Reinforcement Learning for Football Player Decision Making Analysis

Paper Track

Michael Pulis[1] and Josef Bajada[2]
Department of Artificial Intelligence, University of Malta

Abstract

Traditionally, the performance of football players has been evaluated using statistics computed from actions such as goals and assists. However, recent advances in football analytics have yielded Possession Value Models (PVMs) that provide a more granular and objective method that can be used to analyse the decision making abilities of a player. Such models include Expected Threat (xT), Valuing Actions by Estimating Probabilities (VAEP) and On-the-Ball-Value (OBV). Nevertheless, these metrics typically only make use of the data directly related to an event, such as the position of the player with the ball, and the player who receives the ball. These PVMs do not account for the position of teammates and opponents when computing the value of the proposed action. We propose a novel metric called Decision Value (DV) which is computed using Deep Reinforcement Learning. The model is trained on both event and tracking data to allow the model to obtain an optimised decision policy that takes both the positions of the teammates and opponents into account when computing the DV of a particular action. This model can then be used to assess players by their decision making abilities within the context of the game. It can also be applied to help scouts to find players that make the best decisions in particular areas of the pitch, or the ones that make decisions that align the most with a particular team style of play.

1 Introduction

Analysis of the decision making quality of football players has traditionally been centred around the use of metrics such as goals scored and assists provided. However, recent developments in football analytics have resulted in models such as Expected Goals, that have paved the way for more in depth analysis to take place. By evaluating players based on these novel metrics, the role of luck is minimised as the situations that take place throughout the match can be evaluated objectively. The recent surge of football analytics models has been fueled by the increasing volume of data that is recorded for each match

[1] michael.pulis.18@um.edu.mt
[2] josef.bajada@um.edu.mt

that takes place throughout the season, as the first hurdle for developing analytics models is always finding a suitable dataset.

The most common form of data is event data[1], which usually contains information pertaining to the player performing the action, and the player receiving the ball (when applicable). This information can be used to train Possession Value Models (PVMs) that offer insightful results into which players perform actions that increase the team's possession value the most. However, most PVMs found in literature suffer from the issue of not considering the surrounding players when providing the valuation for a particular action. Traditional PVMs only consider factors such as the area of the pitch the ball is being moved into, the distance and angle to goal, the distance covered through the action, and the body part used to carry out the action.

In this work, we propose a novel model for evaluating player decisions called **Decision Value** (DV). The model will make use of both the event data and the tracking data to consider the action within the context that they were taken. The model is trained using Deep Reinforcement Learning, and the main advantage of the model is that it can take into account the position of the teammates and opposition players when valuing the decision made by players.

2 Background and Related Work

2.1 Event Data

Event-based datasets contain the events that occur throughout an entire match [2][3], [4]. For each event in the match the dataset contains information such as the type of action that was carried out, the location of the action, the related player that carried out the action, the body part used, and the event outcome. Event data can be used to generate visualisations such as heat maps [5] and passing networks [6]. PVM models such as Expected Threat (xT) [7], Valuing Actions by Estimating Probabilities (VAEP) [8] and the StatsBomb On-the-Ball Value (OBV) [9] can be trained on this type of event-based data. The xT model assumes that actions are usually taken with the sole aim of increasing the team's chance to score. The pitch is split into zones, and an iterative process is used to find the value of each zone, to identify which zones tend to lead to the ball being transitioned to more valuable zones.

The VAEP and OBV models work on the slightly more nuanced observation that players tend to make decisions based on two factors: to increase their team's chance of scoring, and to decrease their team's chance of conceding. Several features are extracted from the moment the action is performed to train two separate models that learn to predict these values. A comparison between xT and VAEP [10] found that the xT model produced more consistent results, while VAEP captured the risk-reward tradeoff better than the xT model. OBV performs similarly to the VAEP model, although it makes use of StatsBomb's latest xG models to value shot actions, and performs better in certain edge cases [9].

2.2 Tracking Data

Tracking data typically captures the position of all the players on the pitch, consisting of both the teammates and the opposition. This data is collected throughout the match, with the positions being recorded at regular time intervals. This type of data can also be used to generate Pitch Control Models (PCMs) [11], [12]. These models aim to represent which areas of the pitch belong to each team. The features used to generate the model vary, but they typically include the location and the velocity of each player. The output of a PCM for a given scenario is a cartesian representation of the pitch with values ranging from 0 to 1 for each coordinate. A value of 0 indicates that should the ball be dropped at that particular location it would be expected to be given to the opposition, and vice versa should the value be 1. Thus the degree with which that particular coordinate on the pitch belongs to the team with possession of the ball can be extracted directly by reading the output from the PCM.

2.3 Reinforcement Learning

Reinforcement Learning (RL) is a branch of Artificial Intelligence that studies algorithms that solve sequential decision making problems. An RL-based framework typically makes use of a *state*, which is deduced from the observations sensed from the environment, a set of possible actions that an agent can take, and a *reward* obtained for taking an action. This reward can be both positive or negative, the latter indicating that the reward is actually a cost. When an action takes place, a new state is generated from a new environment observation, and the process repeats itself, in what is known as the RL Interaction Loop.

The objective of RL algorithms is to find a *policy*, a function that maps states to actions, that maximises the sum of all rewards obtained from the sequence of subsequent alternating states and actions. One of the strengths of RL is that rather than considering just the immediate value of an action in the current state, it estimates the long term effect of the action. RL can be used in both online mode, where the system learns from interacting directly with the environment, or in offline mode, where the system learns from historic data.

2.3 RL in Football Analytics

RL techniques have also been used on historic football event data for player decision analysis [13]. In this work, player behaviour was modelled as a Markov Decision Process (MDP). The authors consider the possible actions to be “shoot”, or “move” to another location on the pitch. The policy was represented as a probability distribution over the possible actions, to allow for different possibilities to be considered. The reward function was defined such that a reward of 1 is returned each time a goal is scored, or 0 otherwise.

The event data was then used to learn the transition function and the policy. From the results, the authors determined which players made the most 'risky' passes, whilst also identifying the possible actions particular teams could take advantage of to add a few goals to their total throughout the course of an entire campaign.

Another example of RL for football analysis was carried out in [14]. In this work, the authors set out to develop a system that could predict the ideal action to take in *critical scenarios*. A critical scenario is defined as the action that takes place before the team loses possession of the ball (such as the moment before a shot is taken) , or the moment possession is lost itself. First, a CNN-LSTM model was trained on data obtained from 104 European matches to predict the next action in critical scenarios. The prediction is made as a probability distribution over the possible actions *(pass, foul, clear, and shoot)*. Policy Gradient learning was then used to optimise the probability distribution such that it maximises a reward function that was derived from the Expected Goals (xG) of the attacking team [14]. The results showed that DRL could also be applied to real world football event data to learn decision analysis in critical scenarios.

