International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8, Issue-1, May 2019

Machine Learning Approach for Identification of Peer Quality Factors among Sportsman

Bh.V.RamaKrishna, Vadali Srinivas, B. Sushma

Abstract*: *The process of identification multi-talent game players improves the chance of substitution of players among different games when situation demands. The application of machine learning and knowledge engineering techniques over player's statistical data is a novel approach using Data mining techniques for this purpose. In this paper some standard machine learning techniques applied over training data collected from two different games (Volley-Ball and Basket-Ball). The physical characteristics are used for identification of quality factors among players which helps to estimate the player's correlation in abilities among two games. The strong ARM (Association Rule Mining) applied for selecting highly cohesive qualities which improves player skills suitable for both games. When in National or International championship games substitutions for players is in scarcity for specific games this approach provides multigame players and increases the chance of winning trust in teams.

Index Terms*: *ARM, Characteristics, Substitution, Multi-Game Players, Machine Learning.

I. INTRODUCTION

Sports historical data mining using Neural Networks predict the future trends of data [6]. There is a need of substitution for players when players belonging to a game are injured seriously such that some other player must represent his role [4][8]. Also selecting multi-game talented players improves the success rate in National/International selected teams. The decision tree based classification uses testing criteria highly useful to split data into various branches of hierarchy by maintaining good Gain ratio at leaves (Majority class value) [6][7]. The sports world holds large amounts of data related to players, teams, tournaments, seasonal games and other statistics related to games [9]. The application of supervised or unsupervised learning techniques over sports data forecasts the performance prediction and supporting decision making over criteria based analysis with suggestive guidelines for improvement [5]. The KDD process used to classify players based on their roles and identifies classification rules for future data analysis [1][7]. The ability to discover similarity patterns among player communities possible by Machine Learning techniques [7][9]. The winning team statistics are mined with data mining tools to effectively identify the contribution of each player for team success [8]. The player ratings are quantified by Neural Networks and Bayes Theorem scores effectively [3][10]. The Outlier Analysis technique helps in sports data mining to reveal the weaknesses among teams, improvement factors and player inabilities leading to team failure [10].

In Section 2 player statistics for Basketball and Volleyball players along with role of players in team are described. In Section 3 training data sets considered for this paper and data mining tools used for analysis work overviewed. The methodologies used in this work discussed in Section 4 followed by Section 5 and 6 result analysis and conclusion respectively.

Revised Manuscript Received on 30 May 2019.
***** Correspondence Author
Dr B.V. Ram Krishna*, Professor, CSE- Department Name, KIET, Kakinada, India.
Mr. Vadali Srinivas, Assoc. Professor, CSE- Department, KIET, Kakinada, India.
Smt. B. Sushma, Asst. Professor, CSE- Department, Vardhaman Engineering College, Hyderabad, India.

II. PLAYERS STATISTICS

A. Volley Ball players Statistics

The Volley ball game is a team oriented strategic game. Players need some physical characteristics criteria decided by national and international sports authorities. Table 1 represents standard physical characteristics consideration for team selection by global sports authorities.

Table1. Volley ball Player Physical Characteristics

Characteristic(s)	Acceptance limit
Chest (cm.)	26.5 – 29.0
Weight (Kg.)	152 – 175
Height (Cm.)	178 – 215
Core Strength	26 – 35
Agility	62.5 – 76
BFA (%)	7.0 – 9.50
Pulse Rate	82 – 102
Jump Rate (Inch.)	30.0 – 34.50

All players in team of 6 players play different roles during game play. Each Player assigned to a role by coach based on their physical characteristics. The role assignment criteria are shown in Table 2.

Table2. Role Assignment Criteria (*Volleyball*)

Role(s)	Player(s) Assignment
Setter	1
Middle Blocker	1
Wing Spiker	1
Libero	1
Defender	1
Opposite Hitter	1

B. Basket Ball Player Statistics

Basket ball game is another team oriented strategic game having some similarities with Volley ball game. The player's physical characteristics criteria decided by National and International sports authorities given in Table 3. Players are validated according to the characteristics for selection into team.

Characteristic(s)	Acceptance limit
Chest (cm.)	26.5 – 29.0
Weight (Kg.)	152 – 175
Height (Cm.)	178 – 215
Core Strength	26 – 35
Agility	62.5 – 76
BFA (%)	7.0 – 9.50
Pulse Rate	82 – 102
Jump Rate (Inch.)	30.0 – 34.50

Table3. Basket ball Player Physical Characteristics

Role(s)	Player(s) Assignment
Point Guard	1
Shooting Guard	2
Small Forward	2
Power Forward	2
Center	2
Swing Man	1
Stretch Four	2

Table4. Role Assignment Criteria (*Basketball*)

The players in a team of 12 players have to take different roles during game play. The Table 4 shows the roles and allotment of players for each role. The principle physical abilities are identified as Core Strength, Agility, Jump rate and Leaping Ability for both game players. There is a commonality between roles among Basketball players and Volleyball players as shown below.

