Playoff Probability Based on NHL Teams Stats

Marty Knuth

The question I want to answer is: can I predict the probability of making the playoffs based on team statistics from past seasons? This research is interesting because I am very interested in sports analytics particularly in hockey and the NHL. I would consider this data to be useful for teams to determine how likely they are to make the playoffs. The direct beneficiary of my research would the front offices of NHL teams or fans of teams on the cusp of making the playoffs. General Managers could use the information to determine rosters moves at the deadline and what weaknesses they need to address. My plan for analyzing the data would be as follows: I need to gather all available team statistics from the last 10 seasons, remove all statistics directly correlated to wins and losses, create a binary representation of what teams made the playoffs so I can use a probabilistic machine learning model to determine what statistics best correlate to making the playoffs.

Feature Table

|  |  |
| --- | --- |
| **Feature** | **Description** |
| Goals for per game at all events | This is how many goals that team scored in the season. Divide by the number of games so I can predict teams that haven’t played a full season. |
| Goals against per game at all events | This is how many goals that team allowed in the season. |
| Power Play % | Percentage of times scored on the powerplay |
| Penalty Kill % | Percentage of times they didn’t get scored on the penalty kill. |
| Shots for per game | Average number of shots on goal this team has |
| Shots against per game | Average number of shots on goal this team allows |
| Shooting percentage | Percentage of shots that become goals |
| Team save percentage | Percentage of shots against that are not goals |
| Even Strength Goals For % | Percentage of goals for at even strength – even strength means the teams have an even number of players on the ice |
| Even Strength Shots For % | Percentage of shots for at even strength |
| Even Strength Fenwick For % | Percentage of unblocked shot attempts for at even strength |
| Even Strength Corsi For % | Percentage of shot attempts for at even strength – typically used to measure possession of the puck |
| Expected Goals For % | Percentage of expected goals for – expected goals uses shot location data to predict the expected number of goals a team will score |
| Goals for per 60 | Goals scored for every 60 minutes of ice time played at even strength |
| Goals against per 60 | Goals scored against for every 60 minutes of ice time played at even strength |
| Shots for per 60 | Shots on goal for every 60 minutes of ice time played at even strength |
| Shots against per 60 | Shots on goal for the opposing team for every 60 minutes of ice time played at even strength |
| Fenwick for per 60 | Unblocked shot attempts for at even strength for every 60 minutes of ice time |
| Fenwick against per 60 | Unblocked shot attempts against at even strength for every 60 minutes of ice time |
| Corsi for per 60 | Shot attempts for at even strength for every 60 minutes of ice time |
| Corsi against per 60 | Shot attempts against at even strength for every 60 minutes of ice time |
| Expected goals for per 60 | Expected goals for at even strength for every 60 minutes of ice time |
| Expected goals against per 60 | Expected goals against at even strength for every 60 minutes of ice time |
| Even strength shooting percentage | Percentage of shots converted to goals at even strength ( may drop because it is very similar to regular Sh% ) |
| Even strength save percentage | Percentage of shots saved by goalie at even strength ( may drop because it is very similar to regular Sv% ) |

\*\* please note that I may drop the following features: GF%, CF% xGF% FF% because they could be too similar to CF/60, CA/60, xGF/60 etc but I left it in for this