Hockey Rebound Prediction

Andrew Simonds, Mike McColm, Kyle Air, Jake Zeldin

QMIND – Queen's AI Hub
Queen's University, Kingston, Ontario K7L 3N6, Canada

Andrew Simonds: andrew.simonds@queensu.ca
Mike McColm: mike.mccolm28@gmail.com
Kyle Air: 16kaa11@queensu.ca
Jake Zeldin: 17jjz@queensu.ca

Abstract: *With goalies being so skilled in the NHL, one of the best ways to score a goal comes from a second chance on a rebound into a high danger area of the ice. Our team has developed a tool to accurately predict where a rebound will be directed after a shot. The tool will be used to train players to shoot in certain locations to produce the highest scoring chance following the shot. The team trained a deep neural network, with help from TensorFlow libraries, using sample data consisting of shot location on ice, and shot location on net to then predict the rebound angle in 30-degree bins surrounding the net. The final product consists of a GUI which allows a user to choose where the puck is being shot from and the location on net and output a spray chart of probabilities of where the resulting rebound will be.*

---

1. INTRODUCTION

1.1 Motivation

While certain sports have been infatuated with data analytics, the hockey analytics awakening is still in its infancy [1]. Companies are beginning to emerge with the ability to extract vast amounts of live data from a hockey game, such as player and puck location on the ice. Although National Hockey League (NHL) franchises have yet to adopt data analytics as a major form of scouting and coaching, it is evident that these practices will soon be implemented.

In 2016, NHL coach John Chayka became the youngest general manager in the NHL's history, and he also happens to be the founder of Stathletes, a hockey analytics company [2]. He is one of the main driving forces in proving the validity of data analytics in professional hockey.

One of the specific problems in the game today is the lack of knowledge regarding rebounds, and tendencies associated with certain goalies. This is what our team within QMIND aims to explore further.

1.2 Related Works

Similar projects have been completed within the realm of hockey analytics, such as the investigation into how the pace of play in a game affects how well a team will perform. The research paper, "Playing Fast Not Loose: Evaluating team-level pace of play in ice hockey using spatio-temporal possession data", dives into this problem by assessing various pace metrics and analyses of players and teams. Their findings suggest that team-level pace is beneficial, but only to a certain point [3]. Some of the main issues within their project involved the randomness of the sport, which will prove to be a similar challenge for our project. The group working on this project plans to expand their findings into similarly structured sports such as soccer, basketball, and rugby.

1.3 Problem Definition

In a similar manner to the team mentioned in Section 1.2, our team plans to tackle a difficult to grasp aspect of hockey, in the form of rebound prediction. Through our general passion for the sport and research before starting the project, we came to a consensus that a fantastic problem to focus on would be creating a way to take all of the events leading to a rebound event, and format insights from these events in a useful fashion to be used by players and coaches. Our main problem we decided on was to take in a shot location on the ice, and shot position on the goalie, and output accurate predictions for where the rebound angle would be. We believe that with this information, players will be able to exploit certain goalie tendencies, or goalies will be able to acknowledge and work to fix some of their weaknesses.

2. METHODOLOGY

2.1 Preprocessing

As one of the main issues seen in similar hockey analytics projects was the randomness of the game, we tried to reduce this element by focusing on clean shots on goal to analyze rebounds for. This took out shots on goal involving deflections or other unique shots. The randomness still remains in the project, but due is further touched upon in the discussion of our project.

Through the partnership with Iceberg Sports Analytics, we were provided with a very large dataset of shot attempts through a series of American Hockey League (AHL) games, specifically focused on the Bridgeport Sound Tigers, the AHL affiliate of the New York Islanders. Each of these data points included a number of features, including shooter and goaltender coordinates, shot type, player information, and a number of other characteristics. Our first step was to filter the data down to a more manageable size. With such a large dataset, we had the luxury of being able to remove datapoints with missing or incomplete information. Categories unrelated to the target (the output rebound angle) were discarded to simplify the process. We also used binning on continuous data (i.e. splitting the offensive zone into twelve zones, the net into six zones, and the rebound bins into nine zones of 30° angles), as seen in Figure 1.

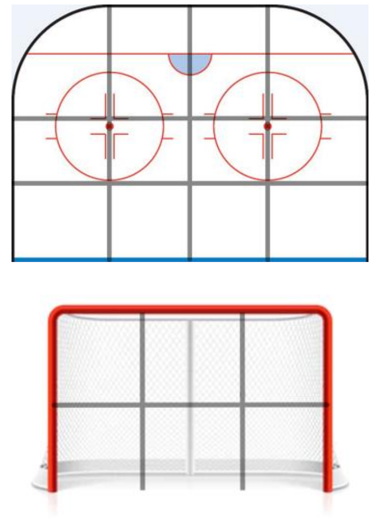

Figure 1: A view of the team's Graphical User Interface displaying the binning performed on both the ice surface and the net.

