

Dick, Uwe; Brefeld, Ulf

Article — Published Version

Action rate models for predicting actions in soccer

AStA Advances in Statistical Analysis

Provided in Cooperation with:

Springer Nature

Suggested Citation: Dick, Uwe; Brefeld, Ulf (2022) : Action rate models for predicting actions in soccer, AStA Advances in Statistical Analysis, ISSN 1863-818X, Springer, Berlin, Heidelberg, Vol. 107, Iss. 1, pp. 29-49,
<https://doi.org/10.1007/s10182-022-00435-x>

This Version is available at:

<https://hdl.handle.net/10419/310999>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Action rate models for predicting actions in soccer

Uwe Dick¹ · Ulf Brefeld¹ 

Received: 17 December 2020 / Accepted: 7 February 2022 / Published online: 2 March 2022
© The Author(s) 2022

Abstract

We present a data-driven approach to predict the next action in soccer. We focus on passing actions of the ball possessing player and aim to forecast the pass itself and when, in time, the pass will be played. At the same time, our model estimates the probability that the player loses possession of the ball before she can perform the action. Our approach consists of parameterized exponential rate models for all possible actions that are adapted to historic data with graph recurrent neural networks to account for inter-dependencies of the output space (i.e., the possible actions). We report on empirical results.

Keywords Soccer analytics · Football · Elite sports · Actions · Rate models · Game flow · GRNNs

1 Introduction

In team sports, the coordination of players is often of utmost interest as unorganized teams expose weaknesses and open spaces that can be exploited by the opponents. However, viewed from an action-centric perspective, everything depends on the decisions of the player who is currently in ball possession. Depending on whether she decides on a safe pass, risks a through ball or continues to dribble onwards, changes the future of the game. Obviously, her decision is not independent of the overall coordination of her team, including their actual positions, velocities, roles and tasks in the game, etc.

Passes have been frequently studied in the literature (Spearman et al. 2017; Peralta Alguacil et al. 2020), for instance focusing on effectiveness (Goes et al. 2019) or in risk vs reward analyses (Power et al. 2017). Additionally, since football is a dynamic game, there has been a lot of research on predicting players movements

✉ Ulf Brefeld
brefeld@leuphana.de

Uwe Dick
dick@leuphana.de

¹ Machine Learning Group, Leuphana University of Lüneburg, Lüneburg, Germany

(Brefeld et al. 2019) and rating collective movements (Dick and Brefeld 2019). In this paper we focus on the player in ball possession and propose a model that simultaneously predicts (i) the player's next action, (ii) the point in time of that action, and (iii) the player's position at the time of action. To the best of our knowledge, this is the first approach that integrates player and ball movement models with predictions over player actions to address the flow of the game.

Our model grounds on parameterized exponential distributions and assigns positive *rates* to all future player actions. The proposed *action rate model (ARM)* allows to not only predict the next ball action but also when it will be performed. Since decision making of the ball possessing player depends on the positions and trajectories of all other players and the ball, we propose to represent the three learning tasks (i–iii) as a joint optimization problem using graph recurrent neural networks (GRNNs) to account for inter-dependencies between players. We empirically evaluate our approach on trajectory data from 54 professional soccer matches recorded at 25 frames per second.

The remainder is organized as follows. Section 2 reviews related work. We present our graph recurrent action rate models in Sect. 3. Section 4 reports on empirical results and Sect. 5 concludes.

2 Related work

There is a wide variety on previous work on data-driven analysis of sports and, in particular, on soccer. The analysis of passes is the topic of several papers in soccer analytics and different aspects of passes are investigated in depth. For example, Spearman et al. (2017) present a model to compute the likelihood of successful passes. They make use of a physics-inspired movement model that predicts whether a player can reach a certain position in time and augment ball dynamics to estimate probabilities that possible receiving players can reach a ball or whether it can be intercepted by a defending player. Power et al. (2017) compare the probability of an interception (the risk of a pass) versus its reward. The latter is given by the likelihood that the attacking team will take a shot at goal within 10 s after the pass. The effectiveness of passes is studied by Goes et al. (2019) who propose to quantify how defensive players move and coordinate to intercept passes.

Other publications rate general game states according to different measures and can also be used to rate the effectiveness of passes. For example, Link et al. (2016) propose to estimate the dangerousity of a situation by combining positions, pressure, control, and a predicted density of future positions. Other interesting metrics deal with the quantification of space and its correlation with scoring opportunities (Spearman 2018) and defensive strategies (Fernández et al. 2019). Combining these concepts with Spearman et al. (2017) and Fujimura and Sugihara (2005) allows to identify potential runs of attacking players by maximizing a combination of pass probability, pitch impact and pitch control (Peralta Alguacil et al. 2020). Their approach also allows to compute optimal positions for attacking players and to compare these predictions to historic data.

Estimating future positions of soccer players is another aspect that has been widely investigated. A general problem when learning coordinated movements of several agents, like players in a team, is that trajectories come as unordered sets of individuals. When learning from several games, there is a need to incorporate different teams and players and, consequentially, the model has to work without a natural ordering of the players. Le et al. (2017b, 2017a) learn future positions of players by estimating their roles in a given episode and then use these role assignments to predict future movements given the role. A similar approach has been taken by Zhan et al. (2019, 2018) who also use role assignments to predict future positions. Their results for basketball players are computed using a variational recurrent neural network that allows for the inclusion of macro goals. Similarly, Felsen et al. (2018) compute tree-based role assignment to predict trajectories using a conditional variational autoencoder.

In general, graph representations suggest themselves to model interactions of players and ball. In one way or another, players and ball are identified with nodes in a fully connected graph, where edge weights correspond to their interaction and are learned in the training process. For example, Yeh et al. (2019) propose to leverage graph neural networks (GNN) which are naturally suited to model coordinated behavior because of their invariance to permutations in the input. The authors propose a graph variational recurrent neural network to predict future positions of soccer and basketball players. Similarly, Hoshen (2017) and Kipf et al. (2018) propose graph-related attention mechanisms to learn player trajectories.

Graph neural networks have been widely used to model structured or relational data, see (Battaglia et al. 2018) for an overview. In cases where data is sequential in nature, graph recurrent neural networks (GRNN) have been widely deployed, for example to mix graph representations with recurrent layers (Sanchez-Gonzalez et al. 2018), such as gaited recurrent units (GRU, Cho et al. 2014).

Due to the complex nature of movements in soccer, a natural assumption on the distribution of future positions is its multi-modality. Trivially, any probabilistic model that aims to predict future positions in team sports needs to reflect the multi-modal nature in some sense. Hence, the previously mentioned contributions (Zhan et al. 2019, 2018; Yeh et al. 2019; Felsen et al. 2018) make use of (conditional) variational models (CVM) with Gaussian emission functions to account for multi-modality in the data. However, Graves (2013) has shown that combining recurrent neural networks (RNNs) with mixture density networks (MDNs) (Bishop 1994) that learn a Gaussian mixture model (GMM) as output distribution, yields good results for spatiotemporal tasks. In fact, Rudolph et al. (2020) recently showed that combining GMM emissions with recurrent graph networks performs on par with more complex conditional variational models.

