A success metric for individual player encounters in Ice Hockey videos

William Jiang[1], Ken M. Nsiempba[2]
John Zelek[2], David Clausi[2], Yuhao Chen[2],
[1] University of California, Los Angeles
[2] Vision and Image Processing Group, Systems Design Engineering, University of Waterloo
{jiangwil}@g.ucla.edu
{kmnsiemp, jzelek, dclausi, yuhao.chen1}@uwaterloo.ca

Abstract

Player scouting is a critical component of hockey, enabling teams to identify standout players, devise strategies, and assess performance. However, traditional scouting methods are resource-intensive, requiring substantial time and travel expenses for organizations. An approach to performance assessment has been video footage analysis; however, in the context of ice hockey, challenges arise primarily due to the fast-paced nature of the sport and the limited field of view in broadcast footage.This makes it difficult to consistently track all players on the ice simultaneously. A solid performance metric that would make the most out of the information at hand is thus required. We present the 1v1 Success Metric Pipeline, an autonomous system designed to analyze video footage and calculate a success metric for individual player encounters. By automating the tracking and evaluation process, this tool not only streamlines scouting but also provides coaches with valuable insights for player performance, which may come into play during line matching and other strategic decision-making during games. Results show that the ISM metric properly reflects the danger level of a one on one encounter and that despite the sample size being small, the model achieves low root mean square error.

1 Introduction

The development of a rigorous performance metric is partially motivated by shortcomings in existing metrics. They might only focus on physical characteristics rather than performance. For example, [1] discusses how traditional scouting has often prioritized player characteristics like height and weight over actual performance metrics. When not properly defined, existing metrics can be biased. The studies [2, 3] highlight how bias can significantly impact scouting and player selection. Despite French Canadian players outperforming English Canadian players in offensive metrics, biases rooted in geopolitical tensions and stereotypes about their defensive abilities created additional barriers for them. This suggests that subjective perceptions, rather than just performance, can influence scouting decisions in the NHL. These examples underscore the need for a rigorous, performance-based metric focused solely on gameplay, free from biases and irrelevant characteristics.

Sports analytics involves capturing and analyzing data to improve athlete performance, refine scouting, and optimize team strategies. Video footage provides a primary source for live or post-game analytics, offering a rich dataset to analyze player actions and team dynamics. However, human analysis is limited by factors such as attention span, fatigue, the difficulty of tracking multiple concurrent events, and the long duration of games. These limitations highlight the need for computer-based algorithms capable of automatically and consistently identifying actions, game events, and team strategies from video footage.

In hockey, video analysis faces specific challenges due to the fast pace of the game, frequent motion blur, and limited field of view in broadcast footage. Several studies have attempted to detect actions and events within these constraints, yielding mixed results [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. These challenges complicate the consistent detection of actions or events, such as puck possession or defensive plays. This paper proposes a new method designed to work effectively despite these limitations.

In hockey, each team has six players on the ice, assigned to specific positions: three forwards (center, left wing, and right wing), two defensemen (left and right), and one goaltender. Each position has distinct responsibilities that contribute to the team's overall strategy.

The defensemen play in front of their team's goal and focus on blocking shots, clearing the puck, and disrupting the opposition's offensive efforts. They often support forwards in transitioning the puck up the ice, initiating attacks, and maintaining control in the offensive zone.

The primary objective of this paper is to develop a specialized metric for evaluating defensemen's performance in hockey, particularly in one-on-one situations. This metric is intended to help teams and coaches more precisely assess a defenseman's ability to isolate and neutralize offensive threats. To support this objective, a dedicated data generation and labeling pipeline has been developed, enabling the creation of comprehensive defensive performance metrics.

The paper is organized as follows:

In section 2, we'll discuss related work and review key performance indicators used in previous studies. Section 3 details our approach to processing footage, ensuring accurate player detection, and extracting player positions. We also explain the labeling process, outline relevant geometrical variables used in our metric, and discuss the implementation of the extreme gradient boosting method for training the model.

The 4 section presents our qualitative and quantitative findings.

