YUVA-SQ: A Cognitive Scouting Model for The Beautiful Game

Sadanand Venkataraman*, Sundharakumar KB*, Bharathi Malakreddy A†, Santhi Natarajan*
*Cognition Lab, Dept. of CSE, Shiv Nadar University Chennai, Kalavakkam, Tamil Nadu, India 603110
†Research & Consultancy Cell, BMS Institute of Technology and Management, Bengaluru, Karnataka, India 560064

2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT) | 979-8-3503-8681-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICITIIT61487.2024.10580784

***Abstract*—The "Beautiful Game" of football has adapted itself to the world of modernisation and globalisation. Collecting game intelligence data through talent scouting and tactical scouting has emerged as a key factor in boosting the football club economy. From being an instinct-based identification of the right talent to the right moves, scouting models have now grown to be a technologically advanced platform for unlocking the right player talent pool. With huge unstructured imaging and numerical data gathered from across multiple clubs and games, this analytics needs advanced Artificial Intelligence (AI) based analytical models to work on big data sets, for an accurate set of predictive outcomes. As football provides a massive temporal strategy space, highly rational on-field negotiations and team-to-team interactions decide the next move in the mind game. With the help of a case study, our paper drives in the need for a cognitive scouting model like YUVA-SQ that go a long way in on-field and off-field high level decision making towards winning a game.**

I. Introduction

A. *The Beautiful Game!*

Football, the sport that brings together the world as “the beautiful game”, has been in existence in some form or the other for almost 3000 years now. The world-wide broadcasting of games from the top leagues of Europe and more Asian leagues and competitions through the Asian Football Confederation (AFC) identified leagues have made the sport a global favourite. With the increased involvement of the whole world, different bodies have been setup to improve and analyse the scientific part of the sport, whose main goal is for the improvement of players’ training and in aiding managers tactically.

B. *Talent Scouting Model: To Win!*

Club football being a multi-billion-dollar market today, has a major share of revenue fuelled by the Scouting Networks of the world [1], [2], [3]. However, lack of standardization of scouting and selection process is a challenge world wide. While performing team selection and tactic implementation, coaches consider different factors such as team morale, player availability, game schedule, performance track record of the player etc alongside many external factors. In addition, coaches and scouts tend to need upon their knowledge in the game and their instincts whilst scouting and selection.

C. *The Behavioral Science of Football: True Challenge in Scouting*

A very crucial element in scouting is to model the mental game of a footballer. It is challenging to quantify or qualify on-field decision making during the game. This hugely depends upon team-to-team interactions, player-to-player interactions, highly granular game strategies in a massive temporal strategy space, strictly rational on-field negotiations, player-ball paired trajectory modelling, etc.

Player benchmarking, player chemistry, player performance index, team performance index, etc are key aspects of behavioural sports research, which are very important to give the most relevant on-field and off-field feedback to the players to potentially improve the player performance. This is the key in the scouting model, which is not sufficiently looked into by the current models [4], [5], [6], [7].

D. *Our Contributions*

This paper is going to explore the key aspects of developing an international model for scouting in football, with the help of our novel AI based behavioral analytics. We present the results of our predictive and prescriptive behavioral models based on a case study. We also discuss how it can be achieved in India and extend the model to the rest of the world.

II. Performance Analysis: Player Performance Comparison from A Statistical Perspective

This section highlights our findings and results based on the comparative football analytics of two proficient players representing two different teams. Kevin De Bruyne, who is Player 1 and referred to as KDB, is the preferred midfielder playing for the blue side of Manchester who has been playing for them from 2015. Bruno Fernandes, who is Player 2 and referred to as Bruno, is Manchester United’s first choice attacking midfielder who in 2023 was handed the captain’s armband after joining the club in 2020. This comparison is done only for the **Barclays Premier League 2020-21 Season**. Table I presents the cumulative performance data of the two players for the 2020-21 season in the Premier League. Table VI is a consolidated plot listing for comparative performance visualization.

