

PAPER • OPEN ACCESS

Predicting Player Position for Talent Identification in Association Football

To cite this article: Nazim Razali *et al* 2017 *IOP Conf. Ser.: Mater. Sci. Eng.* **226** 012087

View the [article online](#) for updates and enhancements.

You may also like

- [Meeting report](#)
- [AUTOMATIC SELECTION OF H-ALPHA EMISSION-LINE GALAXIES](#)
Oscar Alonso Lasheras
- [On double-layer and reverse discharge creation during long positive voltage pulses in a bipolar HiPIMS discharge](#)
Andrea Dagmar Pajdarová, Tomas Kozak, Tomáš Tölg et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Predicting Player Position for Talent Identification in Association Football

Nazim Razali¹, Aida Mustapha¹, Faiz Ahmad Yatim²,
Ruhaya Ab Aziz¹

¹ Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

² Co-Operative College of Malaysia, 103 Jalan Templer, 46700 Petaling Jaya, Selangor, Malaysia

E-mail: nazim.uthm@gmail.com, aidam@uthm.edu.my, faiz@mkm.edu.my, ruhaya@uthm.edu.my

Abstract. This paper is set to introduce a new framework from the perspective of Computer Science for identifying talents in the sport of football based on the players' individual qualities; physical, mental, and technical. The combination of qualities as assessed by coaches are then used to predict the players' position in a match that suits the player the best in a particular team formation. Evaluation of the proposed framework is two-fold; quantitatively via classification experiments to predict player position, and qualitatively via a Talent Identification Site developed to achieve the same goal. Results from the classification experiments using Bayesian Networks, Decision Trees, and K-Nearest Neighbor have shown an average of 98% accuracy, which will promote consistency in decision-making through elimination of personal bias in team selection. The positive reviews on the Football Identification Site also indicates that the framework is sufficient to serve as the basis of developing an intelligent team management system in different sports, whereby growth and performance of sport players can be monitored and identified.

1. Introduction and Related Work

Talent identification is a very crucial avenue in effort to achieve top class sporting excellence and efficient guidance in any sports. Because the process of discovering talents is by attending regular practice program, looking for talented people and giving them a privileged position is very important [1]. Talent search programs have long existed in many countries using different methods and models. The application of the most advanced sports talent discovery has existed in the former East Germany and the Soviet Union [2, 3]. However, their approach relies heavily on the ability of fully experienced experts. Discovering talents in sports using Artificial Intelligence (AI) methods is still at the infancy stage. Among the works, [4] developed a web-oriented expert system with a fuzzy module to predict the most suitable sports for the person who has taken the test. [5] used a fuzzy immune algorithm to learn sports training pool, while Peter et al. [6] used artificial neural networks to analyze the sport of golf.

In similar vein, one of the main challenges in professional football is talent management, which concerns on how to ensure the right player for the right position at the right time. These tasks involve a lot of experience in decision management and coaching. In general, there is no



scientific equations or formula adopted to recognize the most suitable position for each player in a particular team formation. The assignment is carried out by coaches using his experiences and observations of their players [7]. This paper attempts to fill the gap by proposing a talent identification framework to assign individual players into different positions in football such as goalkeeper, sweeper, wing backs, right and left back, defensive midfielder, wingers, wide midfielder, center midfielder, attacking midfielder, secondary striker, and finally forward like in the work of [8]. The goalkeeper, however, is a special position that is different in terms of the qualities required such as "establishing connection" and "the ability to repulse overhead shoots" [9]. Therefore, this research ignores this position in the dataset used during the classification experiments.

In this research, the positions are assigned based on individual player skills, which covers three aspects; their physical, mental, and technical skills. For each skill, the literature has identified certain qualities required, which are usually being measured personally by the coach. The specific skills required in football such as speed, agility, passing or tackling are shown in Table 1. Next, to select talented players for a specific position, the combination of skills in Table 2 should be taken into account. For example, a defensive midfielder needs high quality in attacking play, physically strong, high shooting and backing ability, high potential of game reading, as well as highly vigorous [10, 11].