2.5 Combining Event and Tracking Data

More recently, work has been carried out that combines both the event and tracking data, such as that provided by the StatsBomb 360 dataset [4]. This dataset solves the issue of manually aligning event and tracking datasets by providing both the event and tracking data aligned within the same dataset. In their work, StatsBomb proposed a metric called Line-Breaking-Passes [15]. This model requires a definition of what constitutes a line-breaking pass, together with the different types of passes that can be carried out. The main advantage of this metric is that the insights obtained from this model are informed by both the value obtained from the PVM model, as well as the positions obtained from the tracking data.

A metric for evaluating player pass decision making was also developed by [16]. In this work, a heuristic that made use of a PVM was used, in order to quantify the value obtained from a specific pass. A metric was then devised to predict the likelihood that a pass is intercepted before it arrives at its destination. This also allowed the authors to study the risk-reward tradeoff taken by each player when passing the ball. Similar work was also carried out [17], where a risk-reward metric was developed to analyse paired tracking and event data.

While having different research objectives, these works reinforce the hypothesis that combining event and tracking data together can lead to a more informed situation analysis than that drawn by using either in isolation.

3 Methodology

3.1 Dataset

The dataset used to carry out this work was provided by StatsBomb, and contains paired event and tracking data from 580 games from the English Premier League. The data is provided from the 2020/21 and 2021/22 seasons of the competition. The events provided are taken from the fixtures played by 10 teams over both seasons. Each of the 10 teams play against 19 opponents, providing enough coverage to analyse the results that pertain to all the teams in the league.

The events were filtered by type, and only "Pass", "Carry", "Take On", "Shot" and "Clearance" were included. This resulted in around 870,000 actions extracted from the 580 games. The **socceraction** library [18] was used to load and pre-process the data.

In Reinforcement Learning, a *terminal state* is a state from which no other action can be taken. An *episode* is a sequence of alternating states and actions, starting from some initial state and ending with a terminal state. In our case, an episode starts when the team gains possession of the ball and ends when possession is lost.

The events in the dataset were processed to identify which of them correspond to initial states, and which correspond to terminal states, delineating the separate episodes in each match. A total of roughly 120 thousand episodes were extracted from the data.

3.2 Possession Value Model (PVM)

To value events within the dataset, two main PVMs were considered, namely the VAEP model, and StatsBomb's OBV model. These two were preferred over the Expected Threat model as it operates on the assumption that the primary motivating factor at all times for players is to increase their team's chance of scoring. The VAEP and OBV models operate on the slightly more nuanced assumption that players are motivated by increasing their team's chance of scoring, whilst also decreasing the team's chance of conceding. The OBV model used within, and it was implemented through the precalculated values provided by StatsBomb within the dataset itself.

3.3 Possession Control Model

Implementations of PCM can be found online, however they are not directly applicable for datasets that do not contain the position of all 22 players at all times. Thus, a custom implementation was written that can generate a PCM given a frame that contains any amount of players on the pitch. To generate the PCM model for each event in the dataset, a Voronoi Diagram was generated [19]. This was done through two main libraries. First, the **voronoi_plot_2d** function from the scipy library was used to generate the polygons that correspond to each section within the final PCM. Then, the **cv2** library was used to render an image from the polygons, whilst also adding a blur around the border of polygons belonging to different teams, thus allowing for the grey area between sections to exist. This was done to ensure that the model does not produce wildly different values for pixels which are at close proximity to each other.

3.4 Deep Reinforcement Learning Model

Our objective is to generate a model which, given a specific scenario, determines the best action to take. This is done by learning the value of all the possible actions from historical data. Given that the size of all possible states is very large, and not fully covered by the data, we used a function approximation approach which generalises across the data and extrapolates the value to unseen but similar states. This was implemented using a Deep Reinforcement Learning (DRL) model, which combines RL with a conventional Deep Learning model to learn the value function.

Since we are learning from historical data, and the agent is not interacting directly with the environment, we perform Offline RL to train our model. We used the **d3rlpy** library, which incorporates implementations of the most recent RL algorithms and is highly optimised to perform offline DRL.

3.4.1 Observations

The first step required to frame a problem as an RL task is to determine how to represent each state in a way that would allow the algorithm to learn from it effectively. A common form of representation used in previous work is to feed direct data obtained from the scenario as an observation, alongside some additional preprocessed features [20] [21]. In our case, the game state is represented as a set of four images that are generated from the combined event and tracking data provided by StatsBomb, as shown in Figure 1.

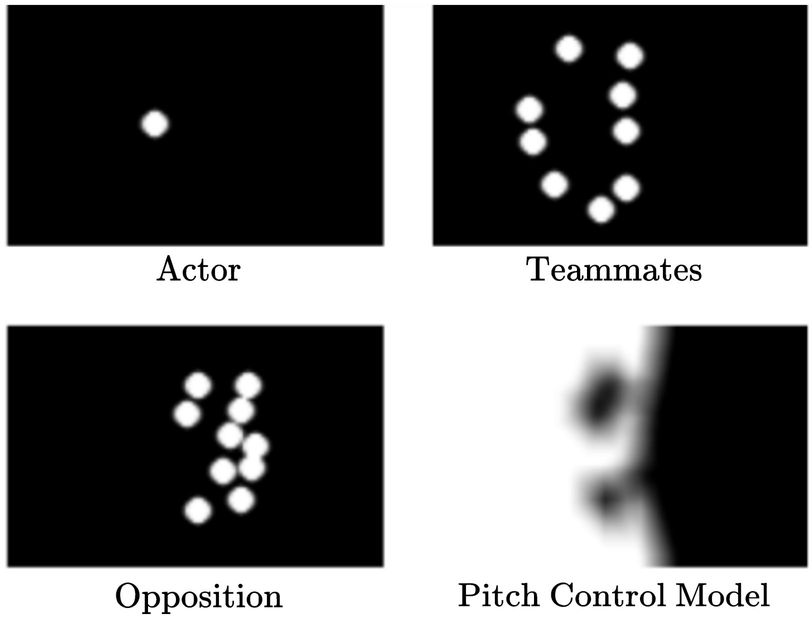

Figure 1: Visualisation of the representation of a given game state

Each image is 105 by 68 pixels. The first one contains the location of the actor (the player with the ball at their feet), the second one contains the location of the teammates, and the third one provides the location of the opponents. The fourth one is a PCM representation that provides the model with additional contextual information that identifies which areas belong to the team with possession of the ball.