3 Graph recurrent action rate models

3.1 Preliminaries

In this paper, we focus on two-dimensional representations of players and ball, given by their x/y positions on the pitch at a given time t . We enumerate players by the set

$\mathcal{I} = \{1, \dots, 22\}$ and reserve index 0 for the ball such that the position and additional attributes of player $i \in \mathcal{I}$ and the ball at time t are given by $x_t^i \in \mathbb{R}^d$ and $x_t^0 \in \mathbb{R}^d$, respectively. Dimensionality $d \geq 2$ depends on the number of additional attributes such as whether the player is currently in ball possession. The set containing positions of all players and ball at time t is denoted by $x_t = \{x_t^0, x_t^1, \dots, x_t^{22}\}$. Since soccer is not a static game, we furthermore focus on consecutive timesteps (episodes) and write $x_{:,t}^i = x_0^i \dots x_t^i$ to denote the trajectory of positions of player i until time t . Analogously to x_t , we denote the set of trajectories of all players and the ball until time t by $x_{:,t} = \{x_{:,t}^0, x_{:,t}^1, \dots, x_{:,t}^{22}\}$. Additionally, we distinguish the ball possessing player by a superscript b and denote her position by x_t^b .

We aim to devise a model that predicts (i) the next action of the ball possessing player b as well as (ii) the time to that action. Since these outcomes are not independent, the model needs to simultaneously predict all possible futures and requires a ground-truth that differentiates between action (y_t^A) and time to this action (y_t^T). Every episode $x_{:,t}$ in the training set is therefore annotated with label pair $y_t = (y_t^T, y_t^A)$ that is defined as follows.

The label $y_t^T \geq 0$ encodes the time until the player b loses control of the ball, either by playing a pass, losing the ball to another player, or because the referee interrupts the game, e.g. due to a foul. The label $y_t^A \in \mathbb{R}^{11}$ distinguishes between passes to teammates and loss of ball. Note that there are several reasons for a loss of ball, including tackles, failed dribbles, and sheer bad luck. Thus, y_t^A is represented as an 11-dimensional vector where every element corresponds to one outcome. If the ball is successfully passed to teammate $k \in \{1, \dots, 10\}$, we have $y_t^A[k] = 1$ and $y_t^A[i] = 0$ for all $i \neq k$. In case of an unsuccessful pass attempt, the notation allows for $y_t^A[k] \in [0, 1]$ to quantify the probability that teammate $k \in \{1, \dots, 10\}$ was the intended receiver of that pass.¹ If the player loses the ball, the action is indicated by $y_t^A[11] = 1$ and $y_t^A[i] = 0$ for $1 \leq i \leq 10$. It follows that $\mathbf{1}^\top y_t^A = 1$ whenever a pass is played or the ball is lost.

In case the referee interrupts the game, e.g. due to a foul, none of the above actions can be performed and the action label is $y_t^A = \mathbf{0}$.

3.2 Player action rate models

We now present our main contribution, a graph recurrent action rate model (ARM) that predicts the next action and simultaneously estimates the time to this action. To account for the complex nature of the problem, we pursue a model that ranks all possible futures and commits to the most likely one.

We begin by assuming that the time T to an action follows an exponential distribution with rate parameter f . To be able to contextualize the distribution of time to an actual situation on the pitch, we employ a positive-valued θ -parameterized *rate function* $f(x_{:,t}; \theta) \in \mathbb{R}_+$ instead of a constant rate. The rate function thus forecasts

¹ We elaborate on how to quantify these probabilities in Appendix A.

the rate of an action at time t for an episode $x_{:t}$ and changes the distribution of time to an event to

$$p(T; f(x_{:t}; \theta)) = f(x_{:t}; \theta) \exp \{ -f(x_{:t}; \theta)T \}.$$

We deploy a dedicated rate model for every player action k , where k ranges from a pass to one of the $k \in \{1, \dots, 10\}$ teammates to a loss of the ball $k = 11$. We indicate the respective action by a superscript and call

$$f^{1:11} = (f^1, \dots, f^{11})^\top = \text{ARM}(x_{:t})$$

the *action rate model*. Accordingly, the time to an action k is governed by an exponential distribution with pdf

$$p(T^k; f^k(x_{:t}; \theta)) = f^k(x_{:t}; \theta) \exp \{ -f^k(x_{:t}; \theta)T^k \}. \quad (1)$$

and the expected time until action k is given by $E[T^k] = 1/f^k(x_{:t}; \theta)$.

Our approach assumes that the rates of actions are mutually independent given the parameters θ . This is a sensible assumption since an open player k in a game situation x_t should receive a high rate $f^k(x_{:t}; \theta)$ and this rate should be independent of whether she is the only available open player or not. However, the *likelihood* that player k actually receives the pass depends on other available pass options and hence the likelihood of the next action being action k is the probability that action k occurs *before* all other actions $i \neq k$

$$p(k | f^{1:11}(x_{:t}; \theta)) = p(T^k < T^{i \neq k} | f^{1:11}(x_{:t}; \theta)) = \frac{f^k(x_{:t}; \theta)}{\sum_{i=1}^{11} f^i(x_{:t}; \theta)}. \quad (2)$$

Note that the r.h.s. of Eqn. 2 is the softmax function of $\log f^{1:11}$. It follows that our model reduces to standard softmax classification of the next action if we ignore the time to the next action. The time until the next action depends on all action rates and is given by the expectation of the exponential distribution with rate parameter $\sum_{i=1}^{11} f^i(x_{:t}; \theta)$,

$$\begin{aligned} E[T^{\min} = \min\{T^{1:11}\} | f^{1:11}(x_{:t}; \theta)] &= E \left[p(T^{\min}; \sum_{i=1}^{11} f^i(x_{:t}; \theta)) \right] \\ &= \frac{1}{\sum_{i=1}^{11} f^i(x_{:t}; \theta)}. \end{aligned} \quad (3)$$

3.3 Deriving the loss function

To finally learn the parameters θ of the action rate models, we need to derive appropriate loss functions that allow for a meaningful training process. In the following, we differentiate three cases and show that all three give rise to the same loss.

