BISHOP OR ROOK'D: A CREDIBILITY ESTIMATION FOR NHL GOALTENDER CAREER PERFORMANCE

PROJECT PROPOSAL

Jeff Faath
School of Computer Science
Georgia Institute of Technology
jfaath3@gatech.edu

Saiem Gilani
School of Industrial & Systems Engineering
Georgia Institute of Technology
sgilani9@gatech.edu

Brian Haley
School of Computer Science
Georgia Institute of Technology
bhaley9@gatech.edu

Michael Parkatti
School of Industrial & Systems Engineering
Georgia Institute of Technology
mparkatti3@gatech.edu

Spring Smith
School of Industrial & Systems Engineering
Georgia Institute of Technology
ssmith494@gatech.edu

Problem Description

We seek to determine an appropriate sample size of data needed for a Buhlmann credibility estimation on complement of NHL Goaltender Save percentage (or Goals Allowed percentage). The goal is to determine at what point in a Goaltender's career can you establish a 95% confidence interval that the player's true mean save percentage will not include the league average (or perhaps at a top 10 level). The Goaltender metrics will be incorporated into the GA portion of the Pythagorean win expectation,

$$\frac{(\text{Goals Scored (or GS)})^2}{(\text{Goals Scored (or GS)})^2 + (\text{Goals Allowed (or GA)})^2}$$

to determine Goaltender Win Shares (or Goaltender Wins Above Replacement) which will be among our baselines used.

Literature Survey

Hockey Goaltender Analysis
Andrew Thomas mentions goalies can be measured by their save rate using a Poisson distribution. This accounts for the low number of goals, but we alleviate that by looking at shots rather than just goals [15]. Schuckers goes into detail about the old ways of measuring goalies (save proportion) to the newer ones (linear models using shot type, location, and angle from previous shot). None of these models incorporate machine learning or quality of the shooter [11]. Roith and Magel use stepwise analysis to determine important factors for winning a single NHL game and making it to the playoffs. The analysis shows stopping shots is the most important factor. We plan to further analyze the importance of goalies and determine when a team should consider replacement [9]. Schuckers measures shot difficulty and uses average distribution of shots and each goalie's spatially smoothed shot performance map as a basis for goaltender comparison (as in Figure 1 below) [10].

General Hockey Analysis
Macdonald uses ridge regression to model how valuable an NHL player is to their team based on their expected contribution to goals/hour. Macdonald focuses only on 5v5 playing stances whereas we plan to consider all playing stances [6]. Schulte, Liu, and Li analyze junior league players with seasonal data to determine success in the NHL for drafting decisions. They use logistic regression to determine whether a player would play in an NHL game, but do not account for how well the player would perform [14]. Shuckers and Curro delve into topics such as home-ice effect and possession changes and what effect they have on probabilities of shots becoming goals. They used these factors to create a rating for forwards and defenseman, whereas we intend to use similar factors for rating goaltenders [12]. Using SPORTLOGiQ's spatio-temporal dataset, team-level pace metrics were developed in even-strength situations across offensive, neutral, and defensive zones by tracking the path of the puck and possession events over a spatial polygon grid and applying Gaussian kernel smoothing baselined to league

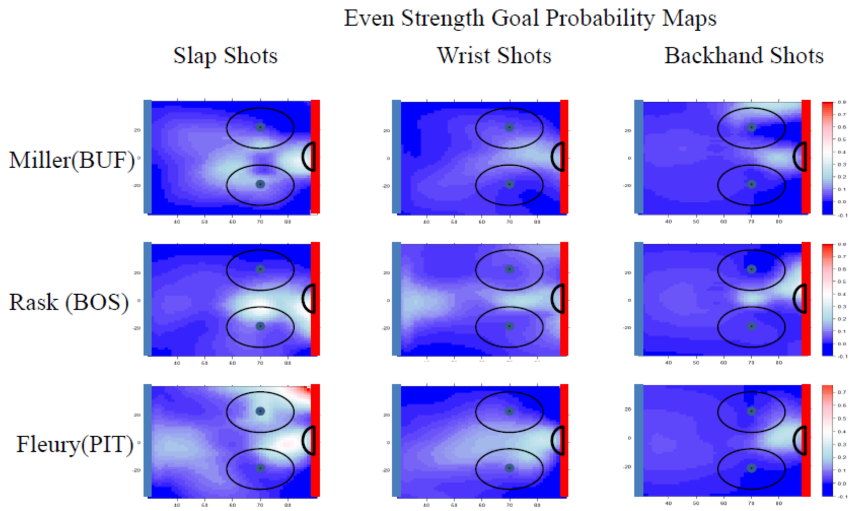

Figure 1: Goal Probability Maps at Even Strength of Different Shot Types for Selected Goalies
Scale is Blue (low) to Red (high)

average. This SPORTLOGiQ spatio-temporal dataset was also used to create player clusterings using an affinity propagation algorithm and activity location data. A Markov model was built to measure their relative scoring impact [16].

