# Analysis of NHL Goalscoring in Critical Situations

## Insights into Critical Goalscoring Behaviour and Player Classification

**Shakeel Jivraj**

**University of Toronto**

**May 2024**

# Table of Contents

## Abstract

Traditional National Hockey League (NHL) statistics simply focus on the number of goals scored by players, which overlooks their performance in critical situations, such as close or tied games. This research aims to identify NHL players who excel in scoring goals during pivotal moments and establish a classification system to discern their performance in critical game scenarios. Using data retrieved from the NHL API, such as Shot Attempts Percentage (SAT%) and Unblocked Shot Attempts Percentage (USAT%) observed during close and tied game situations, alongside game-winning goals, a ranking system was constructed to evaluate NHL players based on their performance in critical game contexts. In addition, a random forest binary classification model was developed to categorize players based on their performance in critical situations. As a result of the high negative imbalance in the dataset, various metrics such as precision and recall, as well as class weightings, were used to assess the accuracy of the model. While the model was reliable in correctly identifying top NHL goalscorers during pivotal game situations, further research is needed to determine its predictive accuracy in classifying players.

# 1. Introduction

During the 2023-2024 season, Auston Matthews scored 69 goals, the highest single-season total since Alexander Ovechkin’s 65-goal season in 2007-2008 (“Most goals in a season between 1999-00 to 2023-24”). Matthews’ exceptional scoring ability garnered widespread attention within the hockey world, since goalscoring is obviously a very important aspect to winning hockey games. However, an issue with traditional goalscoring statistics is that they treat all goals equally regardless of the game context. For example, a goal scored when a team is leading 5-1 late in the third period carries the same weight as a game-tying goal in the third period. It is thus important to not only analyze NHL players with the most goals, but also the players that teams can rely on to score goals in close or tied games.

This research examines goalscoring in critical situations, often referred to as "clutch" moments. Through analysis of NHL data, including Shot Attempts Percentage (SAT%) and Unblocked Shot Attempts Percentage (USAT%) observed during close and tied game situations, as well as game-winning goals, the aim is to identify the most effective players when the game hangs in the balance. Due to the subjective nature of a player’s performance in pivotal contexts, the NHL statistics must be combined to establish an accurate ranking system. In addition, this paper seeks to develop a random forest binary classification model determines if an NHL player can be considered a clutch goalscorer. Various statistical tests will be applied to determine the accuracy of the model.

# 2. Metrics Used to Analyze Performance in Critical Situations

## 2.1 Overview

This section covers the statistics that were selected to measure the goalscoring of players in pivotal moments of NHL games. Only forwards were considered due to their primary responsibility of scoring goals compared to defensemen and goaltenders. All data was retrieved from the NHL API.

## 2.2 SAT% in Critical Moments

The formula for Shot Attempts (SAT) % is provided below ("NHL Player Stats"):

$$\text{SAT \%} = \frac{Shots\ on\ Goal + Missed\ Shots + Blocked\ Shots}{Shots\ Against + Missed\ Shots\ Against + Blocked\ Shots\ Against} \times 100$$

SAT% measures the shot differentials of players. A higher SAT% implies that the player is controlling the puck for longer periods of time and generating more shot attempts than opponents, which can contribute to increased goalscoring opportunities.

All SAT% data was collected from critical moments, such as close or tied game situations.

## 2.3 USAT% in Critical Moments

The formula for Unblocked Shot Attempts (USAT) % is provided below ("NHL Player Stats"):

$$\text{USAT \%} = \frac{Shots\ on\ Goal + Missed\ Shots}{Shots\ Against + Missed\ Shots\ Against} \times 100$$

USAT% is a similar metric to SAT% but excludes blocked shots. USAT% evaluates a player's ability to direct shots towards the opposing team's goaltender and create goalscoring chances.

All USAT% data was collected from critical moments, such as close or tied game situations.