Table 1. Skills required among football players.

Physical Skills	Mental Skills	Technical Skills
speed	ability to read game	finishing
agility	calmness	passing
jumping	creativity	shooting
acceleration	boldness	heading
height	confidence in decision	tackling
strength	leadership	carrying
vigor		dribbling
		single play
		crossing
		backing
		attacking

The remaining of this paper is organized as follows. Section 2 presents the proposed framework for talent identification in predicting player positions for football. Section 3 presents the experimental setup with details of the dataset. Section 4 presents the results for classification experiments as well as a the user acceptance test on the system produced based on the framework. Finally, Section 5 concludes the research with some indication for future works.

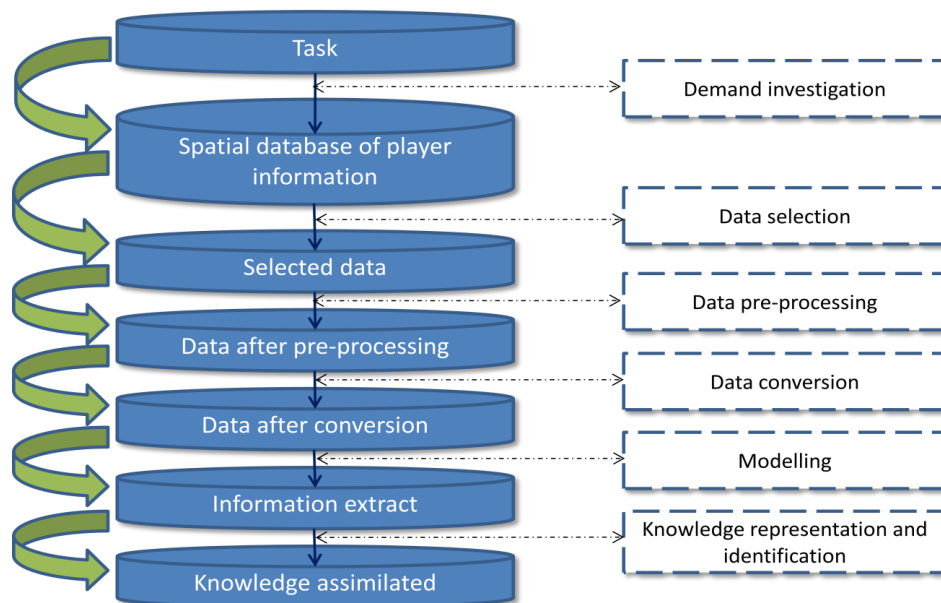
2. Talent Identification Framework

Pundits believe that every footballer has certain abilities that make them a better attacker than a defender and vice versa. However, football talent discovery relies on personal experience and sensation by which the performance of skills are measured by the coach or team managers. One of the key factors in measuring abilities in individuals such as in the game of football is by forecasting their best position on the field. While the discovery of the ability to participate in an organized game such as football is very important, currently there is no practical approach to allocate the true talent of the football players on the football field.

Table 2. Skills for different positions in football.

Position	Skills
sweeper	ability to read game, calmness, leadership, agility, speed [10, 12]
wing backs	backing, boldness, height, physical strength, heading ability [11, 13]
right/left backs	speed, acceleration, tackling, crossing, agility [10, 12]
defensive midfielders	attacking play, physical strength, shooting, backing, ability to read game, vigor [10, 12]
center midfielders	vigor, passing, shooting, creativity, confidence and decision making, leadership [11]
wingers	vigor, crossing, passing, tackling, speed, carrying ball [11, 13]
attacking midfielders	creative, carrying the ball, passing, calmness, dribbling, shooting [11, 13, 12]
wide midfielders	crossing, carrying the ball, passing, dribbling, speed, acceleration, agility, vigor [11, 10]
secondary strikers	height, ability of single play, passing, agility, physical strength, heading ability [11]
forward	finishing, shooting, heading ability, speed, calmness, dribbling, creativity [11, 12]

In this research, a framework for talent identification in football is proposed. The framework is used to identify talents for positions such as sweepers or wingers in a football team based on individual player's skills, so the team has the highest probability to win with such strengths and abilities placed at the correct position. The proposed framework for talent identification is shown in Figure 1 followed by the brief descriptions of each step in the framework.