3.4.2 Actions

The actions that the player can take in a particular situation are represented by a vector of seven numeric values. The first five elements of the vector are a one-hot encoding of the chosen action, while the last two elements are the scaled x and y coordinates of the target location of the performed action, where (-1,-1) represents the top left corner of the pitch, and (1, 1) represents the bottom right corner. Figure 2 shows an example of a "pass" action, where the scaled target location was (0.9, 0.5), which corresponds to (120,60) in actual pitch coordinates.

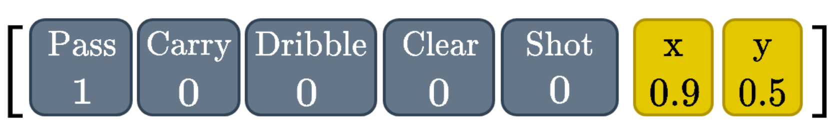

Figure 2: The action vector, with the first five elements representing the possible actions, and the last two elements representing the target location.

Since we are using continuous numeric values for our action representation, we use *continuous control* [21] DRL algorithms to predict the values of this vector. The action with the highest predicted value will be chosen, with the scaled coordinates as its target location.

3.4.3 Reward

The reward function, shown in Figure 3, was devised to capture two important aspects of the chosen action, a, when in some state s:

1. the value of the action if it is successful, and
2. the risk of the action not being successful (and thus losing possession).

$$R(s,a) = \begin{cases} \mathrm{xG}(a) & \text{if } a \text{ is a shot} \\ R_p(s,a) & \text{if possession is retained through } a \\ -n & \text{if possession is lost through } a \end{cases}$$

Figure 3: The reward function.

If the action chosen by the player is a shot, then the corresponding reward is the xG value of the shot. Thus, the higher the likelihood of scoring the shot, the higher the reward. On the other hand, if the action is not a shot, and the possession of the ball is lost, then the reward given for the action is a negative constant, indicating that loss of possession is highly undesirable. In our case, $n = 1$. Since most clearances result in possession being lost, the reward for clearances is always calculated using R_p , as otherwise the model will simply learn to always value clearances negatively without room for nuance.

If the action is not a shot, but possession is still retained, then the Possession Reward, R_p, is calculated from the change in possession value, denoted by $\delta V(a)$, and the pitch control value, denoted by the function $pcm(s, a_{end})$, of the target location of the action, a_{end}, when taken in state s.

$$R_p(s,a) = \begin{cases} \delta V(a) \cdot pcm(s, a_{end}) & if\ \delta V(a) >= 0 \\ \delta V(a) \cdot [1 - pcm(s, a_{end})] & otherwise \end{cases}$$

Figure 4: The possession reward function.

Figure 4 shows how the Possession Reward is computed. If $\delta V(a)$ is positive, this indicates that the action has increased the possession value. This is weighted by the value obtained from $pcm(s,\ a_{end})$, which indicates the probability of retaining possession. On the other hand, if $\delta V(a)$ is negative, this indicates a decrease of possession value. In this case, the decrease is weighted by the complement of the probability of retaining possession, $1 - pcm(s, a_{end})$.

This mechanism uses the pitch control value to make the reward for an action more context aware. If an action has a substantial increase in possession value but ends up in a riskier position, the likelihood of losing possession is higher. Conversely, if the action has a decrease in possession value, but ends up in a more secure position of the pitch, the likelihood of retaining possession is higher.

Figure 5 illustrates two actions, with their respective rewards shown in Table 1. In Case 1, the positive $\delta V(a)$, is scaled down due to the fact that the area poses a slight risk of possession loss. In Case 2, the negative $\delta V(a)$ is weighted by the probability of losing possession. Since the ball is expected to end up in an area where possession is likely to be retained, the decrease in possession value is also scaled down.

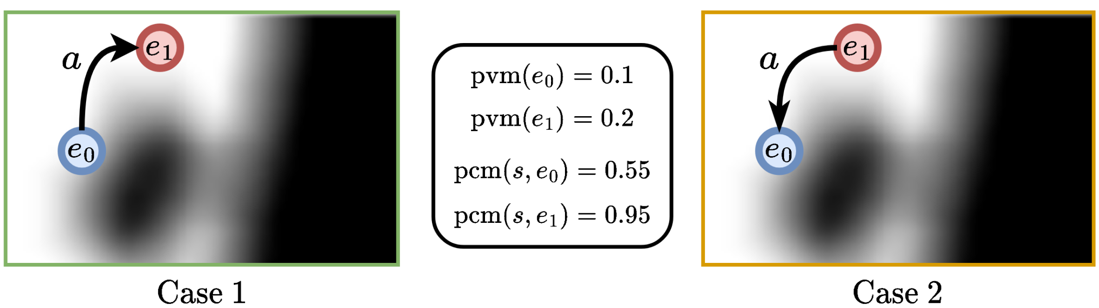

Figure 5: Two example actions with their respective possession and pitch control values.

Case 1	Case 2
$R(s,a) = \delta V(a) \cdot pcm(s, a_{\text{end}})$	$R(s,a) = \delta V(a) \cdot [1 - pcm(s, a_{\text{end}})]$
$= [\mathrm{pvm}(e_1) - \mathrm{pvm}(e_0)] \cdot \mathrm{pcm}(s, e_1)$ $= [0.2 - 0.1] \cdot 0.95$ $= 0.095$	$= [\mathrm{pvm}(e_0) - \mathrm{pvm}(e_1)] \cdot [1 - \mathrm{pcm}(s, e_1)]$ $= [0.1 - 0.2] \cdot [1 - 0.55]$ $= -0.045$

Table 1: Reward calculation for sample cases

3.4.4 DRL Algorithm

The DRL model was trained using Implicit Q-Learning (IQL) [22]. IQL was specifically developed to learn the state-action value function, also known as the Q-function, from offline data. The four images from the observation were combined together into one 4-channel image, which was then passed to a NatureDQN-based encoder [23]. The hyperparameters used are listed in Table 2. Default values[3] were used for hyperparameters that are not stated. The reward obtained from the formula shown in Figure 3 was also scaled to a value in the range (-1, 1). This was done using the **minmax_scaler** from the **sklearn** library [24].

Hyperparameter	Value
Epochs	1073
Actor Loss	6e-8
Critic Loss	3e-5

Table 2: Hyperparameters used to train the DRL model.

3.5 Expected Goals

To value the shots within the dataset, the StatsBomb xG model [25] was used. The dataset already includes the precalculated xG values for each shot. So these values were used as the rewards for "shot" actions in the DRL model, without any need to train a separate xG model.