## 2.4 Game-Winning Goals

A game-winning goal is the goal that was scored by a player to put their team ahead and win the game. For example, if the score of the game was tied 2-2, but the final score was 4-2, the goal that resulted in the score becoming 3-2, would be the game-winning goal. Game-winning goals demonstrate which players can score goals in crucial moments of games.

## 2.5 Available Metrics that were not Considered

Although overtime and shorthanded goals can be indicative of clutch performance, the highest number of goals in these categories for the 2023-2024 season was 3 and 6, which caused difficulties in determining percentiles for these statistics.

## 2.6 Potential Shortcomings with Metrics

While SAT%, USAT% and game-winning goals are useful in determining a player’s goalscoring ability in tied and close game situations, other statistics such as goals scored in the third period as well as game-tying goals would have been beneficial to analyze. However, these metrics are not available on the NHL API and are not provided by any other sources.

In addition, SAT% and USAT% quite literally measure a player’s ability to “get pucks to the net”. They do not account for the quality of the shots and can overstate a player's offensive contribution if the shots generated lack sufficient scoring threat. Despite this limitation, SAT% and USAT% do remain one of the few statistics available to evaluate player performance in critical situations and should be taken into consideration.

# 3. Identifying Top Performers in Critical Situations

## 3.1 Overview

The metrics from the previous section helped to formulate a ranking system that determines the best goalscorers in the NHL during pivotal situations. The percentile rankings of SAT% and USAT% (for close and tied games), as well as game-winning goals, were computed for all players and then combined to formulate the rankings.

## 3.2 Ranking Performance in Close Games

The average of percentiles for SAT% and USAT% from close game scenarios was calculated to determine percentile rankings for generating shot attempts in tight contests. Since there is a small difference in SAT% and USAT% for most players, this provides a balanced assessment of goalscoring when the game is closely contested.

## 3.3 Ranking Performance in Tied Games

This approach follows a similar methodology to section 2.2. The average of percentiles for SAT% and USAT% from tied game situations was calculated to determine percentile rankings for generating shot attempts during tied games.

## 3.4 Weighting Percentile Rankings

The averages of the percentiles for close and tied situations, as well as game-winning goals, were computed to determine a percentile ranking for a player's clutch goalscoring. The following weightings were assigned to each of the percentile rankings:

a) Close Percentile: 30%
b) Tied Percentile: 30%
c) Game-Winning Goals Percentile: 40%

As shown above, a higher percentile was assigned to the game-winning goals. This statistic tracks actual goals scored in critical situations, unlike SAT% and USAT% which may provide inaccuracies due to the quality of shots (see Section 1.6).

## 3.5 Top 10 Rankings

The table below (Figure 1) shows the top 10 goalscorers in critical situations. It displays their percentiles for different metrics, as well as their overall percentile and ranking amongst other players. According to the table, Sebastian Aho is the most clutch goalscorer in the NHL.

| Name | Close Percentile | Tied Percentile | Clutch Percentile | GWG Percentile | Rank |
|---|---|---|---|---|---|
| Sebastian Aho | 97.14% | 96.4% | 97.81% | 99.36% | 1 |
| Seth Jarvis | 96.4% | 95.66% | 96.85% | 98.09% | 2 |
| Zach Hyman | 97.08% | 96.72% | 96.83% | 96.72% | 3 |
| Sam Reinhart | 95.12% | 94.97% | 96.77% | 99.36% | 4 |
| Andrei Svechnikov | 98.3% | 94.76% | 96.61% | 96.72% | 5 |
| Nathan MacKinnon | 93.06% | 91.26% | 94.53% | 98.09% | 6 |
| Mikko Rantanen | 90.26% | 88.46% | 93.15% | 98.83% | 7 |
| Leon Draisaitl | 90.52% | 89.14% | 93.13% | 98.09% | 8 |
| Connor McDavid | 97.56% | 93.91% | 92.91% | 88.67% | 9 |
| Aleksander Barkov | 95.24% | 95.92% | 92.82% | 88.67% | 10 |

Figure 1: Table showing top 10 clutch goalscorers in the NHL

## 3.6 Player Cards

Player cards (Figure 2) on Power BI display each player's percentiles and final ranking amongst other forwards. The cards are available to download in this provided GitHub repository.