A. *Player Impact in the Respective Teams*

Bruno Fernandes and Kevin De Bruyne had a very impactful season for their respective teams through the 2020-21 season, having topped their teams' player statistics. Table IV and Table V shows summarizes Bruno Fernandes' and Kevin De Bruyne's impact in their teams respectively.

III. Our Proposal for a new International Scouting Model

A. Need for Implementing Behavioural Analysis

The statistical inferences from the above case study clearly shows Bruno Fernandes to have out done Kevin De Bruyne in almost every aspect of comparison. Yet, Kevin De Bruyne won the Men's PFA Player of the year 2020-21. This anomaly owes to the difference Kevin De Bruyne made for Manchester City in their memorable 2020-21 Premier League Campaign. He had a significant impact on his team, from a cognitive perspective. His strong mental orientation translated to the team's victory in the league.

The existing analytical methods do not consider a behavioral metric, neither in their season long evaluation nor in their scouting process [8], [9]. Currently, psychometric analytics and test batteries in the sports test the mental state of their players during the medical check of the player during the time of signing. This cognitive evaluation covers the conventional metrics as to how a player will be towards their teammates and general mental state [10], [11], [12], [13], [14]. A continuous evaluation of cognitive well being of players is lacking in the current team evaluation and management process. This gap needs to be fixed from the grassroot level scouting, thereby reducing the effort at a senior level.

We can address this by introducing behavioural analytics. To improve the already vast scouting network of football and utilize it to the fullest, we have formulated a standard scouting metrics, based on the AI assisted advanced data analytics, taking in to consideration the mental and physical abilities of a prospective player. We present YUVA-SQ (Fig. 10), our proposed new International Scouting Model. Fig. 10 gives the workflow the evaluation, selection and development process for Scouting model.

B. YUVA-SQ Evaluation

The Evaluation process consists of Two major assessments. The results of both the tests are cohesively presented as input towards scouting process.

*1) **A Novel Psychometric Assessment: Powered by "Your Offence and Defence Assessment(YODA)**:* YODA, unlike a conventional psychometric tool, serves to be a comprehensive system built on a solid behavioral analytics foundation. Drawing inspiration from the widely recognized Big Five personality traits, YODA evaluates a player over 72 specialized player traits (psychometric attributes) tailored for football, offering a granular insight into a player's cognitive profile.

From ambition and attitude to aggression and mental toughness, YODA's evaluative parameters are exhaustive. These traits, when combined with sub-traits like obsessive passion, visual search, and self-regulation, offer a 360-degree view of a player's mental framework. The following Parameters are assessed through the Psychometric assessment of the prospect:

1) Following Coach's Instructions
2) In-Game Reactions
3) Attitude towards teammates
4) Lifestyle outside training

The assessment through YODA is a meticulously crafted process, evaluating the player in 48 varied scenarios, reflective of real-game situations, offer invaluable insights into a player's mental orientation and potential on-field reactions. YODA seamlessly integrates the esteemed Big Five personality traits—Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism—into its assessment framework. Each trait offers a unique lens, enabling a more nuanced understanding of a player's psyche.

*2) **Coach Assessment**:* Coach's assessment of footballing ability of the prospect is done, and the following abilities of the prospect is quantified:

1) Physical
2) Shooting
3) Passing
4) Ball Control
5) Defense
6) Goalkeeping

Each ability has various sub categories which the coach will be required to assess and quantify, which will cumulatively give a score to the prospect.

C. Selection, Development and Monitoring

The player is further assessed and is taken into the club for further assessment on a trial basis to the club upon meeting the requirements of the club through the initial evaluation. During the trial period, we evaluate the player for:

1) Inclusion in Youth Team: Prospect is put into the youth system of the club to inspect the chemistry of the player with the existing club's culture.
2) Video and in-game analysis

Upon completion of the trial period, the initial YODA evaluation is re-done to realize the improvement. The comparison of the initial evaluation and updated evaluation will help the club and coach decide if the prospect fits the culture and style of the club. It further helps in assessing whether it is feasible to include the prospect in the system and nurture the development of the player having yearly evaluations and continuous monitoring of their development.

