Quantifying Pressing Effectiveness and Its Impact on Formations in Football

Kenneth Ssekimpi

Table of contents

Glossary

Abbreviations

AI Artificial Intelligence

CNN Convolutional Neural Network

GAN Generative Adversarial Network

CNN Convolutional Neural Network

GAN Generative Adversarial Network

GRU Gated Recurrent Unit

LSTM Long Short-Term Memory

MLP Multi-Layer Perceptron

RBM Restricted Boltzmann Machine

TCN Temporal Convolutional Network

VAE Variational Autoencoder

WD Wasserstein-Disrtance

cVAE Conditional Variational Autoencoder

GAIN Generative Adversarial Imputation Networks

RMS(L)E Root Mean Squared Error

G-means G-means

NLL Negative Log-Likelihood

MCC Matthews Correlation Coefficient

Parameters

Fe Iron (mg/L)

Al Aluminum (mg/L)

Nomenclature

J cost/loss function

∇ derivative of cost w.r.t. model parameters

w model parameters

b model parameters

x input data

y output data

μ mean

σ^2 variance

Terminology

gradient derivative of cost w.r.t. model parameters

optimizer algorithm used to minimize the cost function

batch subset of training data used in one iteration

epoch one pass through the entire training dataset

overfitting model is memorizing the training data and not generalizing well to new data

underfitting model is not able to learn the underlying patterns in the data

regularization optimisation by adding a penalty term to the loss function

dropout optimise by randomly dropping out neurons during training

1 Introduction

1.1 Background

Association football, also known as soccer, is a game played by two teams of 11 players each. Teams compete by advancing a ball into their opponent's goal, adhering to established rules that govern gameplay, player conduct, and scoring, with the aim to score more goals than their rivals (Sumpter, 2016; Memmert & Raabe, 2018). Although the fundamental simplicity of football contributes significantly to its global popularity and appeal, the game simultaneously possesses incredible complexity characterised by movement patterns, match plans, playing philosophies, and creativity (Bornn, Cervone & Fernandez, 2018; Vicente et al., 2024). These qualities are collectively referred to as football tactics (Memmert & Raabe, 2018).

Football tactics involve strategically positioning players on the field and co-ordinating their movements to maximise the chances of winning matches. This encompasses both the formation that a team adopts (that is, the spatial arrangement of players on the pitch) and their overall style of play (Wilson, 2010). Furthermore, Rein & Memmert (2016) describe football tactics as the actions and strategies implemented by a team and its players during a match to achieve specific goals, primarily winning the game. These actions are typically adaptations to dynamically-changing situations in the match and the behaviour of the opposing team, managing space, time, and individual actions on the pitch.

Fundamental to understanding these tactical deployments are the distinct player positions and the common tactical formations they operate within. Football teams are typically composed of players in four main categories: Goalkeepers (GK), primarily responsible for preventing goals; Defenders (DEF), who protect their goal and prevent opposing attacks (e.g., Centre-Backs, Full-Backs); Midfielders (MID), who link defence and attack and control central areas (e.g., Defensive Midfielders, Attacking Midfielders, Wingers); and Forwards/Strikers (FWD), whose main role is to score goals. These positions are strategically arranged into tactical formations, which dictate a team's general shape and approach during different phases of play. A phase of play in football refers to a segment of a match, typically defined for analysis or coaching purposes, characterised by the actions and organisation of one or both teams (Ghezzi & Sotudeh, 2024). It can be understood as a sequence of consecutive events or actions that fit together (Decroos, 2020)

References

Bornn, L., Cervone, D. & Fernandez, J. 2018. Soccer analytics: Unravelling the Complexity of "The Beautiful Game". *Significance*. 15(3):26–29.

Decroos, T. 2020. Soccer Analytics Meets Artificial Intelligence: Learning Value and Style from Soccer Event Stream Data. PhD thesis. KU Leuven. Available: https://tomdecroos.github.io/reports/thesis_tomdecroos.pdf.

Ghezzi, E. & Sotudeh, H. 2024. Hudl StatsBomb conference 2024. Old Trafford, Manchester, United Kingdom: StatsBomb. 13. Available: https://statsbomb.com/wp-content/uploads/2024/10/Match-Phases-In-Practice-Ghezzi-and-Sotudeh.pdf.

Memmert, D. & Raabe, D. 2018. *Data Analytics in Football: Positional Data Collection, Modelling and Analysis*. 1st ed. Taylor & Francis. Available: https://books.google.co.za/books?id=O9tdDwAAQBAJ.

Rein, R. & Memmert, D. 2016. Big data and tactical analysis in elite soccer: Future challenges and opportunities for sports science. *SpringerPlus*. 5(1):1410. DOI: 10.1186/s40064-016-3108-2.

Sumpter, D. 2016. *Soccermatics: Mathematical Adventures in the Beautiful Game*. 1st ed. Bloomsbury Publishing. Available: https://books.google.co.za/books?id=CoZVCwAAQBAJ.

Vicente, L.N., Alleck, T., Giovannelli, T., Mitchell, R. & Remen, O. 2024. Why is soccer so popular: Understanding underdog achievement and randomness in team ball sports. *arXiv preprint arXiv:2404.06626*.

Wilson, J. 2010. *Inverting the Pyramid: The History of Football Tactics*. 1st ed. Orion Books Ltd. Available: https://books.google.co.za/books?id=qYSAhJn-srwC.