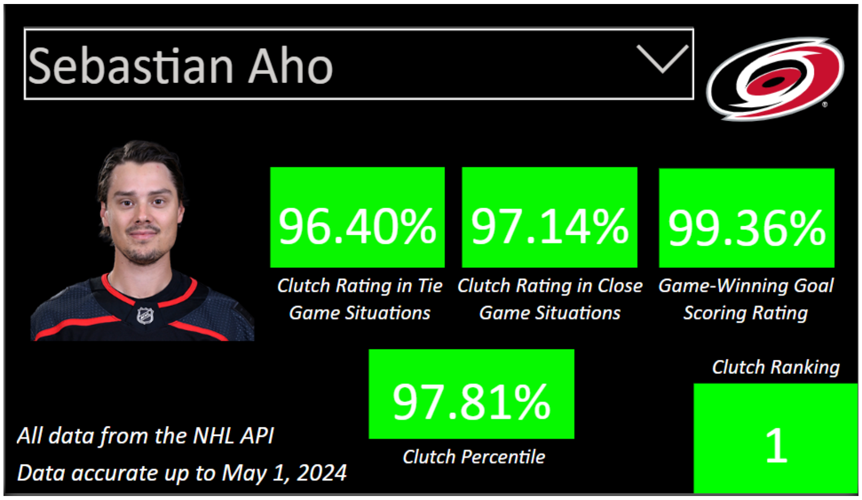


Figure 2: Player card showing percentiles and final ranking for Sebastian Aho

# 4. Developing a Random Forest Classification Model

## 4.1 Overview

The most statistically involved section of this study was developing a classification model that determines if an NHL player can be considered a clutch goalscorer. Once an appropriate algorithm was selected, many statistics were employed to evaluate the model's effectiveness.

## 4.2 Selection of Algorithm

After considering different algorithms for the classification task of identifying clutch goalscorers in the NHL, a random forest model was selected. This is because random forest's ensemble approach captures non-linear relationships between SAT%, USAT%, game-winning goals and clutch goalscoring, which provides an accurate assessment of clutch goalscoring behaviour.

## 4.3 Defining Classification for Goalscoring in Pivotal Moments

Since determining whether a player is an effective goalscorer in critical moments can be subjective, it was important to develop a method that concretely assesses SAT%, USAT%, and game-winning goals.

### 4.3.1 Player Classification Using SAT% and USAT%

The distributions of SAT% and USAT% in close and tied situations are shown below (Figure 3).