4 Results

Our DRL model was evaluated using different criteria. This included evaluating different possible actions taken within real world scenarios, as well as identifying which players obtained the highest DV in each main position category *(Goalkeepers, Defenders, Midfielders and Attackers)*. The average DV obtained by the players was also grouped by the team for which the action was performed, to obtain a DV League Table which was compared with the real world league positions obtained by each team. Finally, the DV metric was used to evaluate how well the transfer activity of clubs in the Premier League aligns with the areas of the pitch highlighted by the DV model.

4.1 Valuing Different Actions within the same scenario

In the following examples, the DV model's ability to consider the position of the surrounding teammates as well as the opposition players will be evaluated, by comparing the value of different actions within the same scenario. The values obtained in these examples have been scaled to the range between 0 and 1 to increase interpretability of the results.

[3] https://d3rlpy.readthedocs.io/en/latest/references/generated/d3rlpy.algos.IQL.html#d3rlpy.algos.IQL, Accessed 30th August 2022

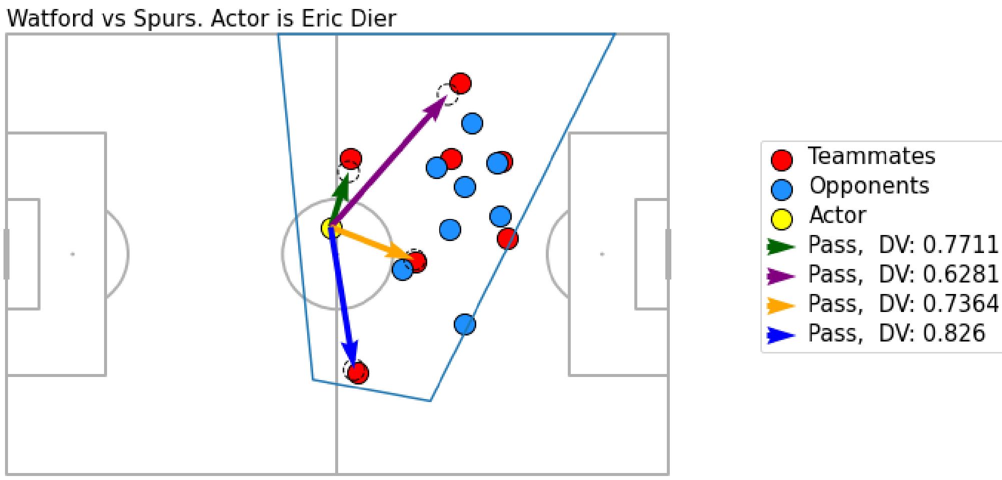

Figure 6: Valuing Eric Dier's Possible Actions

In this first example, Eric Dier has possession of the ball close to the middle of the pitch. The DV for 4 different possible passes was calculated. The pass towards the left wing received the lowest value, which is understandable since it presents a high risk of losing possession. Passing towards the middle of the pitch to a teammate that is pressed by two opposition players also receives a lower score since the area presents a greater risk. The actions that receive the highest DV in this case are the passes in green and blue. The latter receives the highest DV value, which makes tactical sense, as the teammate at the bottom of the screen is in more space, and in a better position to move the team forward.

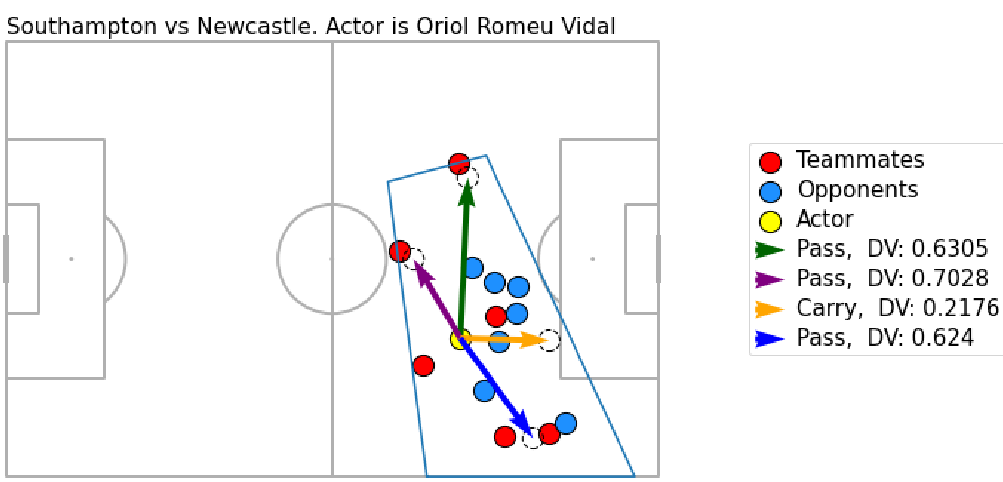

Figure 7: Valuing Oriel Romeu's Possible Actions

In this second example, Oriol Romeu has possession of the ball in a threatening position. Carrying the ball directly towards the edge of the box receives the lowest DV in this case, as that area is surrounded by opposition players, and the chance of success is reasonably low. The action which receives the highest DV in this case would be to pass it back to the teammate indicated with the purple arrow. This is due to the fact that the pass is easy to execute, and the teammate is also in a valuable position, as he can perform an easy pass to the teammate operating on the left wing.

This result is particularly interesting, as traditional PVMs such as xT might assign a negative reward to this action, since the ball would be moving to a less valuable location on the pitch. Passing directly to the teammate on the left wing also receives a higher DV. The lowest DV pass is to pass it to the right wing, as this area has significantly more opposition players surrounding the teammates.

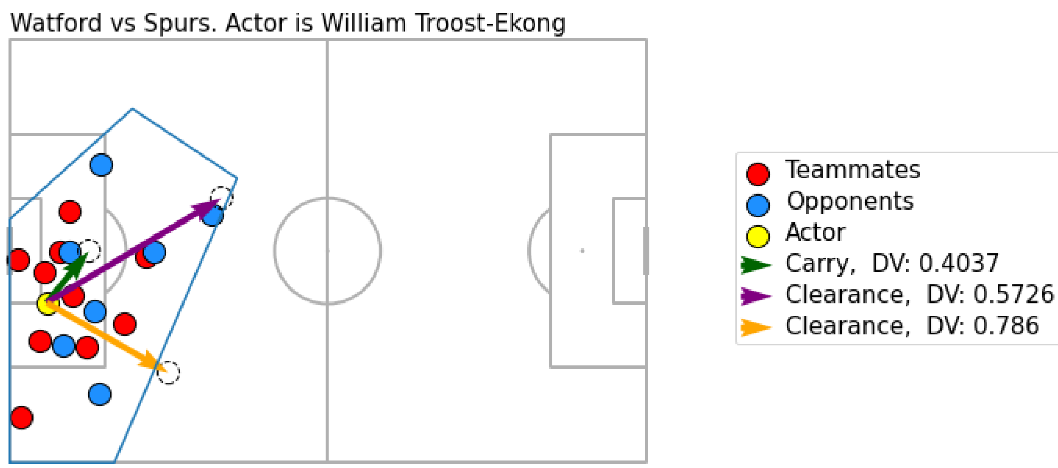

Figure 8: Valuing Troost-Ekong's Possible Actions

In this example, Watford's Troost-Ekong has possession of the ball inside his own penalty area. The first DV to be calculated is for him to proceed to carry within his own box, which received a low value. This indicates that it has a higher chance of resulting in possession loss. The other two simulated actions are clearances. The clearance marked in purple is directed towards the opposition, whilst the clearance marked in orange is directed towards an area with no opposition. This is reflected in the fact that clearing the ball towards opposition is less valuable than clearing it towards an area controlled by your teammates.