2 Background

There are various key performance indicators (KPIs) in sports that can significantly impact both team and individual performance. These KPIs can be broadly categorized into technical, tactical, physical, interpersonal, and combinations of these factors:

Technical KPIs involve game elements such as passes, recoveries, rebounds, and shots, which directly relate to the execution of skills during gameplay [14, 15, 16]. Studies have shown that passing patterns can be used to develop indices for evaluating player performance. For instance, [17] utilized passing patterns to create a performance index that was validated against traditional scoring metrics, offering practical insights that reduce the reliance on subjective coaching assessments and improve training and competition analysis.

Tactical KPIs focus on strategic elements like player positioning and team formations. Tactical analysis provides insights into how well a team maintains structure and adapts to various scenarios during a game, influencing overall team performance [18].

Physical KPIs include metrics such as the distance covered, player speed, and playing time. These indicators assess the physical contributions of players and their fitness levels. For example, [19] examined how the allocation of playing time among players impacts the likelihood of winning, highlighting the importance of physical performance in game outcomes.

Interpersonal and psychological KPIs examine the social dynamics and mental aspects of players and teams. These factors can involve personality traits, communication, and the quality of relationships between players and coaches. Studies like [20] explored the influence of personality traits on team sports performance, while [21] analyzed the effects of interpersonal dynamics, such as closeness, commitment, and complementarity, within the coach-athlete relationship.

In many cases, a comprehensive evaluation of team and individual performance involves a combination of these factors. Studies like [22] and [23] emphasize the importance of integrating multiple KPIs to gain a holistic understanding of what contributes to success in sports.

3 Methodology

3.1 Data Collection

Video subsequences ("shots") were extracted from the NHL dataset, resulting in 150 clips, each lasting 10 seconds, featuring one-on-one encounters between players. Each clip was processed with our custom-trained YOLOv10 object detection model, enabling accurate tracking of players and generation of bounding boxes. The technic described in [24] was utilized. The bounding boxes were used to determine player footprints, defined as the midpoint of the lower bounding box edge. Homography allowed for the transformation of video frames to account for perspective distortions. Player tracking was then performed on the warped video frames, and key measures like distances and angles were computed for each frame. For each clip, we documented player IDs, the frame at which the forward initiated a shot, the calculated instantaneous success metric (ISM) values, and the defensive net's side on the rink. This annotated dataset provides the foundational data we can use to extract the geometrical features required for training our extreme gradient boosting model as seen in figure 1:

1. Distance between the forward and the net, denoted as d_1.
2. Distance between the forward and the defenseman, denoted as d_2.
3. Distance between the defenseman and the shooting lane, denoted as d_3.

4. Angle of the forward with respect to the center line of the net, denoted as θ.

To obtain the features, we generated a filtered dataset containing only the bounding box coordinates for the two players of interest, derived from the original tracking output. We then applied homography transformations to map the player footprints onto the rink's spatial coordinates to determine the position of the forward and defenseman. Additionally, knowing the orientation of the net (left or right side of the rink) allowed us to accurately define the net's center position for each clip.

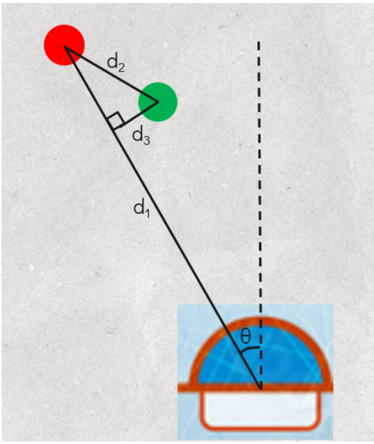

Figure 1: Diagram showing the features for our samples: the forward is in red while the defenseman is in green

The distances between the forward and the net d_1 and between the forward and defenseman d_2 are calculated using the Euclidean formula, which finds the square root of the squared differences in x and y coordinates between each pair of points. To calculate d_3, the perpendicular distance from the defenseman to the shooting lane between the forward and the net, we used vector projections to locate the nearest point on this line segment. Finally, the Euclidean distance between this closest point and the defenseman gives us d_3, indicating how close the defenseman is to the shooting lane. Finally, the angle θ between the forward and the net is found using the arctangent of the y and x differences between the the forward and the net.

