Learning to rate actions in multi-agent scenarios

Uwe Dick and Ulf Brefeld

Leuphana University of Lüneburg, Germany
{uwe.dick,ulf.brefeld@leuphana.de}

1 Introduction

We investigate how to learn functions that rate game situations on a soccer pitch according to their potential to lead to successful attacks. We follow a purely data-driven approach using techniques from deep reinforcement learning to valuate multi-player positionings based on positional data.

From a conceptual point of view, the execution of a successful (good) attack pattern should raise the likelihood of scoring a goal. That is, the likelihood of scoring before the execution of the pattern should be smaller than afterwards. We follow this conceptual line and explore likelihoods of 'being successful' from arbitrary game settings. We argue that, by being able to derive such valuations, we will be able find good (and bad) attacking patterns by measuring differences in likelihoods.

2 Contribution

We use tracking data consisting of a sequence of x/y-coordinates of all players and the ball for a set of soccer games, sampled at 2 frames per second [1, 2]. The data is split into episodes of continuous ball possessions of either team. Every episode is assigned a binary target value that is 1 if the ball possessing team carries the ball into the final 25m of the opponent's half and 0 otherwise.

The goal is to compute a value function that maps an arbitrary positioning of players and ball to its expected return, that is the average target value of all possible subsequent episodes. We take a Deep Reinforcement Learning-based approach and learn the value function with convolutional neural networks (CNNs) [4, 3, 5]. The CNN treats every positioning of players and ball as a 2D image with nine channels such that the resulting 3D tensor encodes the positions as well as the velocities of the players and the ball and a ball possession indicator. The model is then optimized with stochastic gradients using backpropagation.

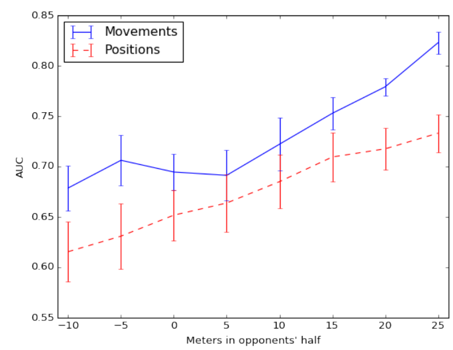

Figure 1: Averaged AUC values and standard errors.

3 Empirical Results

We extract episodes from 10 soccer games. The binary target variables render the evaluation a binary ranking problem and we use the area under the ROC curve (AUC) to quantify the quality of the scoring functions. We report on leave-one-game-out cross validation results.

The AUC scores are shown in Figure 1. Our approach is able to learn value functions of player posi-

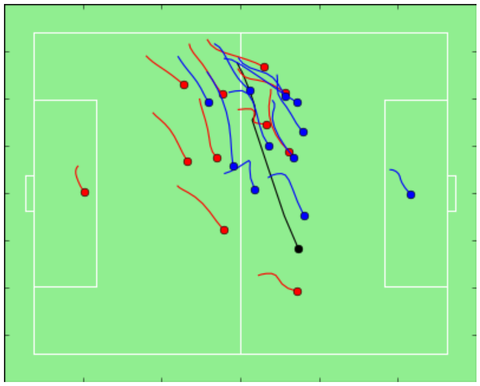

Figure 2: Example

tioning for all areas of the pitch. Not surprisingly, the closer the ball possessing team is to the opponent's goal, the higher the average AUC. The closer a team gets to the dangerous zone, the less chance influences entering into it. Figure 2 shows a pattern for which the learned valuation changed the most. Red broadens the game with a long diagonal ball to an open player on the right wing.

References

[1] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews. Large-scale analysis of soccer matches using spatiotemporal tracking data. In *Data Mining (ICDM), 2014 IEEE International Conference on*, pages 725–730. IEEE, 2014.

[2] P. Lucey, A. Bialkowski, M. Monfort, P. Carr, and I. Matthews. quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. In *Proc. 8th Annual MIT Sloan Sports Analytics Conference*, pages 1–9, 2014.

[3] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International Conference on Machine Learning*, pages 1928–1937, 2016.

[4] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[5] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.