4.2 Average DV obtained by Position

4.2.1 Goalkeepers

Player Name	Average DV	Count
Ederson de Moraes	-0.354565	741
Alisson Becker	-0.460242	916

Aaron Ramsdale	-0.495122	704
Edouard Mendy	-0.503006	571
Kasper Schmeichel	-0.51757	367
Hugo Lloris	-0.519072	436
Robert Sanchez	-0.539296	987
Emiliano Martínez	-0.544221	298
David de Gea	-0.55074	431
Alex McCarthy	-0.568185	266

Table 3: Average DV obtained for Passes in the Premier League by Goalkeepers (at least 200 passes)

This data shows that Manchester City and Liverpool have the two best performing goalkeepers in the league when it comes to passing the ball. This is not surprising, given how comfortable the two Brazilian goalkeepers are on the ball. Aaron Ramsdale was purchased by Arsenal at the start of the 2021/22 season, in part due to his ability on the ball, thus it might seem counterintuitive that he does not achieve a higher score than Ederson or Alisson. A possible reason for this is that since the model learns the DV from the outcomes of actions taken by an average premier league player, if the passes were to be carried out by an average player, they would have a lower success rate than what Ramsdale achieves.

4.2.2 Defenders

Player Name	Avg DV	Count
Séamus Coleman	0.089674	38
Robin Koch	0.065084	44
Cheikhou Kouyaté	0.063654	28
Kieran Tierney	0.05986	43
Jamaal Lascelles	0.059058	79
Lucas Digne	0.057613	43
Joël Veltman	0.055783	93
Rob Holding	0.054375	48
Ethan Pinnock	0.054067	87
Andreas Christensen	0.053472	75

Player Name	Avg DV	Count
Rúben Dias	-0.176835	3516
Nathan Aké	-0.202367	1255
Kyle Walker	-0.204106	2322
Oleksandr Zinchenko	-0.215581	1358
Aymeric Laporte	-0.238361	4061
Trevoh Chalobah	-0.278672	1749
João Cancelo	-0.27929	4475
John Stones	-0.280319	1296
Davinson Sánchez	-0.283704	1562
Andreas Christensen	-0.290068	1819

a) Average DV for clearances by Defenders (at least 25 clearances).

b) Average DV obtained from all actions by Defenders (at least 900 actions).

Table 4: Defender DV analysis.

Table 4 shows the average DV obtained from defenders using two different criteria. The first one is the average DV obtained from performing clearances. The second one is the average DV obtained for all actions. Table 4a highlights which defenders are making the best decisions with regards to clearing the ball away from danger during the game. It is interesting to note that in this regard, the highest ranking players hail from various different clubs, as the list contains players from both relegation threatened teams, as well as from other well established outfits. When considering all actions, shown in Table 4b, one team in particular stands out, as Manchester City employed 7 out of the top 10 players with the highest average DV throughout the 2021/22 season. It is also interesting to note that despite having a very high number of actions performed, the average DV obtained by Manchester City players is very high, indicating that the values are the result of long term tactical planning, as opposed to one off high value decisions.

4.2.3 Midfielders

Name	Average DV	Count
İlkay Gündoğan	-0.26434	292
Adam Lallana	-0.294683	307
Fernandinho	-0.300958	290
Bernardo Silva	-0.311137	286
Mateo Kovačić	-0.341066	307
Oliver Skipp	-0.341172	238
Jordan Henderson	-0.359192	496
Jorginho	-0.360534	600
Rodri	-0.365304	915
Yves Bissouma	-0.380784	458

a) Average DV from passes from defensive midfield by Midfielders *(at least 180 passes)*

Name	Average DV	Count
Rodri	-0.227742	647
Bernardo Silva	-0.238212	213
İlkay Gündoğan	-0.255801	242
Jorginho	-0.275102	363
James Ward-Prowse	-0.286655	240
Kevin De Bruyne	-0.299761	235
Pierre-Emile Højbjerg	-0.30889	311
Mateo Kovačić	-0.321072	206
N''Golo Kanté	-0.323576	261
Fred	-0.336103	223

b) Average DV from passes from attacking midfield by Midfielders *(at least 180 passes)*

Table 5: Midfielder DV analysis.

Table 5 shows the midfielders that received the highest average DV for two different categories, passes into defensive zones on the pitch, and passes into attacking zones of the pitch. Table 5a shows that amongst other things, Gundugan and Bernardo Silva are two industrious players that contribute effectively both going forwards and backwards. It is also interesting to note that whilst most of the players hail from the traditional Top 6 teams, Brighton employed 2 out of the 10 players within the first table during the 2021/22 season.

Table 5b is once again dominated by Manchester City players, indicating that their players manage to undertake less risk when making forward passes to teammates. In similar fashion to Aaron Ramsdale's inclusion when discussing the Goalkeepers in Section 4.2.1, by simply observing his actions or reading most analytics, one would conclude that Kevin de Bruyne is consistently one of the best performing midfielders in the league. However, the DV assigned by this model does not take individual player ability into account when performing the calculation, and thus, it could be argued that de Bruyne overperforms his DV due to his exceptionally high skill.

4.2.4 Attackers

Player Name	Avg DV	Count
Riyad Mahrez	-0.263441	579
Jack Grealish	-0.293717	642
Gabriel Jesus	-0.299774	501
Raheem Sterling	-0.312263	609
Callum Hudson-Odoi	-0.328971	260
Philip Foden	-0.355336	644
Nicolas Pépé	-0.360095	215
Bukayo Saka	-0.372763	690
Heung-Min Son	-0.374812	559
Hakim Ziyech	-0.382391	412

a) Average DV obtained from passes by Attacking players in attacking zones *(at least 180 such passes)*.