3.2 Instantaneous Success Metric (ISM)

Each one-on-one encounter was evaluated using expert reviews, which assigned a score known as the Instantaneous Success Metric (ISM). This score provided a quantifiable measure of the likelihood the forward were to score if they were to shoot the puck at that instant.

3.3 Extreme Gradient Boosting

The collected data was used to train an XGBoost model. XGBoost (Extreme Gradient Boosting) is a robust machine learning algorithm designed for supervised learning tasks such as classification and regression. It operates based on the principle of gradient boosting, combining an ensemble of weak prediction models (typically decision trees) to enhance predictive accuracy [25].

$$\mathcal{L}(\phi) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

Where $\mathcal{L}(\phi)$ represents the overall objective function to be minimized, with n denoting the number of training examples. The term $l(y_i, \hat{y}_i)$ is the loss function that quantifies the difference between the predicted value y_i and the true label $\hat{y}_i$.

The loss function is given by:

$$l(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2$$

The regularization term, $\Omega(f_k)$, penalizes the complexity of each tree f_k to mitigate overfitting:

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

Where T is the number of leaves in the tree, w_j represents the weight of leaf j, γ is a parameter that penalizes the number of leaves, and λ is a parameter that penalizes the L2 norm of the leaf weights.
The model's prediction for each instance is represented as:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i)$$

To model the Individual Scoring Metric (ISM), we applied an XGBoost regression algorithm, utilizing game-specific features to capture the impact of spatial dynamics in one-on-one encounters. Our input dataset included four predictor variables: distances d_1 (forward to net), d_2 (forward to defenseman), d_3 (defenseman to shooting lane), and the angle θ between the forward and the net. The ISM served as the response variable, representing the scoring potential for each scenario. To preprocess the data, we included organizing features (X) and target variables (y), which we loaded into a DMatrix, an optimized data format for efficient memory usage and computational speed in XGBoost.

We configured the XGBoost model with parameters suitable for regression tasks. The objective of the model was configured to minimize squared errors, with model performance assessed using root mean squared error (RMSE). Additional parameters were selected based on initial experimentation and insights from previous predictive modeling research, aiming to achieve a balance between model complexity and generalization. These included:

- **Learning rate** ($\eta = 0.1$): Controlled the step size for each boosting iteration.
- **Maximum tree depth** (6): Set to capture interactions between variables without overfitting.
- **Subsampling rate** (0.8): Enhanced model robustness by training on 80% of the dataset per iteration.

Model training proceeded through 100 boosting rounds, where each iteration optimized predictive accuracy by minimizing residual errors. This iterative boosting process enabled the model to capture complex interactions among features, leveraging XGBoost's strengths in handling structured data with nonlinear relationships.

This approach allowed us to predict ISM values effectively, integrating spatial metrics that reflect player positioning and engagement dynamics in one-on-one plays, contributing meaningful insights into scoring potential.

4 Results

In the qualitative results found in figure 2, our trained XGBoost algorithm predicts the ISM for each frame based on distances and angles, visualized over the hockey rink. We generated a time series graph of ISM to identify sections of the clip where the defenseman allowed the highest scoring opportunities. Another feature of the visualization is that, after the forward shoots the puck, the ISM display changes from white to red font, clearly indicating the moment of the shot. Currently, we must manually input the frame number when the forward shoots the puck, but this could be automated using further research of hockey shot detection.

Figure 3 illustrates the model's RMSE (Root Mean Square Error) during training over 100 boosting rounds. The training RMSE steadily decreases, indicating that the model continues to learn from the data. However, the validation RMSE plateaus after an initial decline, suggesting that additional rounds might not improve performance on the validation set and could lead to overfitting.

Figure 4 shows the importance of various features used in the model, measured by the F-score. Feature d_1 has the highest importance, with a significantly higher F-score than other features, indicating it contributes the most to model predictions. Other features (θ, d_2, and d_3) also play a role but are less influential in comparison.

