Improving the Evaluation of Defensive Player Values with Advanced Machine Learning Techniques

Michał Zaręba
Faculty of Mathematics and Computer Science at Adam Mickiewicz University
Poznań, Poland *michal.zareba@amu.edu.pl*

Tomasz Piłka
Faculty of Mathematics and Computer Science at Adam Mickiewicz University
Poznań, Poland *tomasz.pilka@amu.edu.pl*

Tomasz Górecki
Faculty of Mathematics and Computer Science at Adam Mickiewicz University
Poznań, Poland *tomasz.gorecki@amu.edu.pl*

Bartłomiej Grzelak
Faculty of Mathematics and Computer Science at Adam Mickiewicz University
Poznań, Poland *bartlomiej.grzelak@lechpoznan.pl*

Krzyszof Dyczkowski
Faculty of Mathematics and Computer Science at Adam Mickiewicz University
Poznań, Poland *krzysztof.dyczkowski@amu.edu.pl*

Abstract

Quantifying defensive actions, which offensive indicators have historically overshadowed, is challenging in football analysis. This study presents a novel approach using XGBoost and neural networks to evaluate defensive play using On-Ball Value (OBV), Valuing Actions by Estimating Probabilities (VAEP), and eXpected Threat (xT) indicators. The proposed evaluation of Defensive Player Value using machine learning techniques is presented. A comparative assessment of expert ratings and market values in a Polish PKO BP Ekstraklasa case study highlights the method's effectiveness. The research contributes to the development of sports analytics by addressing the long-term challenge of evaluating the defensive play of football players.

Keywords: football, player evaluation, machine learning, deep learning, XGBoost

1. Introduction

In football analytics, the main focus has been on player performance metrics, specifically expected goal (xG) and expected threat (xT) models. The xG metric aims to quantify shot quality, and numerous studies and data companies continue to refine methods for calculating shot quality [6]. In contrast, the xT model, introduced by Singh [14], uses a Markov model to assess the dynamics of ball possession, providing insight into how individual actions on the pitch contribute to creating goal-scoring opportunities. However, these models often focus on personal actions and overlook the interconnected nature of the events that lead to those actions. Some researchers advocate a more holistic approach, analyzing play sequences to understand the game's dynamics better [13, 16]. This approach highlights the intricacies of football, where actions that directly lead to goals or assists are only a fraction of the total. It emphasizes the importance of player creativity and strategic decision-making. Recent methodologies have broadened this scope of analysis to include aspects such as player creativity [12], team performance evaluation, and pattern recognition in games [5].

Significant AI-driven developments have assessed player quality and on-field activity in recent years [17]. In particular, metrics such as OBV [15], VAEP [3], and xT have emerged to assess different facets of football play [11, 14]. However, existing metrics do not adequately evaluate defensive actions. Therefore, our work focuses on improving defensive assessment by integrating established methods and deep learning techniques. This research underscores the ongoing quest to refine football analytics to ensure a holistic evaluation of player performance, including defensive skills.

2. Dataset and methods

2.1. Dataset

The dataset utilized for this research was sourced from StatsBomb[1], covering the 2021/2022 and 2022/2023 PKO BP Ekstraklasa league seasons (612 games). It includes detailed event data (Table 1), such as team and player information, possession chains, individual player actions, and event locations, providing a holistic view of the game's dynamics. This event data logs every pass, shot, tackle, and dribble, including the time and location, and is crucial for deriving metrics like xT, VAEP, or OBV. To compute xT values, we defined a set of moving actions and refined our dataset to include these specified actions. Adopting the sequence definition from the DaxT article [7], a sequence consists of two consecutive successful events followed by an unsuccessful third event, succeeded by a defensive action from the opposing team.

Table 1. Information for different seasons used in experiments

	2021/2022	**2022/2023**
Events	579,229	572,942
Moving actions	330,904	339,630
Interceptions	6,177	4,278
Tackles	2,696	2,471

2.2. Football metrics

The xT [14] metric assesses a player's likelihood to score or create a chance based on position, action type, and ball location. VAEP [3] measures the value a player adds by assigning probabilities to actions leading to a goal, considering the game's context. OBV [15] evaluates the impact of a player's actions on scoring or preventing goals, focusing on the initiator. OBV's calculation details are proprietary and limited to StatsBomb data.