Case I: The first case ignores bad passes and non-actions such as stop of play and focuses on well-defined actions. Let $x_{:,t}$ be an episode that ends in action k , i.e. we have $y_t^A[k] = 1$ and $y_t^A[i] = 0, \forall i \neq k$. Recall that the second label y_t^T denotes the observed time until action. Using the results from the previous section and exploiting the memorylessness of the exponential distribution shows that the negative log-likelihood of event k can be written as

$$\begin{aligned} & -\log \left[p \left(y_t^A, y_t^T \mid f^{1:11}(x_{:,t}; \theta) \right) \right] \\ &= -\log \left[p \left(k \mid y_t^T; f^{1:11}(x_{:,t}; \theta) \right) \times p \left(y_t^T \mid f^{1:11}(x_{:,t}; \theta) \right) \right] \\ &= -\log \left[\frac{f^k(x_{:,t}; \theta)}{\sum_{j=1}^{11} f^j(x_{:,t}; \theta)} \right] - \log \left[\sum_{j=1}^{11} f^j(x_{:,t}; \theta) \exp \left\{ -\sum_{i=1}^{11} f^i(x_{:,t}; \theta) y_t^T \right\} \right] \quad (4) \\ &= -\log f^k(x_{:,t}; \theta) + \sum_{i=1}^{11} f^i(x_{:,t}; \theta) y_t^T. \end{aligned}$$

Thus, the log-likelihood of an action k being performed exactly T time steps into the future scales linearly with the log-rate of action k . On the other hand, the rates of all actions are penalized proportionally to the time it takes until the action is observed. This is an important characteristic of the score function and implies that low scores of all actions also lower the likelihood of a quick pass.

Case II: We now consider the case of interrupted or bad passes where the ground truth contains a distribution of likely pass receivers with $y_t^A[k] \in [0, 1]; k \in \{1, \dots, 10\}$ and $\sum_{k=1, \dots, 10} y_t^A[k] = 1$, rather than a one-hot encoded true receiver. It follows that turnovers are ignored, so that $y_t^A[11] = 0$. Thus, the target y_t^A is a probability distribution over the teammates and instead of minimizing the negative likelihood of the observed action as above, we propose to minimize the Kullback–Leibler-divergence

$$\text{KL} \left(y_t^A \parallel p \right) = \sum_k y_t^A[k] \log \left[y_t^A[k] / p(k \mid f^{1:11}(x_{:,t}; \theta)) \right]$$

between the ground-truth likelihood for receivers and the action distribution p . Focusing on a single trajectory-label pair $(x_{:,t}, y_t)$, the corresponding loss is given by

$$\begin{aligned}
& KL(y_t^A \parallel p) - \log p\left(y_t^T \mid f^{1:11}(x_{:t}; \theta)\right) \\
&= \sum_k \left[y_t^A[k] \log y_t^A[k] - y_t^A[k] \log p\left(y_t^T \mid f^{1:11}(x_{:t}; \theta)\right) \right] \\
&\quad - \log \left[\sum_{j=1}^{11} f^j(x_{:t}; \theta) \exp \left\{ - \sum_{i=1}^{11} f^i(x_{:t}; \theta) y_t^T \right\} \right] \\
&= c(y_t^A) + \sum_k -y_t^A[k] \log \frac{f^k(x_{:t}; \theta)}{\sum_j f^j(x_{:t}; \theta)} \\
&\quad - \log \left[\sum_{j=1}^{11} f^j(x_{:t}; \theta) \exp \left\{ - \sum_{i=1}^{11} f^i(x_{:t}; \theta) y_t^T \right\} \right] \\
&= c(y_t^A) + \sum_k -y_t^A[k] \log f^k(x_{:t}; \theta) + f^k(x_{:t}; \theta) y_t^T
\end{aligned} \tag{5}$$

The constant $c(y_t^A) = \sum_k y_t^A[k] \log y_t^A[k]$ is not affected by parameters θ and can be dropped from the final equation as it does not change the optimum. Interestingly, for an action k , where $y_t^A[k] = 1$ and $y_t^A[i] = 0$ for $i \neq k$, the loss functions in Eq. (4) and (5) agree.

Case III: Finally, ball possession phases may end without a defined action, e.g. because the referee interrupts the game due to a foul. In these cases, the play stops at time y_t^T before any action could be performed. The likelihood of such a non-action is the likelihood that the expected time of each action happens only *after* y_t^T . Since the complementary cumulative distribution of the exponential distribution is given by

$$p\left(T^k > y_t^T \mid f^k\right) = \exp\{-f^k y_t^T\},$$

it follows that the negative log-likelihood of such an event can be written as

$$-\log \left[p(T^1 > y_t^T \wedge \dots \wedge T^{11} > y_t^T) \right] = \sum_{k=1}^{11} f^k(x_{:t}; \theta) y_t^T, \tag{6}$$

which is again identical to the previous results in Eq. (4) and (5) for $y_t^A = 0$.

Loss function $\mathcal{L}_A$: To sum up, the losses in Eqs. (4), (5), and (6) are identical. We thus define the loss function for the next action which is given by

$$\mathcal{L}_A = \sum_{(x_{:t}, y_t)} \sum_k -y_t^A[k] \log f^k(x_{:t}; \theta) + f^k(x_{:t}; \theta) y_t^T. \tag{7}$$

3.4 Movement model

To accurately predict the next action on the pitch, the model needs to predict the position of the player in ball possession in the near future. Since it can be beneficial

to run with the ball instead of passing the ball immediately, either to open up new passing lanes or to simply gain space, our model estimates where the player runs to in the near future. This way, in addition to predicting which action will be observed next and when it will be performed, we also predict where it will take place.

We assume that the distribution $p_M(x_{t+T}^b | T, x_{:t})$ of the future position of the ball possessing player b can be expressed by a mixture density network (MDN, Bishop 1994; Rudolph et al. 2020) for time horizons $T > 0$ seconds. Due to the spatial setup, we deploy a mixture of G two-dimensional Gaussians, leading to the density

$$p_M(x_{t+T}^b | T, x_{:t}; \theta) = \sum_{g=1}^G \pi(g | T, x_{:t}; \theta) \mathcal{N}(x_{t+T}^b | \mu_g(T, x_{:t}; \theta), \sigma_g(T, x_{:t}; \theta) \cdot \mathbb{I}) \quad (8)$$

with mixing weights $\pi(g | T, x_{:t}; \theta) \in [0, 1]$, means $\mu_g(T, x_{:t}; \theta)$ and variances $\sigma_g(T, x_{:t}; \theta) \cdot \mathbb{I}$ of the $1 \leq g \leq G$ Gaussians, where $\mathbb{I}$ denotes the two-dimensional unit matrix. Figure 2 (right) shows a graphical representation.