Player Name	Avg DV	Count
Riyad Mahrez	-0.135018	447
Jack Grealish	-0.174115	543
Nicolas Pépé	-0.17483	166
Paul Pogba	-0.191623	217
Gabriel Martinelli	-0.194877	189
Philip Foden	-0.202489	468
Gabriel Jesus	-0.204194	437
Raheem Sterling	-0.214814	583
Cristiano Ronaldo	-0.219422	296
Hakim Ziyech	-0.219596	282

b) Average DV obtained from carries by Attacking players *(at least 130 such carries)*.

Table 6: Average DV from Passes and Carries obtained by Attacking Players.

Table 6 shows the ability of Attacking players to move the ball forward through carries and passes as measured through the DV metric. These results show that most of the highest performing attackers were highly technical players who mostly play on either wing. Most of the players with the highest average DV in this aspect were also employed by Manchester City during the 2021/22 season, with Riyad Mahrez and Jack Grealish being the two best performing players in both aspects. This might offer an explanation as to why Manchester City chose to sell two of their attacking players, as they had 5 players in Mahrez, Grealish, Foden, Jesus and Sterling that perform similarly.

Player Name	Avg DV	Count
Raheem Sterling	0.268763	51
Heung-Min Son	0.240843	81
Christian Pulisic	0.238608	28
Chris Wood	0.198477	36
Sadio Mané	0.193774	96
Timo Werner	0.192978	43
Daniel Welbeck	0.192021	43
Che Adams	0.186486	49
Jarrod Bowen	0.186106	40
Bryan Mbeumo	0.176171	33

a) Average DV obtained from shots by attacking players that did not play in ST *(at least 25 shots)*

Player Name	Avg DV	Count
Jamie Vardy	0.249256	27
Diogo Jota	0.237689	86
Romelu Lukaku	0.215954	39
Kai Havertz	0.213712	56
Alexandre Lacazette	0.201098	41
Ollie Watkins	0.193582	38
Edward Nketiah	0.181173	27
Cristiano Ronaldo	0.17919	92
Harry Kane	0.179186	117
Roberto Firmino	0.178918	29

b) Average DV obtained from shots players who played in ST *(at least 25 shots)*

Table 7: Shooting DV statistics for attacking players.

This final set of results, shown in Table 7, is concerned with the DV obtained from shooting. In these shooting categories, the model is estimating whether shooting was the right action to perform considering the surroundings. In table 7A, the average DV obtained from shooting by attacking players who are not central strikers was calculated. In this category, there is a mixture of players that are associated with prolific scoring in Sterling, Son and Mane. However, there are also players that do not score as prolifically as the aforementioned players, such as Werner and Welbeck. This indicates that whilst these players are deciding to shoot in the right moments, they might be experiencing a run of unfortunate finishing or goalkeepers who overperform. Further insight into the finishing capabilities of the players could be looked into through the PSxG model. These results, alongside the previous results show that Raheem Sterling has a considerably well rounded game, as his decision making process is rewarded quite highly from several different aspects, which might explain why Chelsea were so eager to purchase him.

The players who played primarily in ST and obtained the highest average DV also contains players that are typically associated with scoring consistently in Vardy, Lukaku, Ronaldo and Kane. However, a high average score is also obtained by other players that are not typically associated with such prolific scoring, indicating that they are placing themselves in the right positions to take high quality shots.

4.3 DV League Table

To achieve the values shown in Table 8, all the actions that were of the type Carry, Pass, Take On, Clearance or Shot were processed by the DV model to predict their value alongside the scenario they were performed in. The values assigned to each action were then grouped by the team for which the action was performed. The results are shown in descending order, starting with the team that obtained the highest average DV.

Premier League - 2021/22				
Team Name	**Average DV**	**DV Order**	**Final League Position**	**Difference**
Manchester City	-0.243	1	1	0
Chelsea	-0.333	2	3	-1
Arsenal	-0.340	3	5	-2
Liverpool	-0.354	4	2	+2
Tottenham	-0.360	5	4	+1
Leicester City	-0.379	6	8	-2
Manchester Utd	-0.382	7	6	+1
Brighton	-0.393	8	9	-1
West Ham	-0.406	9	7	+2
Aston Villa	-0.421	10	14	-4
Newcastle Utd	-0.431	11	11	0
Southampton	-0.433	12	15	-3
Crystal Palace	-0.451	13	12	+1
Wolves	-0.457	14	10	+4
Leeds United	-0.478	15	17	-2
Watford	-0.482	16	19	-3
Everton	-0.488	17	16	+1
Norwich City	-0.496	18	20	-2
Brentford	-0.501	19	13	+6
Burnley	-0.521	20	18	+2

Table 8: DV League Table 2021/22

(Red indicates underperformance over DV, Green indicates overperformance over DV)

The results within this table show that when ordering the teams by how much DV they received on average, the resulting list aligns well with the actual team standings at the end of the season. It is interesting to note that the model correctly predicts which teams would be playing a UEFA competition next season, barring the inclusion of Leicester instead of West Ham. It is also noteworthy that Leicester achieved a higher average DV than a more established team such as Manchester United, which is indicative of how many of their players have good decision making qualities on the ball. Brighton also achieved a score that put them quite close to a European position with regards to the DV league table.

Most fans would argue that Manchester City and Liverpool are the two best teams in the Premier League, and thus at first glance this table seems to contradict this observation. It is important to note that this table is valuing actions taken on the ball. The results suggest that whilst Liverpool do in fact take decisions on the ball that correlate with a top 4 position, some of their best actions also come when they are not in possession of the ball. This could be through their highly effective pressing that they deploy when possession is lost to force the opposition to surrender possession in dangerous areas of the pitch. It could also be the case that some of the most valuable actions taken by the players such as Trent-Alexander Arnold's long passes that are critical to Liverpool's attacking output would not receive a high valuation by the model since the average player would not be as successful as he is with passes of such high difficulty.

At the other end of the table, the model also correctly predicts two of the three relegated teams, Norwich and Burnley. Interestingly, the results show that Brentford made worse decisions when compared to Norwich. This could be due to the fact that Brentford achieved a significant portion of the goals from set-pieces, as well as how well they are able to press opposition players when they don't have possession of the ball. Both aspects are not currently captured by our DV model.

Norwich, on the other hand, finished at the very bottom of the table in the Premier League. However, their average DV indicates that on-the-ball, they were of a higher level than dead last. This indicates that possibly, they were let down by other aspects of their game, pressing, their defensive transitions as well as achieving the 2nd worst underperformance of their xG in the entire league[4]. When summing the total xG achieved by their opponents against how many goals they actually achieved, the data also shows that they might have been unlucky to concede as many goals as they did, indicating that their season might not have been as bad as it looks at face value.