D. YODA results

As a case study, The BMS Institute of Technology & Management, Bengaluru Men's football team was restructured post the 2020-21 season's break using the findings of YODA Assessments performed on the interested players. We present our YODA assessment test results, performed over a 50 minutes session, for Player ID 'YODA035'. He is an 18 year old, trained to play as a striker. In his initial evaluation of player YODA035, the coach identified the player to be clear-headed and who knows how good he is and will work towards becoming a better player.

TABLE I: Player Information.

Sl. No	Feature	Description	Player 1	Player 2
1	Player	Name of the Player whose data is displayed along the rest of the row.	KDB	Bruno
2	Games	Number of games [Player] has played throughout the [Season] Premier League Season	25	37
3	Goals	Total Number of Goals Scored by player in the [Season] More the better	6	18
4	Shots	Total number of attempted shots by [Player] in the [Season] season	81	121
5	Time	Total number of Minutes [Player] has been on the pitch for the [Season]	2008	3117
6	xG	Total number of expected goals [Player] should have scored in the [Season] season More the better	9.98	16.02
7	Assists	Total number of assists [Player] provided during the [Season] Season More the better	12	12
8	xA	Total number of expected assists [Player] should have provided to other players in his team during the [Season] season More the better	10.96	11.47
9	Key_passes	Total number of Key Passes completed by [Player] in the [Season] Season More the better	79	95
10	Season	The season in which [Player]'s data is displayed along the row	2020	2020
11	Team	The team for which [Player]'s data is displayed along the row	Manchester City	Manchester United
12	Yellow	Total number of Yellow Card bookings recorded by [Player] in the [Season] season Lesser the better	1	6
13	Red	Total number of Red Card bookings recorded by [Player] in the [Season] season Lesser the better	0	0
14	npg	Non-Penalty Goals scored by [Player] in the [Season] season More the better	4	9
15	npxG	Non-Penalty Goals expected to be scored by [Player] in the [Season] season More the better	7.69	8.41
16	xGChain	The Expected goals to be scored after [Player] had possession of the ball More the better	22.52	26.91
17	xGBuildup	All the attacking actions by [Player] excluding key passes and shots on goal. More the better	11.29	11.93

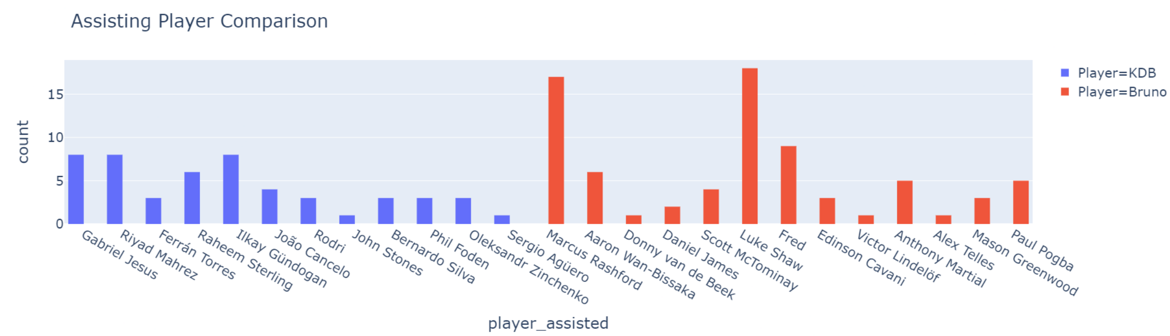

Fig. 1: Assisting Player Comparison.

TABLE II: Shot Types

Shot Type	Description
Direct Free-kick	This shot is taken by a team that was fouled from the spot the foul was committed.
Penalty	If a illegal infringement(handball) or foul is committed inside the D Box, the team that was fouled gets to take a shot from the Penalty mark with only the goalkeeper to beat.
Open Play	This is when the ball is struck from open play, at any part of the field.
From Corner	This is a metric when a shot is taken after the ball is delivered from a corner kick.