Loss function $\mathcal{L}_M$: The distribution in Eq. (8) can be learned by minimizing the negative log-likelihood of the trajectory data w.r.t the MDN model. Let again $(x_{:t}, (y_t^T, y_t^A))$ be a trajectory-label pair where a ball possession phase $x_{:t}$ ends after y_t^T time steps with action y_t^A . The log-likelihood of that sample is given by

$$\log p_M(x_{t+y_t^T}^b | y_t^T, x_{:t}; \theta), \quad (9)$$

and, consequentially, the loss function for the movement model that is to be minimized using the entire training set is given by

$$\mathcal{L}_M = - \sum_{(x_{:t}, y_t)} \log p_M(x_{t+y_t^T}^b | y_t^T, x_{:t}; \theta). \quad (10)$$

3.5 Joint optimization

The previous sections introduced loss functions for the next action and player movement, respectively. Naturally, the tasks could be learned separately from one another. However, since the tasks are clearly correlated, it will prove beneficial to address them simultaneously. Thus, we propose to jointly minimize the losses $\mathcal{L}_A$ in Eq. (7) and $\mathcal{L}_M$ in Eq. (10). The joint loss is given by

$$\min_{\theta} \sum_{(x_{:t}, y_t)} \left(\underbrace{\sum_k -y_t^A[k] \log f^k(x_{:t}; \theta) + f^k(x_{:t}; \theta) y_t^T}_{\mathcal{L}_A} - \underbrace{\log p_M(x_{t+y_t^T}^b | y_t^T, x_{:t}; \theta)}_{\mathcal{L}_M} \right). \quad (11)$$

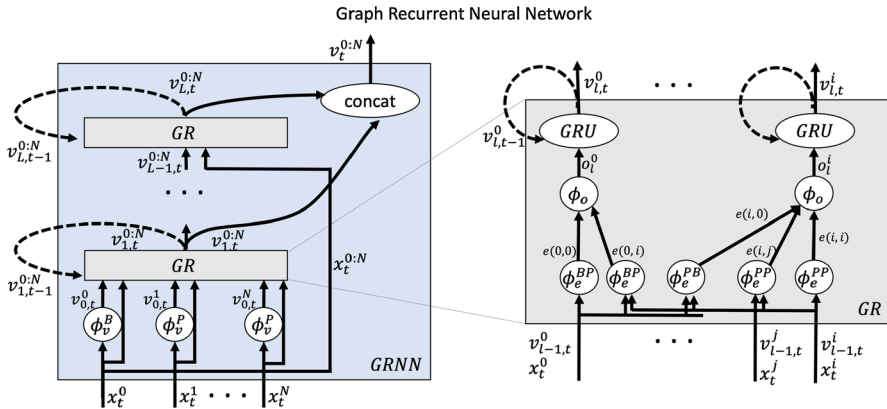


Fig. 1 Left: Graph recurrent neural network. Right: Exemplary layer

3.6 Network architecture

In this section, we describe the full network architecture of our model. To account for inter-dependencies among players and ball, we represent players and ball by a fully connected graph and resort to graph recurrent neural networks (GRNN) (Sanchez-Gonzalez et al. 2018) as the underlying technique. The GRNN computes a contextualization of the situation on the pitch and passes this summary to the action rate model (ARM) component responsible for computing action score functions $f^{1:11}(x_{:,t})$ and to a mixture density network (MDN) that computes the distribution $p_M(x_{t+T}^b | T, x_{:,t}; \theta)$ of future positions.

GRNN: The GRNN part of the network represents the interaction between players using a fully connected graph. Players and ball correspond to nodes in that graph and edges represent their relations. This part of the model is depicted in Fig. 1 (left) and consists of several layers. One layer of the model is shown in Fig. 1 (right).

Recall that x_t^i consists of the position of player/ball i on the pitch, as well as information on the team and ball possession which are one-hot encoded. The input to the layer l are feature vectors $v_{l-1,t}^i$ and x_t^i for $0 \leq i \leq 22$ representing every player and the ball. The layer connects every player/ball i to all other players/ball j via edges $e_t^X(i, j)$, $X \in \{PP, BP, PB\}$. Since we have three types of edges, namely edges between two players (PP), between ball and player (BP), and between player and ball (PB), the edge features are realized by different functions given the type of the edge. All types are computed via an attention function $\alpha_l^X(\cdot; \theta_l)$.

$$e_l^X(i, j) = \phi_e^X(x_t^i, x_t^j, v_{l-1}^j; \theta_l) = \alpha_l^X(x_t^i - x_t^j; \theta_l)^\top v_{l-1}^j$$

$$o_l^i = \phi_o(\{e(i, j) : j \in \{0, \dots, N\}\}) = \sum_j e_l(i, j).$$

The intermediate features o_l^i are fed into a standard gated recurrent unit (GRU) (Cho et al. 2014) to compute the output of the layer as $v_{l,t}^i = \text{GRU}(v_{l-1,t}^i, o_l^i)$. To sum up, the layer of the GRNN shown in Fig. 1 (right) is denoted by

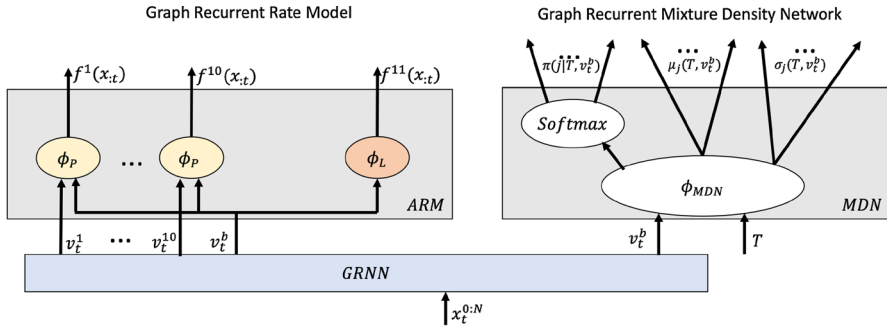


Fig. 2 Left: Depiction of graph recurrent rate model. Right: Graph recurrent mixture density network

$$v_{t,t}^{0:22} = GR(v_{t-1,t}^{0:22}, v_{t,t-1}^{0:22}, x_t^{0:22}; \theta_l).$$

The full GRNN, displayed in Fig. 1 (left), concatenates L such layers.

$$v_t^{0:22} = (v_{1,t}^{0:22}, \dots, v_{L,t}^{0:22}) = GRNN(v_{0,t}^{0:22}, v_{t-1}^{0:22}, x_t^{0:22}; \theta)$$

Inputs $v_{0,t}^i$ to the first layer are computed by a single layer fully connected network $\phi_v(x_t^i)$.

ARM: The action rate model (ARM) layer of the network is responsible for predicting the next action. The idea is to compute the score function $f^{1:11}(x_{:,t}; \theta)$ of the action rate models introduced in Sect. 3.2 from the output of the GRNN. Figure 2 (left) shows a visualization. We have $f^{1:11}(x_{:,t}) = ARM(v_t^{1:N})$ and note that pass actions $k \in \{1, \dots, 10\}$ depend on both the passer b and receiving player k . Consequently, score functions for those actions take both representations v_t^b and v_t^k as input such that $f^k(x_{:,t}; \theta) = \phi_p(v_t^k, v_t^b; \theta)$. By contrast, a lost ball ($k = 11$) is considered a 'local' action and treated as a function of player b only (though information on other players is inherent due to the graph structure), leading to $f^{11}(x_{:,t}; \theta) = \phi_L(v_t^b; \theta)$. Both functions ϕ_p, ϕ_L are implemented as a fully connected layer with a single output and exponential activation function. The desired probabilities $p(k | f^{1:11}(x_{:,t}; \theta))$ and time distribution $p(T^{\min} | f^{1:11}(x_{:,t}; \theta))$ are obtained using Equations (2) and (3), as described in Sect. 3.2. Figure 2 (left) shows a visualization.

