EXPECTED PRESSING SUCCESS (xPS)

Abstract

Pressing in football has evolved significantly from the unstructured 'Total Pressing' of the 1970s Dutch teams to the more systematic 'Gegenpressing' of the modern era. Balancing aggressive pressing with defensive restraint presents a critical challenge for today's coaches. Our research seeks to identify the key indicators of successful pressing and explore how visual maps and predictive models can enhance the execution of pressing strategies. We will investigate situational and spatio-temporal factors such as available passing channels, distance from the goal, player movement, and pressing intensity to evaluate the effectiveness of pressing actions. The core objective of our study is to address the optimal balance between pressing and defensive posture based on an opponent's style of play. We created a classification model to identify the expected pressing success of a pressing event in terms of forcing a pass backward and changing possession. Based on this model prediction values we will apply prescriptive analytics to create a spatiotemporal time series model to output a value between [0,1] indicating (0, no press) and (1, press) at every interval of the game based on the risk and return value of the situation. We achieved an ROC AUC of 0.702 for our best model using the Gradient Boosting Classifier on a dataset of more than 30,000 records of pressured passes from Statsbomb's Open 360 data. We then used these predictions to apply prescriptive analytics to the Euros 2024 tournament.

Keywords: Football, Pressure, Defence Metrics, Machine Learning.

1. Introduction

Football is a global sport, and FIFA has reported that approximately 1.5 billion people tuned in to watch the 2022 World Cup final live on television. This figure highlights the sport's immense popularity, despite the tournament featuring only 32 teams. The field of football analytics has seen significant advancements, particularly with the rise of big data and event data, leading to a transformation in the way we analyze the game. Innovations such as the expected goals metric and other player-focused, event-oriented metrics have emerged. However, one area that has not received as much attention is understanding the factors and situations that define an effective press.

Pressing is a coordinated defensive strategy where the team without the ball works collectively to pressure the ball carrier, aiming to force mistakes or regain possession. When executed effectively, pressing can lead to the opposing team completing fewer passes and struggling to advance the ball.

Pressing, as we know it, began in 1934 with hockey coach Thomas Patrick Gorman. He told his players to challenge opponents near their own goal and block passing lanes, which led to initial losses but eventually helped them dominate the National Hockey League. In the 1970s, Dutch football teams like Ajax adopted a similar approach, which we today know as “Total Football”. The Dutch Manager, Michels said “For me, it would be better to call my game ‘pressing football’. This is what we wanted to create with our Ajax side and the Dutch national team in 1974: create a basic game where all 10 outfield players push forward even when we don't have the ball. We're always pressing forward". This strategy involves all players pressing forward even without the ball. Today, terms like high block, low block, and Gegenpressing describe different ways teams use pressing to quickly regain possession of the ball.

Code and data are available at: `https://github.com/jainil-kakka/xPS/tree/main`

2. Systematic Literature Review

In this research, we performed a thorough literature review to understand the gaps and what metrics could be developed to calculate pressure and its effectiveness in a match phase. To conduct this literature review we planned and conducted a review which is discussed in detail below:

2.1. Planning the Review:

We first planned the research objective and questions we needed to solve to best help the use of event data in football and derive a defensive metric that could help quantify organized pressing. To fulfill the objective, we formulated three main research questions.

RQ1: What are the key spatiotemporal indicators from match phases that can predict the likelihood of pressing success in football?

RQ2: How can the predicted probabilities of pressing success be used to evaluate and enhance team strategies in terms of pressing organization and effectiveness?

RQ3: In what ways can the predictive probabilities of pressing success inform player recruitment and development, particularly in identifying players with strong anti-press abilities?

2.2. Search Strategy:

We used the PennState LionSearch tool as the search space to look for manuscripts. LionSearch is an integrated search engine of books, e-books, research articles, newspaper articles, and other publications integrated from over 950 database/search engines, including over 80 databases/search engines for healthcare/medicine discipline and 15 for computer/software/information science and engineering. The tool is provided and maintained by the Pennsylvania State University's Library.

2.3. Search Criteria:

For our systematic literature review, we established the following search criteria:

Condition 1 (C1): Involved research on "soccer/football pressing machine learning," focusing on how machine learning is applied to analyze and enhance pressing strategies in soccer or football.