TABLE III: Shot Information

Shot Result	Description
Saved Shot	When the opposing goalkeeper saves the shot from entering the goal.
Goal	When the shot makes it past the goal-line of the opposing team.
Missed Shot	When a shot does not make it past the goal-line of the opposing team, without any attempts to change the trajectory of the ball by other players.
Blocked Shot	When the shot is blocked by a player of the opposing team, from entering the goal-line.

A subsequent YODA evaluation of player YODA035 resulted in the Trait Plot (Fig. 11) and Personality Evaluation (Fig. 12) as outcomes. This translates to the following behavioral analysis summary:

1) He will be a player who will always listen to the coach
2) He will look for the betterment of the team
3) He will not easily mingle with all players
4) He is a very analytical player who has good game knowledge

This evaluation aided the coach to improvise game performance by YODA035 in the subsequent matches by charting a customized training regime as suggested by the YODA evaluation.

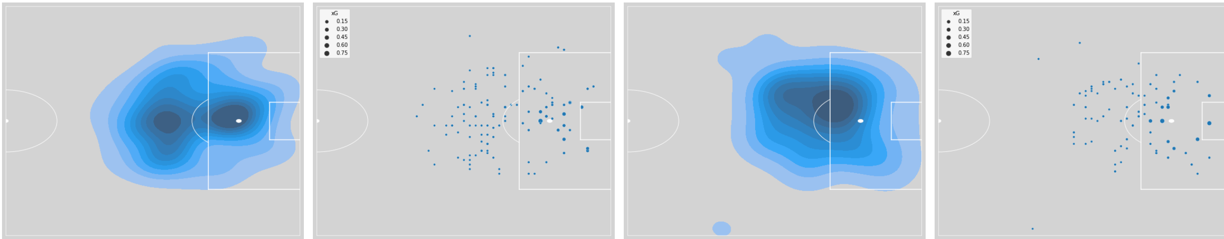

Fig. 2: a)BrunoShotsHM. b)BrunoShotMap c)KDBShotsHM d)KDBShotMap

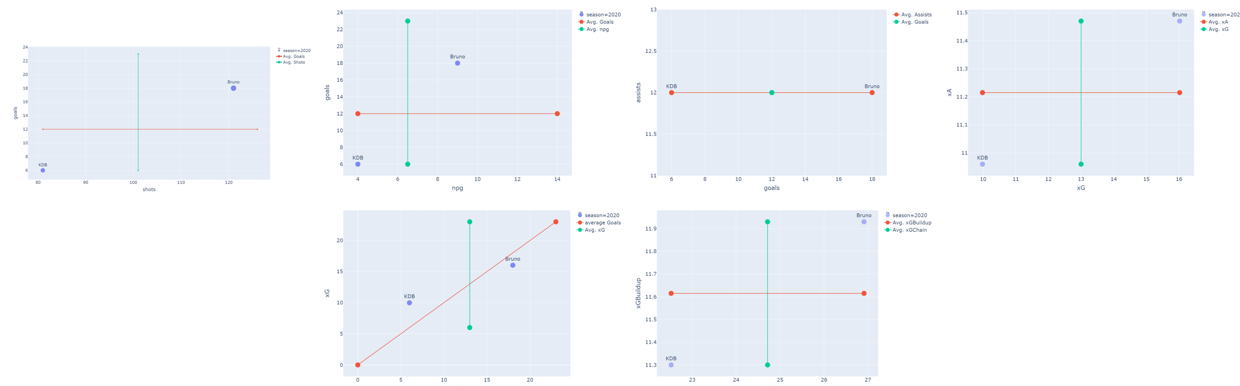

Fig. 3: From Left to Right Top Row: a)GoalsVsShots. b)GoalsVsNPG c)GoalsVsAssists d)xGvsxA. From Left to Right Bottom Row: e)GoalsVsxG f)xGchainvsxGBuildup

TABLE IV: Bruno Fernandes' Impact in Manchester United.