**Figure 1.** Proposed Talent Identification framework.

- Demand investigation: Concerns on understanding the nature of football skills, how they are measured and acquired, and the relationships between the skills with the player positions during the football game.
- Data selection: In this research, the data is taken from the Football Player Information System at Bukit Jalil Sports School (BJSS). The database keeps the rating from 1-10 of physical, mental, and technical skills as well as the position assigned by the coach as the class label.
- Data processing and conversion: Performs data cleansing on data stored in the Football Player Information System, BJSS. Unnecessary information will be deleted. Data conversion change the data into a unified data format.
- Modelling: Constructs the classification models for predicting player positions using 3 algorithms; Bayesian Networks, Decision Tree, and Nearest Neighbor.
- Knowledge representation and assessment: Data from the Football Player Information System together with the knowledge from the classification experiments are displayed in the Football Talent Identification Site. The framework is also assessed based expert evaluation on the website.

3. Framework Evaluation

The proposed framework for talent identification in football sport was validated using both quantitative and qualitative methods. For quantitative methods, a series of classification experiments were carried out in predicting player position based on their individual qualities. Three data mining algorithms were applied, namely, Bayesian networks, decision trees, and nearest neighbor. All classification experiments were carried out using the Waikato Environment for Knowledge Analysis (WEKA) (<http://www.cs.waikato.ac.nz/ml/weka/>). For qualitative evaluation, a web-based Football Talent Identification Site has been developed based on the proposed framework. This research employed expert evaluation based on User Acceptance Test (UAT) by the coaches and team managers in BJSS.

3.1. Dataset

Based on the literature review of [11, 13, 12, 10], this research compiled a list of qualities of three main skills measured by coaches in football teams; physical [11, 13, 10], mental [11, 12], and technical [11]. These qualities are rated between 1 to 10, where 1 is the weakest and 10 is the strongest. The qualities are used as attributes in the experiments as shown in Table 1. From these attributes, an independent test was conducted at the Bukit Jalil Sports School (BJSS) with an experimental group of 100 players between the age of 15 to 17. Each player was manually assigned with one of the 10 positions in a football team by the coach. The positions become the class labels in the dataset: *wing back*, *right back* and *left back*, *defensive midfielder*, *wingers*, *wide midfielder*, *center midfielder*, *attacking midfielder*, *secondary striker*, and *forward*.

3.2. Quantitative Evaluation

The classification experiments were carried out using three algorithms, which are Bayesian networks, decision trees, and nearest neighbor. Due to the small number of samples in the dataset, the experiment used leave-one-out validation technique [14], whereby only 1% of the data is used for testing while 99% used for training. Using this technique, the data was split in all possible ways to divide the original sample into a training and a validation set. The comparison of results from the experiments are shown in Table 3. From the table, the highest classification accuracy achieved during the experiments was 99% for Bayesian Networks, followed by 98% for Decision Tree, and 97% for the Nearest Neighbor. Recall that the objective of the classification experiments was to assign every footballer a position that best matches his physical, mental,

and technical skills. The high classification results showed that the data mining approach is feasible for assigning players to a position best suited based on their individual skills, therefore could assist football coaches in making decisions.

Table 3. Comparison of classification accuracy.

Algorithm	Accuracy
Bayesian Networks	99%
Decision Tree	98%
Nearest Neighbor	97%

3.3. Qualitative Evaluation

Apart from assessing the accuracy in the classification experiments, the Football Talent Site was developed as the interface for coaches and team managers to manually assign the players into positions and execute team formation, or let the system predict the positions for each player according to individual performance. The screenshots of the Football Talent Site are shown in Figure 2 to Figure 4.