Condition 2 (C2): Covered studies related to "soccer/football pressing spatiotemporal features," which examined the spatial and temporal dynamics of pressing tactics.

Condition 3 (C3): Required the intersection of both C1 and C2, aiming to capture literature that integrates machine learning approaches with spatiotemporal analysis of pressing strategies.

This approach ensured a thorough review of relevant research that combines advanced analytical techniques with in-depth tactical insights.

2.4. Inclusion and Exclusion Criteria:

In conducting our systematic literature review, we applied specific inclusion and exclusion criteria to ensure the relevance and quality of the selected studies. We included only peer-reviewed articles that were available in full text and published in English to maintain consistency and reliability. The studies needed to align with our search terms concerning "soccer/football pressing machine learning" and "soccer/football pressing spatiotemporal features." We considered a range of publication types, including journal articles, books, conference proceedings, theses, and dissertations, if they were original research or comprehensive reviews. Conversely, we excluded studies that were not related to soccer or football, or those focusing on analytics for match events or phases that did not directly address pressing strategies. Additionally, non-research content such as editorials, comments, presentation slides, and prefaces were not considered. Any study that failed to meet all these inclusion criteria was also excluded to ensure the review's focus and relevance.

2.5. Conducting the Review:

Several key insights and impacts have emerged based on the review of the 25 papers in the domain, which includes 18 journal articles, 5 conference proceedings, and 1 editorial and book. The data indicates a strong focus on journal articles, reflecting a preference for peer-reviewed and rigorous research. The mean publication year of 2019 suggests that the field is relatively current, with a steady progression in the quantity and relevance of research. The citations range widely, with a mean of 61.12 and a maximum of 568, highlighting significant variability in the impact of individual papers.

Figure 1: The Trend of Paper Types Over Years with Number of Papers

This variability underscores the presence of highly influential studies that have shaped the field, as well as a breadth of research that contributes to a comprehensive understanding of the domain. The skew towards recent journal articles emphasizes the ongoing evolution and growing interest in the subject, suggesting that recent advancements are driving the field forward and shaping current research trends.

3. Related Work

In recent years, machine learning (ML) techniques have increasingly been applied to football analytics, primarily focusing on predicting match outcomes and player performance. These applications range from predicting match results to forecasting injuries and analyzing tactical play. A significant body of work has explored the utility of ML in football, such as the contributions of Berrar et al. (2019) and Hervert-Escobar et al. (2018), who utilized machine learning algorithms and Bayesian approaches respectively to predict match outcomes. However, Stübinger and Knoll (2018) demonstrated the limitations of these methods in consistently outperforming bookmakers, emphasizing the ongoing challenges in predictive accuracy.

Transitioning from match predictions, recent studies have turned their attention to tactical aspects of the game, particularly attacking metrics. For example, Pereira et al. (2019) introduced the Golden Index, a classification system designed to assess player performance during attacking plays, while Kim (2020) developed a taxonomy for classifying the attacking process and creating goal-scoring opportunities. These studies highlight the growing interest in understanding and quantifying attacking strategies, which are crucial for optimizing team performance.

Despite these advancements, defensive metrics, particularly those related to pressing and disruption, remain underexplored. Pressing, a critical aspect of defensive strategy, has been examined through various lenses. For instance, Andrienko et al. (2017) created visualizations to analyze defensive pressure, providing insights into how teams apply pressure during matches. Bojinov and Bornn (2016) further explored this by developing a model for optimal defensive disruption, highlighting how spatial maps can reveal defensive weaknesses.

Addressing this gap, our paper introduces a novel approach by developing an expected pressing success model. This model aims to quantify and understand the effectiveness of pressing strategies, which have traditionally been overlooked in football analytics. By focusing on pressing organizations, we seek to enhance the understanding of defensive tactics and contribute to the broader field of football analytics.

4. Methodology

In alignment with the CMSAC conference's emphasis on using openly available data, our study utilizes the comprehensive and freely accessible dataset provided by StatsBomb.