- Setter – Point guard roles are similar with tasks like team leading, strategy implementation and team coordinating.
- Wing Spiker - Small/Power Forward roles are identical such as controlling game play with defensive nature and making strategic moves around the game court.
- Shooting Guard - Defender roles are meant for blocking the center of game court and avoiding opponent members in reaching our goal zones.
- Opposite Hitter – Swing Man roles are the same where in Basketball he cause obstacle for forwarder and passes ball to team members forwarding other side. In Volleyball hits the ball to opponent side suddenly.
- Center – Middle Blocker roles are for making a strong opposition to other teams who are crossing center zone of court.
- Libero – Stretch Four plays a common role of strategically turning the direction of game play movement.

III. TRAINING DATA SETS

The training data sets used in this work are collected from past National and International game statistics. The collected data represented in MS-Excel sheets and converted into Tab-delimited text files/CSV files which can be used as training data sets in Data Mining tools (TANAGRA, Orange) used in this work. The training data sets are collected on categories like player physical characteristics, winning team characteristics, player abilities and player roles in two popular games Basket-Ball and Volley-Ball.

IV. METHODOLOGY

The player's physical characteristics information summaries are pre-processed and applied to ETL process. The training data set generated hence ready to perform Data mining analysis. The Spearman Correlation technique applied using Orange® Data Mining tool. Around 36 correlation factor combinations resulted along with correlation score values. For both games independently data evaluated. Once all the factors and scoring values obtained a Boolean data sheet constructed based on players and physical characteristics. A-Priori algorithm tool by Orange® Data Mining tool applied on Boolean data sheet for Association Rule Mining. From these two approaches results collected for the analysis work in this paper. The methodology helps to identify the similar skills among players of different games and their interchangeability rate with respect to Support and Confidence values. The Correlation factor rates for combinations of player characteristics classifies the associative dependency into Strong (+ Sign values) and Weak (- Sign values).

V. RESULT ANALYSIS

A. Correlation Analysis

The Spearman Correlation applied over players of two games physical characteristics. Figure 1 represents the data mining tool generated Correlation matrix for Basket ball player's physical characteristics training data. Figure 2 represents the Correlation matrix for physical characteristics training data for Volleyball players.

(Corr)-Matrix	Agility	BFA	Core Strength	Height	Jump rate	Leaping Ability	Pulse	Speed	Weight
Agility	0	+0.662	-0.083	+0.070	-0.251	+0.624	+0.272	+0.678	+0.095
BFA	+0.662	0	-0.054	-0.011	+0.171	+0.566	-0.100	+0.561	+0.151
Core Strength	-0.083	-0.054	0	-0.024	-0.136	+0.079	+0.117	-0.189	+0.290
Height	+0.070	-0.011	-0.024	0	-0.177	-0.292	+0.259	+0.188	-0.288
Jump rate	-0.251	+0.566	-0.136	-0.177	0	-0.031	-0.540	-0.124	-0.028
Leaping Ability	+0.624	+0.566	+0.079	-0.292	-0.031	0	+0.262	+0.534	+0.223
Pulse	+0.272	-0.100	+0.117	+0.259	-0.540	+0.262	0	+0.131	-0.122
Speed	+0.678	+0.561	-0.189	+0.188	-0.124	+0.534	+0.131	0	-0.174
Weight	+0.095	+0.151	+0.290	-0.288	-0.028	+0.223	-0.122	-0.174	0

Fig1. Correlation Matrix for Basketball players

(Corr)-Matrix	Agility	BFA	Core Strength	Height	Jump rate	Leaping Ability	Pulse	Speed	Weight
Agility	0	+0.314	-0.012	+0.356	+0.124	-0.067	-0.264	+0.482	+0.134
BFA	+0.314	0	-0.267	+0.257	-0.033	+0.430	-0.117	-0.024	+0.287
Core Strength	-0.012	-0.267	0	+0.190	+0.532	-0.298	-0.355	-0.002	-0.327
Height	+0.356	+0.257	+0.190	0	+0.249	+0.171	-0.240	+0.264	+0.080
Jump rate	+0.124	-0.033	+0.532	+0.249	0	-0.268	-0.567	-0.176	-0.217
Leaping Ability	-0.067	+0.430	-0.298	+0.171	-0.268	0	+0.270	-0.058	+0.259
Pulse	-0.264	-0.117	-0.355	-0.240	-0.567	+0.270	0	+0.375	+0.447
Speed	+0.482	-0.024	-0.002	+0.264	-0.176	-0.058	+0.375	0	+0.298
Weight	+0.134	+0.287	-0.327	+0.080	-0.217	+0.259	+0.447	+0.298	0

Fig2. Correlation Matrix for Volleyball players

B. Association Rule Mining

The correlation based measures for each physical characteristic metric used in Apriori algorithm to identify frequent item sets.