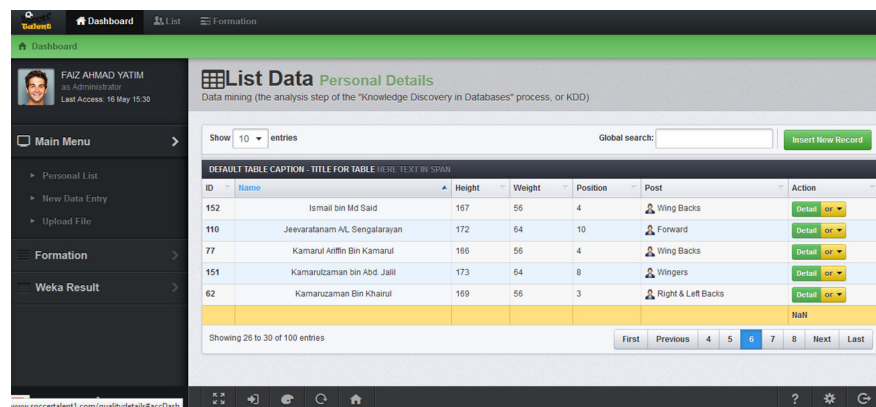


Figure 2. Original position assignment from Football Player Information System, BJSS.

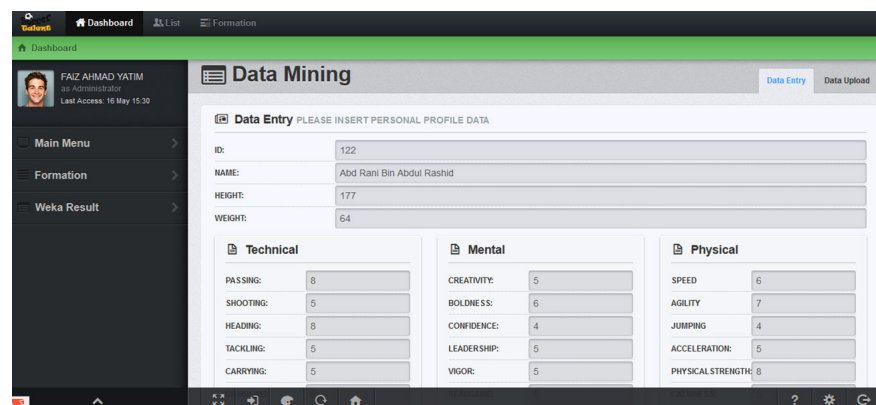


Figure 3. Records on individual player skills.



Figure 4. Recommended positions for each player in a team formation.

To evaluate the site, a User Acceptance Testing (UAT) was conducted on 20 users; coaches and managers at Bukit Jalil Sports School (BJSS). They were required to navigate each functionality in the system. A brief explanation of the objectives and functions of the system was given before the UAT sessions. At the end of the session, questionnaires were distributed to obtain feedback from users. The analysis from the UAT is presented in Table 4 and was analyzed using SPSS.

Table 4. Analysis on UAT.

Questions	Variable	Number	Percentage (%)
System works according to the specifications	Strongly Agree	16	80
	Agree	4	20
	Total	20	100
System has a user-friendly interface	Strongly Agree	15	75
	Agree	5	25
	Total	20	100
System is suitable to be use in the football teams	Strongly Agree	14	70
	Agree	3	15
	Not Sure	3	15
	Total	20	100
System does not has error that will interfere operation	Agree	14	70
	Not Sure	6	30
	Total	20	100
System has achieved the objectives of the development	Strongly Agree	8	40
	Agree	12	60
	Total	20	100