Sl. No	Feature	Team Value	Player Value	Player Impact
1	Games	38	37	97.30%
2	Goals	70	18	25.70%
3	Shots	525	121	23.00%
4	Time	3420	3117	91.10%
5	xG	64.004	16.019	25.00%
6	Assists	51	12	23.50%
7	xA	44.707	11.474	25.60%
8	Key_passes	412	95	23.00%
9	Yellow_Cards	64	6	9.30%
10	Red_Cards	1	0	0.00%
11	npg	60	9	15.00%
12	npxG	55.632	8.407	15.10%
13	xGChain	199.872	26.911	13.40%
14	xGBuildup	124.822	11.932	9.5%

TABLE V: Kevin De Bruyne's Impact in Manchester City.

Sl. No	Feature	Team Value	Player Value	Player Impact
1	Games	38	25	65.70%
2	Goals	82	6	7.30%
3	Shots	600	81	13.50%
4	Time	3420	2008	58.70%
5	xG	79.094	9.983	12.60%
6	Assists	55	12	21.80%
7	xA	56.128	10.958	19.50%
8	Key_passes	457	79	17.20%
9	Yellow_Cards	46	1	2.10%
10	Red_Cards	2	0	0.00%
11	npg	77	4	5.20%
12	npxG	72.244	7.699	10.60%
13	xGChain	266.556	22.516	8.40%
14	xGBuildup	177.706	11.297	6.30%

IV. Conclusions and The Next Phase

Through this paper, we present a descriptive analytics based case study and explored a statistical comparison between two

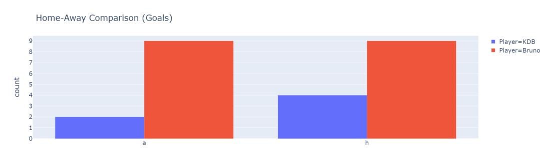

Fig. 4: Home-Away Goals Comparison.

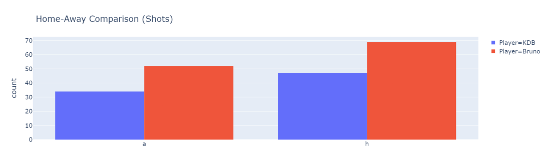

Fig. 5: Home-Away Shots Comparison.

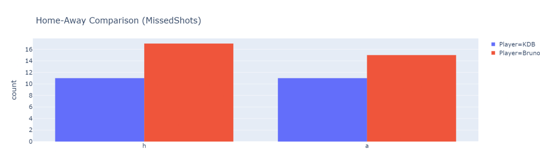

Fig. 6: Home-Away Missed Shots Comparison.

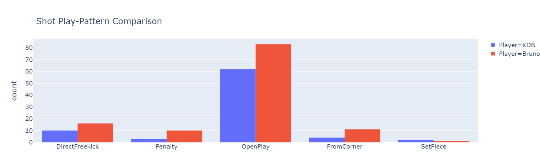

Fig. 7: Shot Play-Pattern Comparison.

