



Beyond Goals: Can We
Quantifiably Predict a
Player's Attacking
Precision?

Project Objective: The Anatomy of a Prediction

- To build a machine learning model that accurately predicts a player's 'On Target Per Average Match'.
- The model will use other key performance metrics as inputs to create a robust predictive tool.
- The goal is to move beyond simple goal counts to understand and quantify the underlying skill of shooting accuracy.

The Blueprint: Profiling Europe's Top Attackers



660

Player Seasons
Analyzed



15

Performance Variables
Captured
(From `Matches_Played` to `Year`)



**Europe's Top
Leagues**

Data Sourced From
(e.g., Spain La Liga, Netherlands
Eredivisie)

Data Quality Statement: A robust dataset with high integrity. Out of 9,900 data points, only 34
Data Quality Statement: A robust dataset with high integrity. Out of 9,900 data points, only 34
(`Club`) were missing and subsequently excluded, ensuring a clean foundation for the model.

The Story in the Numbers: Defining the Extremes

The Prolific Finisher

37

Goals

Average Player: 11.8 Goals

The Marathon Man

4177

Minutes Played

Average Player: 2071 Minutes

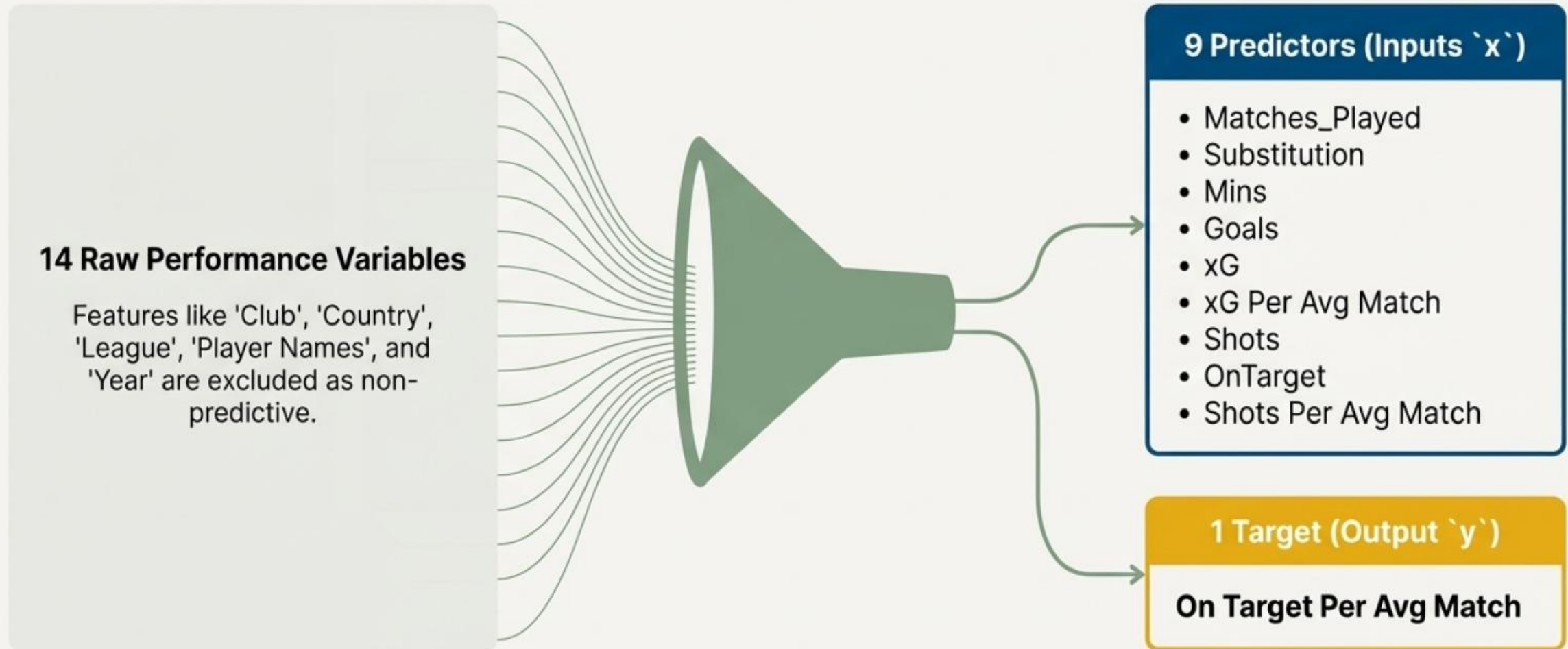
The Relentless Attacker

32.54

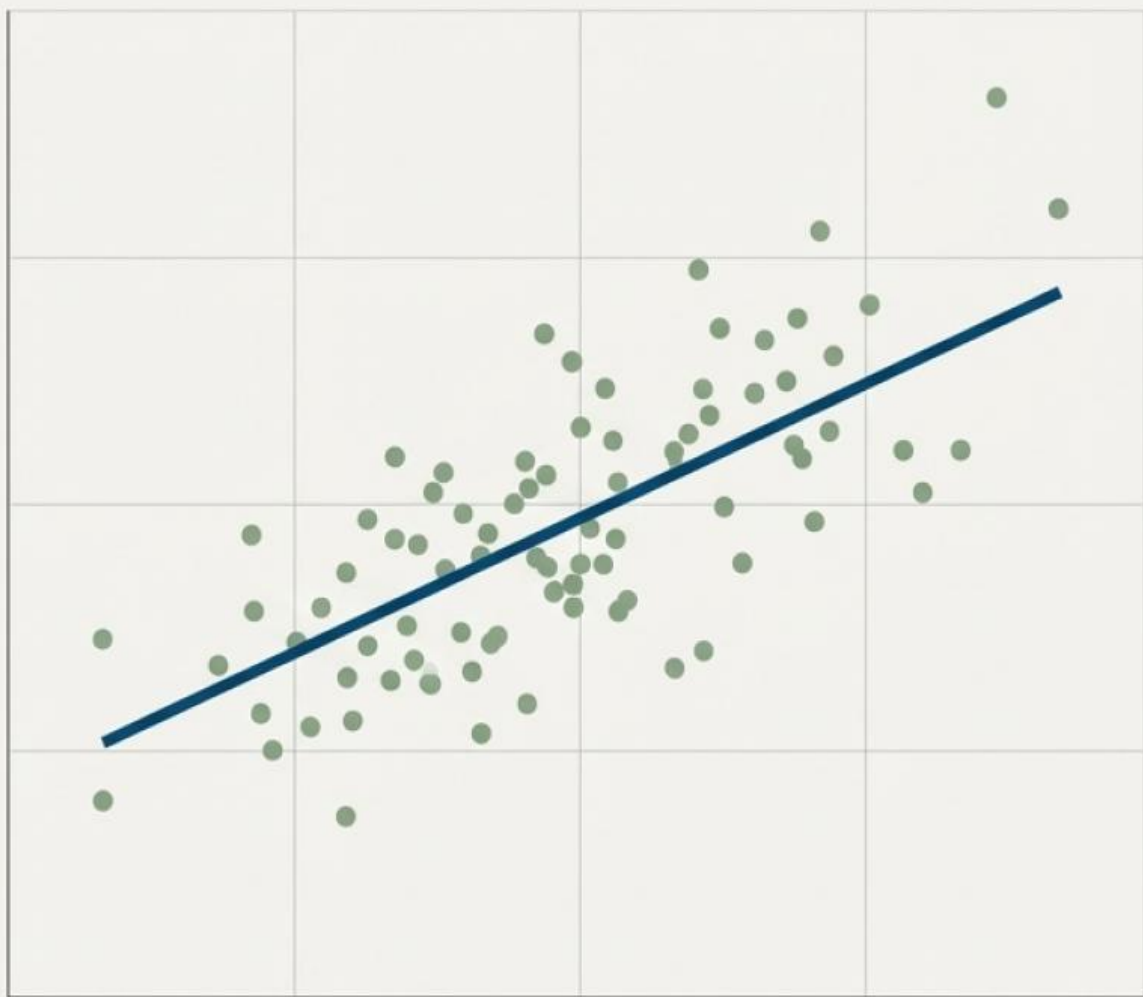
Expected Goals (xG)

Average Player: 10.09 xG

Focusing the Analysis: Selecting Predictors and Defining the Target



The Engine: Applying a Linear Regression Model



We chose Linear Regression—a powerful and interpretable model that finds the straight-line relationship between a set of inputs (player actions) and a desired output (shooting accuracy). Its simplicity makes the results transparent and easy to understand.

