Exploring Inefficiencies in NHL Betting Markets Through the Use of Machine Learning Techniques

Miami University - Department of Sport Leadership & Management

Author: Gus Weyand

Advisor: Dr. Ryan Chen

Abstract

This research paper attempts to build on previous research of examining efficiencies in sports betting markets, specifically the National Hockey League (NHL). The size of the sports betting market, including the NHL market, has increased rapidly in recent years due to several factors, most notably recent legalization of gambling and online gambling in many states. This paper will be using nonlinear machine learning models to find inefficiencies in the market, specifically when it comes to 3-way moneyline bets that offer bettors the opportunity to bet on wins, losses, and draws (includes games that go into overtime/shootout). The most successful models created through the course of this project were Random Forest and XGBoost. These models were able to achieve 62% and 65% returns on investment (ROI) respectively when betting on approximately 10% of all games over the 2022-2023, 2023-2024, and up until February of the 2024-2025 season.

Background

Sports betting has been an increasingly popular pastime for many Americans, and has seen a massive uptick in its popularity recently with its widespread legalization. Currently, sports betting is legal in 39 states and the District of Columbia (DC), an increase of 36 (including DC) since 2018 when the Supreme Court overturned the Professional and Amateur Sports Protection Act, which previously outlawed gambling (RubinBrown, 2019). With new legal status, nearly 20% of Americans have placed wagers on sporting events (Gramlich, 2022). Since 2018, the sports gambling industry has gone from a half billion dollar industry to a $14.3 billion dollar industry in 2024 (RG, 2025). The total dollar amount bet on sports contests came to about $150 billion, with millions of individual bets placed (Ramsey, 2025). About 17% of bettors engage with the NHL, a seemingly small portion of overall bettors, but impactful nonetheless (RG, 2025).

Bettors are generally at a disadvantage when it comes to making a profit, not only because of the hold (essentially the cost of the transaction) placed on bettors by sportsbooks, but because their ideology is not typically data driven (RG, 2025). Often, those placing bets are driven by potentially profitable odds and/or a hunch about a certain outcome, whether that be high odds for an underdog that is perceived to have a higher chance of winning/covering a spread, or an outcome that in the bettor's mind is likely to occur. With more data driven results, those who choose to place wagers on sports could see an increase in profitability, or minimization of losses given that they will be acting on information that is free of bias and personal interpretation. The issue of personal interpretation is most apparent when bettors place wagers on their favorite teams, something that is popular in the sports gambling industry (Na et al., 2018). This phenomena has been labeled “in-group bias”, where superstitions and biased beliefs reign supreme over more logical analysis (Na et al., 2018). In a study conducted of over 500 participants picking on average about 100 soccer games each, their overall accuracy for picking the right outcome was around 42%. However, their prediction accuracy was dramatically lower when picking for their favorite team. On average participants picked their personal favorite team to win nearly 74% of the time, the opposing team to win at about an 11% rate, and draw 15% of games. In reality, participants' favorite team won only 38.7% of the time, lost much more at 30.1%, and drew more at 31.2%. This study highlights a clear bias, as participants consistently overestimated their favorite team's chances of winning and underestimating their chances of losing, despite the actual outcomes suggesting much more balance between the two (Na et al., 2018). With a data driven approach, phenomena like “in-group bias” can be eliminated, achieving much more accurate predictions, and exposing potential inefficiency in the market that can be exploited for monetary gain.

Inefficiencies in a market from an economic perspective occur when “an asset's prices do not accurately reflect its true value” (Hayes, 2024). In the sports gambling industry, this means that if an analysis based on publicly available information can consistently generate a positive profit, it would indicate inefficiency in the market as sportsbooks are not utilizing all publicly available information to set optimal odds (Luxton, Spizman, & Moore, 2022). Inefficiencies can also arise because sportsbook odds are not solely determined by the true probability of an outcome but are also influenced by betting behavior. Sportsbooks often adjust lines to balance the amount of money wagered on each side rather than to perfectly reflect expected probabilities (Luxton, Spizman, & Moore, 2022). Sportsbooks typically count on revenue that comes from “holds”, which represents the built in cost bettors pay, typically in the form of slightly reduced odds compared to the true probabilities (Ramsey, 2025). These holds aim to capture at minimum around 2.5% of the total bet, but can climb as far as 10% with certain bets, such as parlays (Kohers et al., 2021)(Ramsey, 2025). If bets are placed randomly, we would expect to see a negative ROI around the size of the hold. If a model based on public data can exploit these discrepancies to consistently yield positive profit and beat the hold, it suggests that the market is not fully efficient, and is not utilizing all available information to create the best possible lines or odds (Luxton, Spizman, & Moore, 2022).