The final result of the preprocessing stage was a dataset where every datapoint had complete information in the categories that we as a team had deemed most predictive: the shooter's coordinates on the ice, and the shot target location on the goaltender (i.e. high glove, five-hole, etc.), as well as the target feature of output rebound angle.

2.2 Model Creation and Optimization

Once the dataset was refined, we began creating models. After familiarizing ourselves with the different prediction techniques available and working through a trial-and-error process, we decided to proceed with a Deep Neural Network predictor using technology from Python's TensorFlow library. The first iteration of the design returned accuracy rates of between 15-20%. To increase this accuracy, we changed some of the model parameters, including layer sizes and the learning rate. Additionally, we doubled the size of the rebound output bins from their original values of 15% to their final values of 30%. As a result of these changes, we observed drastic increases in their accuracy, to the point where the final model was able to correctly predict the approximate rebound between 45-60% of the time. Once the model was finalized, we also created a Graphical User Interface (GUI) to display the results of the model and allow user interaction.

2.3 Model Evaluation

When evaluating the model, we primarily used logical reasoning and visual inspection to determine how successful the model was. When considering the situation and the unpredictability of hockey rebounds, we were happy with these accuracy rates of around 50%. If the model was guessing at random, the accuracy rates would be around 9% due to the 11 possible bins. As all members of the team were familiar with hockey and typical rebound placement, we were able to confirm that the model's predictions made sense through analysis of output images taken from the team's GUI (shown below in Figure 2).

Figure 2: Display of rebound angle likelihood from the team's GUI.

3. RESULTS AND DISCUSSION

As mentioned earlier, our only real result to validate was the accuracy range, which was between 45-60%. Because of the entropy of the situation, the wide range was expected, and can be attributed to the randomness of the partitioning between training and test data. When compared to the initial projections that we made at the start of this project, we believe our final results outperformed these initial benchmarks. From this, it's evident that our methods did work as expected as all decisions made throughout the process led to these final results. Over the year, we learned a lot, both in our work with new artificial intelligence technologies and techniques as well as our new exposure to the world of sports analytics. We believe that our results and accuracy have reached a level that could make them valuable to hockey teams looking to analyze and optimize their performance and success.

4. CONCLUSIONS AND FUTURE WORK

In conclusion, we were happy with the results and progress made over the year. Our stages of initial research/resource acquisition, data preprocessing, and model creation/refining all helped us build unique skills along the path to creating a working product. If this project were to be continued in the future, several steps could be taken. The first of these would be to make further improvements to the model. This could include acquiring more data from different sources, or perhaps including different predictors to create one ensemble prediction network, which typically can achieve higher results than any other single predictor. Another step would be to present this data to actual teams (i.e. youth hockey) and ask them to try building strategies around it and report the results to confirm that the theoretical results are applicable to a real-world situation. The final future step for our project would be to acquire a mass amount save data from a specific goalie to train the model on to accurately analyze tendencies that the goalie has. With the increased focus on data analytics, including the NHL's plans to accurately track the players and puck at all times [4], we as a team believe that this field of hockey analytics has a very promising future.

REFERENCES

[1] Wheeldon J. *The Future of Hockey Analytics.* TheHockeyWriters.com. September 25, 2017. Accessed on: March 24, 2020. [Online]. Available: https://thehockeywriters.com/future-of-hockey-analytics/

[2] Basen N. *How Analytics Will Help Win the Next Stanley Cup*. TheWalrus.ca. November 13, 2019. Accessed on: March 25, 2020. [Online]. Available: https://thewalrus.ca/how-analytics-will-help-win-the-next-stanley-cup/

[3] Yu D., Boucher C., Bornn L., Javan M. *Playing Fast Not Loose: Evaluating team-level pace of play in ice hockey using spatio-temporal possession data.* SportLogiq.com. March 2, 2019. Accessed on: March 24, 2020. [Online]. Available: https://sportlogiq.com/en/news-views/sportlogiq-s-research-papers-at-sloan-sports-analytics-conference

[4] Cotsonika, N. *Puck, Player Tracking in final testing stage before Stanley Cup Playoffs.* NHL.com. March 3, 2020. Accessed on: March 24, 2020. [Online]. Available: https://www.nhl.com/news/puck-player-tracking-technology-unveiled-during-2020-postseason/c-315806398