The results obtained from the DV metric were also compared with the traditional xG metric, the total xG conceded by a team, labelled as xGA (Expected Goals Against) and the difference between the xG generated by a team, and the xGA, called xGD (Expected Goal Difference):

[4] https://understat.com/league/EPL/2021, Accessed on August 27th, 2022

Total xG	Total xGA	Total xGD	Mean DV	Mean OBV
-0.82	0.80	-0.86	-0.90	-0.78

Table 9: Correlation between actual team rank and each metric

Table 9 shows that all the different metrics have a very high correlation with the final league ranking obtained. This indicates that all the metrics have significant explanatory power. The table also shows that although DV is trained with both xG and OBV, it has the highest correlation with the team's final rankings. This is despite the fact that the model is never made aware of the concept of different teams during the training process, since each event is treated separately, making use of only the coordinates of the teammates, opponents and actor on the ball, together with the performed action.

4.4 DV difference per zone of the Pitch

In this section, the average DV obtained per zone for all Premier League teams was calculated *(for the 2021/22 season)*. For each particular team, the difference between the average DV per zone and the DV obtained by said team was calculated. The results can be seen below in Figure 9 and 10 respectively.

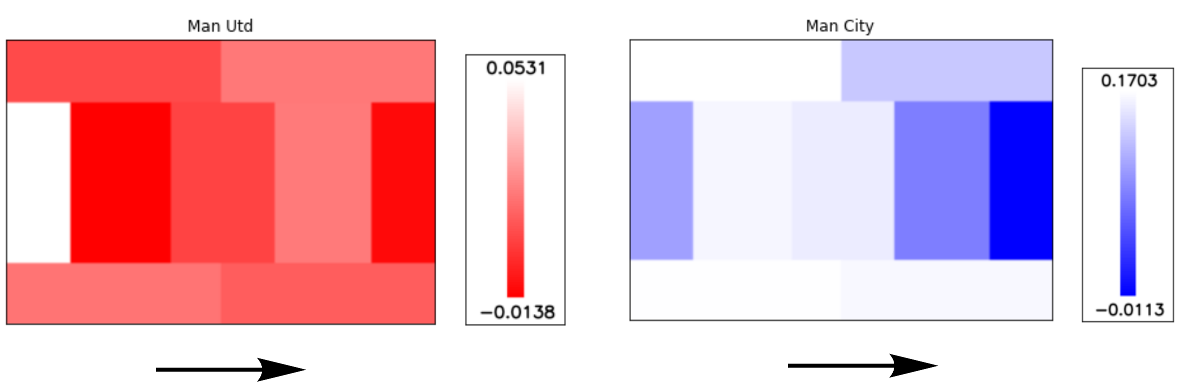

Figure 9: DV difference per zone for Manchester United and Manchester City

The results in Figure 9 contain the areas where Manchester United and Manchester City performed better or worse than the premier league average per zone. The darkest regions for Manchester United are in the zones typically occupied by the left back, defender, defensive midfielder, central midfielder, right winger and the striker. This aligns quite closely with the transfer activity of the club at the start of the 2022/23 season, as with the exception of the striker, they purchased players to address each of the darker zones shown in Figure 9.

Manchester City's is mostly covered with lighter colours, which is positive and indicates that they tend to achieve higher DVs across most areas of the pitch. The darkest region of the pitch is in the position typically occupied by the striker, which is also where they invested most heavily at the start of the 2022/23 season.

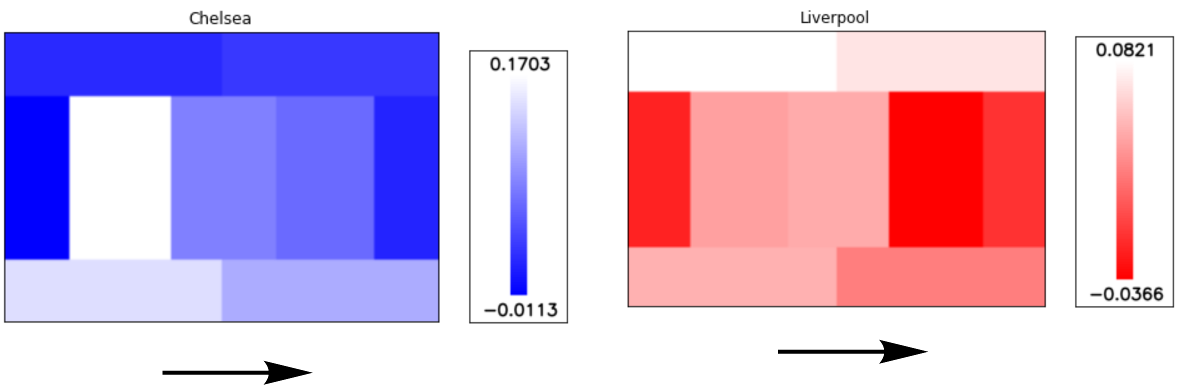

Figure 10: DV difference per zone for Chelsea and Liverpool *(attacking towards the right)*

In Figure 10, the same charts have been drawn up for Chelsea and Liverpool respectively. Whilst some eyebrows were raised at Chelsea's acquisition of Marc Cucurella, a position that most did not think required heavy investment, the chart indicates that reinforcements in this position are quite sensible. Chelsea also purchased a young promising keeper, as well as forwards Pierre-Emerick Aubameyang and Raheem Sterling. These acquisitions are in line with the areas that Chelsea should have looked to improve, according to the chart in Figure 10.

For Liverpool, the chart shows that some of the worst performing areas were the attacking zones, with zones typically occupied by attacking midfielders and strikers. This aligns well with the areas that Liverpool chose to reinforce this year, as they purchased striker Darwin Núñez, as well as midfielder Artur on loan. The darker zones on the right hand side might also indicate that Trent-Alexander Arnold tends to make risky actions that would yield a much lower rate of success were they to be attempted by players with a lower skill level.

5 Conclusion

In this work, we propose a new framework, named Decision Value, through which player decisions are assessed within the context of the game. While traditional PVMs lack the ability to consider the surrounding players, our model combines both event and tracking data to evaluate an action. The StatsBomb 360 data from 10 teams over 2 seasons of the English Premier League was used to train a DRL model, which uses the current position of the player with the ball, the position of his teammates and those of the opponents. A pitch control model that predicts the likelihood of retaining or losing possession was also developed, which was also used as input to the model.

The results show that our model accurately differentiates which actions have a higher Decision Value within the context of the game. We also analysed the model in a more quantitative manner, by comparing it with other football analysis metrics. The average DV of players had a high correlation with their actual performance, and the average DV of the whole team also correlated highly with the DV League table. Finally, an analysis of the average DV difference between teams' performance per zone over the Premier League average showed explanatory power with regards to the transfer activity of Premier League clubs.