Training the Model,

Testing the Truth

80% Training Set
528 Players

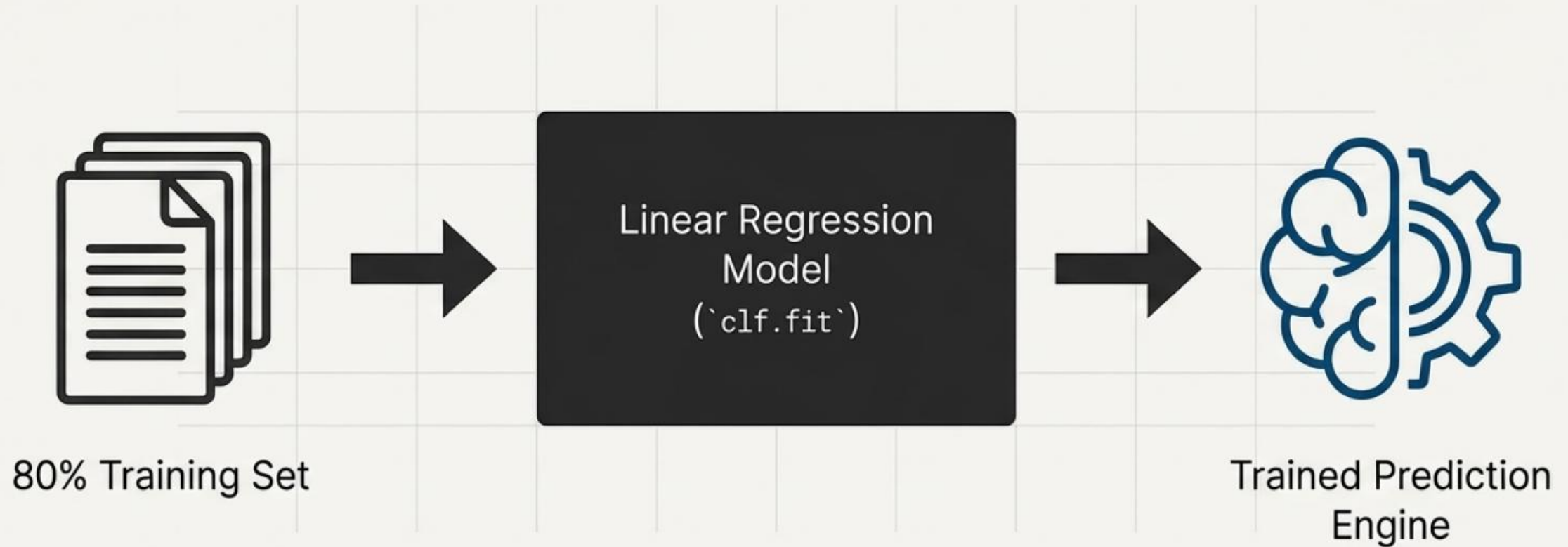
20% Test Set
132 Players

Used to teach the model the patterns between performance metrics and shooting accuracy.

Held back for a final, unbiased evaluation of the trained model.

Methodology: ``train_test_split`` with ``test_size=0.2`` and ``random_state=42`` for reproducible results.