TABLE VI: Visualization

Plot	Description
Fig. 1	No of assists comparisons where we show the list and number of assists other teammates of the players have provided to them. It is interesting to notice both of the top assist providers to both Bruno Fernandes and Kevin De Bruyne were players playing on the left side of the pitch (Luke Shaw for Bruno Fernandes; Gabriel Jesus, Riyad Mahrez and Gundogan for Kevin De Bruyne).
Fig. 2	Here, we present Shot maps that show the effectiveness (higher values of xGChain and xGBuildup) of the players (1b and 1d). The heatmaps (1a and 1c) show all the shots taken by each player, side by side towards the opponent's goal. The darker the spot on the heat-map, higher the xG of each shot.
Fig. 3a	Goals versus shots comparisons where we plot a graph between the number of goals scored by the players against the number of shots taken by the players. From this graph, Bruno Fernandes comes out to be the better player with 18 goals from 121 shots whereas Kevin De Bruyne scored 6 goals from 81.
Fig. 3b	Goals versus Non Penalty Goals comparisons where we plot a graph between the number of goals scored by the players against the number of those goals not to be scored from a penalty. From this graph, Bruno Fernandes comes out to have 50% of his goals to be scored from open play whereas Kevin De Bruyne has 66.66% of his goals from open play.
Fig. 3c	Goals versus total assists comparisons where we show a comparison of all the goal involvements the players have had throughout the season with a graph plotted between the number of assists they provided against the number of goals they scored. Bruno Fernanes has the higher contribution with 18 goals and 12 assists, a total of 30 goal involvements whereas Kevin De Bruyne has 6 Goals and 12 Assists, making his contribution to be 18 goal involvements.
7Fig. 3d	Total Expected Goals(xG) vs Total Expected Assists(xA) comparisons where we show the total number of goals and assists both the players were expected to achieve. Bruno Fernandes was expected to provide 16.02 goals and 11.07 assists and Kevin De Bruyne was expected to provide 9.98 goals and 10.96 assists.
Fig. 3e	Total Goals vs Total Expected Goals(xG) comparisons where we show the total goals each player was expected to score against the actual goals they scored. Here Bruno Fernandes has scored 18 goals having an xG of 16.02 whereas Kevin De Bruyne ended up scoring only 6 Goals out of an expected 9.98.
Fig. 3f	xGChains vs xGBuildup comparisons where we show the expected goals for the team in which the player was involved in the possession of the ball, including the final pass or shot against the goals for the team in which the player was involved in the possession of the ball, excluding the final pass or shot against. Here Bruno Fernandes has a xGChain of 26.91 and xGBuildup of 11.93 whereas Kevin De Bruyne has a xGChain of 22.52 and xGBuildup of 11.3.
Fig. 4	Home-Away Goals comparisons where we we have split the total goals scored by the players into the goals scored in home and away games. Usually, this metric shows how a player is affected by the presence and absence home supporters. However, in this case study, there were no attendance in the stadiums where the games were played.
Fig. 5	Home_Away shot comparisons where we can see the number of total shots by the players both in home games and away games. Here, greater is better, especially if the players' xG is higher.
Fig. 6	Home_Away Missed shot comparisons where we can observe the number of missed shots by the players both in home games and away games. Here, lesser is better.
Fig. 7	Shot play-pattern comparisons where we can see the different types of shots the players took throughout the season. The shot types and descriptions are listed in Table II.
Fig. 8	Shot result comparisons where we can see the different results of shots the players took throughout the season. The shot types considered here, along with their descriptions, are listed in Table III
Fig. 9	Final Comparision Per90 consolidates all the metrices of Table I, in a radar to see how the player was expected and performed through a concept of Per90 which averages out the entire season's performance into 90 minutes. From the figure, we notice Kevin De-Bruyne has had more Key Passes, shots, assists and xGChain per 90 minutes played whereas Bruno Fernandes has done better in only the goals per 90 minutes stat. This is because Kevin De Bruyne has played for lesser minutes than Bruno Fernandes through out the Premier League Campaign.

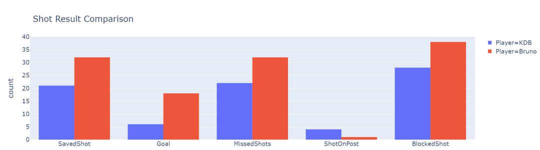

Fig. 8: Shot Result Comparison.

of the best midfielders in the premier league through the 2020-21 Season. The results highlighted how only statistical quantification of players is insufficient to realize the full impact a player has in the campaign of a team. We introduce our new scouting Model YUVA-SQ, that works from the grassroot level and aid in improving the development of young prospects shaping them from early stages in their career to benefit them as well as the success of a football club through our model. This can be modified to fit the requirements of various other team sports such as cricket and basketball in improve the development and fit of a young prospect in those sports as well. Our solution is a small step towards All India Football Federation's (AIFF) Vision 2047, in analyzing and improving skill-sets to put the Indian football name on the world stage.

