VAEP: An Objective Approach to Valuing On-the-Ball Actions in Soccer
(Extended Abstract)

Tom Decroos¹, Lotte Bransen², Jan Van Haaren¹,² and Jesse Davis¹
¹KU Leuven, Department of Computer Science & Leuven.AI, Belgium
²SciSports, Netherlands
{tom.decroos, jesse.davis}@kuleuven.be, {l.bransen, j.vanhaaren}@scisports.com

Abstract

Despite the fact that objectively assessing the impact of the individual actions performed by soccer players during games is a crucial task, most traditional metrics have substantial shortcomings. First, many metrics only consider rare actions like shots and goals which account for less than 2% of all on-the-ball actions. Second, they fail to account for the context in which the actions occurred. This work summarizes several important contributions. First, we describe a language for representing individual player actions on the pitch. This language unifies several existing formats which greatly simplifies automated analysis and this language is becoming widely used in the soccer analytics community. Second, we describe our framework for valuing any type of player action based on its impact on the game outcome while accounting for the context in which the action happened. This framework enables giving a broad overview of a player’s performance, including quantifying a player’s total offensive and defensive contributions to their team. Third, we provide illustrative use cases that highlight the working and benefits of our framework.

1 Introduction

A fundamental question for soccer clubs is:

How will a soccer player’s actions impact his or her team’s performances in games?

Understanding the answer to this question informs crucial tasks such as player acquisition, player evaluation, and scouting. Typically, this question was addressed in a subjective manner via video analysis by human scouts. The availability of new data sources such as optical tracking and event stream data has spurred interest in automated approaches for objectively quantifying the impact of the individual actions performed by a player during a match.

Unfortunately, the extremely low-scoring and dynamic nature of soccer matches makes this a difficult task. Most existing approaches for valuing actions in soccer suffer from three important limitations. First, the vast majority of these approaches focus on the concept of the expected value of a goal attempt [Lucey et al., 2014; Caley, 2015; Altman, 2015; Mackay, 2017], which, while more frequent than goals, still only occur rarely. Second, existing approaches ignore much of the action’s context, that is, the circumstances under which the action was performed. For example, some approaches only consider the action’s location [Singh, 2019; Van Roy et al., 2020] and ignore important factors such as the goal difference, time remaining, and speed of play (e.g., if a pass occurs during a rapid counter attack). Third, most approaches only consider immediate effects and fail to account for an action’s effects a bit further down the line.

This extended abstract of our KDD’19 paper [Decroos et al., 2019] describes our data-driven framework for valuing actions in a soccer game. Unlike most existing work, it considers all types of actions (e.g., passes, crosses, dribbles, take-ons, and shots) and accounts for an action’s context as well as its possible longer-term effects. Intuitively, an action value reflects the action’s expected influence on a match’s score-line. That is, an action valued at +0.05 is expected to contribute 0.05 goals in favor of the team performing the action, whereas an action valued at -0.05 is expected to yield 0.05 goals for their opponent. This extended abstract summarizes five important contributions:

1. A language for representing player actions;
2. A framework for valuing player actions;
3. A model for predicting short-term scoring and conceding probabilities at any moment in a game;
4. A number of use cases showcasing our most interesting results and insights; and
5. A Python package¹ that implements the first four contributions and is becoming widely used in the soccer analytics community.

2 SPADL: A Language for Describing Actions

Event stream data annotates the times and locations of specific events (e.g., passes, shots, and cards) that occur in

¹https://github.com/ML-KULeuven/socceraction
Do:

Learn a function that assigns a value to a state. That is, an action
The effect of every on-the-ball action is that it alters the game
Given:

A fundamental task in soccer analytics is attempting to under-
3 V AEP: A Framework for Valuing Actions
automatic analysis tools.
optional information snippets enables us to more easily apply
for each action addresses the third challenge, as excluding
(e.g., Opta records weather changes).

The need for a uniform representation language is a typical
AI problem. To this end, we propose SPADL (Soccer Player
Action Description Language), a language to describe soccer
actions that unifies event stream formats from different ven-
dors (e.g., WyScout, StatsBomb, Opta) into a common vocab-
ulary that enables subsequent data analysis. SPADL describes
player actions, as opposed to the formats by commercial ven-
dors that describe events. The distinction is that actions are
a subset of events that require a player to perform the action.
For example, a passing event is an action, whereas an event
signifying the end of the game is not an action.

