Statistical Evaluation of Outdoor Field Hockey Penalty Corners

Samuel R. Hughes[a], Gregory J. Matthews[b], & Jun Yan[a]*

[a]*Department of Statistics, University of Connecticut, Storrs, CT*
[b]*Department of Mathematics and Statistics, Loyola University Chicago, Chicago, IL*

https://doi.org/10.33697/ajur.2025.145

Student: samuel.hughes@uconn.edu
Mentor: gmatthews1@luc.edu, jun.yan@uconn.edu*

ABSTRACT
Penalty corners stand out as pivotal goal-scoring opportunities in field hockey, crucial to a team's triumph. This study harnesses data from women's collegiate field hockey games to formulate a statistical model predicting the likelihood of scoring a penalty corner, contingent on the strategies deployed. Various machine learning algorithms are compared to ascertain the most predictive model and to dissect the paramount factors influencing penalty corners. The XGBoost model emerges superior, boasting an area under the curve (AUC) score of 0.667 on out-of-sample observations. With other predictors held constant, the model reveals that drag flicks, sweep shots, and deflections are positively associated with goal occurrences, while, intriguingly, direct shots—despite their prevalence—are negatively associated with scoring probability.

KEYWORDS
College Sports; K-Nearest Neighbor; Lasso; Quantitative Analysis; Random Forest; Sporting Strategy; Sports Analytics; XGBoost

INTRODUCTION

Field hockey, originating from ancient Greece, has sustained its prominence through centuries and currently boasts over 250 women's collegiate programs actively competing[1]. This sport engages two teams, each consisting of eleven players, in a strategic contest to score by maneuvering a ball into the opponent's goal using specially crafted sticks. Unlike ice hockey, field hockey is played on either natural or artificial grass surfaces, within a field measuring 91.4 meters by 55 meters. Notably, goals are only valid if scored from within a defined area, termed the striking circle[2]. The strategic complexities and widespread popularity of field hockey underscore the significance of analytical studies within the sport.

Our analysis focuses on penalty corners. With approximately one-third of goals occurring on penalty corners in high-level women's field hockey matches, and the North Carolina Tar Heels, winner of five of the last seven national championships, dedicating twenty percent of their practice time to them, it is clear that penalty corners are pivotal to match outcomes[3, 4]. A penalty corner is awarded to the attacking team when a defender commits a foul within the striking circle. During this play, an inserter hits the ball from the end line into the field, while teammates, positioned above the penalty circle by rule, endeavor to control the ball and score. Conversely, the defensive team positions four outfield players and a goalkeeper behind the goal line, who may move only after the inserter has struck the ball, aiming to thwart the attackers[2].

Some studies have examined strategic approaches to penalty corners in outdoor field hockey. Laird and Sutherland[5] analyzed the 1998 field hockey World Cup and found most successful goals resulted from straight shots, either flicked

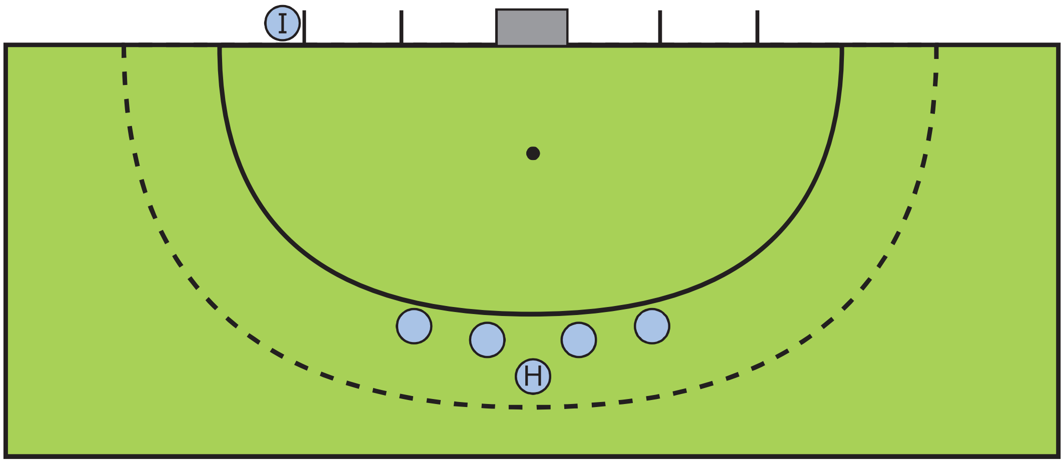

Figure 1. A diagram of a possible offensive setup during a field hockey penalty corner, with blue circles representing attackers. The inserter ("I") puts the ball into play by passing it to a teammate positioned outside the shooting circle, as required by the rules. The ball typically flows to the hitter ("H"), who then plays the ball toward the goal, often after a teammate stops the ball just outside the circle. Other offensive players execute roles such as stopping the ball, screening defenders, or positioning for rebounds or deflections.

