Predicting and Betting on NHL Games Using Neural Networks

Pravin Fonseka [*] [1]

Abstract

Professional Hockey is one of the major sports in the United States, but has been the most difficult for statisticians and betting organizations alike for predicting the outcomes of games. Thanks to recent developments in machine learning and the arrival of open-source libraries such as TensorFlow and PyTorch, new methods can be used to possibly increase the maximum accuracy of NHL game predictions and beat Vegas betting odds as they currently stand consistently. A simple artificial neural network was used to predict the outcomes of NHL regular season games. The neural network model was trained with 22 post-priori features per team from multiple statistics websites. Out of 50 games from the 2018-2019 NHL regular season, the neural network on average predicted 29 of the games correctly, making an accuracy of 58%, with plenty of room for improvement. This prediction rate led to a slight loss on average of between 0.8 and 1.5%.

1. Introduction

Professional Ice Hockey, commonly just referred to as "Hockey" is the 4th most popular professional sport in the United States based off TV ratings with an average of 4.8 million viewers in playoffs and generating 4.86 billion dollars in revenue for the last season.(Statista, 2018) It is expected that with the National Hockey League (NHL)'s current growth, it will generate an extra 65 million dollars in additional revenue from betting operators and data providers just this year. (AGA, 2018)

The recent increase in hockey's popularity has led to more information sources on games, players, teams and statistics, ranging from goals scored per game to faceoff win percentages. Experts from NHL affiliated organizations from ESPN, TSN, CBS and gambling institutions try to predict the outcome of NHL games. Numerous attempts have been made to outperform these experts and gambling institutions in predicting those outcomes using mainly data-analytics and some Machine Learning methods.

1.1. Prediction Accuracy

Of the major sports, Hockey is the least predictable. No matter what is used, it seems that none of them have been able to beat the maximum accuracy of 62% given by sports statisticians, compared to 90%+ for the NBA. (Weissbock, 2014) There are many reasons for Hockey being so difficult to predict, compared to the other sports, especially Basketball. The two sports both have games that last sixty minutes, with hockey having three periods of twenty minutes and basketball four quarters of fifteen, but that is all they have in common.

One major reason for the prediction inaccuracy is how much of an impact star players have. A Hockey superstar such as Alexander Ovechkin will only play for twenty minutes at most as he shares time with three other lineups, while a Basketball superstar like Lebron James can play more than 48 minutes in regulation. Another is the sample size of the scoring. In basketball, it is normal for teams to score over 200 points combined, while it is rare in hockey for the teams to score over 10 combined in one game. Each mistake made that gives away a goal in hockey is far harder to makeup than a mistake that gives a basket worth two or three points in basketball. Finally, hockey has goalies. The issue with goalies is that they can be inconsistent and therefore a major variable that accounts for the unpredictability of games. Although there could be more confounding variables, these are the major reasons for the 62% maximum accuracy and what makes Hockey predictions so challenging yet interesting.

1.2. Prediction Methods

Those publicized projects that have attempted to predict game-winners using Machine Learning methods used classification algorithms, such as logistic regressions, random forest, or the most common Support Vector Machines. Still, there has been no method that has successfully and consistently beaten or at least hit the 62 percentage in prediction accuracy. More recently, thanks to google and other organi-

[*]Equal contribution [1]Courant Institute of Mathematics, New York University, New York, USA. Correspondence to: Pravin Fonseka <pf984@nyu.edu>.

Proceedings of the 35^{th} International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.

zations releasing open-source libraries for machine learning such as TensorFlow or PyTorch, ML methods are easily accessible. The method focused on here is Deep Learning, which has not yet been widely used for predicting the outcomes of Hockey games.

This is inspired by a Dutch Data Science researcher, who did test Deep Learning and compare it to classification algorithms in predicting NFL games. He showed that Neural Networks can at least equal the performance of common ML methods with even a simple Deep Network only being off the best performing methods by 1.5%.(Bosch, 2018) That is why the purpose of this project is to see how a simple Deep Network performs with the NHL in terms of game prediction accuracy and performance against real Las Vegas betting systems for its practicality to serve as a baseline for future Deep Learning work on Hockey.