4.1 Data Source:

StatsBomb is renowned for its commitment to advancing football analytics by providing high-quality, openly accessible data. Their dataset includes detailed spatial data captured via 360-degree tracking and comprehensive event data. The spatial data offers a nuanced view of player positions and movements throughout the match, while the event data records specific actions such as passes, shots, and defensive maneuvers.https://statsbomb.com/what-we-do/hub/free-data/

Furthermore, StatsBomb's 360 data provides a significant advantage with its inclusion of teammate and opposition locations for over 3,400 events per match. This advanced tactical event data, known as StatsBomb 360, enriches each event with contextual information, including more than 50 exclusive metrics. This comprehensive spatial context is essential for advanced performance and recruitment analysis, enabling us to build a robust expected pressing success model. By utilizing this cutting-edge data, our study benefits from a detailed and nuanced view of player interactions and tactical dynamics, aligning with the high standards of analysis promoted by the CMSAC conference.

4.2 Data Acquisition and Storage:

We began by fetching the data from StatsBomb's GitHub repository, which provides comprehensive football analytics data in JSON format. The data, which includes both event and spatial information, was downloaded from the repository and subsequently stored in MongoDB. This choice of database was driven by the need to efficiently manage and query large volumes of JSON data. MongoDB's flexibility in handling nested structures made it an ideal choice for storing the complex, hierarchical data provided by StatsBomb.

4.3 Data Engineering:

The data engineering phase involved several critical steps to prepare the data for analysis. First, we focused on extracting events related to passes that occurred under pressure. Specifically, we filtered the dataset to include only those events where the event type was "Pass" and the pass was executed under pressure. This subset of data was then enriched with spatial context by integrating StatsBomb's 360 data. The StatsBomb 360 data, which includes detailed positional information of teammates and opponents for each event, was crucial for understanding the dynamics of each pass under pressure.

We defined the dependent variable for our analysis as a function of pass direction, distinguishing between forward and backward passes. This binary classification allowed us to analyze how different pressing situations influenced the success and direction of passes.

4.4 EDA for one sample match (ENGLAND VS SPAIN EURO FINAL):

For a real-world example, we chose a match such as England vs Spain Euro final, we had a total of 3312 events containing where the top 4 most occurring events Pass, Ball Receipt (when the player receives the ball), Carry (when moving player receives the ball), and Pressure constitute almost 87% of the events. They are distributed evenly between the halves of the match. Here is the heat map of the various events and where they occurred on the pitch.

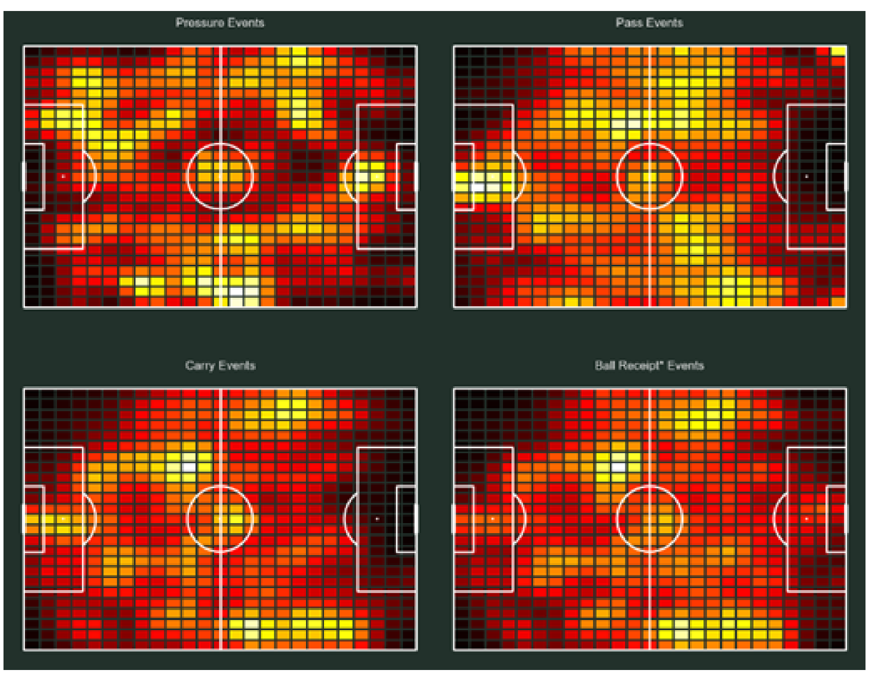

Figure 2: Heatmap plots of Pressure, Pass, Carry, and Ball Receipt Events

The pressure event has a property called duration. All the events that fall under the same time frame have a tag called under_pressure which is set as true indicating while this event occurred the player/actor was under pressure.