MDN: The future distribution of whereabouts of the ball possessing player is computed in the mixture density network (MDN) layer, compare Fig. 2 (right). The subnetwork $\phi_{MDN}(v_t^b, T; \theta)$ takes as input the output v_t^b of the GRNN as well as the predicted time until the next action T . The MDN-part of the network computes the G Gaussians with means $\mu_j(T, x_{:,t}; \theta) \in \mathbb{R}^2$ and variances $\sigma_j(T, x_{:,t}; \theta) > 0$, $j = 1, \dots, G$ and, using an additional softmax layer, also the mixing weights. During training, the time until the next action T is simply the ground truth y_t^T from the respective training instance. However, at prediction time, the predicted time until the next action is computed from the output of the action rate model (ARM) using Eq. (3).

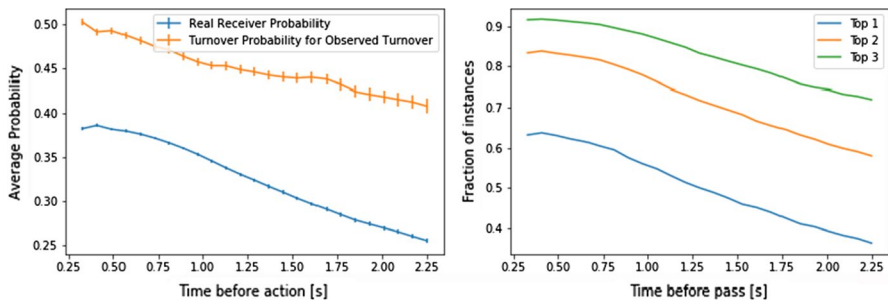


Fig. 3 Left: Predicted probabilities for true turnovers and true receiving players. Right: Accuracies for true player being among the top- k predicted ones

4 Empirical evaluation

The empirical evaluation is conducted on synchronized tracking and event data from 54 Bundesliga matches from the 2017/18 season. The tracking data consists of xy -coordinates of all players and the ball sampled at 25 fps. We extract information on successful and unsuccessful passes from the corresponding event data, including passing and receiving players and timestamps. In case of an unsuccessful pass, the receiving player is either missing, or a player of the opponent team if the pass was intercepted. We will thus impute the missing target player by the approach detailed in Appendix A. We make use of additionally available flags like ball possession markers and indicators of whether the ball is in-play or if play is interrupted by the referee. In total, the data contains 3321 direct turnovers, 2607 unsuccessful passes, and 16,379 successful passes. We empirically evaluate the quality of our model using a fivefold cross validation and report average performance over the five runs.

4.1 Predicting the next action

The first experiment evaluates how well our approach can predict the observed action. Naturally, we expect predictions to become better the closer in time this action is since passing and receiving players will be aligned a moment before the pass is being played while they may run into very different directions for longer look-aheads. As the likelihoods in Fig. 3 (left) show, this insight holds true for passes as well as for turnovers where opponents intercept the ball.

In addition, the right-hand side of the figure demonstrates that the true pass receiver is predicted as most likely action in 60% of all cases exactly 0.75 s before the pass is being played. If we extend the evaluation, to the real action being among the three most likely predicted actions, the accuracy increases to 90%. The same message holds for turnovers. Even one second before an observed turnover, our model predicts a turnover with an average probability of 0.45. Figure 4 (left) summarizes the results for turnovers and shows how often turnovers are correctly predicted as most likely next action. For example, one second before an actual turnover,

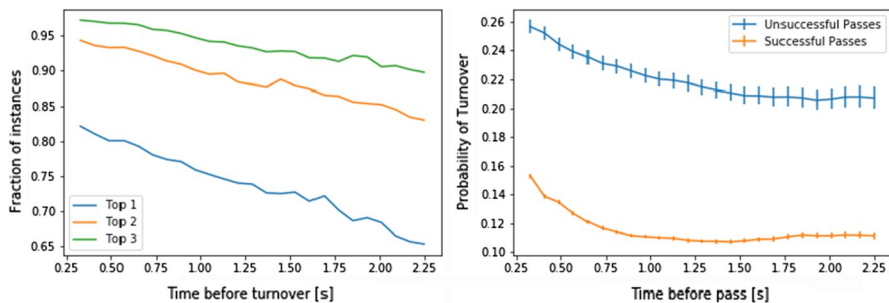


Fig. 4 Left: Accuracies for turnovers being among the top- k predicted actions. Right: Probabilities of a turnover in case of successful and unsuccessful passes

our model predicts a turnover to be the most likely next action in 75% of cases and in 95% of cases it is one of the 3 most likely actions.

Having established the predictive accuracy of the model, we now investigate when and why predictions are likely mistaken. Figure 4 (right) shows an interesting, albeit intuitive, insight. Displayed are predicted probabilities of ball losses in cases where the observed action was a pass. Two things are noteworthy: First, the predicted probability of a turnover increases before the pass. This is to be expected, as usually defenders try to attack the ball carrier (e.g., see also Fig. 7). Secondly, the predicted probability that the ball carrier loses the ball is much higher for unsuccessful passes. This is also not really surprising, as many bad passes are strongly influenced by pressure that is applied to the player by opponents.

4.2 Predicting the point in time of the next action

Figure 5 (left) shows the predicted versus the real observed time until an action. The figure displays a strong correlation between predicted and observed times. However, the predicted times of actions that are less than one second in the future, are generally overestimated whereas times of events that occur further into the future are underestimated. To understand the difficulty of predicting the correct time of the next action, consider the examples in Fig. 5, top right and bottom row. The pictures show three consecutive ball possession phases of the same player within 10 s. The blue team is playing from right to left, the ball position is indicated with a black circle. Numbers and light grey circles indicate the likelihood of a pass to that player according to our model, and the number and circle around the player in possession of the ball indicates the likelihood of a turnover. Orange shades indicates the distribution of positions where the predicted action will take place.