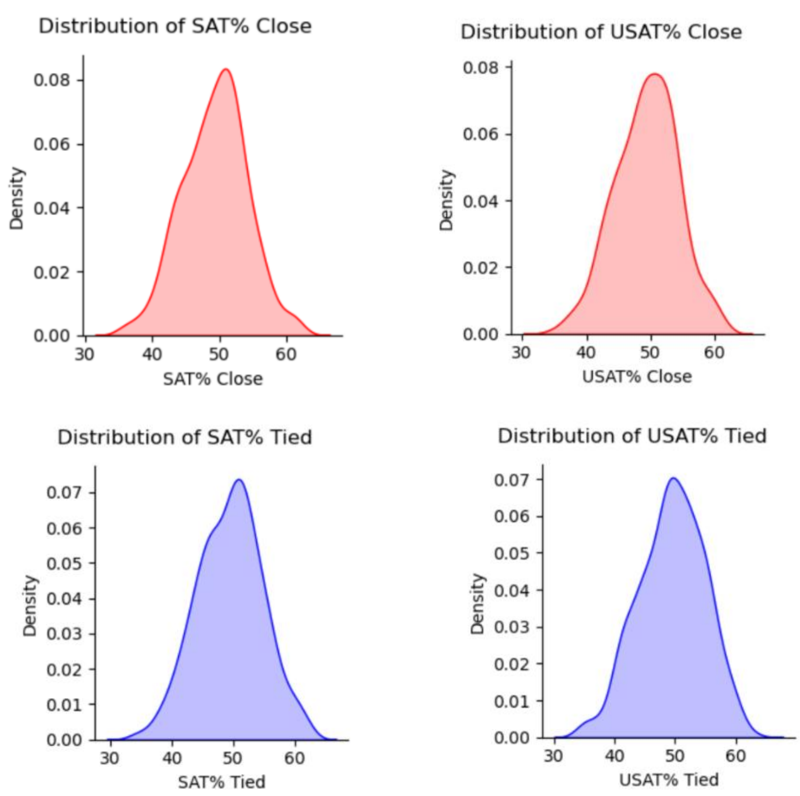


Figure 3: Density plots for SAT% and USAT% in close and tied game situations

The density plots show that each of the metrics follows a standard normal distribution centered at 50%. This seems reasonable because a SAT and USAT of 50% indicate that a player is neither controlling play nor getting outshot, and thus has an average offensive contribution. Considering that SAT and USAT greater than 55% are considered elite (Wilson), one factor in determining if a player can be classified as a clutch goalscorer is if at least one of their SAT% and USAT% metrics in close and tied games exceeds 55%. Given the minor discrepancy between SAT% and USAT%, achieving a value greater than 55% in at least one category was deemed to be sufficient for classification purposes.

### 4.3.2 Player Classification Using Game Winning Goals

The distribution of game-winning goals is shown below (Figure 4).

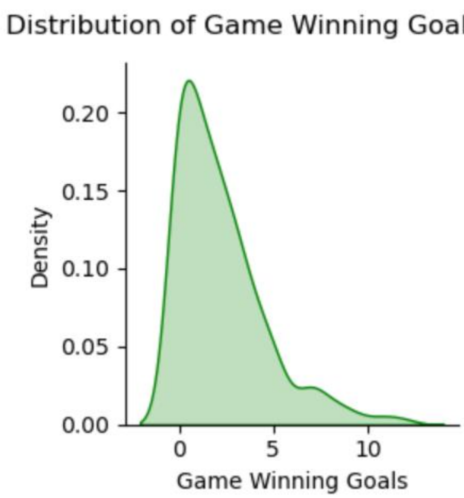


Figure 4: Density plot for game-winning goals

Unlike SAT% and USAT%, the density plot for game winning goals follows a right skewed distribution. Since there is no data available indicating what would be considered an elite number of game-winning goals, the 80th percentile was chosen as the threshold. Therefore, players who have scored more game-winning goals than 80% of other players would be regarded as clutch in terms of game-winning goals.

### 4.3.3 Player Classification Using USAT%, SAT% and Game-Winning Goals

From the analysis of the distributions of USAT% and SAT% in close and tied game situations, as well as game-winning goals, a player will be classified as a clutch goalscorer and assigned a binary value of 1 if they meet the following criteria:

a) They have at least one SAT% and USAT% metric in close and tied game situations that exceeds 55%.
b) They have scored more game-winning goals than 80% of other players in the dataset.

## 4.4 Dataset Imbalance

The binary classification presented a problem of negative class imbalance because over 80% of players in the dataset were classified as non-clutch goalscorers (Figure 5 and Figure 6):

```
clutch_performance
0     385
1      87
```

Figure 5: Number of players classified as clutch (1) and non-clutch goalscorers (0)