General Sports Analysis
Jamieson analyzed home field advantage in sports across multiple sports under various conditions, showing that home teams have a distinct advantage. For NHL, it was shown that home teams win about 59.5% of the time. This can be considered as a weight in our models [5]. Luke, Daniel, Alexander, and Andrew explain that the location of non-shooters on the basketball court are a significant factor to scoring. We could use non-shooters positions in our model and expand on the position data to categorize offensive strategies and goalie's effectiveness [2].

General Statistical Analysis
Our model deals with shots on goal where the success rate is generally less than 10%. This leads to a highly imbalanced data set. Haibo et al provided a review of running learning algorithms with imbalanced data, whose insights can be applied to our development [4]. In two papers, Niculescu-Mizil and Caruana investigate methods to create probability estimations for the positive class label in supervised learning classification models, in a general survey across multiple algorithms and specifically for boosted trees. These will help in correcting biases in our expected goal probabilities, regardless of the selected model type [7][8]. Franks et al proposed a "meta-metric" framework to aid evaluating the stability, discrimination and independence of a sport metric. This framework can be applied to our method to ensure the robustness of the model [3].

Heilmeier Questions

What are you trying to do? Articulate your objectives using absolutely no jargon.
Our team will:

1. Build tools that will help teams and fans to differentiate good goaltenders in the National Hockey League (NHL) from bad goaltenders
2. Create and visualize statistical tests that will allow stakeholders to identify at what point in a goaltender's career they could assert the above point

How is it done today; what are the limits of current practice?

Currently, a goaltender's save percentage is considered the authoritative measure of a goaltender's ability to stop pucks in hockey (saves / shots). While useful, this measure fails to incorporate the context in which each shot is taken, considering them all to be equally challenging. This statistic alone also does not give context to a goalie's peer group.

What's new in your approach? Why will it be successful?

Our approach will utilize high-dimensional machine learning to estimate probabilities for certain shots to become goals (factors such as distance, shot type, angle to the net, etc). This will provide context to rating goaltenders – if a goaltender allows less goals than our model predicted, it's a good signal of his ability.

Who cares?

The primary parties interested in this research would be management of NHL teams – we will be providing them the ability to make relatively early assertions on both their goaltenders' ability (allowing them to commit to longer contracts for good goaltenders or to cut poor goaltenders) and the ability of undervalued goaltenders around the league to target for acquisition.

If you're successful, what difference and impact will it make, and how do you measure them (e.g., via user studies, experiments, ground truth data, etc.)?

The primary benefit of the analysis would be to avoid playing poor goaltenders longer than necessary to properly evaluate their talent level. Playing poor goaltenders invites opposing teams to score more goals, leading your team to lose more games, decreasing your team's chances of making the playoffs. Each playoff home game can mean $5,000,000+ in revenue for the team. The models will be validated through analysis of false positives and false negatives through historical careers.

What are the risks and payoffs?

Risks:

- Our models do not successfully estimate probabilities for shots becoming saves. They may only find noise in the shot data, doing no better than traditional save percentage.
- There are not many goalie careers to test our data on (less than 200 over the last decade). We may not be able to meaningfully decide when you can know a goaltender's ability.
- There are unknown biases in the data, such as transcription errors idiosyncratic to a particular team or arena.

Payoffs:

- Not playing a poor goaltender is likely worth millions in team revenue per season.

How much will it cost?

We will be using cloud services such as AWS to facilitate the machine learning models, likely costing less than $200. Otherwise, the implied costs are time commitments of our five team members.

How long will it take?

Our goal completion date is Nov 24, in advance of the final submission date. This leaves approximately 6 weeks of working time. The five team members expect to dedicate 10 hours per week to the project, contributing to about 300 hours in total.

What are the midterm and final "exams" to check for success? How will progress be measured.

The 'midterm' is when meaningful probabilities for shots to become goals are found through cross-validated machine learning models. Once that predictive power is established, the statistical tests and final visualizations should be relatively straightforward. The 'final' is when our results are successfully in Tableau for public consumption.

Plan of Activities

Phase	Task	Who?	Effort (hrs)	Duration (days)	Goal Start Date	Goal End Date
Planning / Admin	Wireframe development	Michael	2			
Project Proposal	Doc - Literature Survey	All / Spring edit	12	3	10/8	10/11
	Doc - Plan of Activities	Brian / Spring	2	3	10/8	10/11
	Doc - Technical Problem Definition	Saiem	2	3	10/8	10/11
	Doc - 9 questions	Michael	2	3	10/8	10/11
	Slides	Jeff / Saiem	2	3	10/8	10/11
	Video Creation/Narration	Michael/Brian	2	3	10/8	10/11
Data Collection	Download shot dataset from Moneypuck	Saiem	3	9	10/1	10/10
	Get API keys for data download	All	3	9	10/1	10/10
	Download play-by-play JSON files from the SportRadar API for all games in 2015-2018 seasons	Saiem	3	9	10/1	10/10
	Download individual player statistics from Hockey Reference (`https://www.hockey-reference.com`)	Saiem	3	3	10/11	10/14
Data Integration	Join datasets together	Saiem	2	2	10/12	10/14
	QA/Verify cleanliness of data	Brian	2	2	10/12	10/14
Feature Engineering	Performing feature selection and transformation as necessary	Jeff	6	6	10/14	10/20
	Incorporating other advanced metrics from referenced papers	Saiem	4	6	10/14	10/20
ML Modeling	Preprocessing of data	Michael	4	15	10/20	11/4
	Model Selection	Jeff	4	15	10/20	11/4
	Hyperparameter tuning	Spring	3	15	10/20	11/4
	Calibration of Probability Outputs	Michael/All	3	15	10/20	11/4
Progress Report	Construct and submit progress report	All	6	7	11/1	11/8
Visualization Development	Ice Rink Heamap POC	Brian / Saiem	1	2	10/10	10/11
	Create prototype in Tableau	Brian	5			
	Finalize Tableau Presentation Layer	Spring	5			
	Create poster outline and framework	Michael	5			
Final Report and Poster	Compile all results, figures, comparisons for report, poster and presentations	All	20	14	11/8	11/21
Submit Report		Michael	0	0	11/21	11/21