The player first passes to the player to the left of him in direction of play after 2.5 s. The algorithm predicted the next action to be performed after 1.1 s already. The player then receives the ball again (bottom left), and passes after 0.5 s to the player in front of him (algorithm: 0.9 s). The receiving player plays a one-two, and finally passes to the player to the right after 1 s (algorithm: 1.2 s). Thus, the model roughly predicts the next pass to be around 1 s into the future for all three cases,

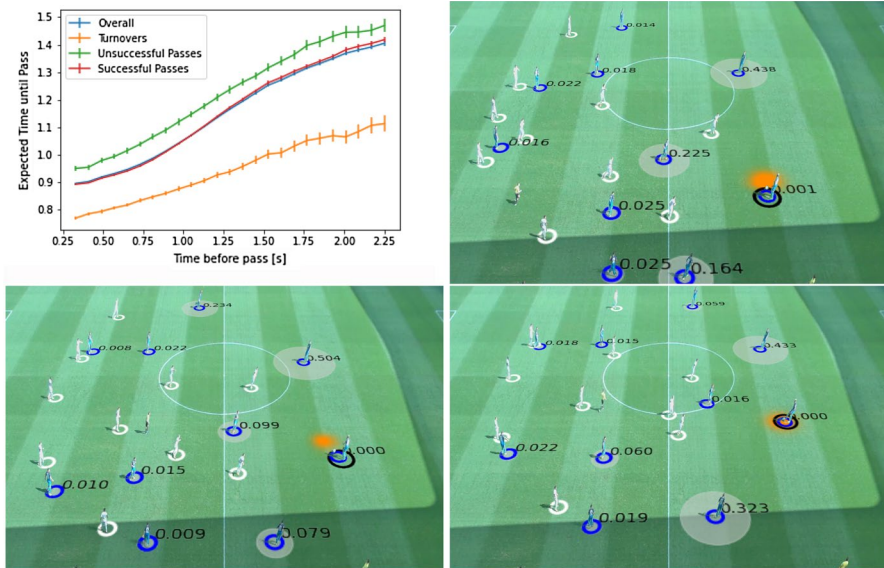


Fig. 5 Top Left: Expected times of the next action in comparison to observed times. Top Right and bottom row: Examples of three consecutive ball possession phases within 10 s. See text for details

whereas the real times were 2.5sec, 0.5sec, and 1sec. Note that our model constantly predicts a pass to the player to the right in all three occasions. Interestingly, all three passes are directly returned by the respective receiving players until the fourth pass is finally the one predicted by the proposed approach and the game continues. We selected this example on purpose as it shows a situation where it is almost impossible to predict the next action even for human experts. Although the predicted player to the right constitutes a good passing option, the near future still depends on individual choices of the ball possessing player.

4.3 Predicting the position of player at time of action

Figure 6 (left) shows an exemplary prediction of future movement of the ball possessing player. The orange area denotes the predicted density of future whereabouts by the model and the light black circle indicates the real observed position of the player at the time of the pass. The model correctly predicts that the player will run for an expected 1.5 s to the light circle until he plays a pass.

The movement model is included by the mixture density network (MDN) component in the optimization. Figure 6 (right) shows that learning the action rate model (ARM) together with the MDN in a single optimization problem leads to smaller overall loss on the test sets, compared to learning each model individually. The reason for is the graph recurrent network (GRNN) that exploits individual traits of the two components in the joint backpropagation steps, where parameters are changed



	MDN Only	ARM Only	ARM + MDN
MDN Loss on Test	2.061	-	2.055
ARM Loss on Test	-	1.159	1.111

Fig. 6 Comparison of MDN and ARM losses on test. The combined model leads to better performance w.r.t. both losses

proportionally to their contribution to the wrong predictions. Since both tasks inter-depend, optimizing both components simultaneously proves beneficial.

Figure 7 shows an exemplary progression of a ball phase that lasts for exactly one second. The images are ordered from top left to bottom right; blue is playing from right to left, the ball position is indicated with a black circle. The displayed sequence starts with a pass to the blue player in the center. Numbers and light grey circles indicate the probability of a pass to that player according to our model, and the number and circle around the ball possessing player indicates the likelihood of a turnover. The predicted movement distribution of this player until the next action is shown in orange.

Just before the player receives the ball in the top right image, the model estimates that the most likely next action is a pass to one of the four indicated teammates, highlighted by larger circles. When the ball is received, the pass to the player on blue's right wing turns out the most likely option. However, the first touch renders a pass to the right wing unlikely but opens another likely passing window to the left wing. Note that, at the same time, the turnover probability increases, mainly because the player decided against the first pass. Nevertheless, as play progresses, the player manages to maintain the ball and pass it on to the left wing which is again the most likely action according to the model.

4.4 Player ranking

4.4.1 Turnovers

The proposed approach could lead to novel key performance indicators to quantify players and teams. In this section we evaluate individual players with respect to their ability to keep possession of ball and continue play in a controlled way while being under pressure from opponents. Additionally, we characterize passing by measuring how 'conventional' or 'expected' their passes are with respect to our model. Note



Fig. 7 Exemplary ball possession phase. See text for details

that the data only comprises a limited amount of games where some teams are over-represented. The results ignore players from underrepresented teams and hence do not reflect the full picture of passing behavior German Bundesliga. In the remainder, we focus on 73 players who had at least 70 ball possessions in the 54 games. These players are affiliated with FC Bayern Munich, TSG 1899 Hoffenheim, Hamburger SV, Borussia Mönchengladbach, FC Schalke 04, and Eintracht Frankfurt.

Table 1 shows the top 20 ranked players with respect to their turnover difference. The latter is given by the difference between predicted turnover probabilities and actually observed turnovers. Larger differences are thus realized by players who often manage to retain the ball in situations where the model predicts a loss of ball.² Note that we only take turnovers into account that end with the opponent team in ball possession.

Bayern Munich clearly dominates the ranking with players like Robert Lewandowski, who ranked 9th in the 2017 Fifa Ballon d'Or vote, Arjen Robben, Kingsley

² Predictions are recorded one second before the end of the ball possession phase.

Table 1 Turnovers

Rank	Player	Turnover difference	Expected turnovers	Observed effective turnovers	Team
1	Max Meyer	0.061	0.131	0.070	FC Schalke 04
2	Robert Lewandowski	0.059	0.389	0.330	FC Bayern Munich
3	Dennis Geiger	0.052	0.142	0.090	TSG 1899 Hoffenheim
4	Lars Stindl	0.048	0.318	0.270	Borussia Mönchengladbach
5	Arjen Robben	0.035	0.283	0.248	FC Bayern Munich
6	Kingsley Coman	0.027	0.325	0.298	FC Bayern Munich
7	Steven Zuber	0.015	0.220	0.206	TSG 1899 Hoffenheim
8	James Rodriguez	0.010	0.202	0.192	FC Bayern Munich
9	Rafinha	0.005	0.124	0.120	FC Bayern Munich
10	Thiago	0.003	0.144	0.140	FC Bayern Munich
11	Javier Martinez	0.002	0.056	0.054	FC Bayern Munich
12	Niklas Süle	0.001	0.041	0.041	FC Bayern Munich
13	Sebastian Rudy	−0.002	0.118	0.120	FC Bayern Munich
14	Corentin Tolisso	−0.002	0.127	0.129	FC Bayern Munich
15	David Alaba	−0.004	0.149	0.153	FC Bayern Munich
16	Arturo Vidal	−0.007	0.119	0.126	FC Bayern Munich
17	Lewis Holtby	−0.019	0.213	0.232	Hamburger SV
18	Franck Ribery	−0.021	0.278	0.299	FC Bayern Munich
19	Joshua Kimmich	−0.028	0.136	0.164	FC Bayern Munich
20	Denis Zakaria	−0.030	0.226	0.256	Borussia Mönchengladbach

Coman, and James Rodriguez. Bayern also won the championship in the 2017/18 season with Schalke being runner-up and Hoffenheim placed third. The remarkable rankings of their youngsters Max Meyer, Dennis Geiger, and Steven Zuber may be surprising at first sight. However, Max Meyer actually became a Germany international in the end of 2016 and Steven Zuber became a regular in the Swiss national team during the 2017/18 season. Dennis Geiger played in the Under19-EuroCup for Germany before the season. Also, Lars Stindl became a Germany international in 2017 and was a starter during their successful 2017 Confederations Cup campaign.