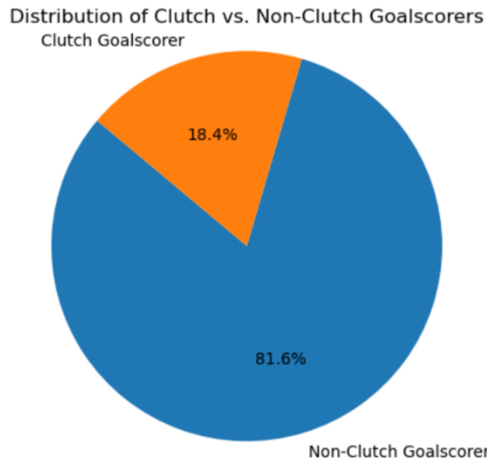


Figure 6: Pie chart breakdown of binary classification

## 4.5 Random Forest Classifier Accuracy

The random forest classifier has an unusually high accuracy of approximately 97%. Due to the high negative imbalance, the model is biased towards predicting the majority class, resulting in inflated accuracy metrics that do not accurately reflect its performance on the minority class.

The model's bias towards the majority class is further shown by the results of the 10-fold Cross-Validation (Figure 7), where the classifier still produces an inflated accuracy.

```
scores = cross_val_score(model, df_players_merged[x_var], df_players_merged[y_var].values.ravel(), cv=10)
scores
```

```
array([1.        , 0.97916667, 0.9787234 , 0.9787234 , 1.        ,
       0.93617021, 1.        , 1.        , 0.9787234 , 0.9787234 ])
```

```
scores.mean()
```

```
0.98302304964539
```

Figure 7: Results of 10-fold Cross-Validation on the classifier

Resampling may not be the best solution to address the issue of negative imbalance in the dataset because it will be difficult to change the thresholds for classifying clutch goalscoring based on USAT%, SAT% and game-winning goals. Therefore, in the next section, other metrics were employed to evaluate the accuracy of the classifier. Please note that the section will contain many statistics, some of which will still be inaccurate.

## 4.6 Additional Statistics to Measure the Classifier's Accuracy

### 4.6.1 AUC under the ROC Curve

The AUC under the ROC curve for this classification model is close to 1. Since there is a negative class imbalance and the classifier is biased towards predicting the majority class, it will have a high true negative rate and a low false positive rate. This means that there is an upward shift of the ROC Curve towards the top-left corner of the plot (Figure 7), leading to an excessively high AUC. Consequently, AUC under the ROC curve would not be a reliable way of measuring the accuracy of the model.

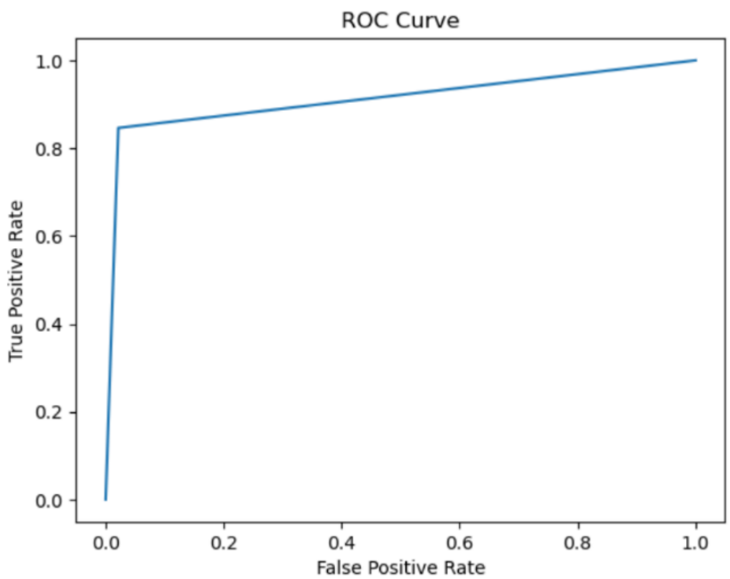


Figure 7: ROC Curve with shift toward top-left corner

### 4.6.2 Precision and Recall

#### 4.6.2.1 Precision

The model has a precision of approximately 88%, which is also influenced by the negative imbalance in the dataset. As explained with the AUC under the ROC curve, bias towards the majority class leads to a high number of true negatives and a low number of false positives. The effect on precision can be seen through the calculation for this statistic:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Regardless of the number of true positives the model identifies, a low number of false positives implies the denominator will be small and thus increase precision.