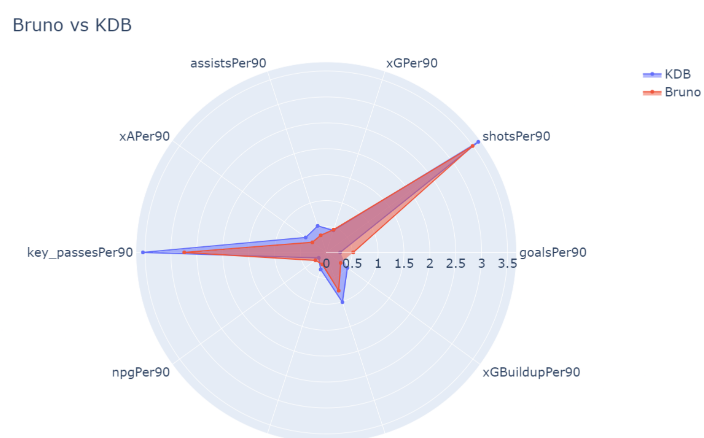

Fig. 9: p90.

References

[1] H. Sarmento, R. Marcelino, J. Campanico, N. Matos, and J. Leitão, "Match analysis in football: a systematic review," Journal of sports sciences, vol. 32, pp. 1831–1843, Dec. 2014, doi: 10.1080/02640414.2014.898852.

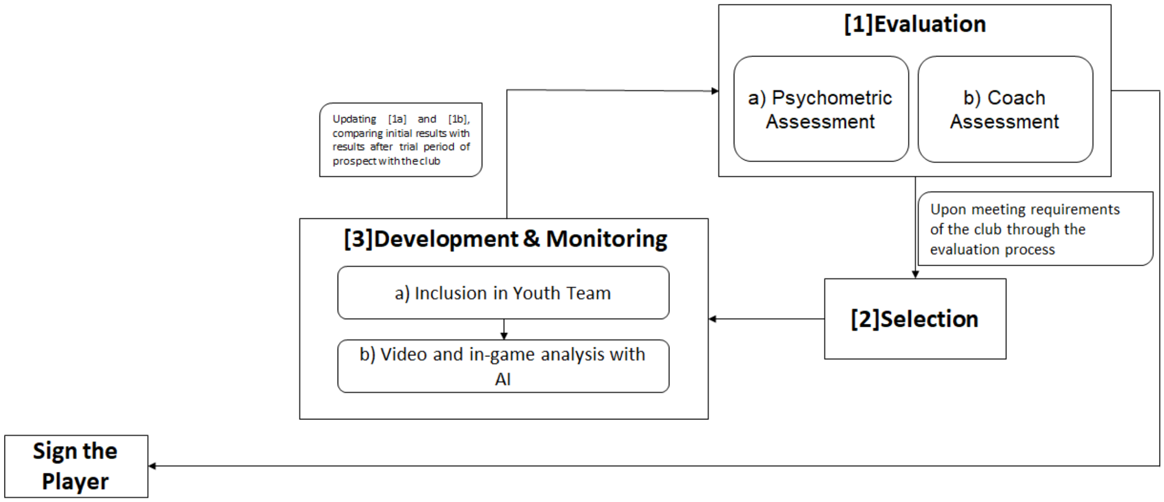

Fig. 10: YUVA-SQ (Evaluation, Selection and Development_Process).

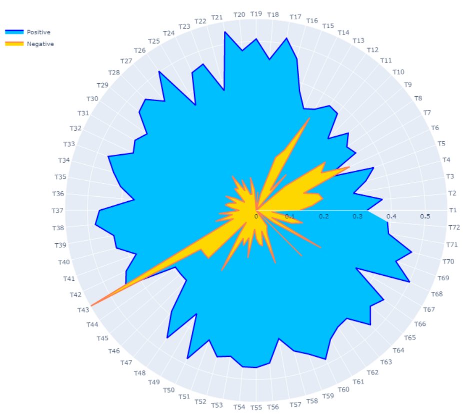

Fig. 11: Trait Plot of YODA035.