Literature Review

There has been significant research that pertains to prediction models for sport, both in a gambling sense and solely for the sake of trying to make the most accurate predictions, as well as those that discuss inefficiencies in different markets. Influence for this project has been taken from these types of prediction models, but are not limited to projects about the NHL. These data driven models all incorporate machine learning in various ways, but are a mix of linear and nonlinear applications, and use a variety of different amounts and quality of data to achieve their goals.

Data-driven prediction models have been created in the past that exposed inefficiencies in the market when betting on a select number of games. Charles Luxton's 2022 paper out of Loyola Marymount University, explored a similar topic to this paper, examining inefficiencies in NHL betting markets, and was able to achieve an ROI of about 8.5 percent when betting on between 8 and 9 percent of all possible bets during the 2020-2021 NHL season (Luxton, Spizman, & Moore, 2022). This research paper attempts to add to Luxton's research by using different/complex machine learning methods, as well as a more comprehensive dataset in order to further isolate inefficiencies in the market, and aims to outperform this study by creating a higher ROI on similar number of games bet over a larger timeline than was originally examined. The model created during Luxton's study was a probit model, a non-linear model used for classification and probability purposes, with likely bets being isolated by selecting only those with the highest probit index (probability of being the correct result) (Luxton, Spizman, & Moore, 2022). This paper focused mainly on very broad and basic statistics such as whether the game was a home game or away game, the team's current win percentage, plus minus, and other basic hockey statistics totalling 8 predictors (Luxton, Spizman, & Moore, 2022).

A similar study was conducted on the NFL betting market, which attempted to find inefficiencies in a range of different scenarios. This study used probit models, like were used in previous paper, to find win probabilities and isolate inefficiencies in the NFL gambling market

(Oswald, 2022). However, this paper also used some less sophisticated methods of analysis including least squares regression in examining how lines shifted from when they were originally posted to when they were closed. This paper did not find significant inefficiencies in the broader market, but were able to find small inefficiencies in a subset of games. This subset included “favorite longshot bias” where teams that are expected to win by an above average amount are typically undervalued, therefore making underdogs overvalued. With these games, there was an opportunity to make money, however this was the only subset where the inefficiencies were not considered negligible (Oswald, 2022). It is likely much harder to find inefficiencies in the NFL betting as its market is one of the largest and most analyzed, with over 35 Billion expected to be wagered on the NFL during the 2024/2025 season (Purdum, 2024).

Other similar projects aiming to predict various facets of Hockey using data-driven approaches in relation to gambling have also been completed in previous years. In 2024, research was conducted at KTH Royal Institute of Technology that covered models to predict goal outcomes in the NHL, which constitutes attempts to predict over/under bets and find the exact final score, a bet offered by many mainstream sportsbooks (Stoor & Nyberg, 2024). The output from models used in this project aimed to create more accurate lines for bookmakers, as well as exploring how the model could be used in conjunction with various betting strategies to find value. This paper covered a much more complex dataset, similar to the dataset that will be used in this project, including game level data about events on the ice. Some of the variables included in this model were past scoring performance for participating teams, common scoring combinations, overall record, and many more. This research analyzed data collected from 2021 to 2024, and was put through Bayesian Poisson Regression and Poisson Regression models, which are generalized linear models. These models were able to achieve an accuracy of greater than 50% for a portion of selected games (Stoor & Nyberg, 2024).

Another key paper that aided the model selection process for this project focused on the use of AI for predictive purposes (Remander, 2020). However, these models were not very accurate, and simply picking home or away teams would have had more prediction power than the model itself had. This is likely due to the fact that there may not have been a sufficient amount of data available to create an accurate neural network. Neural network models have the capability to be the most accurate of all machine learning models when dealing with large and complex datasets, however without a large dataset to train the model, results may be inaccurate (Remander, 2020).

Methods

This project will explore how larger inefficiencies can be drawn out of the NHL betting market, and aims to achieve greater ROI than previous explorations through the use of more complex machine learning methods. This research will also utilize more complex datasets than have previously been used to explore these ideas. The data used for this project comes from MoneyPuck.com, a reputable source of game level data from the NHL. This data contains over 100 variables ranging from how many goals were scored from different areas on the ice (broken into high, medium, and low danger), whether plays were stopped or continued in and outside the zone, the number of hits, freezes, saves, blocked/unblocked shot attempts, and many others including the flip side of these predictors for opponents. Additional columns were added to this data such as a win indicator, which was calculated by subtracting the number of goals scored for the prospecting team from the goals scored against, with a positive number (win) being logged as

1 and negative (loss) as a 0. With this a running win percentage was calculated for each team in each season, serving as a measure of previous success. Numeric variables were converted into 3 week moving averages, with the previous 3 games being used to predict results for each team's next game, in addition to their season winning percentage prior to the game being predicted. Before moving forward with analysis, a number of predictors were removed for being irrelevant to whether a game was won or lost, others were also removed for having high collinearity with other predictors, and therefore redundant to the modeling process.