#### 4.6.2.2 Recall

The model's recall of approximately 85% is not inflated by the data's imbalance. Unlike precision, the denominator of the recall calculation only focuses on how the classifier identifies positive instances in the dataset:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Since the number of false negatives is in the denominator and the recall is quite high, the number of false negatives is quite low. Therefore, the model is accurately identifying positive instances, even if there are more negatives in the dataset.

#### 4.6.2.3 F1 Score

The F1 score of approximately 86% appears higher than expected due to the inaccuracy of the precision. This is shown by the formula for F1:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Precision is multiplied by recall on the numerator but added to recall on the denominator, which means F1 increases when precision is high.

#### 4.6.2.4 Precision-Recall Curve

With a high precision and recall, the AUC for the precision-recall curve is close to 1 (Figure 8). However, this may be overly optimistic because of an artificially increased precision.

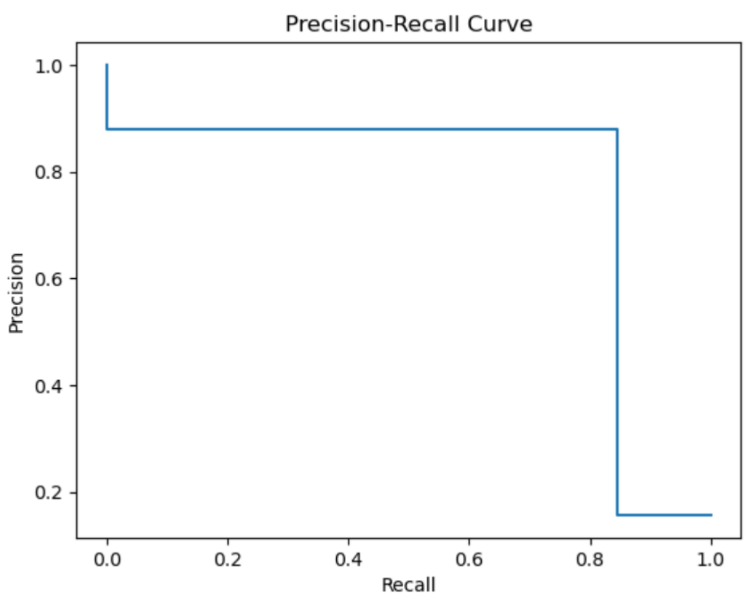


Figure 8: Precision-Recall curve with high AUC

### 4.6.3 Confusion Matrix

The confusion matrix (Figure 9) is an excellent way of analyzing the effect of negative imbalance on test statistics and verifying the previous claims about precision and recall discussed in Section 3.6.2.