1.3. Betting

Besides the accuracy of predictions, one sensible way to evaluate the method would be to compare its predictions to the predictions and odds given by gambling organizations. The simplest bet that can be done for NHL games have a binary outcome, dependent purely on the winner of the game. An example of this would be the San Jose Sharks playing at home versus the Edmonton Oilers. If the betting line for it is -150 for San Jose and +170 for Edmonton and Edmonton wins, then bettors who picked Edmonton would win 170% in addition to whatever amount they bet. If San Jose wins, then those who bet on San Jose would win 67% (100/150) in addition to the amount they bet.(Moskowitz, 2015)

Another close but common bet is based on which team will win or whether the game will go into overtime. In this case there are odds for the two teams and then a third for overtime, which overwrites the winner if overtime does occur. Often the odds for overtime will have very high rewards (+300 at least) due to the low probability of it happening. Since overtime is accounted for and overwrites the winner, then the odds of the other two decrease and have their rewards also increase, like San Jose changing to +110 and Edmonton to +200. There are more specific bets for certain organizations, but this project will only encompass determining which team will win and whether the game will go beyond regulation.

2. Data

The original hope was to use the official API of the NHL to gather necessary statistics in similar fashion to what was done with the previously mentioned NFL Deep Learning project. The NHL API turned out to be outdated with minimal statistics and even missing information for the Las Vegas Golden Knights, which were added as an expansion team for the 2017-2018 season. That meant other sources had to be used for data gathering, such as the official NHL website, ESPN, and other third party sources.

2.1. Statistics

Due to the unavailability of an official or even public third party API that satisfied the needs for data mining, the statistics were found manually for fifty separate data points representing games. Each point would have 22 features for each team so 44 in total. Besides standings data some of the important features included the following:

Acronym	Definition
GF/GA	Goals For/Against
SF/SA	Shots For/Against
PPG	Power Play Goals
PPOpp	Power Play Opportunities
TS	Time Shorthanded
PPGA	PPG Against
FOW/FOL	Face-Off Wins/Losses

For each data point, teams were randomly selected and then a game between games 30 and 80 were selected. The minimum was set to provide enough information and a large sample size to show a team's true strength and the maximum was set since teams often would not use their starters in the last games since those last ones would have no impact on playoff seeding. This was done by using a lengthy process of first using archived webpages of ESPN standings for the day of the game to find the teams' wins, loss, and overtime losses and the same for the last ten games to represent both their overall seasonal and recent strength. Specific statistics for teams from the NHL official website, which did provide them for teams in certain time periods, were used to find out each team's features at the time of their games. The only results recorded for each data point were the winner of the game and whether it went to overtime for the betting.

2.2. Bets

As previously mentioned, two separate organizations were used for finding archived betting information for each game. One was courtesy of Odds Portal and the other was Don Best which would average odds from Westgate, Mirage, Station, Pinnacle, and SIA. The latter would have the binary odds between which team wins, while the first would have a winner or overtime such as the following where the away team won in regulation:

1	X	2
+184	+308	+114

3. Model

3.1. Network Structure

The Neural Network was implemented using PyTorch. (Paszke et al., 2017) Three hidden layers were settled as the ideal number based on the information from the Deep Learning NFL paper as ideal for any kind of Neural Network that was generated. Although not extensively tested, a width of three nodes per hidden layer appeared to have the best performance while not over-fitting so that was decided for the full training and testing.

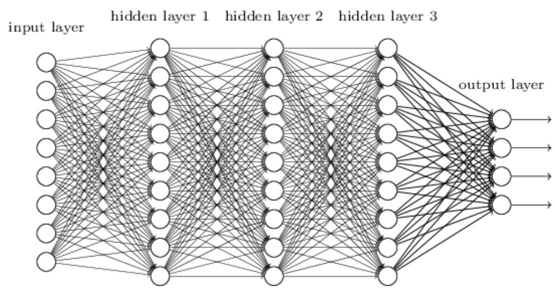

Figure 1.

Figure 1 displays the structure of a three hidden layer network. The network for this however is different in terms of the width of each layer.

With five total layers, there would be four separate weights for propagation between each pairs of layers. They would start off randomized then trained overtime through each phase of backpropagation. The activation function used for each step of backpropagation is the commonly used sigmoid function, to keep all values between 0 and 1.

3.2. Training and Testing

Due to the limited data and having only fifty points total for both training and testing, a method had to be made to maximize the amount of data used for training and minimize the variance of the predictions. Since the network is relatively simple and quick to train, it became evident that many networks could be used for trials and averaging those trials. Each data point would be separately tested, where the other 49 points were used for training. Then, that last point was the only one tested and it was tested five times, with the average being the final prediction, as seen in the predictions spreadsheet. This would be done for every point, so every point would be tested on, while still not being involved in the training for that tested model.

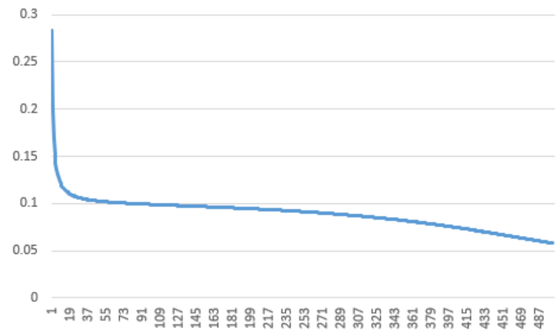

Figure 2.

