

Data Driven Squad Construction and Player Valuation

End to end data science solution combining machine learning and optimization to support high stakes decision making in squad construction.

Overview

This project demonstrates how advanced analytics can be applied to complex, constraint driven decision problems. It combines predictive modeling, optimization, and heuristic methods to replace intuition driven decisions with structured, data informed insights.

Problem Context

Resource allocation decisions often involve uncertainty, tradeoffs, and hard constraints. In talent acquisition scenarios, subjective judgment and market inefficiencies can lead to suboptimal outcomes. This project addresses that challenge by applying data science techniques to identify value, optimize selections, and support strategic decision making.

Approach

The solution follows a multi stage analytical workflow:

- Predictive modeling using machine learning to estimate objective value
- Identification of mispricing by comparing modeled and observed values
- Constrained optimization to select an optimal portfolio under budget and structural constraints
- Heuristic scoring to support tactical decisions and scenario analysis

Techniques and Tools

- Machine Learning: Random Forest, regression techniques
- Optimization: Integer Linear Programming
- Analytics: Feature engineering, model evaluation, heuristic scoring
- Languages and Tools: R, tidyverse, lpSolve, ggplot2

Outcomes

- Demonstrates a production oriented data science workflow
- Shows how machine learning and optimization can be combined for decision support
- Highlights the application of analytics to real world, high impact problems

Repository Structure

- `src/` Core modeling and optimization logic
- `notebooks/` End to end analysis workflows
- `visuals/` Key charts and model outputs
- `report/` Full project documentation and results

Why This Project Matters

This work reflects a senior data science mindset: framing business problems clearly, selecting appropriate modeling techniques, and delivering actionable outputs that support real decisions under constraints.