Effect of Zone Entries and Rush Chances on Goal Scoring

Zack Kehl

Intro

In the ever-changing landscape of hockey strategy, coaches are constantly searching for new ideas to gain a competitive advantage. The ability to effectively navigate the transition from defense to offense is vital and zone entries are the starting point for offensive attacks. Zone entries are important in order to create meaningful puck possession in the offensive zone and one way to do this is with rush opportunities. In this paper, I will be exploring the relationship between zone entries and rush chances on scoring chances and goal scoring. This information can be important for coaches looking to maximize goal scoring opportunities.

Approach (model building)

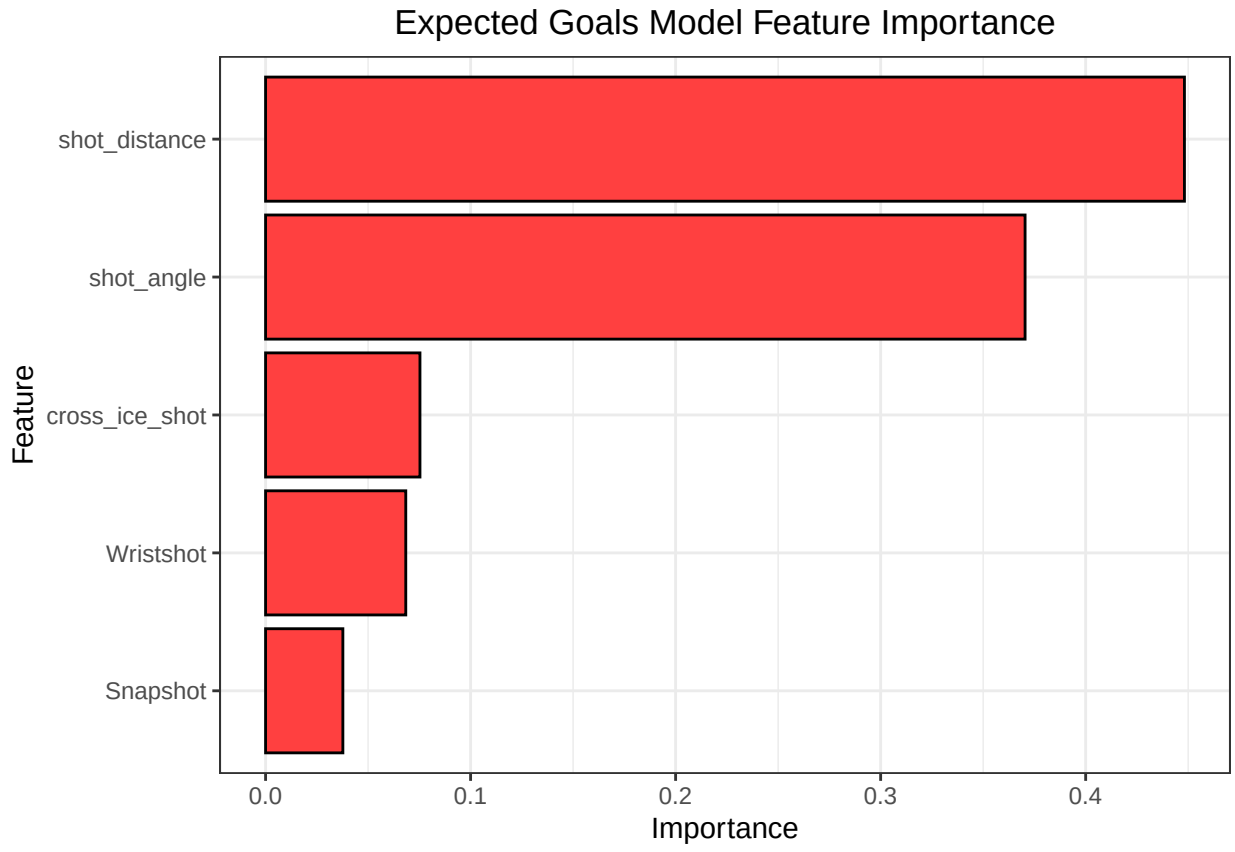
To measure the value of a zone entry on goal scoring and shot quality, I built an expected goals model to determine how dangerous each shot in the data set is. We can then identify shots off the rush, determine the expected value of any given shot and determine whether rush chances are scored more often and at a rate above what is expected of them. I trained an XGBoost binary classifier expected goals model. The data provided was randomly split into a training and testing set as there wasn't a large enough sample size to separate the sets based on a series of games or seasons. I selected the following features as model inputs:

1. Shot Distance - The distance from the shot location to the center of the goal (X coordinate of 189 and Y coordinate of 42.5)
2. Shot Angle - The angle of the shot relative to the center of the goal
3. Cross Ice Shot - Binary classifier of whether the shot came directly from a pass crossing an imaginary line at the Y coordinate of 42.5 (Pass from X coordinate > 42.5 and shot from X coordinate < 42.5 or vice versa)

4. Goal Line Shot - Binary classifier of whether the shot came directly from a pass below the goal line (Pass from X Coordinate > 189 and shot from X coordinate < 189)
5. Rebound - Binary classifier of whether there was a shot attempt in the previous 2 seconds
6. Shot Type - Shot types include Wristshot, Snapshot, Slapshot, Fan, Deflection, Wrap Around, Bat and Poke

Of the 423 shot attempts in the data provided, 338 were in the training set and 85 in the test set. After cross validation and training, the model was evaluated on the test set and the result and feature importance plot can be seen in the figure below.