The sets generated are collection of physical characteristics which having high associativity among each other.

ARM	Item set(s)
Rule_1	{Agility, BFA, Speed}
Rule_2	{Agility, Core Strength, Height, Pulse, Speed, weight}
Rule_3	{BFA, Core strength, Leap Ability, weight, speed}
Rule_4	{Agility, Height, Leap Ability, Pulse}
Rule_5	{Jump rate, Weight}

Table5. ARM Item sets

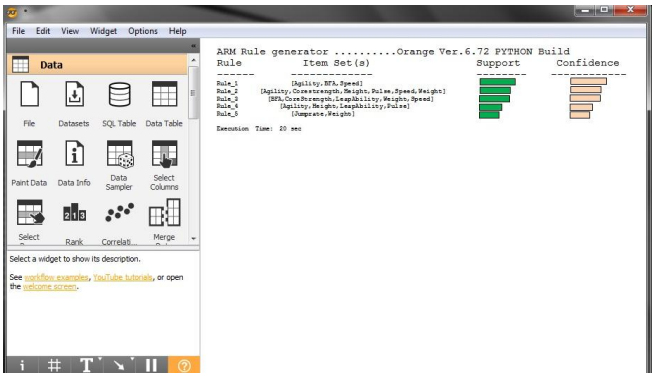

Fig3. ARM Mining for Player statistics

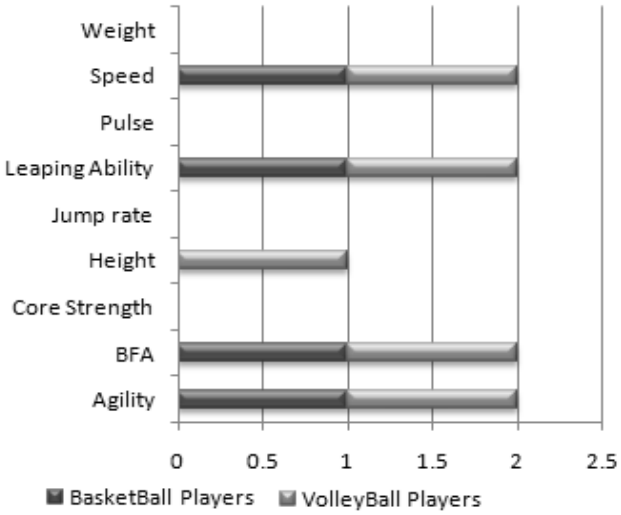

Fig4. Core Factor Weight age Graph

The Table 5 gives cohesive physical factors as item sets depending on Association Rule Mining performed by data mining tool. The Support and Confidence metrics are calculated for generating core factor weight age among Basketball players and Volleyball players. Results shown by Figure 4 indicate high similarity of distribution of skills among two game players. This proves that interchange of players between games is possible and acceptable.

The collection of Figures 5 to 8 are generated using correlation factor measure among principle physical characteristics required for both game players to play effectively in tournaments. The distribution values of player characteristics in collected training data from national and international player physical abilities showing number of intersecting data points on graph. Specifically in jump rate characteristic large scale data points are highly overlapped indicating 'Jump rate' skill is highly essential in both games and the selected players in both games are having almost similar skill rate in this characteristic. Hence 97.3% support shown for interchange of players between these games based on this skill. The remaining principle skills Agility based 77.4% chance, Core Strength based 69.2% chance and Leaping Ability based 97.6% chance of support measure obtained for interchanging of players among two games.