We use these under_pressure column data to conduct the analysis and understand the following using the model

- The movement of the ball (forward or backward)
- The outcome of possession (dispossession or retention).

4.5 Feature Extraction:

4.5.1. Passing Channels and Blockage:

The function *count_ coordinates_ comparison_ m* computes the number of passing channels available and blocked. It does this by measuring the distance between the actor and teammates, and between these teammates and opponents. If the distance is below a certain threshold, it counts as a blocked passing channel.

The function *count_ coordinates_ comparison_ d* enhances this analysis by comparing the distances between actors and opponents with the minimum distance to opponents, providing insights into how effectively the pressing situation blocks passing lanes.

4.5.2. Triangular Areas:

The function *triangles_from_ points* generates all possible triangles formed by the actor and pairs of teammates. This provides a basis for understanding the spatial arrangement of players.

The function *traingle_ area* calculates the area of these triangles. By evaluating the mean, minimum, and maximum areas, we gain insights into the spatial configuration around the player, which influences passing decisions.

4.5.3. Coordinate Extraction:

The *extract_ coordinates_ pressure* function parses freeze-frame data to extract the coordinates of actors, teammates, goalkeepers, and opponents. This data is used to understand player positioning and how it affects passing and pressing dynamics.

4.5.4. Distance Metrics:

The functions *distance_ from_ goal* and *distance_ from_ goal_ pressure* compute the distance of the actor from the goal in both the initial and pressured states. The difference between these distances is used to understand how pressure alters the actor's positioning relative to the goal.

4.5.5. Special Pressing Metrics:

The final feature, *pressing_ intensity*, is derived from the change in opponent distance, the vector difference, and the distance from the goal, offering a comprehensive measure of how pressure affects passing effectiveness. This metric helps in understanding the effect of pressing on the actor's ability to execute a pass.

These features collectively enhance our understanding of how pressing impacts pass execution, providing a robust basis for analyzing and modeling the dynamics of passing under pressure in football.

$$\mathrm{PI} = \frac{\|(P_{\text{actor, pressure}} - P_{\text{actor}}) - (P_{\text{opp, pressure}} - P_{\text{opp}})\| \cdot \|(P_{\text{actor}} - P_{\text{opp, initial}})\|}{\|P_{\text{actor}} - \text{goal}\|}$$

$\mathrm{P}_{\text{actor}}$: Position vector of the actor.

$\mathrm{P}_{\text{actor, pressure}}$: Position vector of the actor under pressure.

$\mathrm{P}_{\text{opp, initial}}$: Position vector of the nearest opponent before the pass.

$\mathrm{P}_{\text{opp, final}}$: Position vector of the nearest opponent after the pass.

$\mathrm{P}_{\text{opp, pressure}}$: Position vector of the opponent under pressure.

P_{goal}: Position vector of the goal (typically at [120,40] for a standard pitch)

In this formula:

- Numerator
  - Vector Difference: $P_{\text{actor,pressure-actor}}$ - P_{actor} = It captures the net change in position due to pressure.
  - Change in Difference: $P_{\text{opp, pressure}}$ - P_{opp} = This measures how the distance to the nearest opponent changes due to pressure.
- Denominator
  - Distance from Goal: P_{actor} - goal = Normalizes the metric by the actor's distance from the goal.

These features were integrated into our model to predict the success of pressing strategies. By leveraging the advanced spatial and event data provided by StatsBomb, our approach offers a comprehensive analysis of pressing dynamics in football, aligning with the open-data principles emphasized by CMSAC.

4.6. Feature Engineering

To extract important features from a list of 70 derived and event variables, we used "FeatureWiz". It implements the SULOV (Searching for Uncorrelated List of Variables) algorithm, method ensures you're left with the most relevant, non-redundant features, followed by recursive XGBoost, to repeatedly identify the best features among the selected variables after SULOV.