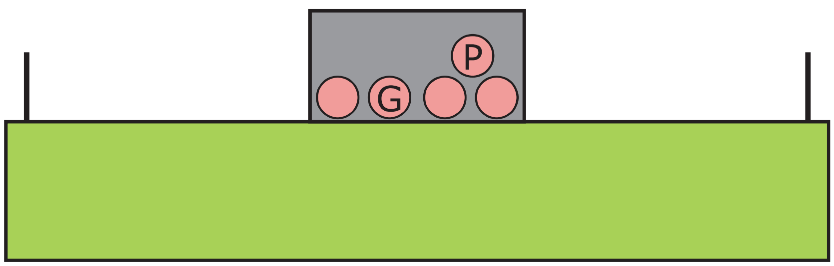

Figure 2. A diagram of a possible defensive setup during a field hockey penalty corner, with red circles representing defenders. By rule, a maximum of five defenders, including the goalkeeper ("G"), begin behind the end line until the ball is inserted. In this instance, all defenders are positioned within the goal. Also shown is the post defender ("P"), who plays a key role in stopping shots near the goal. While most defenders sprint toward the attackers upon insertion, the post defender typically remains near the goal (either just in front or behind the goalkeeper) to provide an additional layer of defense.

or undercut. Kerr and Ness [6] compared push-in performers of different experience and advised coaches to maximize drag distance and utilize a blend of simultaneous and sequential segment rotations to optimize accuracy and ball speed, thereby maximizing drag speed. Lopez De Subijana et al. [7] analyzed the kinematics of drag flicks, a prevalent shot technique during penalty corners, and observed distinct differences in stance width and explosive movement of the stick and pelvis across varying skill levels. Klatt et al. [3] determined that adaptive decision-making after ball insertion can enhance scoring opportunities during a penalty corner. Lord et al. [8] discovered variations in team strategy on penalty corners, contingent upon their physical capabilities, in 2019 men's and women's International Hockey Federation Pro League matches. Notably, no studies have scrutinized penalty corners using both statistics and machine learning to identify areas for improvement. Our analysis will explore a wide spectrum of strategic choices, including shot technique, shot location, and goalie positioning.

In the broader domain of sports akin to outdoor field hockey such as soccer and indoor field hockey, numerous studies have employed machine learning models and statistical testing to extract valuable insights into sporting strategy. Alcock [9] analyzed the probability of scoring a free kick in elite international women's soccer, advocating for shots in situations with minimal goal distance and a favorable angle. Vinson et al. [10] modeled the likelihood of scoring a penalty corner in women's professional indoor field hockey competitions, pinpointing the goalkeeper's strategy—either holding the goal line or confronting attackers—as a crucial predictor. Maneiro et al. [11] discerned notable differences in soccer corner kick tactics between the men's and women's 2018 and 2019 World Cups, respectively. Anzer and Bauer [12] inves-

Variable	Description	χ^2 test p-value
Goal	Indicates if the attacking team scored during the penalty corner	
Deflection	Indicates if the ball changed direction off an attacker toward the goal	0.034*
Direct Shot	Indicates if an attacker hit the ball directly toward the goal	0.130
GK Up	Indicates if the goalkeeper was standing upright while defending	0.214
Post in Front of GK	Indicates if the post defender was positioned in front of the goalkeeper	0.297
Drag Flick or Sweep	Indicates if the shot used a flicking or sweeping motion rather than a standard hit	0.023*
Shot Bottom	Indicates if the shot was aimed at the lower third of the goal	0.052
Shot Info	Indicates if there was a recorded shot on goal during the penalty corner	0.265

Table 1. Descriptions of explanatory variables and the p-values of their association tests with the Goal variable.

tigated momentary scoring probability in top-tier German league men's soccer matches, identifying an extreme gradient boosting model as optimal for unseen data. Of these, Vinson et al.[10] aligns most closely with our work, though it concentrates on professional indoor field hockey rather than collegiate outdoor field hockey. While most pertinent papers have primarily focused on professional sports, our study seeks to shed new light on field hockey and women's collegiate sports.

Our contributions to the field are twofold. First, utilizing qualitative information regarding strategies and techniques employed during a field hockey penalty corner, we hypothesize that the likelihood of scoring a goal can be predicted. Second, we seek to provide insights into the most effective tactics utilized during these set pieces. To provide a brief overview, we utilize data supplied by a women's collegiate field hockey team on penalty corners and compare various machine learning methods to identify the most predictive model. Subsequently, we offer a potential interpretation of the findings within the context of field hockey strategy and gameplay.

DATA

The data for this study was provided by the University of Connecticut NCAA Division I women's collegiate field hockey program. Specifically, the team utilized Hudl SportsCode^{TM}, a sports analysis software, to examine occurrences during a given penalty corner, recording various categorical traits of the play, including goalkeeper positioning, defensive formation, and shot technique, among other variables. The analyses were subsequently converted into a tabular dataset. The penalty corners included in the study span several games from the 2021 season, all involving either the University of Connecticut field hockey team or a scouted opponent. Given that the original intent of the data was to track qualitative traits in a non-statistical analysis, the initial dataset occasionally presented issues for performing statistical methods, such as containing two values in a single data cell. In such instances, we consistently utilized the last value of the cell for analysis. After cleaning, we aimed to use the data to discern strategies that maximize the expected chance of scoring a goal during outdoor field hockey penalty corners.