Analysis on the system functionality showed that 20% of users strongly agreed and 80% agreed that the system worked according to the specifications set forth in its objectives. The user-friendly interface is one of the most important features of each system are developed. The interface is one of the factors that will lead to the continued use of the system by users. Analysis showed that 75% of users strongly agreed and 25% of users agreed that this system has a user-friendly interface. Each system has its construction objectives. The users agreed with two options (i.e. manual vs. prediction) provided in assigning positions to players. However, the most important goal for the system is appropriateness to implement the system in an actual operating environment and the effectiveness of the system in providing options. Analysis showed

that 70% of users strongly agreed, 15% of users agreed and the remainig of 15% of users were not sure that the system is suitable to be implemented in the organization. During discussion, the reasons behind this include uncertainty in management, as well as policy and budgeting.

Error may occur during the process of system development and current operations. A system must have a maintenance schedule to avoid unnecessary mistakes that will interfere with its operation. Analysis of errors in the system showed that 70% of users agreed and 30% of users were not sure that the system has no errors that could disrupt operations. However, this is because the users were not exposed to prolong use of the system that they could identify errors in the system. Nevertheless, as a prototype, the system was working as intended. Finally, analysis of the system's objectives showed that 40% of user strongly agreed and 60% of users agreed that the system has achieved the intended objectives as briefed before the UAT sessions. It can be concluded that there is potential to manage a football player in a systematic and efficient data management system as compared to the current practice.

4. Conclusions and Future Work

In managing a football team, a coach performs player selection as well as team formation based on his personal experience. There is also no special formula for the coach to evaluate and compare between different players. This paper proposed an empirical methodology using data mining approach to quantitatively measure the strengths of each player in different positions based on their physical, mental, and technical skills in football. The classification experiments achieved an average of 98% accuracy, which promises consistency in decision-making though elimination of personal bias in team selection. This research has modestly provided one reliable framework that can be used as a basis for developing an intelligent team management system in different sports, whereby growth and performance of sports player can be monitored and a future champion can be identified. It is hoped that the proposed framework could be extended to cover different sports, in effort to identify the right talent among professional athletes.

Acknowledgements

This research is funded by the Ministry of Higher Education Malaysia under the Research Acculturation Collaborative Effort (RACE) Grant Scheme Vot 1513.

References

- [1] Bazmara M and Jafari S 2013 *Journal of Basic and Applied Scientific Research* **3** 981–986
- [2] Reilly T, Bangsbo J and Franks A 2000 *Journal of Sports Sciences* **18** 669–683
- [3] Williams a M 2000 *Journal of Sports Sciences* **18** 737–750
- [4] Papić V, Rogulj N and Pleština V 2009 *Expert Systems with Applications* **36** 8830–8838
- [5] Mezyk E and Unold O 2011 *Computers in Human Behavior* **27** 1499–1506
- [6] Lamb P F, Bartlett R M and Robins A 2011 *Human Movement Science* **30** 1129–1143
- [7] Sæther S A 2014 Identification of Talent in Soccer - What Do Coaches Look For? URL <http://www.idrottsforum.org/saether140319>
- [8] Bazmara M 2014 *International Journal of Intelligent Systems and Applications* **6** 62–67
- [9] Bazmara M and Jafari S 2012 *International Journal of Computer Science Issues* **9** 318–322
- [10] Ostojic S M 2000 *Facta Universitatis: Series Physical Education & Sport* **1** 23–29
- [11] Ratomir Durašković, Aleksandar Joksimović S J 2004 *Physical Education and Sport* **2** 13–24
- [12] Raven P B, Gettman L R, Pollock M L and Cooper K H 1976 *British Journal of Sports Medicine* **10** 209–16
- [13] McIntyre M C and Hall M 2005 *British Journal of Sports Medicine* **39** 264–6

- [14] Efron B 2010 *In CBMS-NSF Regional Conference Series in Applied Mathematics 1982. Philadelphia, PA: Society for industrial and Applied Mathematics (SIAM)*