teams in 2020, at over 90% accuracy. That particular year, though, was a bit of an outlier. Due to the COVID pandemic, the league suspended the regular season in early March, when most teams had about 10 games left. Instead of resuming the regular season in the fall, the league declared that the top 12 teams in each conference as measured by points percentage would make the playoffs. This decision meant that only 7 teams missed the playoffs that season. As shown by the PCA graph (Figure 4) in the EDA section above, the hardest part of predicting the playoff teams is separating the teams that are quite similar. It is generally not as hard to tell which teams will end up near the bottom of the standings. The 2020 data point confirms that. This trend was consistent for all the models tested.

The next type of model I examined was an elastic net regression model. This strategy combines the penalties from lasso and ridge regression. Several elastic net models achieved the same performance statistics as the linear model but none exceeded it. I noticed that each of these models had very low values for alpha, meaning the penalty term was minimal. The graph for the elastic net models were identical to Figure 5, confirming that the model was effectively just an OLS linear regression model.

I also tried a Bayesian ridge regression model from the family of linear regressors. The best models here were identical in performance and prediction to the OLS linear model as well. After seeing the same result with the elastic net model, this was less surprising.

Random Forest Regression

Next, I steered my efforts away from linear models and built a random forest regressor. I was eager to see if this model could outperform the linear model. I experimented with a variety of hyperparameters, but the best models consistently used just a handful. One of the best performing models is shown in Figure 6 below.

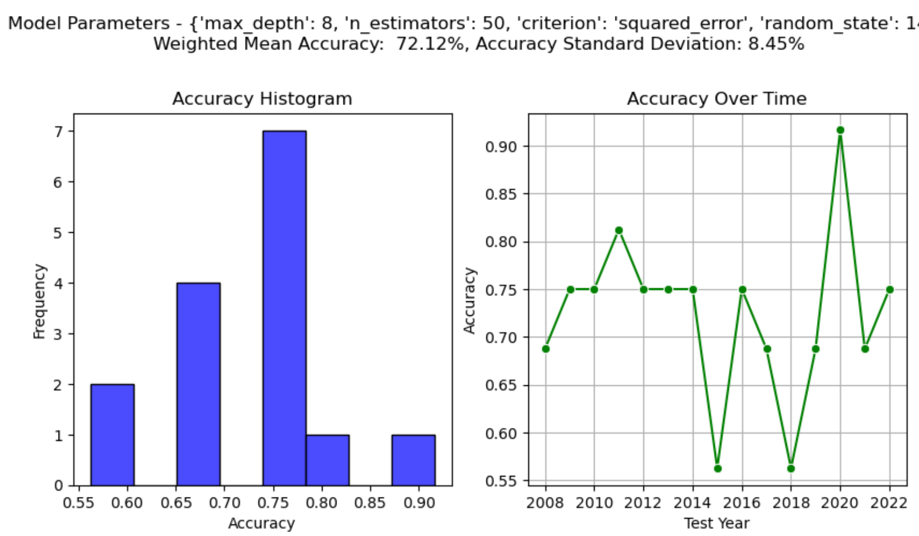


Figure 6: Random Forest Regressor Performance Graphs

Several random forest models achieved similar performance. The above model was best, but by a slim margin. It had a more stable performance in the early years of the training data than the others. I was surprised to see that the model actually had a lower weighted average accuracy than the OLS model, even though it had more points at or above the 75% accuracy threshold.

Support Vector Regression

I tried a support vector regressor in hopes that it might improve upon the linear regression models. In fact, I removed the linear kernel function from the hyperparameter options because it was returning a model that was again, functionally identical to the OLS linear regression model. The performance graphs for one of the support vector regressors is shown below in Figure 7.