2.3. Defensive Player Value (DPV)

To synthesize the individual metrics gathered for interceptions and tackles into a comprehensive assessment of player performance, we employed the aggregation methods to derive a final ranking score for each player. Our method draws inspiration from the established formula presented in the *DaxT* [7] article. This approach splits metrics into defensive and offensive categories, then strategically weighted to reflect their respective contributions to player performance. The weighting scheme is carefully calibrated to ensure that each metric is proportionately represented in the final score, thus providing a balanced evaluation of a player's on-field impact. The formula applied is as follows:

$$\text{DPV} = \frac{1}{2}\left(\frac{I_{\text{xT}} + T_{\text{xT}} + C_{\text{xT}} + I_{\text{VAEP}} + T_{\text{VAEP}} + C_{\text{VAEP}} + I_{\text{OBV}} + T_{\text{OBV}} + C_{\text{OBV}}}{9} + \frac{P_{\text{xT}} + P_{\text{VAEP}} + P_{\text{OBV}}}{3}\right),$$

[1] `https://statsbomb.com/`

where I_{xT} represents the value of xT that the interception event prevented, while T_{xT} corresponds to the xT value that the tackle event prevented. The variables I_{VAEP} and T_{VAEP} denote the VAEP values stopped by interception and tackle events, respectively. I_{OBV} and T_{OBV} represent the OBV values halted by the respective defensive actions. C_{xT} signifies the cumulative xT value from clearance events, and P_{xT} signifies the Pass xT, similarly for VAEP and OBV.

2.4. Baselines

To fairly compare our solution, we used three baseline metrics, which are essential for providing reference points and enabling the quantitative assessment of model performance. Firstly, we utilized an expert system involving 24 experienced scouts/observers who evaluated 45 players from the PKO BP Ekstraklasa league based on their performance during the 2022/2023 season. These experts rated players on a scale of 1 to 10, considering aspects not readily apparent from existing statistical data. Secondly, player market valuation from Transfermarkt[2] was employed as a criterion. Lastly, we used SofaScore[3] ratings.

2.5. Selection of the regression method

To predict the expected threat of football events, we used deep learning (DNN) and XGBoost [1]. The DNN, implemented with `Keras` [2], featured multiple layers, Leaky ReLU activations, and dropout for overfitting prevention. Hyperparameters were optimized using `KerasTuner` [9]. The model, trained with root mean squared error (RMSE) loss and Adam optimizer, underwent 20 epochs with early stopping. The XGBoost model minimized RMSE for predicting xT, VAEP, and OBV. Hyperparameters were optimized via randomized search and 4-fold cross-validation.

3. Results

After comparing the performance metrics of NN and XGBoost models, XGBoost was chosen as the final prediction. Empirical evidence showed that XGBoost outperformed NN in validation loss RMSE across multiple models, as illustrated in Table 2. XGBoost also demonstrated faster training, hyperparameter tuning, and inference speeds [4]. Its efficiency and predictive performance make it ideal for applications requiring frequent model training and re-tuning.

Table 2. Comparison of RMSE for xT, VAEP, and OBV models

Model	DNN	XGBoost
OBV	0.03671	**0.03620**
VAEP	0.03789	**0.03112**
xT	0.00344	**0.00343**

Table 3 shows the results for the top 10 players in the league according to the proposed DPV metric. The results are compared with the expert ratings discussed above, the SofaScore ranking, and the player valuation. The table also includes Information on the player's age and position (CD – Central defender, DM – Defensive midfielder, FB – Full back, AM – Attacking midfielder, WB – Winger), the team he played for, and, in brackets, the position in which the team finished the season. Players with a score in the top 5 for each metric are shown in green. In Figure 1, we present the relationship between DPV and SofaScore rating. We labeled the top 5% of players according to both ratings and restricted our analysis to players with more than 900 minutes played. We can observe that SofaScore tends to favor offensive players, whereas

[2] `https://www.transfermarkt.pl/`
[3] `https://www.sofascore.com/`

DPV accurately identifies the best defensive players.