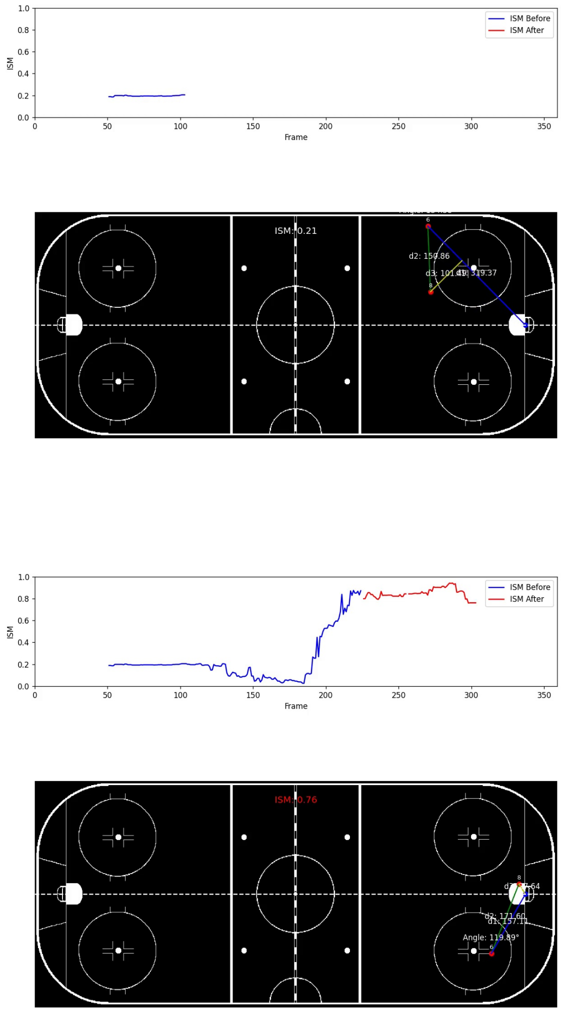

Figure 2: ISM values for various time-frames of a one-on-one encounter

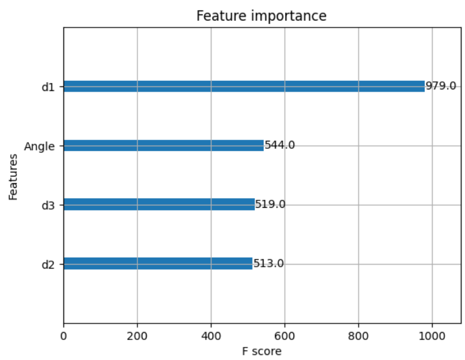

Figure 3: The plot shows the importance scores for each feature used in the model. Feature importance scores reflect how useful each feature was in improving the model's accuracy.

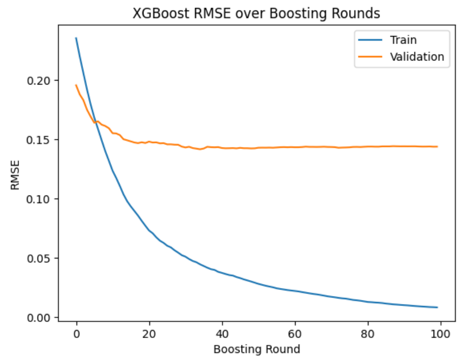

Figure 4: Tracking the model's error reduction over boosting rounds to assess learning progress and generalization

This analysis helps in understanding which features the model relies on most for accurate predictions.

5 Conclusion and Future work

Using the ISM (Individual Scoring Metric) for defensive players throughout a game can offer insights into team performance. This metric could be leveraged in line-matching strategies, where analyzing the peak ISM allowed by each defenseman against various forwards would help identify optimal defensive pairings that collectively minimize team-wide ISM. Such an approach could enhance line compositions, providing coaches with a ranking of the most effective defensemen based on the specific forwards currently on the ice. By optimizing these matchups, teams could reduce opponents' scoring opportunities and improve overall defensive effectiveness.

Effective one-on-one defense is a critical skill in high-pressure situations, requiring a blend of positioning, stick work, and decision-making. A dedicated metric to quantify these abilities would aid in scouting, targeted training, and player development.

With advancements in computer vision, deriving such metrics has become increasingly feasible. Automated techniques now enable the extraction of precise player measurements and behaviors directly from video footage, reducing subjective biases and enhancing the consistency of analyses.