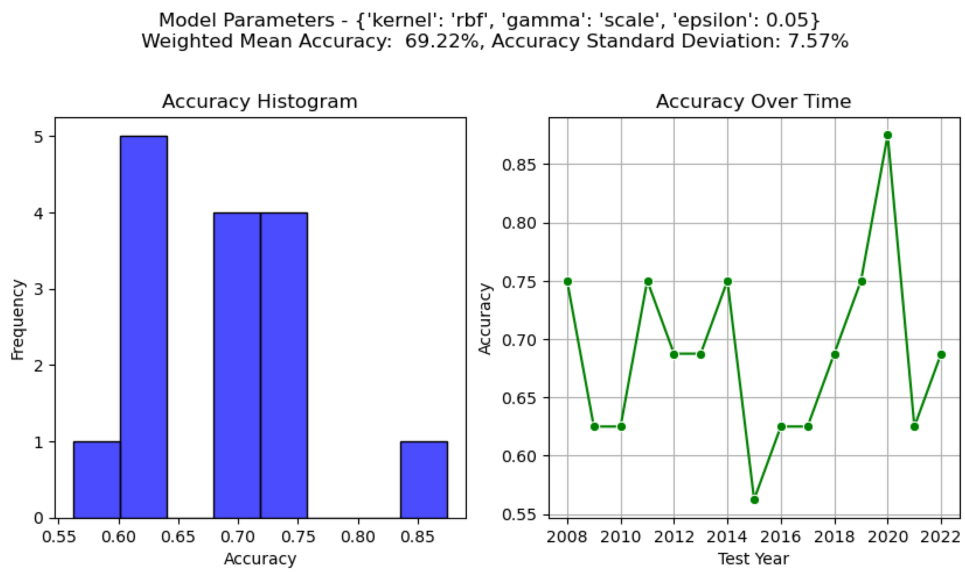


Figure 7: Support Vector Regression Performance Graphs

The non-linear kernels for the support vector regressor performed unimpressively. Like every model, it predicted 2015 poorly. It actually did better than the others in 2018, but was also a poor predictor in the years between, when most other models got back up around 75%.

LightGBM Regression

Although similar to a random forest, I wanted to try a LightGBM regressor in case the extra parameters would help improve the accuracy of the model. I tuned a variety of hyperparameters, including boosting type, number of leaves, depth, and amount of data per leaf. Its performance is shown below in Figure 8.

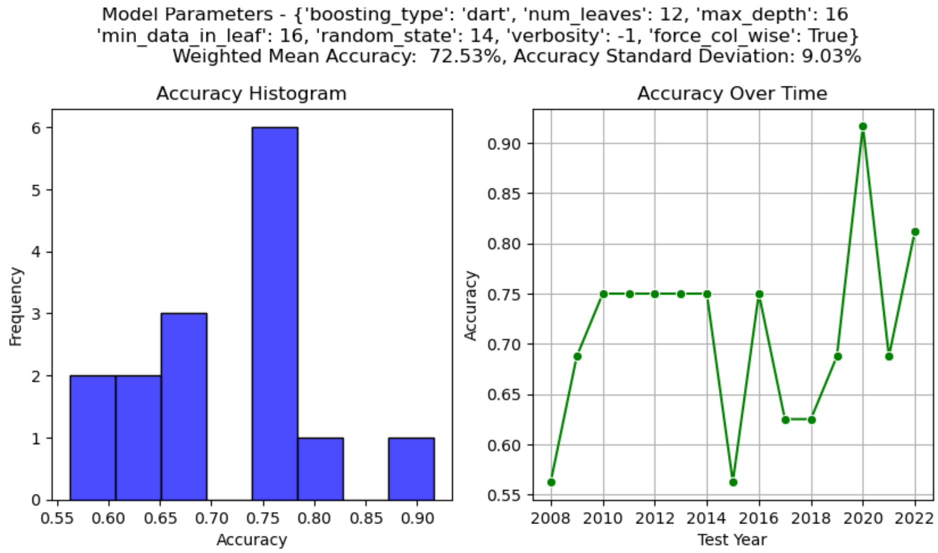


Figure 8: LightGBM Regression Performance Graphs

The LightGBM regressor improved slightly upon the random forest regressor. It held the 75% threshold for several years before the seemingly inevitable 2015 dip. Its performance in more recent years was somewhat encouraging, being over 80% accurate in 2022.

Polynomial Regression

The final individual modeling strategy I attempted was polynomial regression. This was a multi-step process that first involved fitting a polynomial transformer to the training data, then transforming the test data, and then finally using a linear regression model on the transformed data. The results of the best-performing model, a fifth-degree polynomial, are shown below in Figure 9.