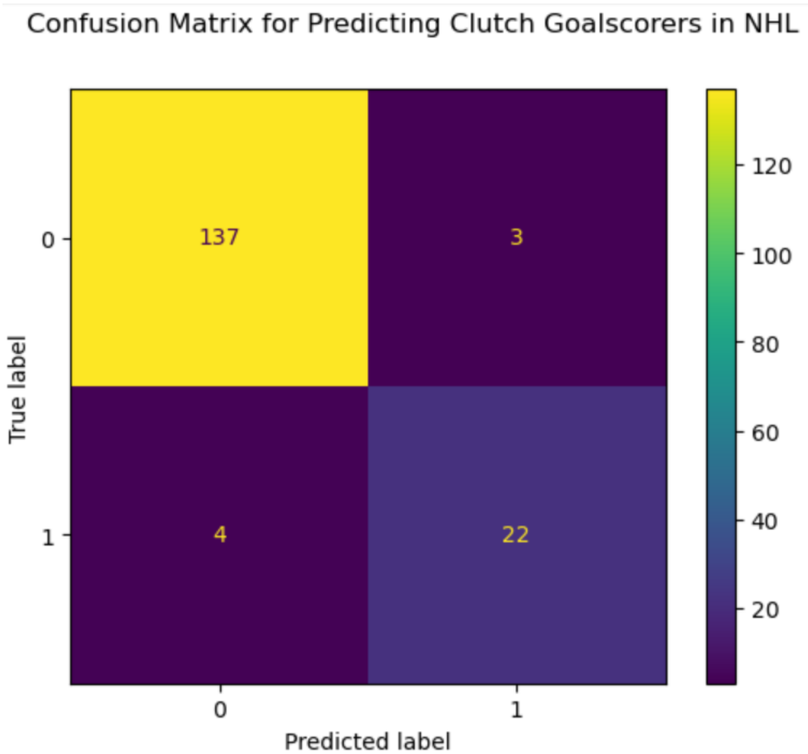


Figure 9: Confusion matrix of classification model

The confusion matrix shows that the dataset has 22 true positives and 3 false positives. Since the number of true negatives is very high (137), the model was biased towards the majority class and identified very few false positives, thus increasing precision. This validates the claim in Section 3.6.2.1 about the model's inflated precision.

The confusion matrix also identifies 22 true positives and 4 false negatives in the dataset. This illustrates that the model is reliably detecting clutch NHL goalscorers and possesses a high recall, as explained in Section 3.6.2.2.

### 4.6.4 Matthew's Correlation Coefficient

The Matthew's Correlation Coefficient (MCC) yielded an accuracy of 83%, which is similar to the F1 Score of 86%. Although the MCC tracks true negatives, unlike the F1 Score, the negative imbalance of the dataset implies that true negatives are abundant. The MCC may provide limited analysis on the minority class, even if it considers all four elements of the confusion matrix. Hence, the MCC was not used to evaluate the performance of the classification model.

#### 4.6.5 Class Weightings

Class weightings were used to limit the bias from the majority class. As demonstrated in Figure 10, higher weight was assigned to the positive class:

```
{1: 2.732142857142857, 0: 0.612}
```

Figure 10: New weightings assigned to each class

After adjusting the weightings, the model had an accuracy of 96% and mean of 97% for the 10-fold Cross-Validation (Figure 11). This proves that the model is consistently identifying clutch NHL goalscorers and further demonstrates a high recall.

```
array([0.97916667, 0.97916667, 0.95744681, 0.9787234 , 1.        ,
       0.93617021, 1.        , 1.        , 0.93617021, 0.95744681])
```

Figure 11: Results of 10-fold Cross-Validation with new weightings

### 4.7 Conclusion

The binary classification model is very effective in correctly identifying clutch NHL goalscorers in the dataset, as shown by its high recall value and the model's performance after changing class weightings to favour the majority group.

However, it is difficult to determine the likelihood of the model correctly classifying an NHL player as a clutch goalscorer because the negative class imbalance increases the precision. It is important to have an accurate precision so that resources are not wasted when predicting the performance of players in critical moments. One potential solution is to analyze data from multiple NHL seasons. While this may not eliminate class imbalance, it can increase training on the minority class and lead to better predictive performance.

# 5. Code

The code for the model described in this paper is available in a GitHub repository. The repository also includes the Power BI dashboard for the player data.

## 6. References

"Most Goals in a Season between 1999-00 to 2023-24." *StatMuse*, www.statmuse.com/nhl/ask?q=most%2Bgoals%2Bin%2Ba%2Bseason%2Bbetween%2B1999-00%2Bto%2B2023-24.

"NHL Player Stats." *PuckPedia*, puckpedia.com/stats.

Wilson, Kent. *Don't Know Corsi? Here's a Handy-Dandy Primer to NHL Advanced Stats*, 6 Oct. 2014, www.calgaryherald.com/sports/Wilson+know+Corsi+Here+handy+dandy+primer+advanced+stats/10265096/story.html.