Due to the size of the dataset being between 80 and 90 predictors, principal component analysis (PCA) was necessary to bring the number of variables down to a reasonable number. Prior to performing PCA, variables were scaled and tests were performed to determine whether data for this process was a good fit for PCA. After PCA analysis, 18 PCA variables were created, capturing around 80% of the total variation in the dataset, with all variables having eigenvalues greater than 1. After performing this step, the data was ready to be put through machine learning models.

The machine learning models that were selected as having the best chance of giving accurate predictions were those that had non-linear properties. In accordance with this, Random Forest, XGBoost, and NGBoost were used to predict wins and losses, specifically their classification sides as this project attempts to find inefficiencies related to moneylines. Random Forest models model data by creating lots of decision trees, which each make a prediction for every observance based on the predictors that have been passed through them, and then vote whether a 1 or a 0 should be predicted. The number of trees that vote for each class can then be formed into a propensity, or the probability of a positive class outcome. If this probability is over .5, the model will predict the game as a win for a certain team, similar to a democratic process.

XGBoost models are similar to Random Forest models in that many decision trees are created, but have the unique ability to “boost” decision trees so that the mistakes of the previous decision trees are remedied by future decision trees (Lev, 2019). NGBoost are similar to both of these, but specialize in probabilistic outcomes rather than “point predictions”, although all 3 models can show both point predictions as well as probabilities/propensities (Duan et al., 2020). As mentioned previously, this project aims to build on previous research by using these models to increase the 8.5% ROI that was seen in previous research over games during the 2020-2021 season. These models will be tested at several different propensity levels to determine their effectiveness over larger and smaller proportions of games, however, comparison will be done based on gambling on around 10% of games.

Total revenue and ROI will be calculated based on average odds for 3-way moneyline bets offered by 6 sportsbooks, including Pinnacle and bet365, Pinnacle being what many other sportsbooks base their odds off of. 3-way moneyline bets are distinct from 2-way moneyline bets, as they offer the possibility for a draw, which in the case of an NHL game would be the game being tied at the end of regulation regardless of whether a team wins in overtime or shootout. The machine learning models used in this project will only predict wins and losses, so those games that are bet on and go to overtime will be classified as lost bets regardless of whether the team wins in overtime.

Results

These three machine learning models were trained on data from every NHL game including playoffs from the 2008-2009 season through the 2021-2022 season, and tested on data from the 2022-2023 season to the current season of 2024-2025, including all games up until

February of this season. All three models demonstrated an ability to identify inefficiencies in the NHL betting market, each with slightly varying degrees of effectiveness. As shown in Figure 1, they achieved similar accuracy rates, which is the overall proportion of correct classifications, averaging around 69%. Among them, the NGBoost model appears to be the most balanced overall, with higher F1, precision, and recall scores (Figure 1). Precision in this case refers to how many games classified as a win were actually wins, where recall measures the number of wins that were called identified by the model (Encord, 2023). F1 mixes both of these metrics to find the balance between the two, typically scores above .5 are acceptable, the 3 models tested in this paper all scored in and around .66 (Encord, 2023). Having high F1, precision, and recall, means that NGBoost is proficient at predicting both wins and losses, however, all models are close in both their overall prediction power and balance.

In addition to this, the levels of ROI at different probability/propensity cutoffs elicit various degrees of capital gain while betting on different amounts of games in total (Figure 2). All wagers are in a standard amount of $100 regardless of individual probabilities, however, other betting strategies could be considered to increase the amount bet for games where teams have a higher probability of winning, but that will not be considered in this paper. Inefficiencies were found at all propensity levels tested, .5, .6, .7, and .8, with ROI increasing as probability of a positive outcome increased and the proportion of games bet on decreased (Figure 3). Bettors looking to utilize this model should evaluate the level of risk they are comfortable taking on when choosing what level or number of games to bet on, however, positive ROI was seen in every category.