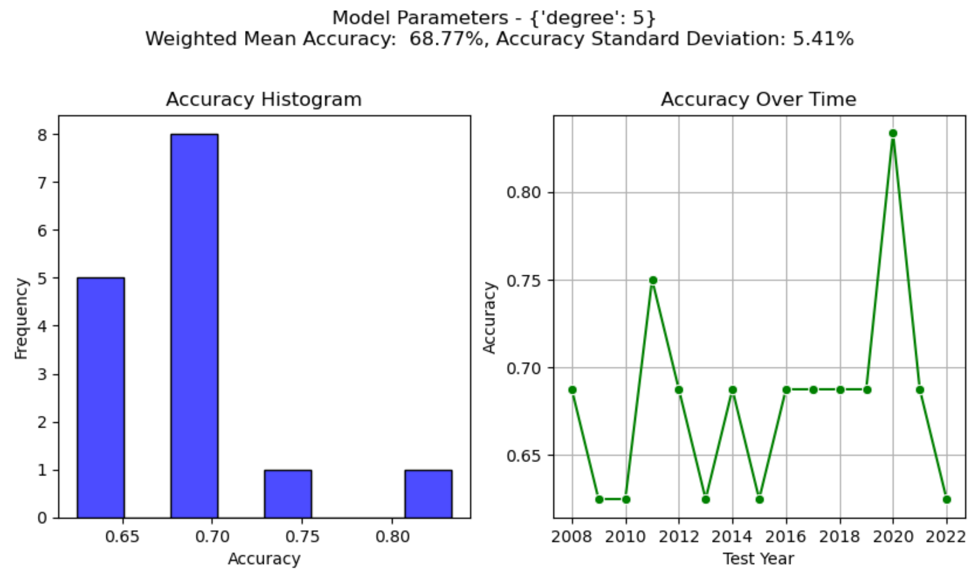


Figure 9: Polynomial Regression Performance Graphs

The model did not perform very well across the board. It only exceeded the 75% accuracy threshold once. The fourth-degree and second-degree polynomials achieved a similar performance. None of them seem to apply to the data very well.

PCA

I reduced the dimensionality of the data through PCA and trained a number of models on the transformed data. None of the models say performance improvements, so I did not utilize the technique in any of the final models. It seems fair to say that the dimensionality reduction corresponded with a loss of important features for this particular problem.

VIII. Results

I used the 2022-23 NHL season data as the holdout set for this project, because I can still evaluate the accuracy of those predictions. It was simple enough to use each model to predict the results of the season and compare. The results were also quite surprising. The linear model, which had the highest weighted mean accuracy at nearly 75%, achieved only 62.5% accuracy on the holdout data. This number corresponded to 10 out of 16 correct picks. This was also true for the elastic net and bayesian ridge models. The fifth-degree polynomial model correctly predicted 11 out of 16 playoff teams, for an accuracy of 68.75%. This was pretty much in line with its average accuracy.

The biggest surprise was that the rest of the models (support vector, random forest, and LightGBM) achieved 75% accuracy on the holdout data. Seeing this, I chose to evaluate the performance of the models as an ensemble, where I averaged the predictions of the three best model types. This ensemble captures the three of the general classes of models I tried, without duplicating votes. The true 2022-23 standings are shown alongside the predictions from the constituent models and the ensemble model below in Table 1.

Table 1: Full 2022-23 NHL Standings with True and Predicted Playoff Status

Conf	Div	Team	True	Linear	SVR	GBM	Ensemble
East	ATL	BOS	1	1	1	1	1
		TOR	1	1	1	1	1
		TBL	1	1	1	1	1
		FLA	1	1	1	1	1
		BUF	0	0	0	0	0
		OTT	0	0	0	0	0
		DET	0	0	0	0	0
		MTL	0	0	0	0	0

Conf	Div	Team	True	Linear	SVR	GBM	Ensemble
	MET	CAR	1	1	1	1	1
		NJD	1	0	0	0	0
		NYR	1	1	1	1	1
		NYI	1	0	1	1	1
		PIT	0	1	0	0	0
		WSH	0	1	0	1	1
		PHI	0	0	0	0	0
		CBJ	0	0	1	0	0
West	CEN	COL	1	1	1	1	1
		DAL	1	0	0	1	0
		MIN	1	1	1	1	1
		WPG	1	0	0	0	0
		NSH	0	1	1	1	1
		STL	0	1	1	1	1
		ARI	0	0	0	0	0
		CHI	0	0	0	0	0
	PAC	VEG	1	0	1	1	1
		EDM	1	1	1	1	1
		LAK	1	1	1	0	1
		SEA	1	0	0	0	0
		CGY	0	1	1	1	1
		VAN	0	1	0	0	0
		SJS	0	0	0	0	0
		ANA	0	0	0	0	0
Accuracy				10/16	12/16	12/16	12/16

The ensemble model was 7-for-8 in the Eastern Conference. This number is aided by the fact that every model nailed the Atlantic Division. The Metropolitan Division held some surprises during the regular season, which also manifested in the model. New Jersey exceeded expectations and made the playoffs while Washington faltered and missed.