Unlike other event stream formats, SPADL always stores
the same eight attributes for each action: time, start location,
end location, involved player, team of the involved player,
the action type (e.g., pass, shot, dribble), the body part used
to perform the action (e.g., foot, head), and the result of the
action (e.g., success or fail). Storing the same eight attributes
for each action addresses the third challenge, as excluding
optional information snippets enables us to more easily apply
automatic analysis tools.

3 VAEP: A Framework for Valuing Actions

A fundamental task in soccer analytics is attempting to under-
stand the value of each action that a player performs during a
game. This task can be formally defined as follows:

Given: An on-the-ball action \( a_i \);

Do: Learn a function that assigns a value \( V(a_i) \) to the action.
The effect of every on-the-ball action is that it alters the game
state. That is, an action \( a_i \) moves the game from game state
\( S_{i-1} \) to game state \( S_i \). VAEP leverages this by measuring the
value of this change in game state:

\[
V(a_i) = V(S_i) - V(S_{i-1})
\]

where \( V(.) \) represents the value of a game state or action.
Hence, the primary question becomes: How should we value a
game state?

Intuitively, a helpful action is one that benefits a player’s
team, either by increasing the chance that his team scores or
decreasing the chance that the opposing team scores. These
actions should be positively valued. In contrast, actions that
do the opposite hurt the player’s team and should be nega-
tively valued. This observation suggests that one way to value
a game state is:

\[
V(S_i) = P_{\text{scores}}^k(S_i) - P_{\text{concedes}}^k(S_i)
\]

where \( P_{\text{scores}}^k(S_i) \) (\( P_{\text{concedes}}^k(S_i) \)) is the probability that the
team possessing the ball in state \( S_i \) will score (concede)
within the next \( k \) actions. Typically, \( k \) is a small number like
five or ten as the effects of most actions are temporally lim-
ited. Practically, the task then becomes estimating these two
probabilities for each state, which can be solved by training a
machine learning model that predicts a probability.

The VAEP framework provides a simple approach to valu-
ing actions that is independent of the representation used to
describe the actions. The framework’s strength is that it trans-
forms the subjective task of valuing an action into the objec-
tive machine learning task of predicting the likelihood of a
future event in a natural way.

4 Estimating Scoring and Conceding
Probabilities

This section describes our method for estimating the scor-
ing and conceding probabilities required by the VAEP fram-
work. This task can be formally described as follows:

Given: A game state \( S_i \);

Do: Estimate \( P_{\text{scores}}^k(S_i) \) and \( P_{\text{concedes}}^k(S_i) \).

For both binary classification problems we train a probabilis-
tic classifier to estimate the probabilities. In principle, any
machine learning model (e.g., logistic regression, random for-
est, neural network) that predicts a probability could be used
to address these tasks. Empirically, we found that gradient
boosted decision tree ensembles such as XGBoost [Chen and
Guestrin, 2016] and CatBoost [Prokhorenkova et al., 2018]
work well in practice.

Applying a standard machine learning algorithm requires
converting the sequence of actions \( [a_1, a_2, \ldots, a_n] \) describ-
ing an entire game into examples in a feature-vector format.
Thus, one training example is constructed for each game state
\( S_i \). We compute 151 features that impact the probability of a
goal being scored in the near future. These features are based
on the last three actions of a game state and can be divided
into three categories.

First, the SPADL features are categorical and real-valued
features based on information explicitly included in the
SPADL representation. We consider categorical features for
an action’s type, result, and the body part used by the player
performing the action. Similarly, we consider real-valued fea-
tures for the \((x, y)\)-coordinates of the action’s start and end
locations, and the time elapsed since the start of the game.

Second, the complex features combine information within
and across consecutive actions. Within each action, these fea-
tures include (1) the distance and angle to the goal for both
the action’s start and end locations, and (2) the distance covered during the action in both the \(x\) and \(y\) directions. Between two consecutive actions, we compute the distance and elapsed time between them and whether the ball changed possession. These features provide some intuition about the current speed of play.