We started with 45 columns out of 70 by removing the fixed effect variables such as team name, player name, position, and others, after implementing SULOV, we left with 21 columns, followed by recursive XGBoost algorithm, which helped us achieve our final dataset of 15 columns of spatio-temporal data.

4.7. Modelling

Classes are almost balanced in the form of 54.8% backward passes and 45.2% forward passes in terms of pressured passes. Despite expecting more backward passes we see an evenly distributed dependent variable which will aid in getting a better classification model. As mentioned in the sections above we used feature extractions in terms of Area (Polygons and Triangles), Length (player distances), and vector information such as movement of actor (ball carrier) and opponent applying most pressure.

Table 1: Performance metrics of different classification models.

#	Classification Model	ROC AUC	Precision	Recall	Accuracy	Support
1	Random Forest Classifier	0.68	0.62	0.63	0.62	6237
2	Decision Tree Classifier	0.57	0.57	0.57	0.57	7796
3	XGBoost	0.69	0.63	0.63	0.63	6237
4	Support Vector Classifier	0.64	0.59	0.60	0.60	6237
5	Logistic Regression	0.68	0.62	0.63	0.63	6237
6	Gradient Boosting Classifier	0.70	0.65	0.65	0.65	7796
7	K Neighbors Classifier	0.57	0.56	0.56	0.56	6237

The classification model's performance was evaluated on both training and test datasets using Gradient Boosting Classification algorithm as shown in Table 1.

Training Performance: The model achieved an overall accuracy of 72.7% on the training data. The precision for predicting backward passes (class 0) was 0.72, while the recall was 0.65, resulting in an F1-score of 0.68. For forward passes (class 1), the model exhibited a recall of 0.79, and an F1-score of 0.76. These results indicate that the model performs slightly better in identifying forward passes compared to backward passes during training.

Test Performance: The performance on the test dataset showed a decrease in accuracy to 64.7%. For backward passes, the recall dropped to 0.56, resulting in an F1 score of 0.59. In contrast, the recall for forward passes was 0.72, and an F1 score of 0.69. The increase in recall shows the test set compared to the training set suggests that the model correctly predicts the positive forward passes.

Overall, the AUC value of 0.7 is generally considered a moderate performance. It suggests that the model has a decent ability to distinguish between pressing success and failure but is not outstanding. For football

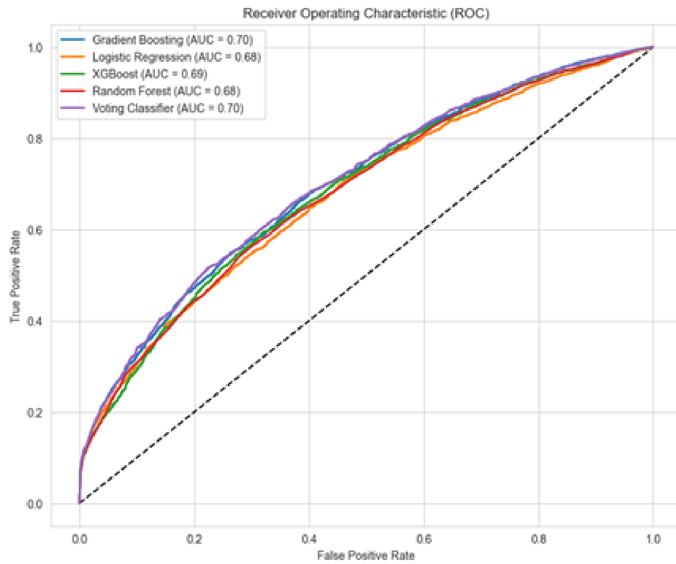

Figure 3: ROC-AUC values of models

analytics, an AUC ROC of 0.7 can be considered quite acceptable, especially given the complexity and variability in spatio-temporal features. Football is inherently noisy and complex, and achieving a higher AUC can be challenging.

5. Discussion

Using the expected success metric defined in the model above we can use it to understand how different teams organize their press and how well they perform in that situation. One such example is presented below where we used the metric to analyze the pressing success of teams when the following teams have possession. To interpret this metric, we can say the higher the xPS the higher the chances of forcing a backward pass.