Table 3. Metrics for top 10 players from the 2022/23 season (sorted in decreasing order of DPV)

No.	Player name (age, position)	Club name (league position)	DPV metric	Expert assessment	SofaScore rating	Player valuation (mln €)
1.	Damian Dąbrowski (31, DM)	Pogoń Szczecin (4)	0.744	7.100 (± 0.85)	7.38	0.90
2.	Joel Pereira (27, FB)	Lech Poznań (3)	0.723	6.765 (± 1.25)	7.34	2.00
3.	Bartosz Slisz (24, DM)	Legia Warszawa (2)	0.696	8.091 (± 0.92)	7.04	1.50
4.	Virgil Eugen Ghiță (25, CB)	Cracovia Kraków (7)	0.692	6.600 (± 1.90)	7.04	1.20
5.	Łukasz Łakomy (22, DM)	Zagłębie Lubin (9)	0.683	6.692 (± 0.95)	7.15	1.50
6.	Fran Tudor (28, FB)	Raków Częstochowa (1)	0.680	7.737 (± 1.10)	6.95	3.00
7.	Robert Ivanov (29, CB)	Warta Poznań (8)	0.680	6.889 (± 1.83)	7.03	0.75
8.	Erik Janža (30, FB)	Górnik Zabrze (6)	0.665	7.333 (± 1.16)	7.32	1.00
9.	Benedikt Zech (33, CB)	Pogoń Szczecin (4)	0.661	7.308 (± 1.11)	6.96	0.30
10.	Krystian Getinger (35, FB)	Stal Mielec (11)	0.653	6.125 (± 1.13)	7.03	0.25

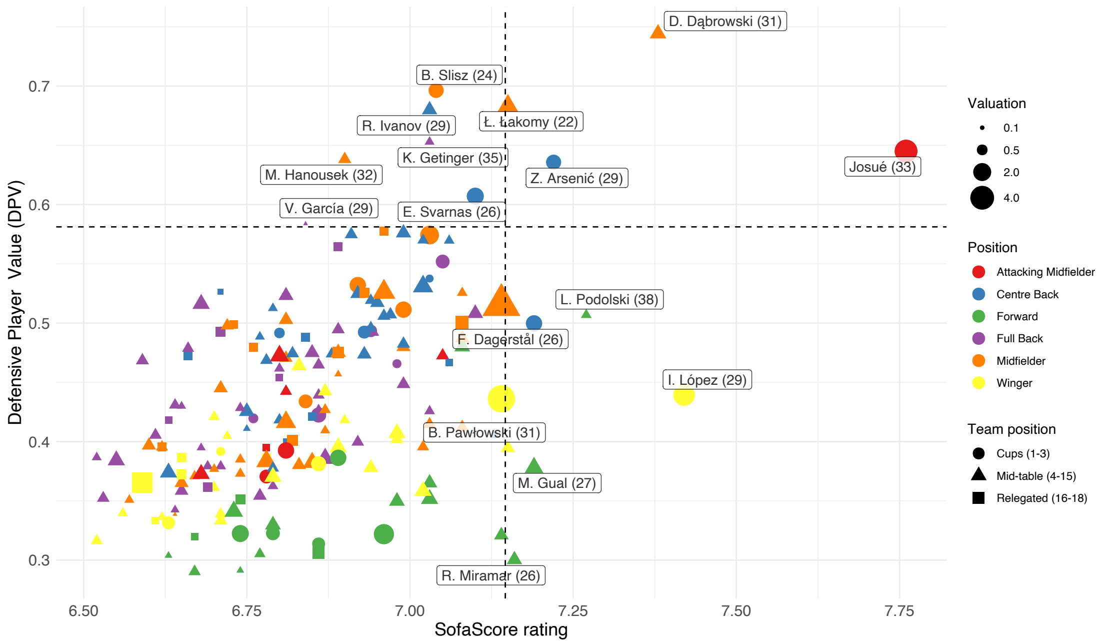

Fig. 1. Relationship between SofaScore Ratings and DPV

4. Discussion and conclusion

Discussions with experts from Lech Poznań revealed that evaluations often focus on central defenders while overlooking full-backs, who frequently take on more offensive roles not accounted for in traditional defensive metrics. Our metric recognizes defenders involved in attacking plays or ball distribution due to the appropriate weighting of metrics like OBV and VAEP. Scouts benefit from identifying promising players, providing a nuanced assessment that includes defensive prowess and offensive contributions. This holistic view is crucial for identifying versatile talents that may otherwise be underrated.