The cleaned data contains $n = 309$ observations, each corresponding to a penalty corner occurrence. Among these 309 penalty corners, 42 (13.59%) resulted in a goal. Each observation contains qualitative insights of the occurrences on the play. These qualitative variables were converted to binary variables, which better allows statistical analysis on the data. The outcome variable of interest is Goal, indicating whether a penalty corner sequence resulted in a goal. **Table 1** presents a list of explanatory variables that can potentially be used to predict scoring.

Figure 3 showcases mosaic plots that effectively visualize the relationship between different strategies and goal scoring in field hockey. These plots use the vertical axis to represent the occurrence of scoring, while the plot width indicates the sample size of each grouping[13]. For enhanced clarity, the total number of observations within each group is also provided. A careful analysis of these plots reveals a discernible positive correlation between the probability of scoring and strategies such as deflections, drag flicks, sweep shots, and targeting the lower part of the net. This suggests that teams employing these tactics can expect a higher chance of scoring. In contrast, a negative relationship is evident between scoring chances and the use of direct shots, indicating that they may be less effective in scoring goals.

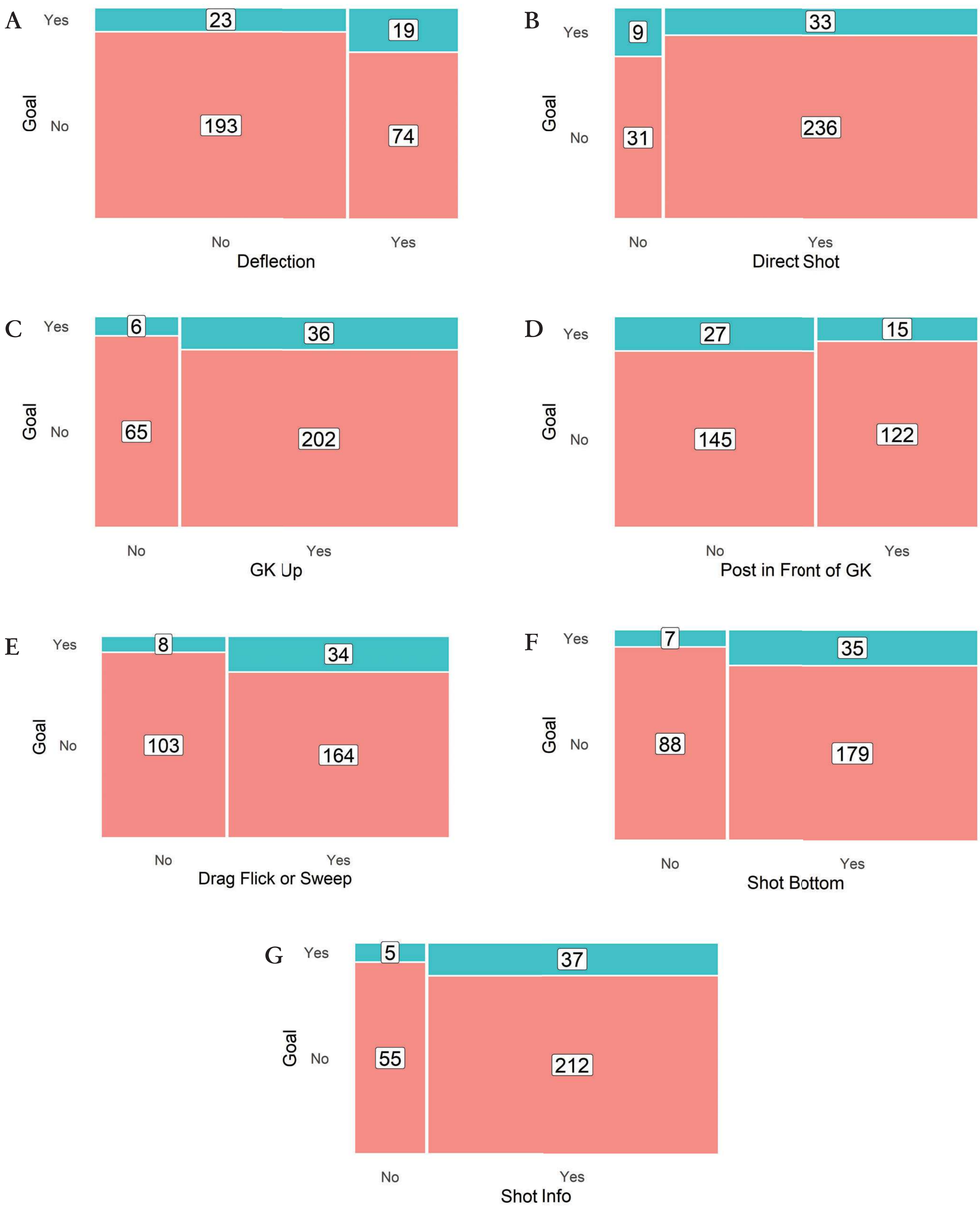

Figure 3. Mosaic plots illustrating the relationship between goal scoring and various explanatory variables. The height of the blue bars represents the proportional difference in goal scoring associated with the presence or absence of a given strategy, while the width reflects how frequently that strategy was employed. Subplots (A–G) correspond to the seven strategies analyzed, with axes labeled in each panel.