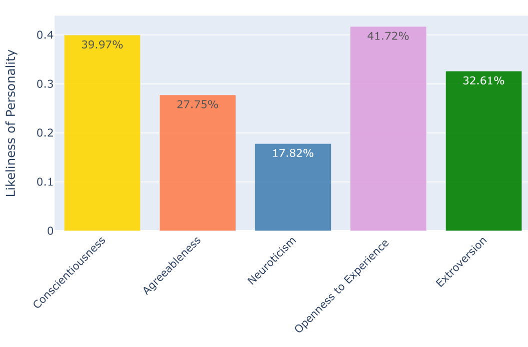

Fig. 12: Personality assessment of YODA035.

[2] A. P. Ayuso, J. P. Vicedo, and O. Jordán, "Content Validity and Psychometric Properties of the Nomination Scale for Identifying Football Talent (NSIFT): Application to Coaches, Parents and Players," Sports (Special Issue: Performance in Soccer), vol. 5, Jan. 2017, doi: 10.3390/sports5010002.

[3] P. Feichtinger and O. Höner, "Psychological diagnostics in the talent development program of the German Football Association: Psychometric properties of an Internet-based test battery," Sportwissenschaft, vol. 44, pp. 203–213, Nov. 2014, doi: 10.1007/s12662-014-0341-0.

[4] O. Höner, A. Votteler, M. Schmid, F. Schultz, and K. Roth, "Psychometric properties of the motor diagnostics in the German football talent identification and development programme," Journal of sports sciences, vol. 33, pp. 1–15, June 2014, doi: 10.1080/02640414.2014.928416.

[5] R. Mackenzie and C. Cushion, "Performance analysis in football: A critical review and implications for future research," Journal of sports sciences, vol. 31, Dec. 2012, doi: 10.1080/02640414.2012.746720.

[6] T. Stolen, K. Chamari, C. Castagna, and U. Wisloff, "Physiology of Soccer," Sports medicine (Auckland, N.Z.), vol. 35, pp. 501–36, Feb. 2005, doi: 10.2165/00007256-200535060-00004.

[7] C. Wright, C. Carling, C. Lawlor, and D. Collins, "Elite football player engagement with performance analysis," International Journal of Performance Analysis in Sport, vol. 16, no. 3, pp. 1007–1032, 2016, Routledge, doi: 10.1080/24748668.2016.11868945, URL: https://doi.org/10.1080/24748668.2016.11868945.

[8] P. Toma and F. Campobasso, "Using data analytics to capture the strategic and financial decision-making of Europe's top football club," Technological Forecasting and Social Change, vol. 186, pp. 122116–122116, 2023.

[9] T. A. Herberger and C. Litke, "The Impact of Big Data and Sports Analytics on Professional Football: A Systematic Literature Review," in Springer, Cham, 2021, pp. 147–171, doi: 10.1007/978-3-030-77340-3_12.

[10] Y. Kim, K.-H. N. Bui, and J. J. Jung, "Data-driven exploratory approach on player valuation in football transfer market," Concurrency and Computation: Practice and Experience, John Wiley & Sons, Ltd, vol. 33, no. 3, 2021, doi: 10.1002/CPE.5353.

[11] R. de la Torre, L. Calvet, D. Lopez-Lopez, A. A. Juan, and S. Hatami, "Business Analytics in Sport Talent Acquisition: Methods, Experiences, and Open Research Opportunities," International journal of business, IGI Global, vol. 9, no. 1, pp. 1–20, 2022, doi: 10.4018/IJBAN.290406.

[12] "Prediction of Credibility of Football Player Rating Using Data Analytics," Springer eBooks, 2022, pp. 775–786, doi: 10.1007/978-3-030-96308-8_72.

[13] P. Rajesh, Bharadwaj, M. Alam, and M. Tahernezhadi, "A Data Science Approach to Football Team Player Selection," in IEEE, 2020, pp. 175–183, doi: 10.1109/EIT48999.2020.9208331.

[14] A. Malakreddy, S. Venkataraman, M. Khan, and S. Padmanabhuni, "Recent Advances in Semantic Segmentation for Sports Analytics," 2023, pp. 275–285, ISBN: 978-981-99-4931-1, doi: 10.1007/978-981-99-4932-8_26.