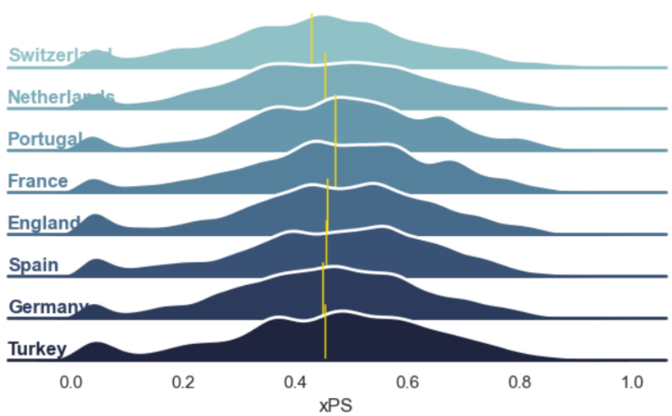

Figure 4: The xPS score of teams

Figure 4 shows how Switzerland, Germany, and Spain were difficult to contain by the opposition as they had the lowest xPS or the lowest chances of forcing a backward pass. Another notable takeaway is Portugal and France who under-performed at the tournament having won the previous versions of this trophy. Both the teams showed high press vulnerability showing that with organization they can be forced to pass backward and make mistakes.

Taking this discussion further we have a scatter plot of teams in the Euros 2024 mentioned in Figure 5 discussing how the teams performed and fared in performing while setting up organized pressing. Austria and Switzerland despite having a good number of pressing actions were not able to take advantage of their pressing

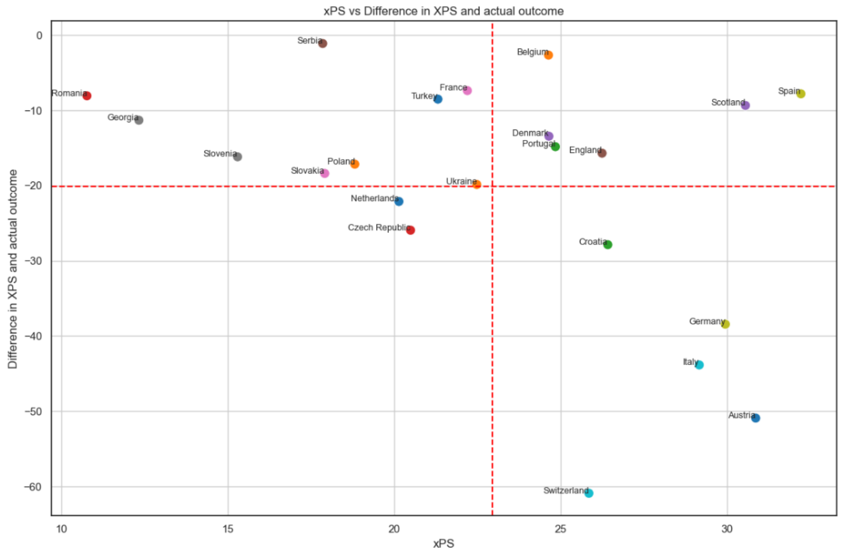

Figure 5: xPS vs Difference of xPS and actual outcome

to force passes backward or regain possession. Whereas teams such as Serbia, Georgia, and Romania despite being good at converting their pressing chances to forcing teams backward did not take enough pressing actions in the first place. The best teams stand out here such that they show that they both completed far more pressing actions as well as converting those to forced passes, Teams such as Spain, England, and to everyone's surprise even Scotland excel in these arts.