Also reported in **Table 1** are the p-values from χ^2 tests for all the explanatory variables, assessing their associations with Goal. "Deflection" and "Drag Flick or Sweep", with p-values of 0.034 and 0.023 respectively, demonstrate statistically significant associations with scoring. Conversely, variables such as "Direct Shot", "GK Up", "Post in Front of GK", "Shot Bottom", and "Shot Info", exhibit p-values ranging from 0.052 to 0.297, indicating no statistical significance in their association with scoring. The variables in this table provide a departure point of using advanced machine learning approaches to predict scoring during penalty corners.

METHODS

In our study, we focused on building machine learning models to accurately predict the occurrence of a Goal during penalty corners in field hockey. We employed a variety of modeling algorithms, each offering a distinct approach to predicting scoring chances. These include lasso logistic regression[14], K-nearest neighbor[15], random forest[16], and extreme gradient boosting (XGBoost)[17].

Each model's tuning parameters were carefully selected through a non-nested 5-fold cross-validation process[18], with stratification based on the Goal outcome. In addition to the stratification, all penalty corners from the same game were assigned to the same fold to prevent data leakage. For the lasso logistic regression, we explored a wide range of penalty parameters C, from 10^{-5} to 10^5, and incorporated second-degree interactions between predictors. The optimal performance was achieved with a penalty of $C = 1$. Similarly, the K-nearest neighbor model was tuned across a spectrum of K values, from 1 to 101, again considering second-degree interactions. The model reached its peak effectiveness when K was set to 49.

We configured the random forest model with 500 trees and undertook a tuning process to optimize various parameters. These parameters included the minimum tree depth sample size (ranging from 1 to 8), the impurity metric (Gini vs. entropy), the minimal number of samples required for a split (either the default value of 2 samples or a range from 10% to 25% of the data), and the maximum number of features to be considered in a tree (options being square root, $\log_2$, or all available features). The optimal performance was achieved using the Gini criterion, a maximum depth of 5, a minimum of 10% of the data samples per split, and considering the square root of the total number of features in a tree.

The XGBoost model underwent a similar tuning process. Similarly to random forest, the model was fixed to 500 trees. We sought the best settings for maximum tree depth (from 1 to 8), learning rate (between 0.05 and 0.30), regularization alpha (between 0 and 10), and regularization lambda (also between 0 and 10). The best results were obtained with a maximum depth of 2, a learning rate of 0.05, a regularization alpha of 0.5, and a regularization lambda of 10.

Out-of-sample predictions from 5-fold cross-validation were computed for the probability of scoring a goal during each penalty corner given the predictors. This method involves dividing the data into five parts, using four for training our model and one for testing its predictions. This approach helps ensure our model's predictions are reliable when applied to new, unseen data. We focused on the best-performing model from each method and examined its effectiveness using Receiver Operating Characteristic (ROC) curves. These curves are crucial for understanding a model's ability to distinguish between two outcomes—in our case, whether a goal is scored or not.

The area under the curve (AUC) of the ROC curve serves as a robust metric to numerically evaluate the predictive performance of classification models[19]. The AUC quantifies the overall ability of the model to discriminate between the binary classification of goal outcome during a penalty corner. An AUC of 1.0 indicates perfect predictive ability, while an AUC of 0.5 suggests no predictive ability, akin to random guessing[20]. In our study, a higher AUC indicates that the model has a higher probability of ranking a randomly chosen positive instance (a scored goal) higher than a randomly chosen negative instance (a missed goal). Thus, in comparing the predictive power of the four methods, the model with the highest AUC would be considered the most proficient in accurately predicting the occurrence of a goal during a penalty corner, providing a valuable metric in the evaluation and comparison of model efficacy.