These models all significantly outperformed past research on the subject of identifying inefficiencies in the NHL betting market. Previously an ROI of 8.5% was achieved betting on roughly 10 percent of all possible moneyline bets during the 2020-2021 season. The Random Forest and XGBoost models in this project were able to yield over 62% and 65% ROI respectively betting on about 11% of all matches, staggering 54% and 57% improvements in ROI. If bettors desire to gamble more often, they may opt to go for other models, particularly the NGBoost model that was able to achieve 36% ROI on about 43% of games. The improvement over previous research could be attributed to several factors including:

1. Models used in the process have a greater ability to model non-linear relationships in data.

2. The level of data used being significantly more complex, where previous research only used 8 predictors being very simple in nature, this project used 18 PCA's that were made up of about 80 predictors that were formed into a 3 week moving average able to capture a teams recent performance, which is indicative of how they will perform in their next contest.

3. Previous research was trained on 4 NHL seasons and tested on 1 season, which was the covid season with severe fan limitation, and saw many teams down players at different times due to illness. The models developed in this project were trained on 14 seasons and tested on 2.5, which did not include the covid season that brought on uncertainty in who would play, potential losses of home-ice advantage, as well as many other factors that are difficult to measure.

4. This project measured ROI by placing bets on 3-way moneylines rather than 2-way moneylines. 3-way money lines offer better odds for wins and losses due to the possibility of games going to overtime and resulting in a tie. There may be more losses in the models developed in this project because of ties, however, the better odds likely cancel this out, especially at higher propensity levels, which indicates teams are more likely to win by a hefty margin rather than it being a close game.

Applications/Barriers to Implementation

Those choosing to gamble on NHL games would be best served implementing a data-driven decision making approach, like the ones described in this paper. Using these data-driven approaches eliminate the possibility of bringing personal bias into decision making, as well as consider additional factors that may be less visible to bettors due to clouded judgement, such as recent performance or shot quality among many others. However, practically speaking, these models are reliant on data, and if data is not available, then models will not work as intended. In order for these models to be able to run and make a prediction for the next game on a teams schedule, they must have that team's previous 3 games to average together. This may be difficult as data used in this project comes from Moneypuck.com, which updates regularly (typically several times a week), but may not have all up to date information, especially if teams play back-to-back nights. Due to this, there may be games where it is not possible to place a bet with these models given that the information is not available. However, if individuals are dedicated enough, this data can be drawn from other sources, such as ESPN, with more advanced metrics calculated by using the Moneypuck.com data dictionary and basic statistics gleaned from ESPN or other platforms. Additionally, due to the use of 3 week moving averages in predictions, bets on games during the first 3 weeks of the season are not possible, because the moving average cannot populate until there are 3 games to average. Even though not conducive to making the best predictions, previous research utilizing simpler datasets do not run into the issue of not having sufficient data available to make predictions.

If these models were to be used long term, users would have to monitor summary statistics (F1, accuracy, precision, and recall) to ensure that the model is still making accurate predictions for both classes. Over time, changes in the game such as strategy or rule changes could make an impact on the models ability to predict accurately. If this is the case, models could be retrained on more recent data so it is more accurately able to model relationships that affect wins and losses. However, this is not a major undertaking, as retraining these 3 models would only take 6-7 hours, assuming data is available, and would then be ready to make new predictions for gambling purposes. The problems tied to up to date data and the use of 3 week moving averages represent the only practical barriers in user implementation, as models trained are ready to make predictions in fractions of seconds when run on new data.

Conclusion

Overall, a combination of factors led to the success of this project exposing additional inefficiencies in the NHL betting market than were discovered previously. These factors included using more complex models and datasets, including much larger training and testing datasets, and the use of a different type of moneyline bet. The best models betting on a similar proportion of data, around 10-12 percent, were able to elicit ROI’s of 65% and 62%. Given that these ROI’s are much higher than the 2.5-10 percent negative return that we expect to see from random choice, there is reason to conclude that there is significant inefficiency in the NHL gambling market. Additionally, 54-57 percent ROI improvement on previous research indicates that the use of non-linear machine learning models are proficient at finding further inefficiency in the market.