The Learning Process: From Data to Prediction Engine



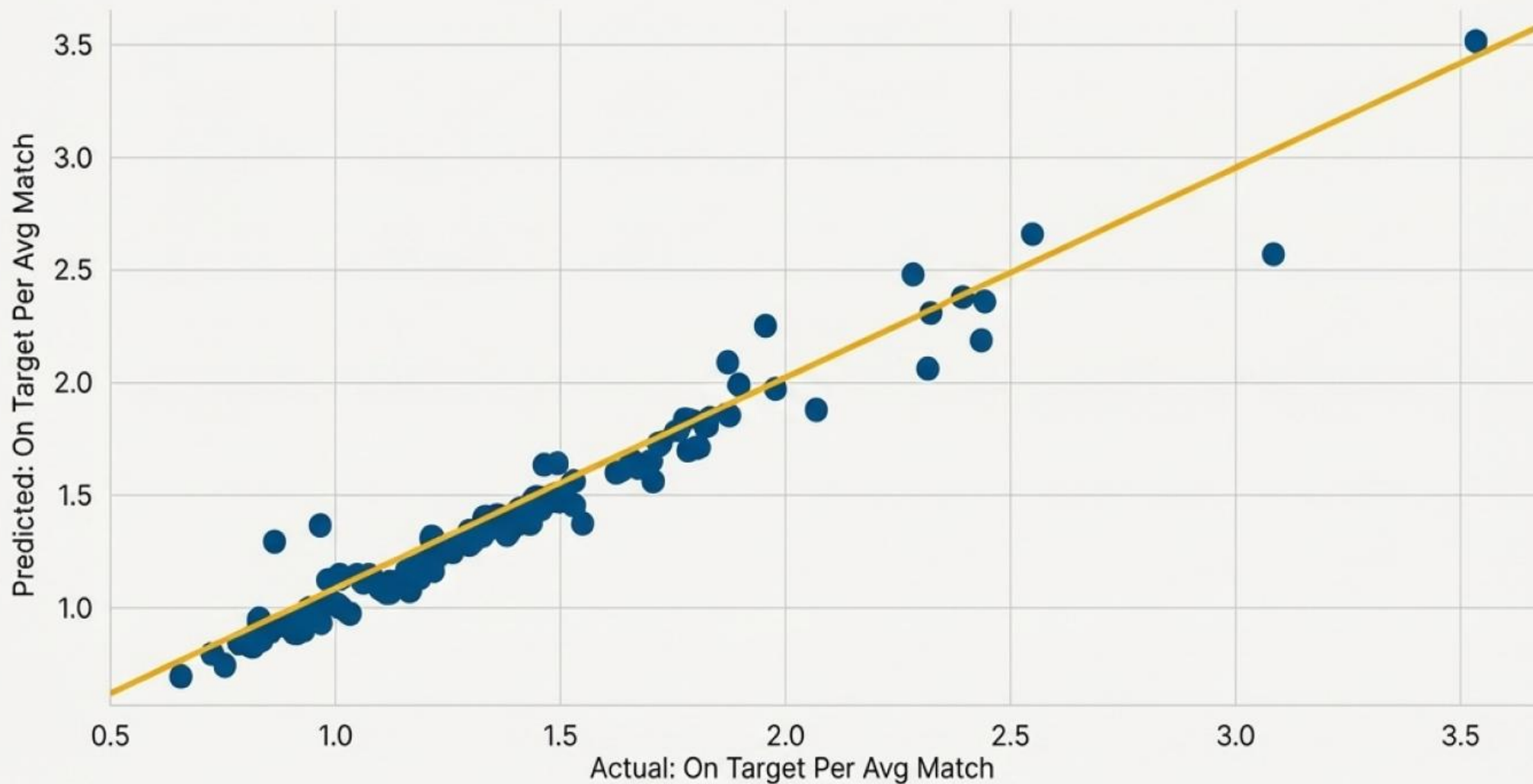
The model analyzes the training data to learn the precise mathematical relationship between the nine predictor variables and the target accuracy metric.

95.65%

R-Squared Score

Our model successfully explains 95.65% of the variance in a player's shooting accuracy using the selected performance metrics. This indicates a near-perfect fit between our model's predictions and the actual data.

Prediction vs. Reality: An Almost Perfect Correlation



Each dot represents a player from the unseen test set. Their close alignment with the diagonal line visually confirms the model's high accuracy.

From Insight to Impact: Potential Applications



Enhanced Player Scouting

Identify players who are highly accurate shooters, even if their goal counts are not yet elite.



Objective Performance Benchmarking

Evaluate player performance against a statistically-backed accuracy model, removing subjective bias.