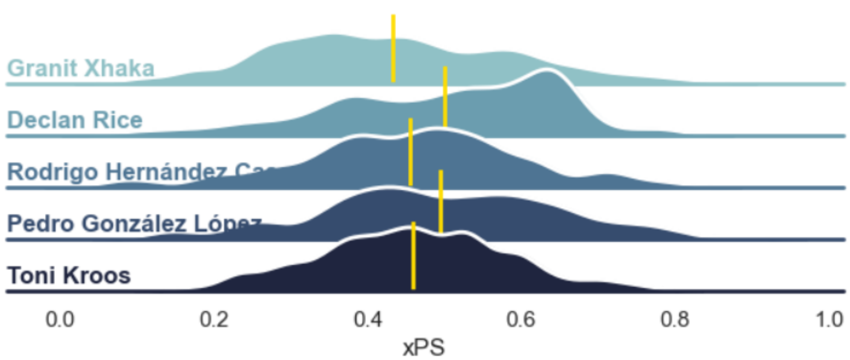

Figure 6: The xPS score of players

Another discussion we can have regarding the use of xPS is the understanding of which player had the least pressing success or the player who cannot be pressed. We found the xPS for the five most frequently pressed players (particularly midfielders) refer Figure 6. The five contained Pedro Gonzalez Lopez (Pedri), Rodrigo Hernandez (Rodri), Declan Rice, Toni Kroos, and Granit Xhaka. Among the five surprisingly Granit Xhaka had the lowest xPS showing that building an organized press against him would not lead to backward passes or even possession spillovers. This is a testament to his latest Ballon D'or nomination and perhaps even the invincible season with Bayer Leverkusen.

We would expect someone like Rodri to be at the top performing charts and the results do not disappoint, showing that the movement, carry, and the art of receiving the ball in tight spaces is something that perhaps every midfielder could learn from him. Toni Kroos too had an amazing season with Real Madrid and was able to perform at a very high level before eventually retiring from the sport. Among the 5 mentioned midfielders Pedri and Rice had the worst xPS showing their vulnerability to the press and why perhaps Dani Olmo should have replaced Pedri in the starting lineup and why England fans criticized Gareth Southgate for non-entertaining football.

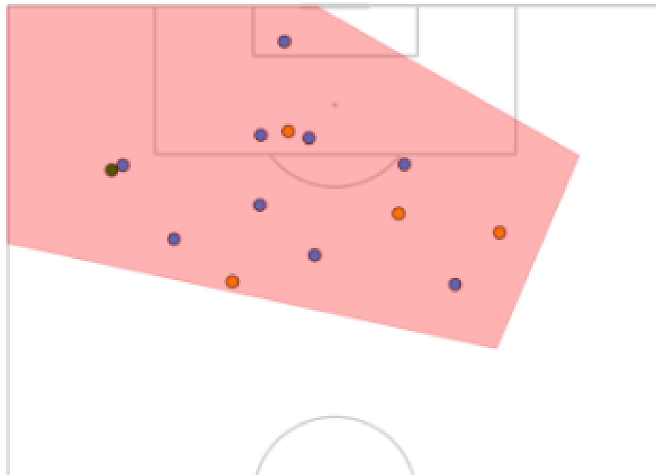

Figure 7: Statsbomb 360 data visualized

Let's talk about a specific match incident which led to Spain getting crowned the Euro Champions that is the Mikel Oyarzabal's goal. Just before the goal Kyle Walker pressed Marc Cucurella outside the box which led to him spraying a low cross deceiving the England center backs and was clinically slotted away by Oyarzabal. The xPS for this event was 0.369 stating that this pressing trigger would 36% lead to a backward pass and 64% a forward pass which eventually happened as the only player in front of Cucurella was Mikel Oyarzabal who was then given a golden opportunity to win his country the trophy. https://www.foxsports.com/watch/fmc-drn5cwb5kar67r4q

References

1. Forcher, Leander, et al. "The keys of pressing to gain the ball–Characteristics of defensive pressure in elite soccer using tracking data." Science and Medicine in Football 8.2 (2024): 161-169.
2. Sotudeh, Hadi. "Potential penetrative pass (P3)." arXiv preprint arXiv:2302.10760 (2023).
3. Herold, Mat, et al. "Machine learning in men's professional football: Current applications and future directions for improving attacking play." International Journal of Sports Science and Coaching 14.6 (2019): 798-817
4. Bauer, Pascal, and Gabriel Anzer. "Data-driven detection of counter-pressing in professional football: a supervised machine learning task based on synchronized positional and event data with expert-based feature extraction." Data Mining and Knowledge Discovery 35.5 (2021): 2009-2049.
5. Berrar, Daniel, et al. "Guest editorial: special issue on machine learning for soccer." Machine Learning 108 (2019): 1-7.
6. Hervert-Escobar, Laura, Neil Hernandez-Gress, and Timothy I. Matis. "Bayesian-based approach learning for outcome prediction of soccer matches." Computational Science–ICCS 2018: 18th International Conference, Wuxi, China, June 11–13, 2018 Proceedings, Part III 18. Springer International Publishing, 2018.
7. Stübinger, Johannes, and Julian Knoll. "Beat the bookmaker–winning football bets with machine learning (best application paper)." Artificial Intelligence XXXV: 38th SGAI International Conference on Artificial Intelligence, AI 2018, Cambridge, UK, December 11–13, 2018, Proceedings 38. Springer International Publishing, 2018.
8. Andrienko, Gennady, et al. "Visual analysis of pressure in football." Data Mining and Knowledge Discovery 31 (2017): 1793-1839.
9. Bojinov, Iavor, and Luke Bornn. "The pressing game: Optimal defensive disruption in soccer." 10th MIT Sloan Sports Analytics Conference. Vol. 11. No. 12.03. 2016.
10. Majumdar, Aritra, et al. "Machine learning for understanding and predicting injuries in football." Sports Medicine-Open 8.1 (2022): 73.