5.1 Future Work

While the model is already yielding accurate results, there are various opportunities for further research and improvements. Since the StatsBomb 360 dataset does not have the positions of all players on the pitch, one improvement would be to train the RL model on a dataset that contains the positions of all the on-field players. Another aspect that could be taken into account is the velocity of the players when computing the PCM.

5.2 Acknowledgements

This work was partially funded by the Malta Digital Innovation Authority (MDIA) Pathfinder Digital Scholarship and the University of Malta Research Seed Fund.

References

[1] J. Van Haaren, P. Robberechts, T. Decroos, L. Bransen, and J. Davis, "Analyzing performance and playing style using ball event data," 2019. https://lirias.kuleuven.be/retrieve/582135 (accessed Aug. 29, 2022).

[2] L. Pappalardo *et al.*, "A public data set of spatio-temporal match events in soccer competitions," *Scientific Data*, vol. 6, no. 1, pp. 1–15, Oct. 2019.

[3] "koenvo/wyscout-soccer-match-event-dataset - GitHub," *https://github.com › koenvo › wyscout-soccer-match-e...https://github.com › koenvo › wyscout-soccer-match-e...*, [Online]. Available: https://github.com/koenvo/wyscout-soccer-match-event-dataset

[4] StatsBomb, "statsbomb/open-data: Free football data from ... - GitHub," *https://github.com › statsbomb › open-datahttps://github.com › statsbomb › open-data*, [Online]. Available: https://github.com/statsbomb/open-data

[5] "Football Heatmaps with Seaborn," *FC Python*, Jan. 12, 2018. https://fcpython.com/visualisation/football-heatmaps-seaborn (accessed Aug. 27, 2022).

[6] "Friends-of-Tracking-Data-FoTD/passing-networks-in-python," *https://github.com › Friends-of-Tracking-Data-FoTDhttps://github.com › Friends-of-Tracking-Data-FoTD*, [Online].

Available: https://github.com/Friends-of-Tracking-Data-FoTD/passing-networks-in-python
[7] K. Singh, "Introducing Expected Threat (xT)." https://karun.in/blog/expected-threat.html (accessed Aug. 27, 2022).
[8] Decroos, Bransen, and Van Haaren, "VAEP: an objective approach to valuing on-the-ball actions in soccer," *Proc. Estonian Acad. Sci. Biol. Ecol.*, 2020, [Online]. Available: https://lirias.kuleuven.be/retrieve/593249
[9] StatsBomb, "Introducing On-Ball Value (OBV)," *StatsBomb | Data Champions*, Sep. 16, 2021. https://statsbomb.com/articles/soccer/introducing-on-ball-value-obv/ (accessed Aug. 27, 2022).
[10] Decroos and Davis, "Valuing on-the-ball actions in soccer: a critical comparison of XT and VAEP," *Proceedings of the AAAI-20 Workshop on Artificial Intelligence in Team Sports*, 2020, [Online]. Available: https://lirias.kuleuven.be/2913207?limo=0
[11] Spearman, Basye, and Dick, "Physics-based modeling of pass probabilities in soccer," *Proceeding of the 11th*, 2017, [Online]. Available: https://www.researchgate.net/profile/William-Spearman/publication/315166647_Physics-Based_Modeling_of_Pass_Probabilities_in_Soccer/links/58cbfca2aca272335513b33c/Physics-Based-Modeling-of-Pass-Probabilities-in-Soccer.pdf
[12] Spearman, "Beyond expected goals," *Proceedings of the 12th MIT sloan sports analytics*, 2018, [Online]. Available: https://www.researchgate.net/profile/William-Spearman/publication/327139841_Beyond_Expected_Goals/links/5b7c3023a6fdcc5f8b5932f7/Beyond-Expected-Goals.pdf
[13] M. Van Roy, P. Robberechts, W.-C. Yang, L. De Raedt, and J. Davis, "Leaving Goals on the Pitch: Evaluating Decision Making in Soccer," *arXiv [cs.AI]*, Apr. 07, 2021. [Online]. Available: http://arxiv.org/abs/2104.03252
[14] P. Rahimian, A. Oroojlooy, and L. Toka, "Towards optimized actions in critical situations of soccer games with deep reinforcement learning," in *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*, Oct. 2021, pp. 1–12.
[15] J. Yorke, "StatsBomb 360: Exploring Line Breaking Passes," *StatsBomb | Data Champions*, Jun. 28, 2022. https://statsbomb.com/articles/soccer/statsbomb-360-exploring-line-breaking-passes/ (accessed Aug. 27, 2022).
[16] Burriel and Buldú, "The quest for the right pass: Quantifying player's decision making," *statsbomb.com*, [Online]. Available: http://statsbomb.com/wp-content/uploads/2021/11/Javier-M-Buldu.pdf
[17] F. Goes, E. Schwarz, M. Elferink-Gemser, K. Lemmink, and M. Brink, "A risk-reward assessment of passing decisions: comparison between positional roles using tracking data from professional men's soccer," *Science and Medicine in Football*, vol. 6, no. 3, pp. 372–380, Jul. 2022.
[18] "socceraction," *PyPI*. https://pypi.org/project/socceraction/ (accessed Aug. 27, 2022).
[19] S. Fonseca, J. Milho, B. Travassos, and D. Araújo, "Spatial dynamics of team sports exposed by Voronoi diagrams," *Hum. Mov. Sci.*, vol. 31, no. 6, pp. 1652–1659, Dec. 2012.
[20] V. Mnih *et al.*, "Playing Atari with Deep Reinforcement Learning," *arXiv [cs.LG]*, Dec. 19, 2013.

[Online]. Available: http://arxiv.org/abs/1312.5602

[21] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," *arXiv [cs.LG]*, Sep. 09, 2015. [Online]. Available: http://arxiv.org/abs/1509.02971

[22] I. Kostrikov, A. Nair, and S. Levine, "Offline Reinforcement Learning with Implicit Q-Learning," *arXiv [cs.LG]*, Oct. 12, 2021. [Online]. Available: http://arxiv.org/abs/2110.06169

[23] D. Shi, F. Tian, and S. Wu, "Energy Efficiency Optimization in Heterogeneous Networks Based on Deep Reinforcement Learning," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[24] "sklearn.preprocessing.MinMaxScaler," *scikit-learn*. http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.MinMaxScaler.html (accessed Aug. 30, 2022).

[25] D. Vatvani, "Upgrading Expected Goals," *StatsBomb | Data Champions*, May 16, 2022. https://statsbomb.com/articles/soccer/upgrading-expected-goals/ (accessed Aug. 27, 2022).