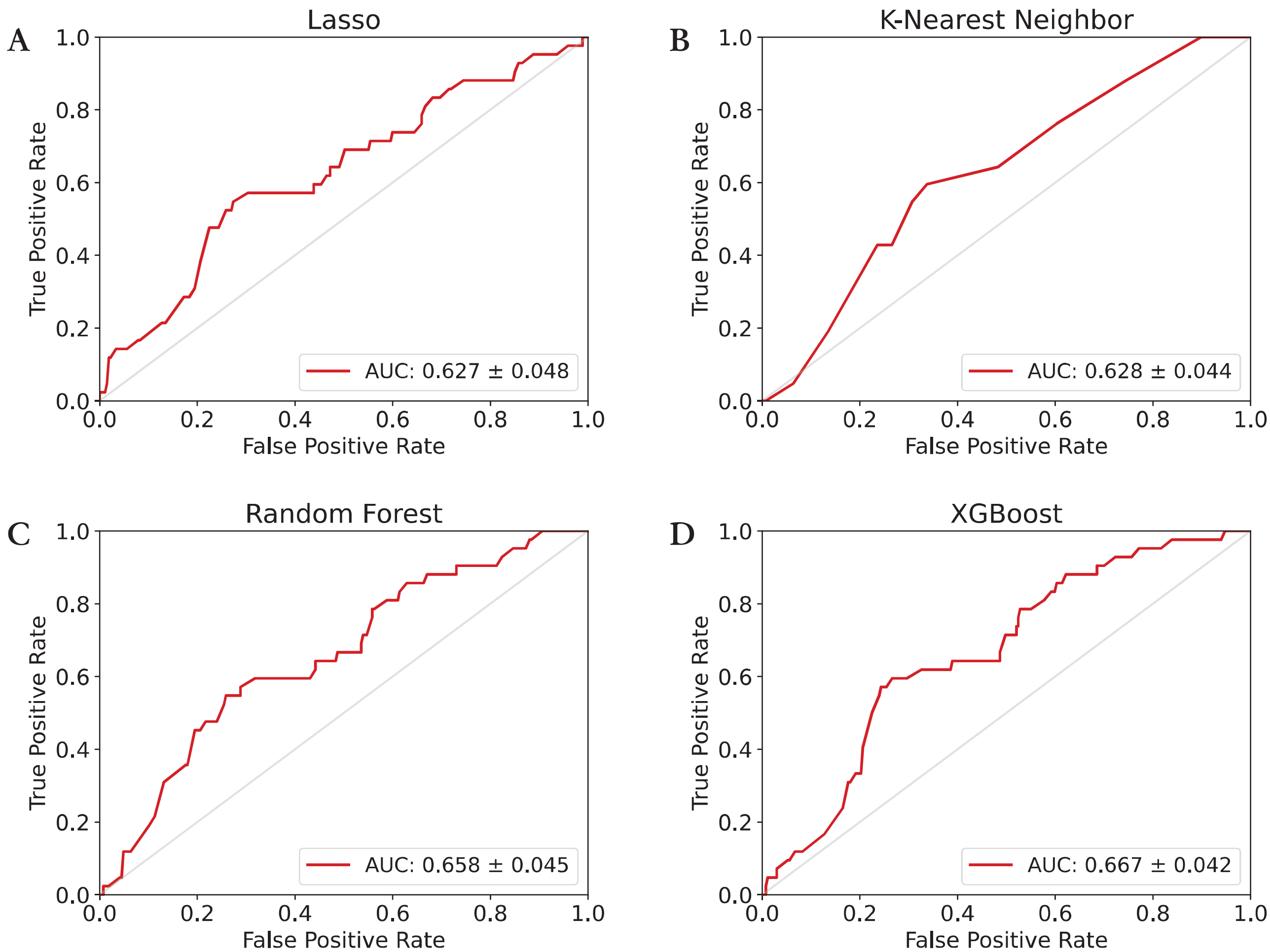

Figure 4. ROC curves for four classification models. Red lines show the overall ROC curve across all out-of-sample predictions, with higher AUC values indicating better classification performance. Each plot also displays the standard error of the AUC, providing a measure of method stability. Subplots (A–D) correspond to: (A) Lasso, (B) K-Nearest Neighbor, (C) Random Forest, and (D) XGBoost, with axes labeled in each panel.

Figure 4 displays the ROC and AUC score derived from out-of-sample predictions using the four methods. In addition to the AUC score, a standard error is also included, calculated using the nonparametric method, to better indicate the stability of each model[21]. XGBoost emerges superior, achieving an AUC of 0.667. Given the relative success of XGBoost compared to the other modeling techniques, it is selected for further analysis. An AUC of 0.667 indicates a moderate ability of the XGBoost algorithm to predict the goal scoring chance of unseen observations. While there may be additional predictors not included in our dataset or model that could enhance the AUC score, the model nonetheless identifies some predictive relationship between the predictors and response, deeming it adequate for the purposes of this study.

RESULTS

We have chosen to present the results from the XGBoost model as it demonstrates superior performance with the highest AUC among all the models we tested. The feature importance scores, derived from XGBoost's built-in functionality, are illustrated in **Figure 5**. In this chart, features are arranged in order of their importance; those with higher importance values are placed on the top, while those with lower importance are on the bottom. A key observation from this analysis is the paramount importance of whether a deflection occurred during the penalty corner sequence, which holds the highest feature importance value at 0.24. Following closely are factors such as the goalkeeper's alignment relative to the post defender (0.23), a direct shot taken (0.18), and shots aimed toward the bottom of the net