Figure 2 displays the loss of an early test. 500 training steps were determined to be a fair balance between improvement of accuracy of the model and not over-fitting.

Once all the predictions are made with five trials for each point, they can be compared to the actual results of the games for both the winner and whether it goes into overtime for an overall prediction accuracy. To keep the betting portion simple, each game would be bet for the same amount of a hundred dollars for the predicted winner. Then the net profit could be calculated as a percentage to see how effective the model is for real world betting.

4. Results

4.1. Predictions

For the overall predictions that were based on the average of five trials for each game, they predicted 29 out of 50 games correct, for an accuracy of 58%. The highest standard deviation of any game predictions was 0.27 and lowest was 0.05, with an average of 0.14. While 29 of the games were wins for the home teams, the trained networks overall favored homes slightly more with an average of predicting that the, predicting that 31 or 62% of those teams would win, though the average of the predictions was actually 0.57 or 57% overall that the home team would win. Also, no game was predicted to go into overtime due to the probability always being less than 50% even when the teams were as even strength as possible.

4.2. Bets

As stated, the betting was done with a hundred dollars being theoretically spent for each predicted winner by the networks for 5,000 overall. For the Odds Portal bets which included the possibility of overtime, the overall net of the bets was a loss of 73 dollars. For the Don Best averaged bets, which were just based on the binary outcome of the winning team, the overall net was a loss of 42 dollars. This meant for respective losses of 1.5% and 0.8%.

5. Conclusion

Given the 62% theoretical maximum accuracy with regular season NHL game predictions, 58% is a good start. The same goes with the bets being so close to being positive for a net gain. Still, this is all done with such a small sample size and much more would have to be done to make a neural network that passes 60% true accuracy and consistently nets a positive profit in betting over many games.

5.1. Future Work

There are many ways this could be improved with the data, features, and model itself. As the least popular of the major four sports in the United States, Hockey is also lacking in the public statistics department, especially when trying to find team statistics for mid-season games, unless they are recorded at the moment or data mining methods are used to calculate them by getting the statistics from each individual game played beforehand. With the proper methods, entire seasons worth of games and data could be used for training and testing as opposed to fifty total for more proper investigations into whether Deep Learning can be effective.

After the time spent on the project, it became clear that there were certain features that were probably unnecessary and important ones that were not used. One that may not have been necessary is face-offs, while ones that may have been helpful or necessary would be to split features based on whether the team is home or away, since that does turn out to have a major impact on performance. For example, the Buffalo Sabres in the 2018-2019 season had a 21-15-5 home an 12-24-5 away record, while the Columbus Blue Jackets had a 22-17-2 home and 25-14-2 away record. The point is that although teams generally have a 55 to 60 percent advantage at home on average, some teams perform much better and some a bit worse at home, which could have a major impact on predictions. Another important feature to have would involve goalie statistics such as save percentage. The issue previously was that the starting goalie was not known. However, if there is enough data to account for who is starting, it could have a major impact as the goalies generally do, as the only players to be playing for all sixty minutes.

Finally, there is the model itself. Three hidden layers were used since they were the most successful in the NFL Deep Learning project, but the features and number of features used widely differ, so parameters such as the width of layers and depth of the network could certainly be tuned with full quantitative experimentation. There could also be the use of more complicated Neural Networks such as LSTMs or even vanilla RNNs. With plenty of ways to improve the work, I am certain that the 62% accuracy could be reached and possibly even beaten. Improvements to the accuracy along with algorithms used for risk-reward management for betting should also make it possible to consistently profit overall from the bets. It would be interesting to see if the project can be improved enough to actually be capable of that for the next NHL regular season.

References

AGA. Nhl could see 216 million in revenue due to legalized sports betting, 2018.

Bosch, P. *Predicting the winner of NFL-games using Machine and Deep Learning*. PhD thesis, Vrije universiteit, Amsterdam, Nederland, 2018.

Moskowitz, T. J. *Asset Pricing and Sports Betting*. PhD thesis, School of Management, Yale University, 2015.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. Automatic differentiation in pytorch. 2017.

Statista. National hockey league stanley cup games average tv viewership from 2008 to 2018, 2018.

Weissbock, J. *Forecasting Success in the National Hockey League using In-Game Statistics and Textual Data*. PhD thesis, School of Electrical Engineering and Computer Science, University of Ottawa, 2014.