This work shows that we can use known metrics (VAEP, OBV, xT) to evaluate players by adopting selected aggregation models. The evaluation employs novel models, including deep learning, to predict the outcomes that defensive actions interrupt, thereby assessing defensive actions. The results enable the assessment of players over a season or part of a season regarding their effectiveness in intercepting the ball and halting the opposing team's attacks. High effectiveness in intercepting and halting attacks contributes significantly to overall defensive organization by reducing the opponent's scoring opportunities and providing stability to the defensive line.

A limitation of our study is the need to develop models separately for each league, considering different play styles. Our solution can include new values as new metrics emerge, such as game context, time remaining, current score, and specific pitch sectors where events occur.

Acknowledgements

KKS Lech Poznań's support and dataset access made this research and publication possible.

References

1. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the KDD (2016)
2. Chollet, F. et al.: Keras. `https://keras.io` (2015)
3. Decroos, T., Bransen, L., Van Haaren, J., Davis, J.: Actions speak louder than goals: valuing player actions in soccer. In: Proceedings of the KDD (2019)
4. Grinsztajn, L., Oyallon, E., Varoquaux, G.: Why do tree-based models still outperform deep learning on typical tabular data? In: Proceedings of the NeurIPS (2022)
5. Grunz, A., Memmert, D., Perl, J.: Tactical pattern recognition in soccer games by means of special self-organizing maps. Human Movement Science 31(2), 334-343 (2012)
6. Lucey, P., Oliver, D., Carr, P., Roth, J., Matthews, I.: Assessing team strategy using spatio-temporal data. In: Proceedings of the KDD (2013)
7. Merhej, C., Beal, R.J., Matthews, T., Ramchurn, S.: What happened next? Using deep learning to value defensive actions in football event data. In: Proceedings of the KDD (2021)
8. do Nascimento, R.F.M., Rios-Neto, H.: Generalized action-based ball recovery model using 360° data. arXiv:2307.04215v1. (2022)
9. O'Malley, T., Bursztein, E., Long, J., Chollet, F., Jin, H., Invernizzi, L., et al.: Keras-tuner. `https://github.com/keras-team/keras-tuner` (2019)
10. Piłka, T., Grzelak, B., Sadurska, A., Górecki, T., Dyczkowski, K.: Predicting injuries in football based on data collected from GPS-based wearable sensors. Sensors 23(3), 1227 (2023)
11. Rathke, A.: An examination of expected goals and shot efficiency in soccer. Journal of Human Sport and Exercise 12(2), 514–529 (2017)
12. Robberechts, P., Van Roy, M., Davis, J.: Un-xPass: Measuring soccer player's creativity. In: Proceedings of the KDD (2023)
13. Van Roy, M., Robberechts, P., Decroos, T., Davis, J.: Valuing on-the-ball actions in soccer: A critical comparison of xT and VAEP. In: Proceedings of the AAAI (2020)
14. Singh, K.: Introducing expected threat (xT). `https://karun.in/blog/expected-threat.html` (2019)
15. StatsBomb: Introducing on-ball value (OBV). `https://statsbomb.com/articles/soccer/introducing-on-ball-value-obv/` (2021)
16. Tylka, M., Wałęsa, S., Girejko, K., Kaczmarek, J., Grzelak, B., Piłka, T.: Valuing passes in actions leading to the third zone on the pitch with machine learning methods. In: Progress in Polish Artificial Intelligence Research 4, 507–512 (2023)
17. Vatavani, D.: Upgrading expected goals. `https://statsbomb.com/articles/soccer/upgrading-expected-goals/` (2022)