The Western Conference had a little more chaos than the models were prepared to account for. In the Central Division, Nashville and St. Louis both missed the playoffs despite appearing to be locks according to the models. In the Pacific Division, Seattle shocked everyone with a strong campaign. Calgary did the opposite and failed to meet expectations despite some big roster moves. Overall, the ensemble model went 5-for-8 in the West.

Ultimately, the ensemble model did achieve the 75% accuracy I was aiming for, with 12 out of 16 correct picks. The next question is whether or not that is an impressive threshold. If the average NHL pundit could predict the season's results to a higher accuracy, the model would be worthless. I assembled another table of predictions below, this time featuring the true playoff teams, the ensemble model's predictions, and predictions from several NHL news outlets and experts.

Table 2: Expert predictions

Team	True	Ensemble Model	The Hockey Guy [5]	The Athletic [6]	Offside [7]	JFresh [8]	The Hockey Writers [9]	Evolving Hockey [10]
BOS	1	1	0	1	1	1	1	1
TOR	1	1	1	1	1	1	1	1
TBL	1	1	1	1	1	1	1	1
FLA	1	1	1	1	1	1	1	1
BUF	0	0	0	0	0	0	0	0
OTT	0	0	0	0	0	0	0	0
DET	0	0	0	0	0	0	0	0
MTL	0	0	0	0	0	0	0	0

CAR	1	1	1	1	1	1	1	1
NJD	1	0	0	0	0	0	0	1
NYR	1	1	1	1	1	1	1	0
NYI	1	1	0	0	0	1	0	0
PIT	0	0	1	1	1	0	1	1
WSH	0	1	1	1	1	1	1	1
PHI	0	0	0	0	0	0	0	0
CBJ	0	0	1	0	0	0	0	0
COL	1	1	1	1	1	1	1	1
DAL	1	0	1	1	1	0	0	0
MIN	1	1	1	1	0	1	1	1
WPG	1	0	0	0	0	1	0	0
NSH	0	1	1	1	1	1	1	1
STL	0	1	1	1	1	1	1	0
ARI	0	0	0	0	0	0	0	0
CHI	0	0	0	0	0	0	0	0
VEG	1	1	0	1	1	0	1	1
EDM	1	1	1	1	1	1	1	1
LAK	1	1	0	0	1	0	1	0
SEA	1	0	0	0	0	0	0	1
CGY	0	1	1	1	1	1	1	1
VAN	0	0	1	0	0	1	0	1
SJS	0	0	0	0	0	0	0	0
ANA	0	0	0	0	0	0	0	0
Accuracy		12/16	9/16	11/16	11/16	11/16	11/16	11/16

While it is impossible to grade every pundit's predictions, I assembled a varied list of trusted figures, both from the NHL news realm as well as the NHL data realm. I did not find any experts who predicted more than 11 out of 16 playoff teams correctly for the 2022-23 season. This result honestly shocked me. I knew the narratives going into the season, but the hockey world really missed the mark with some of these predictions.

The Calgary Flames were coming off an appearance in the 2021-22 Conference Finals, then added Cup winner Nazem Kadri and 115 point scorer Jonathan Huberdeau. My models were high on Calgary to make a run, as were all the experts. Instead, Huberdeau had no chemistry with his new team and the Flames sorely missed Matthew Tkachuk and Johnny Gaudreau. The team was never much of a contender before missing the playoffs.

Seattle, on the other hand, was terrible in its first season and finished near the bottom of the league. They didn't make significant personnel changes, but the team found its footing and put together an inspiring sophomore campaign despite lacking a bona fide superstar. None of my models and only one expert identified Seattle as a playoff team.

Pittsburgh and Washington were similar stories to one another. Both teams featured aging superstars with aspirations of another championship before calling it a career. Many experts had both teams poised to make another run. While Pittsburgh wasn't a unanimous pick among models and experts to make the playoffs, Washington was. Instead Washington never really looked like a playoff team and Pittsburgh narrowly missed with a pair of crushing losses to close out the regular season.

Mirroring the Penguins and Capitals in the East, Nashville and St. Louis were the pair of disappointing aging teams in the West. Both were still considered very safe playoff picks, even if just wildcards. Instead, Nashville was out of contention by the trade deadline before mounting an inspiring surge to narrowly miss the postseason. St. Louis was bogged down by defensive failings and didn't manage to stay competitive by the end of the season.