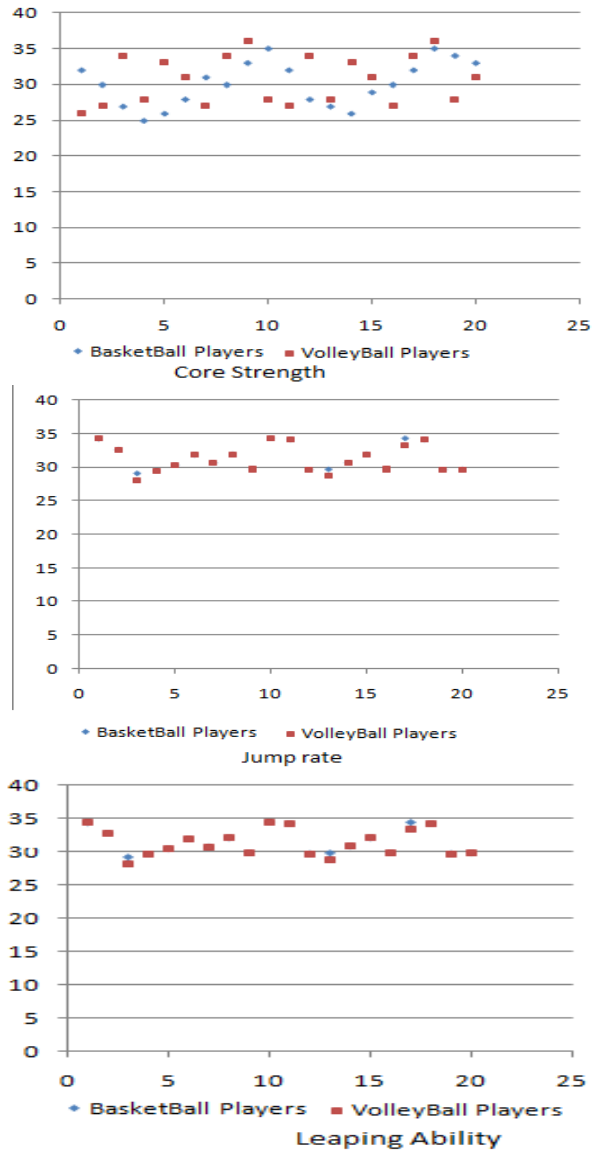

Fig 5-8. Correlation measure graphs

Finally an average of 85.37% of support with 92.4% of confidence obtained for player interchangeability among Volleyball and basketball teams. The similar role based interchanging strategy is acceptable for high success rate among teams.

VI. CONCLUSION

The analysis conducted in this paper results high scale quality factors identification among Basket-Ball and Volley-Ball players. The factors weight age shows similarity score in skills among both game players.

The multigame players identified using approach discussed in this paper selects high ability and good success rate players who can be interchanged between two different games. The work also identified highly influencing cohesive physical factors and their relation among Basket-ball and Volley-ball players. This work provides a view of improvement in physical factors among players. In future on the statistics of National/International winning teams clustering analysis and classification rules to be done. Identification of outliers in both teams to judging the failure causes is another task where we have to study and analyze the results for constructing a better decision support system.

REFERENCES

1. Bruno Almeida Odierna and Ismar Frango Silveira, "Player Game Data Mining for Player Classification", Proceedings of SB-Games 2018, ISSN: 2179-2259, Brazil, Oct29-Nov1, 2018.
2. Edward Nsolo et. al.,"Player valuation in European football", ECML/ PKDD Workshop on Machine Learning and Data Mining for Sports Analytics, Ireland, Sep-2018.
3. Tahir Syed et. al., "Automated Player Selection for a Sports Team using Competitive Neural Networks", IJACSA, Vol. 8, No. 8, 2017.
4. Rehana Mahfuz et. al., "Analyzing Sports Training Data with Machine Learning Techniques", Purdue University, University-Press, Oct-2014.
5. Fanhui Meng and Qingli Li, "Application of Data Mining in the Guidance of Sports Training", ICSEM-13 Conference Proceedings, PP 523-528, China-2013.
6. R. R. Huilgol and S. S. Chhabra, "A Review of Data Mining in Sports", IJCCTE, Issue-9, PP: 121-124, Sep-2012.
7. Anders Drachen et. al., "Game Data Mining", © Springer-Verlag Berlin Heidelberg 2011.
8. Adam Hipp and Lawrence J. Mazlack, "Mining Ice Hockey: Continuous Data Flow Analysis", First International Conference on Advances in Information Mining and Management, ISBN: 978-1-61208-162-5, 2011.
9. Osama K. Solieman, "Data Mining in Sports: A Research Overview", MIS Masters Project, August 2006.
10. Robert P. Schumaker, Osama K. Solieman and Hsinchun Chen, "SPORTS DATA MINING", Text Book, Springer series..

AUTHORS PROFILE

Dr Bh. V. Rama Krishna Currently working as Professor in CSE-department of KIET. His areas of interest are Data Mining, Mobile Computing and Software Engineering. He is an author of few publications and a lifetime member of CSI-India. He received his M.Tech. (CSE) from JNTUK. He received his Ph.D. (CSE) from Acharya Nagarjuna University.

Mr Vadali Srinivas Currently working as Associate Professor in CSE-department of KIET. His areas of interest are Data Mining, Machine learning and Optimization Techniques. He received his M.Tech. (CSE) from GITAM Deemed University. He is pursuing d his Ph.D. (CSE) from JNTU-Kakinada. He is having 10 years teaching experience.

Smt. B. Sushma Completed her M.Tech (SE) from JNTUH. She is a reputed author for few journal publications. She is currently working as Assistant Professor in VCE-Hyderabad. Her areas of interest are Machine-Learning, Image Processing and Computer Networks.