In the future, we would also like to go beyond focusing solely on one-on-one encounters and develop metrics that involve the entire team and their positioning as well.

6 Acknowledgements

We would like to acknowledge Stathletes for the support and the footage provided for these experiments.

References

[1] L. Farah, "Evaluating the Efficacy of Talent Identification and Development in the National Hockey League Entry Draft," Jan. 2022. [Online]. Available: http://hdl.handle.net/10315/39577

[2] M. Lavoie, G. Grenier, and S. Coulombe, "Discrimination and performance differentials in the national hockey league," *Canadian Public Policy/Analyse de Politiques*, pp. 407–422, 1987.

[3] N. Chaikof, K. Clément-Heydra, C. Markey, M. Nahabedian, A. Pilotte, and M. Shaw, "Scouting Automated Ratings Analyzing Habits (SARAH): A Statistical Methodology for Scouting and Player Development," *Linköping Hockey Analytics Conference*, pp. 37–48, Sep. 2022. [Online]. Available: https://ecp.ep.liu.se/index.php/linhac/article/view/478

[4] B. Balaji, J. Bright, H. Prakash, Y. Chen, D. A. Clausi, and J. Zelek, "Jersey Number Recognition using Keyframe Identification from Low-Resolution Broadcast Videos," in *Proceedings of the 6th International Workshop on Multimedia Content Analysis in Sports*, ser. MMSports '23. New York, NY, USA: Association for Computing Machinery, Oct. 2023, pp. 123–130. [Online]. Available: https://dl.acm.org/doi/10.1145/3606038.3616162

[5] K. Vats, M. Fani, D. A. Clausi, and J. S. Zelek, "Evaluating deep tracking models for player tracking in broadcast ice hockey video," May 2022, arXiv:2205.10949 [cs]. [Online]. Available: http://arxiv.org/abs/2205.10949

[6] K. Vats, M. Fani, P. Walters, D. A. Clausi, and J. Zelek, "Event Detection in Coarsely Annotated Sports Videos via Parallel Multi-Receptive Field 1D Convolutions," 2020, pp. 882–883. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w53/Vats_Event_Detection_in_Coarsely_Annotated_Sports_Videos_via_Parallel_Multi-Receptive_CVPRW_2020_paper.html

[7] ——, "Event Detection in Coarsely Annotated Sports Videos via Parallel Multi-Receptive Field 1D Convolutions," 2020,

pp. 882–883. [Online]. Available: https://openaccess.thecvf.com/content_CVPRW_2020/html/w53/Vats_Event_Detection_in_Coarsely_Annotated_Sports_Videos_via_Parallel_Multi-Receptive_CVPRW_2020_paper.html

[8] K. Vats, W. McNally, P. Walters, D. A. Clausi, and J. S. Zelek, "Ice Hockey Player Identification via Transformers and Weakly Supervised Learning," 2022, pp. 3451–3460. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2022W/CVSports/html/Vats_Ice_Hockey_Player_Identification_via_Transformers_and_Weakly_Supervised_Learning_CVPRW_2022_paper.html

[9] K. Vats, H. Neher, A. Wong, D. A. Clausi, and J. Zelek, "KPTransfer: Improved Performance and Faster Convergence from Keypoint Subset-Wise Domain Transfer in Human Pose Estimation," in *Image Analysis and Recognition*, F. Karray, A. Campilho, and A. Yu, Eds. Cham: Springer International Publishing, 2019, pp. 220–231.

[10] K. Vats, M. Fani, D. A. Clausi, and J. Zelek, "Multi-task Learning for Jersey Number Recognition in Ice Hockey," in *Proceedings of the 4th International Workshop on Multimedia Content Analysis in Sports*, ser. MMSports'21. New York, NY, USA: Association for Computing Machinery, Oct. 2021, pp. 11–15. [Online]. Available: https://dl.acm.org/doi/10.1145/3475722.3482794

[11] ——, "Puck Localization and Multi-Task Event Recognition in Broadcast Hockey Videos," 2021, pp. 4567–4575. [Online]. Available: https://openaccess.thecvf.com/content/CVPR2021W/CVSports/html/Vats_Puck_Localization_and_Multi-Task_Event_Recognition_in_Broadcast_Hockey_Videos_CVPRW_2021_paper.html