CV Logloss	CV AUC	Test AUC
0.1784299	0.6326468	0.6722



As seen in the feature importance plot, not all of the intended features were included in the final model due to a lack of observations. There were only 35 instances of goal line shots, 17

instances of a rebound and all shot types other than wrist shots and snap shots had fewer than 27 instances.

After the model was build, I applied it to all of the shot attempts in the data set to give each shot an xG value. I then identified all of the shot attempts that qualified as rush shots. I have defined a rush shot as a shot attempt that occurred within 4 seconds of a zone entry event in the data set.

Overview of Findings

Rush	Shots	xG	G	xG/Shot	Sh%	(G-xG)/Shot	Median xG
No Rush	318	11.2588	12	0.0354	0.0377	0.0023	0.0171
Rush	105	6.3091	8	0.0601	0.0762	0.0161	0.0196

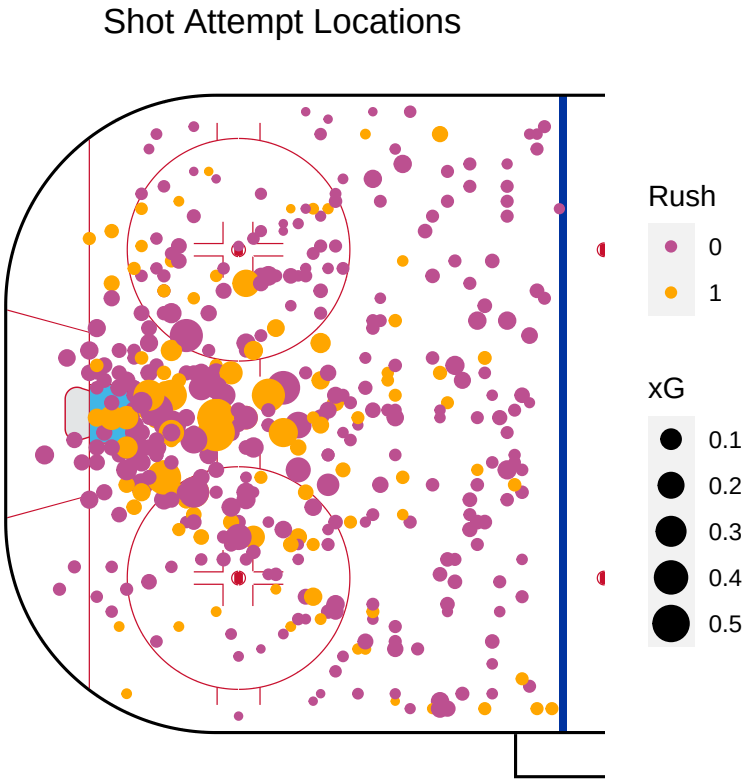
As seen in the above table, shot attempts that come off of the rush have a higher expected goal value per shot attempt, a higher shooting percentage, a higher goals above expected rate and have a higher median expected goal value. We can see that the expected goal value per shot attempt is significantly higher for rush attempts than non rush attempts(0.0601 vs 0.0354). Not only are players creating more dangerous scoring opportunities off of the rush, they are also finishing at a rate well above expected relative to the expected goal value. A shooting percentage of 7.62% off of the rush is well above the shooting percentage of 3.77% of non rush shot attempts. There was also a larger difference in goal to expected goals from rush attempts per shot attempt (1.61%) than non rush attempts (0.23%).

As seen in the table below, of the 11 shots with the highest expected goal value in the data set, 7 of them are rush attempts and of the 14 shot attempts with the smallest expected goal value in the data set, only 3 of them are rush attempts. Having the majority of the highest expected goal values will help increase the average expected goal value per shot attempt for rush attempts however, we can also see that the median value of rush attempts is higher than non rush attempts too.

Highest Expected Goal Shot Attempts

Game ID	shot_distance	shot_angle	cross_ice_shot	Wristshot	Snapshot	rush	xG
20231111	17.01	88.32	1	0	1	1	0.5171822
20231214	17.18	81.63	0	1	0	1	0.4964137
20231108	13.12	49.64	0	1	0	1	0.3998327
20231108	26.23	82.33	1	1	0	0	0.3769870
20231216	24.13	84.05	1	1	0	1	0.3611908
20231111	16.71	51.07	0	1	0	0	0.3604639
20231216	9.01	86.82	1	0	1	0	0.3565393
20231216	8.38	72.65	1	0	1	1	0.2990682
20231108	11.28	77.20	0	1	0	1	0.2983847
20231108	17.50	53.13	0	1	0	1	0.2953499
20231111	17.50	53.13	0	1	0	0	0.2953499

In the figure below, we can see that the shot attempts off of the rush come from closer to the net. The shot attempts off of the rush were from an average of 29.6 feet compared to non rush attempts that came from an average of 33.4 feet. Rush attempts also came from closer to the center of the ice with an average shot angle of 57.5 compared to non rush attempts with an average of 56.3 degrees (a shot from the center of ice would have a Y coordinate of 42.5 in this data set and a shot angle of 90).



Conclusion

In conclusion, this paper has shed light on the relationship between zone entries, rush chances and their affect on creating goal scoring opportunities. These initial results indicate that shot attempts off of the rush are more dangerous than shot attempts that come from extended zone time in terms of expected and actual goals.

Next Steps

1. Include more data points when training the expected goals model. Not all of the intended features of the expected goal model were included due to a lack of observations and having more data points will also improve the model performance in terms of lowering the logloss and increasing the AUC
2. Using player tracking data to identify odd man rushes to find out how this affects the expected goal value of those opportunities relative to rush chances were the defending

team has an equal or larger number of players compared to the attacking team

3. Incorporate rush chances as a feature in the expected goal model