Third, the game context features include the scoreline and goal difference after action \(a_i\). This information is relevant because scoring probabilities depend on these factors [Robberechts et al., 2019].

## 5 Experiments

Evaluating our framework is challenging as no objective ground truth action values or player ratings exist. Lack of ground truth data is an important challenge that frequently occurs when applying AI techniques to real-world data, but is not often discussed in traditional AI papers. In this extended abstract, we provide intuitions into how our framework behaves, show the natural trade-off between the quantity and quality of players’ actions, and list the top-10 players of the 2019/2020 English Premier League season. The full paper contains a number of other use cases [Decroos et al., 2019].

### 5.1 Intuition behind the Action Values

Figure 1 illustrates how our framework works by visualizing the actions and their corresponding values that led to Barcelona’s goal in the 93rd minute of their away game against Real Madrid on December 23, 2017. The attack consists of six actions. The first three actions receive relatively neutral action values (i.e., \(0.00, -0.01, +0.01\)) as they appear further away from the goal and thus do not have a strong effect on Barcelona’s scoring odds. The fourth action is more interesting; Messi receives the ball and dribbles past a Real Madrid defender into the box, which receives an action value of \(+0.05\) for raising the scoring odds from 0.03 to 0.08.

The fifth action is a backwards pass that illustrates both Messi’s genius and the power of VAEP in recognizing this genius. Simple metrics often do not reward backwards passes as they move the ball further away from the goal. In this case however, Messi’s backwards pass is a great action as it moves the ball away from the crowded six yard box and sets up a shot with a better angle for a teammate. In a purely data-driven way, VAEP recognizes that this action increases the scoring odds from 0.08 to 0.17 and awards it a value of \(+0.09\). Finally, Aleix Vidal shoots and scores. For converting a 0.17 scoring chance to a goal, our framework rewards Vidal with an action value of \(+0.83\). If Vidal had missed his shot, he would have been penalized with an action value of \(-0.17\).

### 5.2 Trading Off Action Quality and Quantity

A natural tension exists between the quality and quantity of actions. If a player performs a high number of actions, then it is harder for each action to have a high value. Figure 2 shows the number of actions that players execute on average per 90 minutes (quantity) and the average value of these actions (quality) for those players who played at least 900 minutes during the 2017/2018 season in the Spanish and English leagues. The grey-dotted isoline shows the gap in VAEP rating between top-ranked Lionel Messi and the rest. The isoline is curved as a player’s rating is obtained by multiplying the average value per action (\(x\)-axis) and the average number of actions (\(y\)-axis). As shown by the isoline and more traditional statistics,\(^2\) Messi is clearly in a class of his own.

\(^2\)https://fivethirtyeight.com/features/lionel-messi-is-impossible/
5.3 New Results about the 2019/2020 Season

Table 1 shows the top-10 players in the 2019/2020 English Premier League season up until April 2020 in terms of our VAEP player ratings.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Player</th>
<th>Team</th>
<th>Goals/90</th>
<th>Assists/90</th>
<th>VAEP/90</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R. Mahrez</td>
<td>Manchester City FC</td>
<td>0.42</td>
<td>0.42</td>
<td>0.89</td>
</tr>
<tr>
<td>2</td>
<td>K. De Bruyne</td>
<td>Manchester City FC</td>
<td>0.31</td>
<td>0.59</td>
<td>0.79</td>
</tr>
<tr>
<td>3</td>
<td>S. Mané</td>
<td>Liverpool FC</td>
<td>0.57</td>
<td>0.24</td>
<td>0.69</td>
</tr>
<tr>
<td>4</td>
<td>S. Agüero</td>
<td>Manchester City FC</td>
<td>0.97</td>
<td>0.12</td>
<td>0.68</td>
</tr>
<tr>
<td>5</td>
<td>P. Aubameyang</td>
<td>Arsenal FC</td>
<td>0.62</td>
<td>0.0</td>
<td>0.62</td>
</tr>
<tr>
<td>6</td>
<td>D. Ings</td>
<td>Southampton FC</td>
<td>0.63</td>
<td>0.04</td>
<td>0.62</td>
</tr>
<tr>
<td>7</td>
<td>Mohamed Salah</td>
<td>Liverpool FC</td>
<td>0.61</td>
<td>0.08</td>
<td>0.61</td>
</tr>
<tr>
<td>8</td>
<td>J. Vardy</td>
<td>Leicester City FC</td>
<td>0.72</td>
<td>0.19</td>
<td>0.59</td>
</tr>
<tr>
<td>9</td>
<td>R. Sterling</td>
<td>Manchester City FC</td>
<td>0.46</td>
<td>0.04</td>
<td>0.59</td>
</tr>
<tr>
<td>10</td>
<td>Richarlison</td>
<td>Everton FC</td>
<td>0.36</td>
<td>0.07</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Table 1: The top-10 players who played at least 900 minutes in the 2019/2020 English Premier League season up until April 2020 in terms of our VAEP player ratings.