Appendix

Duan, T., Avati, A., Ding, D. Y., Thai, K. K., Basu, S., Ng, A., & Schuler, A. (2020, June 9). *NGBoost: Natural Gradient Boosting for Probabilistic Prediction*. ARXIV. https://arxiv.org/pdf/1910.03225

Encord. (2023, July 18). *F1 score in Machine Learning*. Encord Blog. https://encord.com/blog/f1-score-in-machine-learning/#:~:text=Typically%2C%20an%20F1%20score%20%3E%200.9,to%20have%20a%20poor%20performance

Gramlich, J. (2022, September 14). *As more states legalize the practice, 19% of U.S. adults say they have bet money on sports in the past year*. Pew Research Center. https://www.pewresearch.org/short-reads/2022/09/14/as-more-states-legalize-the-practice-19-of-u-s-adults-say-they-have-bet-money-on-sports-in-the-past-year/

Hayes, A. (2024, August 15). *What is an inefficient market? definition, effects, and example*. Investopedia. https://www.investopedia.com/terms/i/inefficientmarket.asp#:~:text=According%20to%20economic%20theory%2C%20an,often%20lead%20to%20deadweight%20losses

Kohers, G., Tuttle, M., & Bumpass, D. (2021, December 7). *Examining NCAA/NFL market efficiency*. Allied Academics. https://www.alliedacademies.org/articles/examining-ncaanfl-market-efficiency.pdf

Lev, A. (2019, December 19). *XGBoost versus Random Forest*. JFrog ML. https://www.qwak.com/post/xgboost-versus-random-forest

Luxton, C., Spizman, J., & Moore, D. (2022). *Testing Efficiency in NHL Betting Markets*. Digital Commons - Loyola Marymount University.

Na, S., Su, Y., & Kunkel, T. (2018, September 23). *Do not bet on your favourite football team: the influence of fan identity-based biases and sport context knowledge on game prediction accuracy*. *European Sport Management Quarterly*.

Oswald, R. E. (2022). *Testing the efficiency of the NFL betting market*. University of Northern Iowa. https://scholarworks.uni.edu/cgi/viewcontent.cgi?article=1444&context=hpt/

Purdum, D. (2024, September 3). *Estimated $35 billion expected to be bet on NFL this season*. ESPN. https://www.espn.com/sports-betting/story/_/id/41111932/estimated-35-billion-expected-be t-nfl-season

Ramsey, E. (2025, January 15). *Analysis: How high can US sports betting hold really climb?* Legal Sports Report. https://www.legalsportsreport.com/111012/analysis-2023-us-sports-betting-hold-trend/

Ramsey, E. (2025, January 15). *Turning the page on a record 2024 for US Sports Betting*. LSR. https://www.legalsportsreport.com/218449/2024-us-sports-betting-revenue

Remander, S. (2020). *AI driven NHL predictions - Comparison of classification methods in predicting NHL game event outcomes*. Aalto University. https://aaltodoc.aalto.fi/server/api/core/bitstreams/86ed8e36-759f-42e6-a4ee-a904275cf039/content

RG (Ed.). (2025). *U.S. Sports Betting Revenue 2025 - state-by-state market analysis*. RG. https://rg.org/statistics/us

RubinBrown. (2019). *Focus on gaming: One year since the overruled Professional and Amateur Sports Protection Act*. https://www.rubinbrown.com/insights-events/insight-articles/focus-on-gaming-one-year-sin ce-the-overruled-professional-and-amateur-sports-protection-act/

Stoor, W., & Nyberg, J. (2024). *Beating the odds: A statistical approach to hockey goals*. Digitala Vetenskapliga Arkivet. https://www.diva-portal.org/smash/get/diva2:1894676/FULLTEXT01.pdf

---

Figures

Figure 1

Model	Summary Statistic	Result (%)
XGBoost	Accuracy	69.49%
XGBoost	Precision	68.43%
XGBoost	Recall	65.03%
XGBoost	F1	66.69%
RF	Accuracy	69.02%
RF	Precision	68.95%
RF	Recall	61.91%
RF	F1	65.24%
NGB	Accuracy	69.84%
NGB	Precision	69.25%
NGB	Recall	64.36%
NGB	F1	66.72%

Figure 2

Model	Propensity Cutoff	Games Wagered (%)	Total Wagered	Total Profit	ROI (%)
XGBoost	.5	44.67	$322,100	$113,269.64	35.17
XGBoost	.6	33.09	$238,600	$107,119.94	44.9
XGBoost	.7	21.58	$155,600	$83,558.11	53.7
XGBoost	.8	11.30	$81,500	$53,572.15	65.73
RF	.5	42.17	$304,100	$109,154.98	35.89
RF	.6	23.87	$172,100	$87,289.98	50.72
RF	.7	10.90	$78,600	$49,143.14	62.52
RF	.8	2.90	$20,900	$14,780.60	70.72
NGB	.5	43.67	$314,900	$113,562.05	36.06
NGB	.6	30.02	$216,500	$103,073.02	47.61
NGB	.7	17.72	$127,800	$71,818.22	56.2
NGB	.8	6.85	$49,400	$35,868.96	72.61

Figure 3