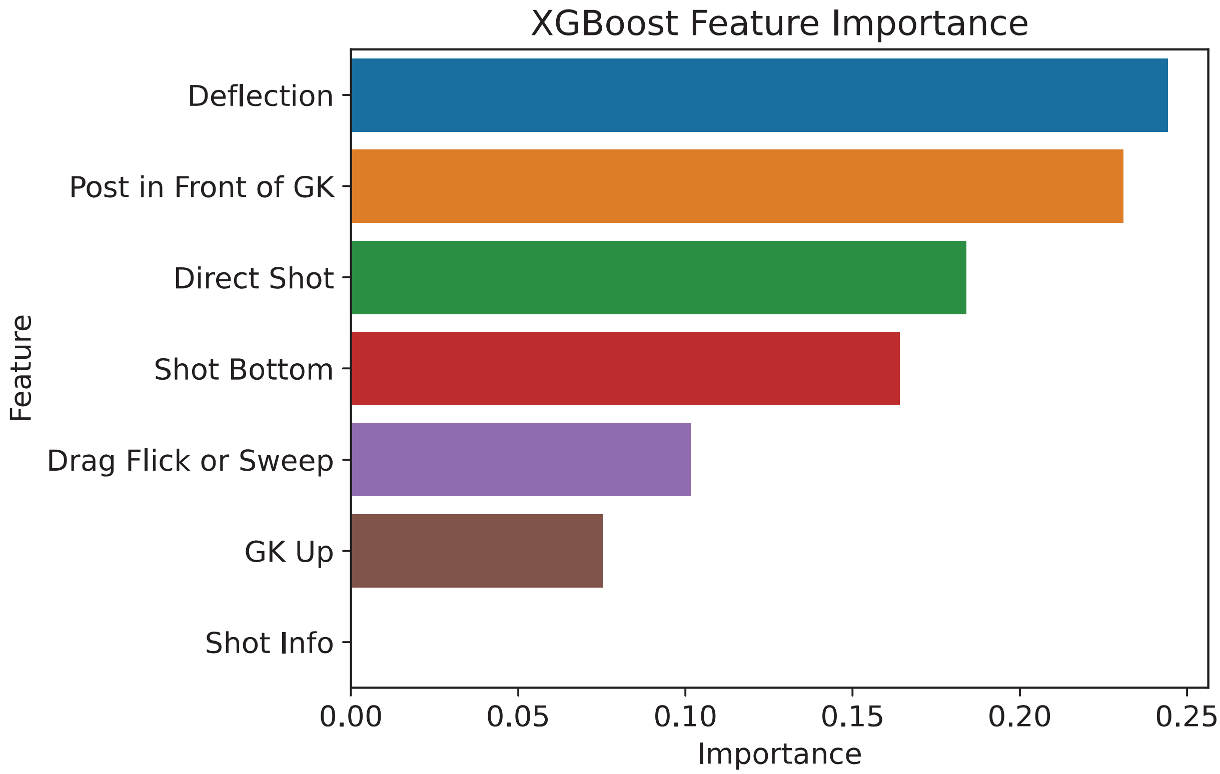

Figure 5. Feature importance scores from the XGBoost model. The plot shows how much each variable contributed to the model's predictions, with higher values indicating greater magnitude of influence on the prediction of whether a goal was scored, regardless of whether the variable increased or decreased the predicted likelihood.

(0.16). Other features with less impact include the use of drag flick or sweep shot techniques (0.10) and whether the goalkeeper adopts an upright stance (0.08). Interestingly, the model did not use the feature that indicates if a shot occurred during the penalty corner (0.00). These results provide insightful perspectives on the factors most influential in determining the success of a shot during a penalty corner in field hockey.

While the feature importance values derived from the XGBoost model offer valuable insights, they fall short in one crucial aspect: they do not indicate the direction of the relationship between each feature and the target variable. Simply put, the feature importance plot does not indicate whether the strategies used in goal scoring have a positive or negative impact on the likelihood of scoring. To address this gap, we turn to SHAP (SHapley Additive exPlanations) values[22]. SHAP values provide a more nuanced view by detailing how each predictor influences the target variable, thereby illuminating the direction of their impacts. This approach allows us to understand not only which features are important, but also how they positively or negatively affect the chances of scoring in field hockey, as indicated by our model.

Figure 6 illustrates the variable impacts on scoring from our analysis. Given that our predictors are binary, a red point on this plot signifies that a particular strategy was employed during the play, while a blue point indicates its absence. This color coding helps us discern that the use of drag flick or sweep shot techniques is strongly positively associated with an increased chance of scoring. When these techniques are executed, they significantly boost the predicted likelihood of scoring. Conversely, their absence is linked to a considerable decrease in scoring probability. Similarly, strategies such as directing shots toward the bottom of the net and deflecting shots toward the net also show a positive relationship with scoring success. Interestingly, the goalkeeper's adoption of an upright stance, while having a weaker influence, still positively correlates with scoring during a penalty corner. On the flip side, strategies that tend to reduce scoring chances include the post defender aligning in front of the goalkeeper and taking direct shots. Both strategies ex-