[12] K. Vats, W. McNally, C. Dulhanty, Z. Q. Lin, D. A. Clausi, and J. Zelek, "PuckNet: Estimating hockey puck location from broadcast video," Mar. 2021, arXiv:1912.05107 [cs, eess]. [Online]. Available: http://arxiv.org/abs/1912.05107

[13] K. Vats, H. Neher, D. A. Clausi, and J. Zelek, "Two-Stream Action Recognition in Ice Hockey using Player Pose Sequences and Optical Flows," in *2019 16th Conference on Computer and Robot Vision (CRV)*, May 2019, pp. 181–188. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8781602

[14] R. Found, "Goal-based Metrics Better Than Shot-based Metrics at Predicting Hockey Success," Jan. 2016.

[15] B. Giovanini, D. Conte, A. Ferreira-Junior, and V. B. Nascimento, "Assessing the key game-related statistics in Brazilian professional basketball according to season phase and final score difference," *International Journal of Performance Analysis in Sport*, vol. 21, no. 2, pp. 295–305, Mar. 2021, publisher: Routledge _eprint: https://doi.org/10.1080/24748668.2021.1881358. [Online]. Available: https://doi.org/10.1080/24748668.2021.1881358

[16] V. Sarlis and C. Tjortjis, "Sports analytics — Evaluation of basketball players and team performance," *Information Systems*, vol. 93, p. 101562, Nov. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306437920300557

[17] W. Zhou, G. Yu, S. You, and Z. Wang, "An Improved Passing Network for Evaluating Football Team Performance," *Applied Sciences*, vol. 13, no. 2, p. 845, Jan. 2023, number: 2 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/13/2/845

[18] D. Brinkjans, D. Memmert, Y. Paul, and J. Perl, "Sciendo," *International Journal of Computer Science in Sport*, vol. 22, no. 2, pp. 77–97, Oct. 2023. [Online]. Available: https://sciendo.com/article/10.2478/ijcss-2023-0013

[19] Y. Li, L. Wang, and F. Li, "A data-driven prediction approach for sports team performance and its application to National Basketball Association," *Omega*, vol. 98, p. 102123, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305048319302002

[20] P. Piepiora, "Assessment of Personality Traits Influencing the Performance of Men in Team Sports in Terms of the Big Five," *Frontiers in Psychology*, vol. 12, May 2021, publisher: Frontiers. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.679724/full

[21] K. Phillips, S. Jowett, A. Krukowska-Burke, and D. J. A. Rhind, "The quality of the coach-athlete relationship predicts objective performance in elite cricket," *International Journal of Sport Psychology*, no. 54, pp. 32–47, 2023. [Online]. Available: https://doi.org/10.7352/IJSP.2023.54.032

[22] Y. Sano and Y. Nakada, "Visualization for Potential Pass Courses and Quantification for Offensive and Defensive Players in Basketball," in *2021 International Conference on Engineering and*

Emerging Technologies (ICEET), Oct. 2021, pp. 1–6, iSSN: 2409-2983. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9659701

[23] A. Martín-Castellanos, M. R. Flores, D. M. Solana, R. L. Del Campo, F. N. Garrosa, and D. Mon-López, "How do the football teams play in LaLiga? Analysis and comparison of playing styles according to the outcome," *International Journal of Performance Analysis in Sport*, vol. 24, no. 1, pp. 18–30, Jan. 2024, publisher: Routledge _eprint: https://doi.org/10.1080/24748668.2023.2262813. [Online]. Available: https://doi.org/10.1080/24748668.2023.2262813

[24] H. Prakash, J. C. Shang, K. M. Nsiempba, Y. Chen, D. A. Clausi, and J. S. Zelek, "Multi Player Tracking in Ice Hockey with Homographic Projections," May 2024, arXiv:2405.13397. [Online]. Available: http://arxiv.org/abs/2405.13397

[25] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 785–794. [Online]. Available: https://dl.acm.org/doi/10.1145/2939672.2939785