We also like to note that Robert Lewandowski has the highest expected turnovers of all players in the top 20, with Kingsley Coman and Lars Stindl also having values over 0.3. Recall that expected turnover describes the likelihood that the average player will lose the ball and, hence, players who usually operate in tight spaces are expected to rank on top. Table 3 in the Appendix B verifies this assumption.

Table 2 Unexpected passes

Rank	Player	Pass receiver (avg. rank)	Team	Position
1	Kevin-Prince Boateng	2.730	Eintracht Frankfurt	Forward/Midfield
2	Mark Uth	2.152	TSG 1899 Hoffenheim	Forward
3	Max Meyer	1.946	FC Schalke 04	Midfield
4	Guido Burgstaller	1.946	FC Schalke 04	Forward
5	Robert Lewandowski	1.914	FC Bayern Munich	Forward
6	Matti Steinmann	1.907	Hamburger SV	Midfield
7	Sebastian Rudy	1.856	FC Bayern Munich	Midfield
8	Lewis Holtby	1.850	Hamburger SV	Midfield
9	Andrej Kramaric	1.785	TSG 1899 Hoffenheim	Forward
10	Florian Grillitsch	1.755	TSG 1899 Hoffenheim	Midfield
11	Thiago	1.721	FC Bayern Munich	Midfield

4.4.2 Passing

Analogously to turnovers, we could deploy the approach with respect to new metrics for passing behavior. To showcase the idea, we locate the true receiver of a pass in the predicted ranking of all possible receivers, computed 1.5 s before the pass happened. From the viewpoint of the model, likely pass actions have been seen many times in the training data and are distinguished accordingly by high likelihoods. By contrast, *surprising* passes are rarely observed in the data and consequentially range in lower likelihood scores. Thus, averaging the predicted ranks of the true receiving players may serve as a proxy for passing behavior of players: Players with small average ranks usually pass to an obvious team member as predicted by the model while players with large averages include many surprising passes that have not been foreseen by our approach.

Table 2 shows the results. The ranking shows that unexpected passes mainly come from forwards and midfielders. Defenders generally take less risks when playing passes and usually choose a simpler pass over a more risky one. A special mention may go to Kevin-Prince Boateng who, according to our model on limited data, shows a distinguished passing behavior.

5 Conclusion/discussion

We presented an action rate model to predict the next action of football players. The approach consisted of parameterized exponential rate models where parameters are shared across different actions and optimized with a graph network. Empirical results showed that the model reliably predicts future actions. Additional empirical evidence supported the inclusion of movement models in the network architecture.

The model also allowed to analyze and compare playing styles and behavior of players and teams. While over the last few years the amount of per-player and per-team statistics that are collected and curated skyrocketed, those statistics often only tell half the story. For instance, the number of committed turnovers per possession³ is certainly an interesting statistic, however, it may not show the full picture. Consider a striker that usually gets the ball in tight spaces. Unsurprisingly, this player loses the ball more often due to a failed dribble or an unsuccessful pass under pressure, than, for example, a defending player who usually has more time and an open lane for a simple pass. But even among strikers different team tactics and player positions influence the likelihood of a player to lose the ball. Our model may help to better assess the individual behavior and characteristics traits of players by comparing their observed actions⁴ to the expected action of the *average player in the same situation*. Thus, a player who concedes ball possession in 30% should be considered better than average at retaining ball control if the average player would concede the ball in 40% of the time. Other possible applications of the model include assessments of how quickly a player or team passes the ball in comparison to others in the same game situations and how "unconventional" or "surprising" a player passes the ball, i.e. how well the model can predict the player's next action in the future.

Predicting intended receivers

Throughout the manuscript, we have assumed that (soft) labels of observed actions are available in the training set. However, this may not necessarily be the case and even currently available commercial data, such as the data we use in our experiments, does not provide information about the most probable receiver of an intercepted pass. Annotating this information by hand is very time-consuming and expensive and raises the question whether intended receivers of passes can be predicted on the basis of tracking data.

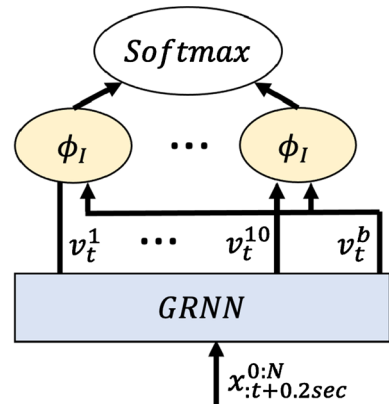
As a solution, we present a data-driven approach for automatically identifying the intended receiver. Although the architecture is very similar to the one presented in Sect. 3.6, this model is not part of the main contribution and cannot be incorporated into Fig. 2 as another output layer as the parameters of the GRNN layer of this network will be very different from the ones of the GRNN in Sect. 3.6. Consequently, the model proposed in this section needs to be trained before the GRNN in Sect. 3.6 to provide estimates for the missing variables in the training data.

Instead of singling out the one most likely receiving player, we aim to predict soft assignments as follows. We define a bad pass to have two possible outcomes: It can either be intercepted by an opposing player or the ball can go out of bounds. To be able to observe the initial trajectory of the ball, we focus on passes where the ball travels for at least $6/25$ s (6 frames) before it is intercepted or out of bounds.

³ <https://statsbomb.com/2018/08/new-data-new-statsbomb-radars/>.

⁴ We consider losing the ball an "action", like passing the ball.

Fig. 8 Model for predicting intended receivers



Training data consists of successful passes, where the intended receiver is the actually observed one. Consider a pass at time t_p and let $y_{t_p}^A$ be the real receiver of that pass as described in Sect. 3.1. This pass is turned into a training instance $(x_{t_p+0.2sec}, y_{t_p}^A)$ and added to the training set. Thus, we consider examples of successful passes where positions and trajectories of players and ball are sampled 0.2 s *after* the ball carrier hit the ball. By doing so, the model learns to predict the receiving player by the initial ball trajectory and the positioning of all other players.

Figure 8 shows an overview of the model. Analogously to the ARM model of Sect. 3.6, let $\phi_I(v_t^k, v_t^b)$ be a score function of outputs of potential receiver k and passer b . The model minimizes the cross-entropy loss between real labels and scores. Note that this is identical to the ARM model, as noted in the remark in Sect. 3.2, when we ignore the time of the pass (which lies in the past according to the data generation process described above).