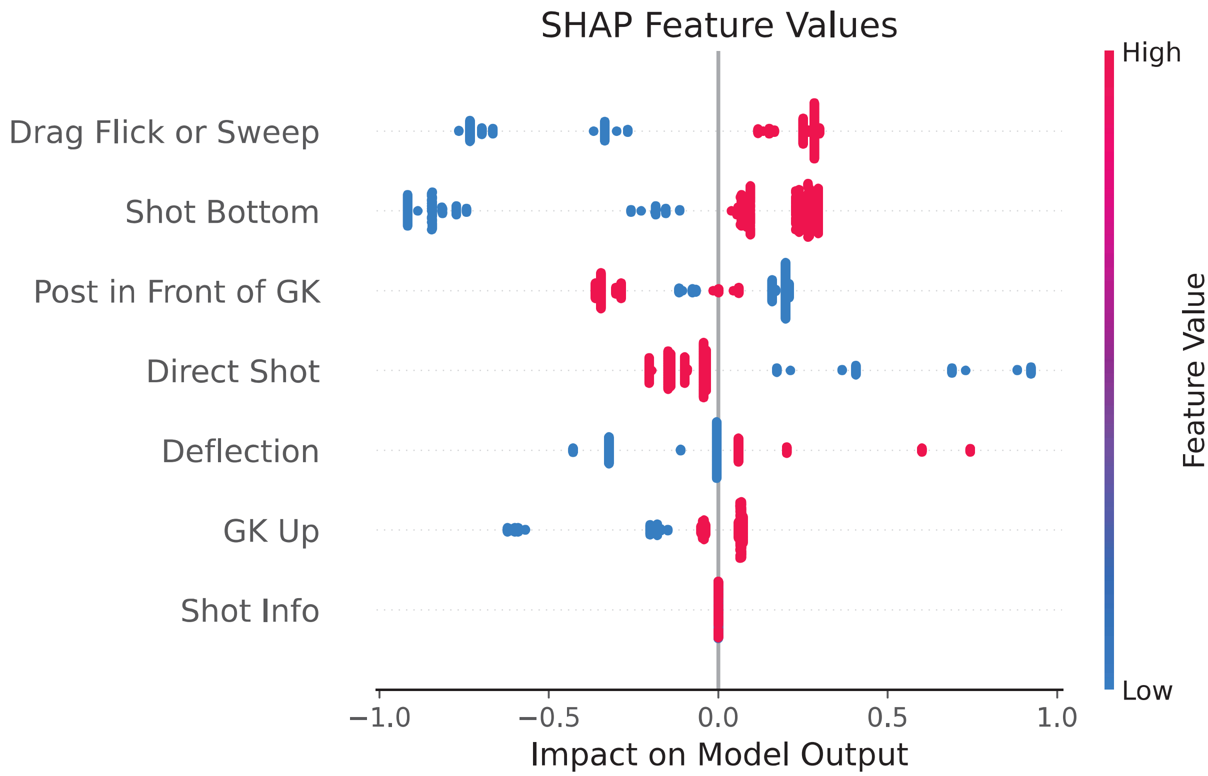

Figure 6. SHAP values from the XGBoost model. The plot illustrates how each binary feature influenced predictions of goal scoring. Each point represents a single observation, with red indicating that the feature occurred and blue indicating that it did not. Points farther from zero indicate a greater impact on the model's prediction, with the direction reflecting whether the feature increased or decreased the predicted likelihood of a goal.

hibit a moderate negative association with scoring. As also noted in **Figure 5**, the feature indicating if a shot is taken during the penalty corner was not used by the model and therefore has no impact on the chances of scoring, according to our model.

Our findings offer valuable insights for field hockey coaches. Teams may improve their scoring effectiveness by strategically focusing on deflections, drag flicks, and sweeps aimed at the lower regions of the net. Conversely, for defense, teams may reduce the opposition's scoring efficiency by aligning their post defender in front of the goalkeeper and the goalkeeper adopting a downward stance. A noteworthy aspect of our analysis is the negative relationship between direct shots and successful goal attempts. Our model suggests, with all other variables constant, a reduced frequency of scoring from direct shots. This may seem surprising, but it is supported by our initial dataset. Goals were scored on 22.5% of the 40 penalty corners without a direct shot, as opposed to only 12.3% of the 269 penalty corners with a direct shot. Therefore, teams might consider exploring potentially more effective strategies if direct shots are identified as the root cause of suboptimal goal scoring opportunities.

DISCUSSION

In the endeavor to predict the likelihood of scoring during a field hockey penalty corner, this study embarked on the creation and comparison of various machine learning models, each tailored to the specific events and strategies deployed during the play. Amongst a spectrum of modeling techniques, XGBoost was distinguished as the most potent predictive model, a claim substantiated by its AUC score of 0.667 on out-of-sample data. This model unveiled a positive association between the probability of scoring and the utilization of the deflection shot type, as well as the drag flick and sweep shot techniques, upright goalkeeper stance, and shots located toward the bottom of the net during penalty corners. Moreover, our findings illuminated a somewhat unexpected and counter-intuitive negative association between the likelihood of scoring and the execution of a direct shot, despite the latter's prevalent deployment in matches. This particular insight accentuates the opportunity for teams to reevaluate and innovate their penalty corner strategies, potentially exploring alternative techniques that might enhance scoring probabilities.

Concluding the analysis, the research into field hockey penalty corners still harbors substantial potential for further exploration and discovery. With data more meticulously tailored for statistical analysis, it is conceivable that numerous additional predictors could be incorporated into a similar model, including the shot distance, precise player locations during the penalty corner, the prevailing score, and the residual time in the game. An additional research avenue could involve addressing a multiclass classification problem, where the response variable encompasses not only the occurrence of a goal, but also other outcomes such as missed shots, saved attempts, or turnovers during the play. It is also pertinent to note that the data only encompasses games from a single season of a specific field hockey program and some scouted opponents. Future studies, armed with data spanning a more expansive array of collegiate women's field hockey teams, may yield even more compelling and robust results.

ACKNOWLEDGMENTS
We are grateful to UConn Field Hockey Head Coach Paul Caddy and Assistant Coach Cheri Schulz not only for their appreciation of sports analytics, but also for graciously volunteering their time to explain field hockey concepts and intricacies.

