Extended xThreat: an explainable quality assessment method for actions in football using game context

van Arem, K.W.; Bruinsma, Mirjam

DOI
10.17028/RD.LBORO.27045427.V1

Publication date
2024

Document Version
Final published version

Citation (APA)
van Arem, K. W., & Bruinsma, M. (2024). *Extended xThreat: an explainable quality assessment method for actions in football using game context*. 135-136. Abstract from 15th International Conference on the Engineering of Sport, Loughborough, United Kingdom. https://doi.org/10.17028/RD.LBORO.27045427.V1

Extended xThreat: an explainable quality assessment method for actions in football using game context

Extended xThreat: an explainable quality assessment method for actions in football using game context

Koen W. van Arem[1] and Mirjam Bruinsma[2]
[1] Delft University of Technology, [2]AFC Ajax

Introduction

In the last decade, systematic collection of data is increasingly being used in the world of football which enables the use of mathematics to improve the performance of individual players and teams [1]. Mathematical models allow for quality assessment for in-game actions by estimating the probability that a goal is scored [2]. These models must be explainable, as they should be reducible to a few clear instructions that a coach can give to a player. Therefore, there is a need for explainable quantification methods for actions in football.

Most models for action quality assessment can be described using either the Valuing Actions by Estimating Probabilities (VAEP) [2] or Expected Threat (xThreat) frameworks [3]. VAEP models typically use machine learning methods [2,4,5], which comes at the expense of explainability. On the other hand, the xThreat framework is explainable but it uses only the position of the player in possession of the ball as input [3]. This means football data analysts using state-of-the-art techniques face a trade-off between an explainable but simplified xThreat model or a VAEP model that takes contextual variables into account. The main goal of this paper is to create an extended xThreat model that can include game context, thus aiming for explainability while taking into account contextual variables.

Methods

The data used for training the model consisted of Ortec event data and Tracab tracking data of three Dutch Eredivisie seasons (’20 to ’23). Two variables were constructed using tracking data to provide game context. The variable *high_ball* is an indicator for the ball being higher than 90 cm at the moment of play and the variable *n_def* is defined as the number of defenders between the player and goal, as shown in Figure 1.

The probability of scoring obtained using the Markov chain can be formulated as

$$xT(s) = P(shot|s) \cdot xG(s) + (1 - P(shot|s)) \cdot \sum_{s'} P(s' \mid s) \cdot xT(s'), \quad (1)$$

where s and s' are possible states of the game [3]. $xT(s)$ and $xG(s)$, are the probability of scoring from state s (xThreat) and the probability of scoring given that a shot is performed, called expected goals (xG), respectively. The xG values were obtained by averaging Ortec xG values for all shots performed at a certain state and the probability of a shot, $P(shot|s)$, was estimated by the empirical frequency.

Instead of defining a Markov state s only by the position of ball possession, the extended xThreat also defines it using the variables *high_ball* and *n_def*. This results in a model that is prone to overfitting. To prevent this, kernel density estimation is applied to obtain a continuous estimation for the transition probabilities $P(s' \mid s)$ and these were

Corresponding author email: koen.vanarem@gmail.com

transformed to discrete probabilities by applying numerical integration. An iterative method was then applied to solve eq.(1) and obtain the values of xT.

Results

The extended xThreat model provides the probabilities of the play ending in a goal for different situations. Figure 2 shows these xThreat values (xT) for different values of *high_ball* and *n_def*. The figure highlights that the probability of scoring is higher with fewer defenders, whereas a high ball decreases the probability of scoring a goal. This means that differences in game context cause differences in the xThreat values demonstrating the need for game context variables.

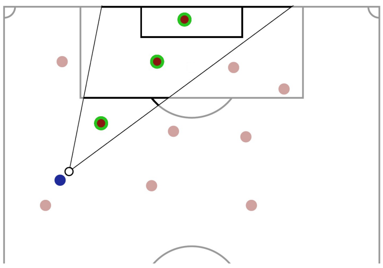

Figure 1: An example showing how the variable *n_def* is defined. The highlighted opponents are considered active defenders.

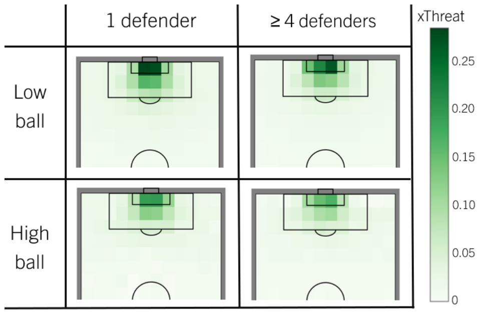

Figure 2: xThreat values on the opponent half the field given different values of *high_ball* and *n_def*.

Conclusion

This study was aimed at creating an extended xThreat model that includes game context. Different values in game context variables lead to a clear distinction in xThreat values. It can be concluded that the extended xThreat has incorporated game context variables allowing a more realistic prediction of the game of football and better instructions of the players.

References

1. Rein R, Memmert D (2016) Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. SpringerPlus 5:1410. https://doi.org/10.1186/s40064-016-3108-2
2. Decroos T, Bransen L, Van Haaren J, David J (2019) Actions speak louder than goals: Valuing player actions in soccer. Proc 25th ACM SIGKDD Int Conf Knowl Discov & Data Min. https://doi.org/10.1145/3292500.3330758
3. Van Roy M, Robberechts P, Decroos T, Davis J (2020) Valuing on-the-ball actions in soccer: a critical comparison of XT and VAEP. Proc. AAAI-20 Workshop Artif Intell Team Sports
4. Fernández J, Bornn L, Cervone D (2019) Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. Proc 13th MIT Sloan Sports Anal Conf
5. Stöckl M, Seidl T, Marley D, Power P (2021) Making offensive play predictable-using a graph convolutional network to understand defensive performance in soccer. Proc 15th MIT Sloan Sports Anal Conf

Corresponding author email: koen.vanarem@gmail.com