11. Eggels, H., R. van Elk, and M. Pechenizkiy. "Expected goals in soccer: Explaining match results using predictive analytics." The machine learning and data mining for sports analytics workshop. Vol. 16. 2016.
12. Umami, Izzatul, Deden Hardan Gautama, and Heliza Rahmania Hatta. "implementing the Expected Goal (xG) model to predict scores in soccer matches." International Journal of Informatics and Information Systems 4.1 (2021): 38-54.
13. Hewitt, James H., and Oktay Karakuş. "A machine learning approach for player and position adjusted expected goals in football (soccer)." Franklin Open 4 (2023): 100034.
14. Pereira, Tiago, et al. "The Golden Index: A classification system for player performance in football attacking plays." Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology 233.4 (2019): 467-477.
15. Kim, Jongwon. The attacking process in football: a taxonomy for classifying how teams create goal scoring opportunities. Diss. Middlesex University, 2020.
16. Goes, F. R., Matthias Kempe, and Koen APM Lemmink. "Predicting match outcome in professional Dutch football using tactical performance metrics computed from position tracking data." (2019).
17. Low, Benedict, et al. "A systematic review of collective tactical behaviours in football using positional data." Sports Medicine 50 (2020): 343-385.
18. Memmert, Daniel, et al. "A tactical comparison of the 4-2-3-1 and 3-5-2 formation in soccer: A theory-oriented, experimental approach based on positional data in an 11 vs. 11 game set-up." PloS one 14.1 (2019): e0210191.
19. Ric, Angel, et al. "Timescales for exploratory tactical behaviour in football small-sided games." Journal of sports sciences 34.18 (2016): 1723-1730.
20. Stafylidis, Andreas, et al. "Key Performance Indicators Predictive of Success in Soccer: A Comprehensive Analysis of the Greek Soccer League." Journal of Functional Morphology and Kinesiology 9.2 (2024): 107.
21. Ali, Ajmol. "Measuring soccer skill performance: a review." Scandinavian journal of medicine science in sports 21.2 (2011): 170-183.
22. Clemente, Filipe M., et al. "Developing a tactical metric to estimate the defensive area of soccer teams: The defensive play area." Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology 230.2 (2016): 124-132.
23. Forcher, Leander, et al. "The use of player tracking data to analyze defensive play in professional soccer A scoping review." International Journal of Sports Science & Coaching 17.6 (2022): 1567-1592.
24. Wilson, Robbie S., et al. "Predicting the defensive performance of individual players in one vs. one soccer games." Plos one 13.12 (2018): e0209822.
25. Santos, Pedro, Carlos Lago-Peñas, and Oscar García-García. "The influence of situational variables on defensive positioning in professional soccer." International Journal of Performance Analysis in Sport 17.3 (2017): 212-219.
26. Llana, Sergio, et al. "The right place at the right time: Advanced off-ball metrics for exploiting an opponent's spatial weaknesses in soccer." Proceedings of the 14th MIT Sloan Sports Analytics Conference. 2020.
27. Memmert, Daniel, and Dominik Raabe. Data analytics in football: Positional data collection, modelling and analysis. Routledge, 2018.
28. Ram Seshadri, GitHub AutoViML/featurewiz: Use advanced feature engineering strategies and select the best features from your data set fast with a single line of code, 2020, `https://github.com/AutoViML/featurewiz`.