REFERENCES

1. Harris, J. (2016). Everything You Need to Know About One of the World's Oldest Sports `https://www.ncaa.com/news/fieldhockey/article/2016-08-19/field-hockey-everything-you-need-know-about-one-worlds-oldest`. NCAA.
2. Ansari, A. (2022). Field Hockey Rules: A Guide to Understanding the Sport `https://olympics.com/en/news/field-hockey-rules-how-to-play`. Accessed: 2025-08-20.
3. Klatt, S., Noël, B., Schwarting, A., Heckmann, L., and Fasold, F. (2021). Adaptive gaze behavior and decision making of penalty corner strikers in field hockey, *Frontiers in Psychology* 12, 674511. *https://doi.org/10.3389/fpsyg.2021.674511*
4. Willis, C. (2024). Penalty Corner Goals Lift No. 1-Seeded UNC Field Hockey to ACC Championship Game `https://www.dailytarheel.com/article/2024/11/sports-field-hockey-syracuse-acc-tournament-semifinals-charly-bruder-erin-matson-november-2024`. The Daily Tar Heel.
5. Laird, P.Sutherland, P. (2003). Penalty corners in field hockey: A guide to success, *International Journal of Performance Analysis in Sport* 3(1), 19–26. *https://doi.org/10.1080/24748668.2003.11868270*
6. Kerr, R.Ness, K. (2006). Kinematics of the field hockey penalty corner push-in, *Sports Biomechanics* 5(1), 47–61. *https://doi.org/10.1080/14763141.2006.9628224*
7. Lopez De Subijana, C., Juárez, D., Mallo, J., and Navarro, E. (2010). Biomechanical analysis of the penalty-corner drag-flick of elite male and female hockey players, *Sports Biomechanics* 9(2), 72–78. *https://doi.org/10.1080/14763141.2010.495414*
8. Lord, F., Pyne, D. B., Welvaert, M., and Mara, J. K. (2022). Identifying and analysing game styles and factors influencing a team's strategy in field hockey, *Journal of Sports Sciences* 40(8), 908–919. *https://doi.org/10.1080/02640414.2022.2037839*
9. Alcock, A. (2010). Analysis of direct free kicks in the women's football World Cup 2007, *European Journal of Sport Science* 10(4), 279–284. *https://doi.org/10.1080/17461390903515188*
10. Vinson, D., Padley, S., Croad, A., Jeffreys, M., Brady, A., and James, D. (2013). Penalty corner routines in elite women's indoor field hockey: Prediction of outcomes based on tactical decisions, *Journal of Sports Sciences* 31(8), 887–893. *https://doi.org/10.1080/02640414.2012.757341*
11. Maneiro, R., Casal, C. A., Arda, A., and Losada, J. L. (2019). Application of multivariant decision tree

technique in high performance football: The female and male corner kick, *PLoS One* 14(3), e0212549. *https://doi.org/10.1371/journal.pone.0212549*
12. Anzer, G.Bauer, P. (2021). A goal scoring probability model for shots based on synchronized positional and event data in football (soccer), *Frontiers in Sports and Active Living* 3, 624475. *https://doi.org/10.3389/fspor.2021.624475*
13. Theus, M. (2012). Mosaic plots, *Wiley Interdisciplinary Reviews: Computational Statistics* 4(2), 191–198. *https://doi.org/10.1002/wics.1192*
14. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B (Methodological)* 58(1), 267–288. *https://doi.org/10.1111/j.2517-6161.1996.tb02080.x*
15. Laaksonen, J.Oja, E. (1996). Classification with learning k-nearest neighbors In *Proceedings of International Conference on Neural Networks (ICNN'96)*, 1480–1483. *https://doi.org/10.1109/ICNN.1996.549118*
16. Breiman, L. (2001). Random forests, *Machine learning* 45(1), 5–32. *https://doi.org/10.1023/A:1010933404324*
17. Chen, T.Guestrin, C. (2016). Xgboost: A scalable tree boosting system In *Proceedings of the 22nd ACM SigKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. *https://doi.org/10.1145/2939672.2939785*
18. Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society: Series B (Methodological)* 36(2), 111–133. *https://doi.org/10.1111/j.2517-6161.1974.tb00994.x*
19. Fawcett, T. (2006). An introduction to ROC analysis, *Pattern Recognition Letters* 27(8), 861–874. *https://doi.org/10.1016/j.patrec.2005.10.010*
20. Hanley, J. A.McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve., *Radiology* 143(1), 29–36. *https://doi.org/10.1148/radiology.143.1.7063747*
21. DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach, *Biometrics* 44(3), 837–845. *https://doi.org/10.2307/2531595*
22. Lundberg, S. M.Lee, S.-I. (2017). A unified approach to interpreting model predictions In *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 4765–4774. *https://doi.org/10.48550/arXiv.1705.07874*

ABOUT THE STUDENT AUTHOR

Samuel R. Hughes graduated from the University of Connecticut in December 2022 with a Bachelor of Arts in Statistics and Data Science double major, complemented by a concentration in Economics. He then earned a Master of Science in Data Science from Northeastern University in December 2024. He currently works in data science at Hologic, Inc.

PRESS SUMMARY

Penalty corners stand out as pivotal goal-scoring opportunities in field hockey, crucial to a team's triumph. This study harnesses data from women's collegiate field hockey games to formulate a statistical model predicting the likelihood of scoring a penalty corner, contingent on the strategies deployed. We find that drag flicks, sweep shots, and deflections are positively associated with goal occurrences, while, intriguingly, direct shots—despite their prevalence—are negatively associated with scoring probability.