

From Ratings to Rising Stars: Predicting Player Potential Using FIFA Attributes

Jonas Schrade (01/—) & Niklas Bacher (01/1081503)

Introduction

This project uses the European Soccer Database with a specific focus on player-level information. A central component of the dataset is the set of player attributes derived from the FIFA video game series. These attributes summarize players' technical skills, physical characteristics, and overall ability as assessed by professional scouts and analysts. Although the ratings originate from a simulation game, they are widely used as standardized and comparable measures of player quality and have been shown to correlate strongly with real-world performance. As such, they represent an accessible and realistic information source for clubs, analysts, and researchers interested in player evaluation.

From an applied perspective, accurately predicting player performance based on such attributes has clear practical relevance. Clubs can use these predictions to support recruitment decisions, contract negotiations, or talent development, while analysts may rely on them to benchmark players across leagues and seasons. Understanding the predictive value of FIFA attributes therefore has direct implications for real-world decision-making in professional football.

The goal of this project is to assess whether complex models such as deep neural networks provide meaningful improvements over simpler statistical learning approaches when predicting not only current player performance but also future potential. Player potential is inherently a latent concept: it reflects expected development and long-term ability rather than directly observed outcomes. Capturing such nonlinear and potentially complex relationships may favor flexible models like neural networks. To evaluate this, a traditional benchmark model is used as a baseline and compared against a neural network using the same predictors and outcome. This comparison allows us to determine whether deep learning provides added value in identifying players with high potential, or whether simpler models already capture most of the relevant signal contained in the FIFA-based player attributes.

Analysis

Any exploratory analysis of your data, and general summary of the data (e.g. summary statistics, correlation heatmaps, graphs, information about the data...). Tell the reader about the types of variables you have and some general information about them, Plots and/or Tables are always great. This should also include any cleaning and joining you did.

If you want a table you can make one with [this website](#) and paste the markdown table here. For example:

A	B	C	D	E
a	b	c	d	e
a	b	c	d	e
a	b	c	d	e

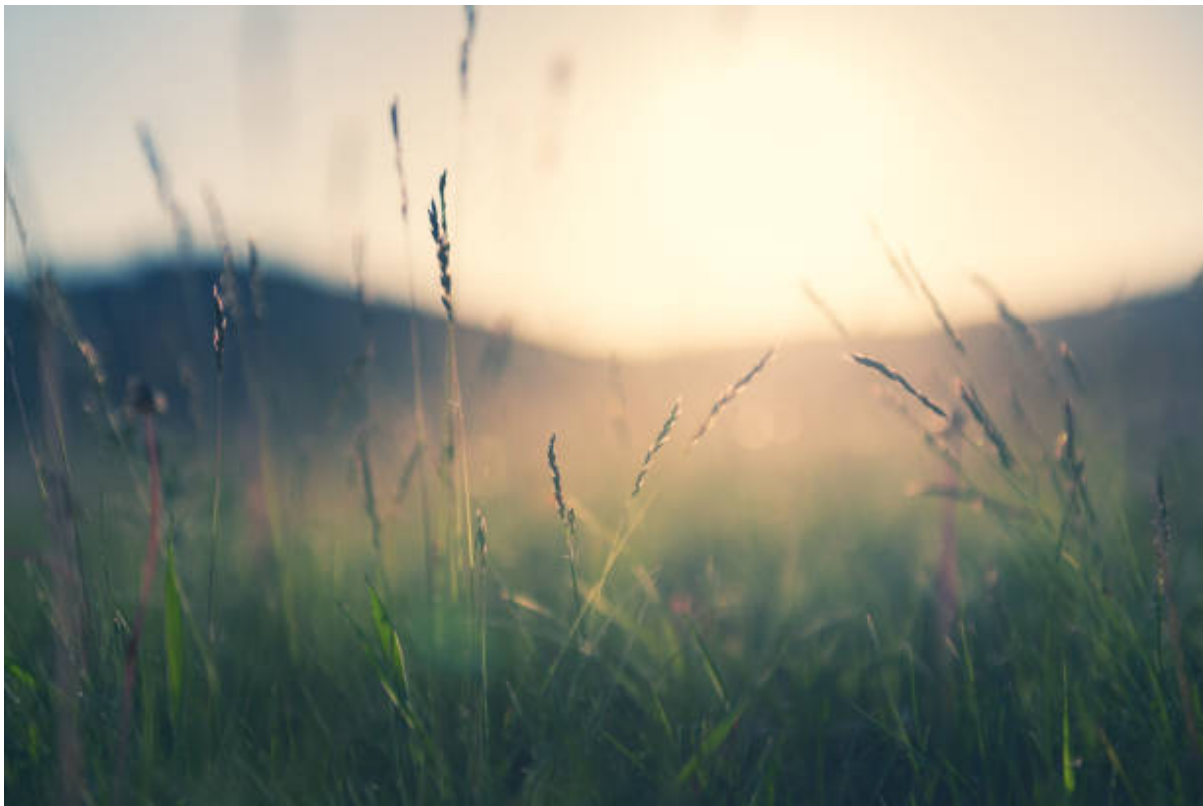


Figure 1: My Caption Here

(Note that the `width=300` argument controls how wide your image will be.)

Methods

Explain the structure of your model and your approach to building it. This can also include changes you made to your model in the process of building it. Someone should be able to read your methods section and *generally* be able to tell exactly what architecture you used. However REMEMBER that this should be geared towards an audience who might not understand R code.

Results

Discussion of how your model performed. Include a discussion about whether or not Deep Learning was necessary in this situation. Note: I do not want you to include large chunks of R output, just summaries that explain your model and its performance sufficiently. One of the marking criteria is whether you can do this in a structured way.

Reflection

Reflections on what you learned/discovered in the process of doing the assignment. Write about any struggles you had (and hopefully overcame) during the process. Things you would do differently in the future, ways you'll approach similar problems in the future, etc.