Distribution of Work

All team members contributed equally and are earnestly participating with the project.

References

1. Albert, J. (2017). Situational Statistics, Clutch Hitting, and Streakiness. Handbook of Statistical Methods and Analysis in Sports (pp. 89-112). Boca Raton, FL: CRC Press
2. Bornn, L; Cervone, D; Alexander, F; Miller, A. (2017). Studying Basketball through the Lens of Player Tracking Data. Handbook of Statistical Methods and Analysis in Sports (pp. 245-269). Boca Raton, FL: CRC Press
3. Franks, A. M., D'Amour, A., Cervone, D., Bornn, L. (2016). Meta-Analytics: Tools for Understanding the Statistical Properties of Sports Metrics. arXiv:1609.09830v1
4. Haibo H., Garcia E. A. (2009). Learning from Imbalanced Data. IEEE Transactions On Knowledge And Data Engineering, vol. 21, no. 9, pp. 1263-1283
5. Jamieson, J. P. (2010). The Home Field Advantage in Athletics: A Meta-Analysis. Journal of Applied Social Psychology, 40, 7, pp. 1819–1848.
6. Macdonald, Brian. (2012). An Expected Goals Model for Evaluating NHL Teams and Players. MIT Sloan Sports Analytics Conference, 2012. Retrieved from `http://www.sloansportsconference.com/wp-content/uploads/2012/02/NHL-Expected-Goals-Brian-Macdonald.pdf`
7. Niculescu-Mizil, A; Caruana, R. (2005). Obtaining calibrated probabilities from boosting. UAI'05 Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence (pp. 413 - 420). Arlington, VA: AUAI Press
8. Niculescu-Mizil, A; Caruana, R. (2005). Predicting Good Probabilities With Supervised Learning. ICML '05 Proceedings of the 22nd international conference on Machine learning (pp. 625 - 632). Bonn, Germany
9. Roith J, Magel R. (2014). An analysis of factors contributing to wins in the National Hockey League. International Journal of Sports Science, 4(3), 84–90. doi:10.5923/j.sports.20140403.02
10. Schuckers, M. (2011). DIGR: A Defense Independent Rating of NHL Goaltenders using Spatially Smoothed Save Percentage Maps. MIT Sloan Sports Analytics Conference. Retrieved from: `http://www.sloansportsconference.com/wp-content/uploads/2011/08/DIGR-A-Defense-Independent-Rating-of-NHL-Goaltenders-using-Spatially-Smoothed-Save-Percentage-Maps.pdf`
11. Schuckers, M. (2017). Statistical Evaluation of Ice Hockey Goaltending. Handbook of Statistical Methods and Analysis in Sports (pp. 307-325). Boca Raton, FL: CRC Press
12. Schuckers, M; Curro, J. (2013). Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. MIT Sloan Sports Analytics Conference, 2013. Retrieved from `http://www.sloansportsconference.com/wp-content/uploads/2013/Total%20Hockey%20Rating%20(THoR)%20A%20comprehensive%20statistical%20rating%20of%20National%20Hockey%20League%20forwards%20and%20defensemen%20based%20upon%20all%20on-ice%20events.pdf`
13. Schulte, O., Zhao, Z., & Javan, M. (2017). Apples-to-Apples: Clustering and Ranking NHL Players Using Location Information and Scoring Impact. MIT Sloan Sports Analytics Conference. Retrieved from `http://www.sloansportsconference.com/wp-content/uploads/2017/02/1625.pdf`
14. Schulte, O; Liu, Y; Li, C. (2018). Model Trees for Identifying Exceptional Players in the NHL Draft. arXiv:1802.08765v1
15. Thomas, A. (2017). Poisson/Exponential Models for Scoring in Ice hockey. Handbook of Statistical Methods and Analysis in Sports (pp. 271-285). Boca Raton, FL: CRC Press
16. Yu, D., Boucher, C., Bornn, L., & Javan, M. (2019). Playing Fast Not Loose: Evaluating team-level pace of play in ice hockey using spatio-temporal possession data. MIT Sloan Sports Analytics Conference. Retrieved from `http://www.sloansportsconference.com/wp-content/uploads/02/HockeyPace.pdf`