**Project Starlight: Robot-eye for (Turkish) football player-talent-identification**

Zsolt Kontra (https://orcid.org/0000-0002-9280-0560) - László Pitlik (<https://orcid.org/0000-0001-5819-0319>)

e-Mails: zskontra00@gmail.com, pitlik@my-x.hu / Kodolányi János University and MY-X research team Hungary

Keywords: sport-economics, talent-management, artificial intelligence, optimization, forecasting

**Abstract**

Present status: Moneyball (2011), the American sports drama film can be interpreted as a massive initialization towards data-driven decision making in the sport-economics. Parallel, (inter-)national rankings in all sport disciplines should support the decision making e.g., in talent management (of footballers) and/or in case of sport bets. On the other hand: the talent-management is and will also be a kind of forecast-oriented activity.

Goal/Task: The goal is simple and trivial: creating a robot being better than the human benchmarks concerning talent-management (c.f. identification of football players here and now having a massive potential in the future, where this potential can be interpreted as the trend of the monetary value development of a football player). If the goal is simple, so the task is also trivial: searching for new knowledge representation forms (models) being capable of better filtering young talents than the human experts do it.

Solution: The human solution (the benchmark) can be automated - like the searched AI. The human expert can derive correlation values between the economic value of the focused footballer set and each descriptive phenomenon (like Aggression, Corners, Crossing, Dribbling, Finishing, FirstTouch, Freekicks, Heading, LongShots, Longthrows, Marking, Passing, PenaltyTaking, Tackling, Technique, Anticipation, Bravery, Composure, Concentration, Vision, Decisions, Determination, Flair, Leadership, OffTheBall, Positioning, Teamwork, Workrate, Acceleration, Agility, Balance, Jumping, LeftFoot, NaturalFitness, Pace, RightFoot, Stamina, Strength, Consistency, ImportantMatches, Versatility, Adaptability, Ambition, Loyalty, Pressure, Professional, Sportsmanship, etc.). The source of the data is: <https://www.sigames.com/> / SI Games. The human expert selects a relatively small number of these attributes (e.g., 5) having high correlation levels in the past (year-by-year). Then, the entire set of footballers will be filtered based on age and the level of the selected attributes. Thresholds can be constant or even pre-defined through formulas (like MODUS). The human expert (the talent-fisherman) is good, if the ratio of the players with value-increasing-trend in the selection is high. The own AI should optimize the human-like analytical process.

The goal of the task is therefore to create a robot that is better than human experts at talent management in football, specifically in identifying young football players with a high potential for future monetary value development. The task involves searching for new knowledge representation forms (models) that can better filter young talent than human experts. The human solution, which serves as the benchmark, involves selecting a relatively small number of attributes that have high correlation levels with the economic value of a set of football players. The human expert selects these attributes based on past data and filters the entire set of players based on age and the level of the selected attributes. The goal is to optimize the human-like analytical process to make the AI even better than the human expert.

Already closed experiments:

Based on the data concerning Turkish footballers in 2017, 129 players were analysed with a focus on 46 professional attributes related to the players' values from 2017 to 2022 each year. The top 5 attributes that had the most influence on a player's value were Composure, Passing, Anticipation, Decisions, and Stamina. After further filtering for players under the age of 23, 28 desired players were selected from the original 129. The experiment then started with setting the values of the top 5 attributes (Composure, Passing, Anticipation, Decisions, and Pace) to 9, 10, 11, and 12, which resulted in the following outcomes. Results:

66 players with 21.5% accuracy and 313,083 euros in earnings. / 39 players with 20.5% and 490,883 euros in earnings.

15 players with 20.0% and 250,579 euros in earnings. / 4 players with 50.0% and 377,351 euros in earnings.

Another approach was to calculate the mode of each of the top 5 attributes and add them up, giving a total of 52 points out of 100 (as the attributes rating goes from 1 to 20, with 20 being the best). This resulted in 36 players with 13.9% and 472,479 euros.

There was yet another approach where the top attribute (Stamina) was replaced by another positively correlated attribute, Pace (rank 15). The following results were obtained using the same filtering method as before:

20 players with 80.0% and 1,276,149 euros in earnings. / 14 players with 71.4% and 1,957,090 euros in earnings. / 3 players with 100.0% and 1,252,899 euros

Finally, the robot-eye method (similarity analysis - <https://miau.my-x.hu/myx-free/>) resulted in 2,393,055 euros in earnings, while the regression method resulted in 1,957,744 euros in earnings.

Discussion: The goal of this project is to create a robot capable of outperforming human experts in talent management for football, specifically in identifying players with high potential. The human method involves selecting a few attributes, such as anticipation, composure, and stamina, that have a high correlation with the player's monetary value, and then filtering players based on age and the level of these attributes. However, this approach can result in errors or low success ratios if the number of attributes chosen, or the minimum level of these attributes is too high. The robot-eye, with its ability to see complex patterns in raw data, can increase the success ratio through solver-based processes. The publication focuses on player scouting using data obtained from football and aims to lay the foundation for a talent recognition system to help young players who are stuck in the grid, based on their position or playing style. The research is based on six years of data from over 150,000 real football players and involves determining development patterns and prominent developments through anti-discrimination analysis. This research also seeks to answer the question of whether it's possible to determine five essential attributes for being an excellent football player. The target audience for this research is vast, as football is the most popular sport worldwide, with billions of people watching it according to FIFA, and the research has applications for a variety of groups, including clubs, scouts, media platforms, journalists, and IT developers and startups.

Future: In the future of talent management in football, it will be crucial to not only evaluate a player's individual attributes but also their contribution to the team's overall capabilities. This calls for a more comprehensive approach to data analysis that takes into account not just the player's individual attributes, but also their ability to fit within the team dynamic. This shift from solely analysing a player's attributes as an individual to considering their impact on the team will be necessary for effective talent management in football.

Demo-URLs: <https://miau.my-x.hu/miau/297/starlight_v1.xls>