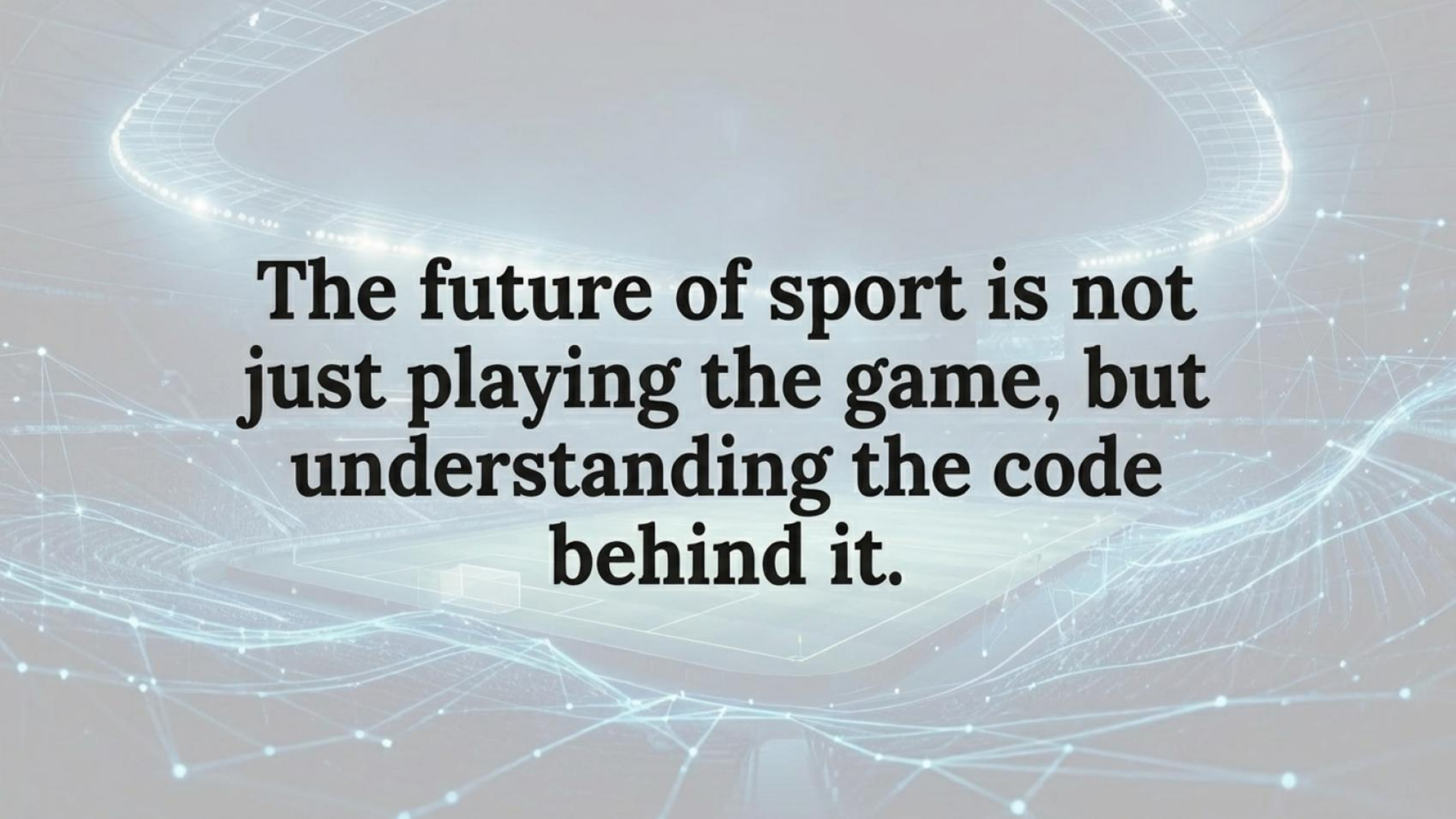
Advanced Fantasy Football Analytics

Create more sophisticated player valuation models based on predictive accuracy, not just past results.



Identifying Over/Under-Performing Players

Pinpoint athletes whose on-target performance is significantly better or worse than their other stats would predict.

The background of the image is a conceptual illustration of a futuristic sports arena. A large, glowing blue wireframe structure represents the stadium's seating and architecture. Overlaid on this is a complex network of thin, light-blue lines and dots, resembling a data or communication network. The overall color palette is cool, with various shades of blue and white against a dark background.

**The future of sport is not
just playing the game, but
understanding the code
behind it.**