6 Discussion and Perspectives

This paper falls within a growing line of data science research for analyzing sports data (e.g., [Liu and Schulte, 2018; Routley and Schulte, 2015; Decroos et al., 2018; Pappalardo et al., 2019; Bransen et al., 2019; Fernández et al., 2019; Rudd, 2011]) which has been fueled by the availability of data such as

- event stream data, which describes specific events that occur during a match;
- tracking data, which records the position of each player multiple times per second;
- athlete monitoring data, which records information such as heart rate and accelerations.

These data and the questions posed by domain experts (e.g., clubs, sports scientists, biomechanists) yield an incredibly rich and broad set of problems from an AI perspective. The following non-exhaustive list includes some illustrative examples. Pattern mining [Decroos et al., 2018] and non negative matrix factorization [Decroos and Davis, 2019] have been used to understand tactics and playing style. Reinforcement learning and Markov decision processes have been used to analyze technical performance in a match [Liu and Schulte, 2018; Schulte et al., 2017]. Multi-agent learning has been used to learn the value of teamwork from outcomes of pass interactions between players [Beal et al., 2020]. Predictive modeling has been used to understand and manage training load [Jaspers et al., 2018]. Deep learning [Kim et al., 2018] and integer linear programming [Shirrit et al., 2011] have been used for tracking players during a match.

Similarly, on a technical level these problems touch upon a number of core areas within AI:

**Representation** Sports data of various sources always needs to be preprocessed and put into a format compatible with AI techniques. For example, SPADL is a great way to represent soccer event stream data when training a classifier, but a more suitable representation is needed when applying a different AI technique such as pattern mining [Decroos et al., 2018].

**Interpretability** Ultimately, experts are interested in translating the findings arising from analytics into practice. For example, an effect of the insights from expected goals models is that the number of long range shots have dramatically decreased in recent years.³

**Decision making** Decision making plays a crucial role in sports, and AI can often offer support. E.g., What strategy should a coach employ to maximize the odds of winning a game? Which action should a player perform in a specific situation? How long should an athlete train for?

**Understanding behavior** Professional athletes exhibit specific behaviors on the pitch that are often a combination of their innate knowledge of the sport, their role within their team, and their own personality. As sports data often essentially describes the actions of players, this is a great data source to understand their behavior. For example, [Bransen et al., 2019] aim to understand how professional soccer players adapt their behavior to high mental pressure situations.

**Experimental evaluation** Often, real-world data lack ground truth labels. These may be hard to obtain or simply not exist. This is frequently the case in sports analytics as there are no objectively best players. Moreover, a team’s tactical plan is only known to the team itself. However, the richness of the sports domain encourages researchers to devise creative solutions for this problem. Some examples are measuring correlation with proxy signals such as player salary [Liu and Schulte, 2018] and evaluating by asking domain experts [Bransen et al., 2019; Pappalardo et al., 2019]).

For all of these reasons, we believe that sports will remain a fruitful source of real-world problems and a test bed for AI in the coming years.

**Acknowledgments**

Tom Decroos is supported by the Research Foundation-Flanders (FWO-Vlaanderen). Jesse Davis is partially supported by KU Leuven Research Fund (C14/17/07, C32/17/036) and Research Foundation - Flanders (EOS No. 30992574, G0D8819N). The authors thank Wyscout for supplying the event stream data used in this extended abstract.

References