New Jersey was yet another massive surprise of the season. Led by a young core, they were expected to need a couple more seasons to mature into a contender. They took the league by storm and were a truly complete team by the playoffs. Just like the Winnipeg Jets, another surprise, only one expert picked them to even make the playoffs.

IX. Predictions

In light of these major storylines and others, it is remarkable that so many experts predicted even 11 teams correctly. It would be quite easy to see how someone would pick a sensible list of teams only to end up around 50% correct. I'm left with a stronger impression of my model, knowing that it got more teams correct than most experts. Now all that is left to do is predict the 2023-24 season and await the results. The predictions are in Table 3 below:

Table 3: 2023-24 NHL Playoff Predictions

East		West	
Atlantic	**Metropolitan**	**Central**	**Pacific**
Boston Bruins	New Jersey Devils	Colorado Avalanche	Edmonton Oilers
Tampa Bay Lightning	New York Rangers	Dallas Stars	Vegas Golden Knights
Toronto Maple Leafs	New York Islanders	Minnesota Wild	Los Angeles Kings
Buffalo Sabres	Carolina Hurricanes	Winnipeg Jets	Seattle Kraken

The storylines are numerous for the upcoming season. Many people expect the Pittsburgh Penguins to push back into the playoff picture, though my model has them narrowly missing. Another notable team is last year's Finals runner-up, the Florida Panthers, who are estimated to miss the playoffs by my model. My model doesn't expect young, upstart teams like Detroit, Arizona, Vancouver, and Ottawa to make the playoffsIt does have the Buffalo Sabres making the cut, so perhaps there are some underlying statistics pointing toward Buffalo becoming a contender. Boston is still seen as a lock despite losing first line center Patrice Bergeron. There are so many wrinkles and variables that it's impossible to say for sure which factors make a team good. I am just excited to watch the season and revisit my predictions in April to find out how I did.

X. References

[1] "NHL league revenue 2022," Statista. Accessed: Oct. 20, 2023. [Online]. Available: https://www.statista.com/statistics/193468/total-league-revenue-of-the-nhl-since-2006/

[2] "Hockey is luckiest major North American sport: Researchers," torontosun. Accessed: Oct. 20, 2023. [Online]. Available: https://torontosun.com/2016/04/14/hockey-is-luckiest-major-north-american-sport-researchers

[3] "The Strong/Weak Link Framework, and How It Can Guide Your Prioritization," Versett. Accessed: Oct. 20, 2023. [Online]. Available: https://versett.com/writing/undefined

[4] D. Davies, "The story behind the sports betting boom," *NPR*, Apr. 06, 2023. Accessed: Oct. 20, 2023. [Online]. Available: https://www.npr.org/2023/04/06/1168349259/the-story-behind-the-sports-betting-boom

[5] *Looking Back at 2021-22 and Trying to Predict Who Makes the 2023 NHL Playoffs*, (Sep. 08, 2022). Accessed: Oct. 24, 2023. [Online Video]. Available: https://www.youtube.com/watch?v=2Bj0I0jAEYw

[6] dom 📈 [@domluszczyszyn], "2022-23 NHL standings projections Updated daily at @TheAthleticNHL: https://theathletic.com/3666576/2022/10/07/nhl-2022-23-stanley-cup-playoff-chances-and-projected-standings/ https://t.co/cXKTyqUYiH," Twitter. Accessed: Oct. 24, 2023. [Online]. Available: https://twitter.com/domluszczyszyn/status/1578392207723208704

[7] "The 2022-23 NHL standings as predicted by oddsmakers | Offside." Accessed: Oct. 24, 2023. [Online]. Available: https://dailyhive.com/vancouver/2022-2023-nhl-standings-predicted-oddsmakers

[8] JFresh [@JFreshHockey], "final pre-season standings projection. so help me god. https://t.co/XzyStg4M2q," Twitter. Accessed: Oct. 24, 2023. [Online]. Available: https://twitter.com/JFreshHockey/status/1578450034886336512

[9] K. Gipe, "2022-23 NHL Season Preview," The Hockey Writers. Accessed: Oct. 24, 2023. [Online]. Available: https://thehockeywriters.com/2022-23-nhl-season-preview/

[10] "Evolving-Hockey." Accessed: Oct. 24, 2023. [Online]. Available: https://evolving-hockey.com/blog/2022-2023-nhl-team-point-projections/