To determine optimal hyperparameters (e.g., number and width of layers) we perform a 5 fold cross-validation where data is split along games. The final model constitutes of 2 graph layers of width 256 and predicts the real receiver with an average accuracy of 0.89 ± 0.03 . As an out-of-sample test, we manually annotated all unsuccessful passes in one game (100 passes) with the intended receiver using video footage. On these 100 bad passes, the model performed similarly and achieved an accuracy of 0.87 ± 0.1 .

Expected turnover and player position

Table 3 shows expected turnover and the main position of the player. The table shows that large expected turnover values are generally predicted for forwards as they usually operate in tight spaces. Hence, defenders naturally realize smaller values.

Table 3 Expected turnover and player position

Rank	Player	Expected Turnovers	Team	PlayingPosition
1	Guido Burgstaller	0.429	FC Schalke 04	Forward
2	Filip Kostic	0.412	Hamburger SV	Forward
3	Robert Lewandowski	0.389	FC Bayern Munich	Forward
4	Mark Uth	0.373	TSG 1899 Hoffenheim	Forward
5	Raffael	0.350	Borussia Mönchengladbach	Forward
6	Thomas Müller	0.338	FC Bayern Munich	Forward
7	Tatsuya Ito	0.332	Hamburger SV	Forward/Midfield
8	Thorgan Hazard	0.331	Borussia Mönchengladbach	Forward
9	Amine Harit	0.330	FC Schalke 04	Forward
10	Kingsley Coman	0.325	FC Bayern Munich	Forward
11	Lars Stindl	0.318	Borussia Mönchengladbach	Forward
12	Aaron Hunt	0.310	Hamburger SV	Forward
13	Andrej Kramaric	0.299	TSG 1899 Hoffenheim	Forward
14	Marius Wolf	0.286	Eintracht Frankfurt	Forward/Midfield
15	Arjen Robben	0.283	FC Bayern Munich	Forward
16	Franck Ribery	0.278	FC Bayern Munich	Forward

Acknowledgements We would like to thank Hendrik Weber as well as DFL and Sportec Solutions for providing the data for this study.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Battaglia, P.W., Hamrick, J.B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., et al.: Relational Inductive Biases, Deep Learning, and Graph Networks. arXiv preprint [arXiv:1806.01261](https://arxiv.org/abs/1806.01261) (2018)
- Bishop, C.M.: Mixture Density Networks. Tech. Rep. NCRG_94_004, Aston University (1994), https://publications.aston.ac.uk/id/eprint/373/1/NCRG_94_004.pdf
- Brefeld, U., Lasek, J., Mair, S.: Probabilistic movement models and zones of control. Machine Learning, Special Issue on Soccer Analytics **108**, 127–147 (2019)
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder–decoder for statistical machine translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).

- pp. 1724–1734. Association for Computational Linguistics, Doha, Qatar (Oct 2014). <https://doi.org/10.3115/v1/D14-1179>, <https://www.aclweb.org/anthology/D14-1179>
- Dick, U., Brefeld, U.: Learning to rate player positioning in soccer. *Big Data, Special Issue on Sports Analytics* 7(1), 71–82 (2019)
- Felsen, P., Lucey, P., Ganguly, S.: Where Will They Go? Predicting Fine-Grained Adversarial Multi-Agent Motion using Conditional Variational Autoencoders. In: *European Conference on Computer Vision*, pp. 732–747 (2018)
- Fernández, J., Bornn, L., Cervone, D.: Decomposing the Immeasurable Sport: A Deep Learning Expected Possession Value Framework for Soccer. In: *MIT Sloan Sports Analytics Conference* (2019)
- Fujimura, A., Sugihara, K.: Geometric analysis and quantitative evaluation of sport teamwork. *Systems and Computers in Japan* 36(6), 49–58 (2005). <https://doi.org/10.1002/scj.20254>, <https://onlinelibrary.wiley.com/doi/abs/10.1002/scj.20254>, _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/scj.20254>
- Goes, F.R., Kempe, M., Meerhoff, L.A., Lemmink, K.A.: Not every pass can be an assist: a data-driven model to measure pass effectiveness in professional soccer matches. *Big data* 7(1), 57–70 (2019)
- Graves, A.: Generating Sequences With Recurrent Neural Networks. *arXiv preprint arXiv:1308.0850* (2013)
- Hoshen, Y.: Vain: Attentional Multi-Agent Predictive Modeling. In: *Advances in Neural Information Processing Systems*, pp. 2701–2711 (2017)
- Kipf, T., Fetaya, E., Wang, K.C., Welling, M., Zemel, R.: Neural Relational Inference for Interacting Systems. In: *International Conference on Machine Learning*, pp. 2693–2702 (2018)
- Le, H.M., Carr, P., Yue, Y., Lucey, P.: Data-Driven Ghosting Using Deep Imitation Learning. In: *MIT Sloan Sports Analytics Conference* (2017)
- Le, H.M., Yue, Y., Carr, P.: Coordinated Multi-Agent Imitation Learning. In: *International Conference on Machine Learning*, pp. 1995–2003 (2017)
- Link, D., Lang, S., Seidenschwarz, P.: Real time quantification of dangerousity in football using spatiotemporal tracking data. *PloS one* 11(12), e0168768 (2016)
- Peralta Alguacil, F., Fernandez, J., Piñones Arce, P., Sumpter, D.: Seeing in to the future: using self-propelled particle models to aid player decision-making in soccer. In: *Proceedings of the 14th MIT sloan sports analytics conference* (2020)
- Power, P., Ruiz, H., Wei, X., Lucey, P.: Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data. In: *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1605–1613 (2017)
- Rudolph, Y., Brefeld, U., Dick, U.: Graph conditional variational models: Too complex for multiagent trajectories? In: *ICBINB@NeurIPS workshop* (2020)
- Sanchez-Gonzalez, A., Heess, N., Springenberg, J.T., Merel, J., Riedmiller, M., Hadsell, R., Battaglia, P.: Graph networks as learnable physics engines for inference and control. In: *ICML* (Jun 2018), <http://arxiv.org/abs/1806.01242>, *arXiv: 1806.01242*
- Spearman, W.: Beyond expected goals. In: *Proceedings of the 12th MIT sloan sports analytics conference*, pp. 1–17 (2018)
- Spearman, W., Basye, A., Dick, G., Hotovy, R., Pop, P.: Physics-based modeling of pass probabilities in soccer. In: *Proceeding of the 11th MIT Sloan Sports Analytics Conference* (2017)
- Yeh, R.A., Schwing, A.G., Huang, J., Murphy, K.: Diverse Generation for Multi-Agent Sports Games. In: *IEEE Conference on Computer Vision and Pattern Recognition* (2019)
- Zhan, E., Zheng, S., Yue, Y., Lucey, P.: Generative Multi-Agent Behavioral Cloning, p. 10 (2018)
- Zhan, E., Zheng, S., Yue, Y., Sha, L., Lucey, P.: Generating Multi-Agent Trajectories using Programmatic Weak Supervision. In: *International Conference on Learning Representations* (2019)