International Journal of Multimedia and Ubiquitous Engineering
Vol. 9, No. 11 (2014), pp. 293-302
http://dx.doi.org/10.14257/ijmue.2014.9.11.29

Evaluating Technical and Tactical Abilities of Football Teams in Euro 2012 Based on Improved Information Entropy Model and SOM Neural Networks

Minjian Wang

Jiaxing Vocational Technical College, Jiaxing, Zhejiang, China, 314036
25092779@qq.com

Abstract

In this paper, improved information entropy model and SOM neural networks are proposed for evaluating technical and tactical abilities of football teams in Euro 2012.Information entropy model is an effective performance evaluation tool. As a fine clustering tool, SOM has been used widely. The approach make input space cluster automatically and the learning process of clustering doesn't need supervision. At the same time, eleven indicators that reflect technical and tactical ability of football teams are selected. Then, by means of entropy method and clustering analysis with statistic data of Euro 2012, characteristic and law is discovered between competition results and technical and tactical abilities, and a method for comprehensive assessment of technical and tactical abilities of football teams was proposed.

Keywords: *Euro 2012, entropy method, clustering analysis, technical and tactical abilities*

1. Introduction

The 2012 European Football Championship (hereinafter referred to as the "Euro 2012 ") has come to close. The overall skill level of each team was relatively high and the gaps among those teams were narrow, which allowed the audiences to enjoy high-quality football games. It can be said that headed by Spain, the European top teams represent the development trend of the world football games. Therefore, we need to study Euro 2012 team in aspects of technical and tactical abilities to identify the success factors of the European teams so as to provide some references and lessons for the development of football in other countries [1, 2].

2. Research Objective and Evaluation Indicators

Our research targets the 16 teams participated in Euro 2012. First adopts group single round robin scoring rule, which means that two teams winning out every group will become the final eight, then through playoffs, the final four, champion and runner-up teams will produced, which involve a total of 31 games. In order to comprehensively evaluate the technical and tactical ability of a team, we need to consider the factors that reflect the team technical and tactical ability, and take the numbers of game played into consideration. So we choose each team's average value of goal, shoot, corner kick, offside, free kick, pass kick ball, possession percentage, crossing, fumble, being shot and foul as the indicator to reflect technical and tactical ability. Then through statistics, we can obtain the statistic data of technical and tactical ability of the16 teams in Euro 2012, as shown in table 1 and Table 2.

ISSN: 1975-0080 IJMUE

Table 1. Offensive Statistic Data of Euro 2012[1]

Team	Goal w_1	Shoot w_2	Corner kick w_3	Offside w_4	Free kick w_5	Shoot kick w_6	Possession Percentage w_7	Crossing w_8
Germany	2	16	7	2.4	12.6	12.6	59.1	25.2
Spain	2	17	7.33	3.17	17.17	13.5	65.73	15.83
England	1.67	13.33	5.33	2	14.33	8.33	39.98	24.33
Czech	1	9.75	5.25	2	13.25	8.25	48.08	14.75
Portugal	1.2	16.4	8.2	2	14.6	11.6	43.84	25.2
Italy	0.67	18.33	5	2.67	16.83	15.7	51.78	17.33
Greek	1.25	8	2.5	3	17	4.75	38.55	15.25
Russia	1.67	19.67	7	1.33	11.33	15.7	59.93	16.67
Croatia	1.33	10.67	4.67	0.67	13.67	8	43.1	25
France	0.75	16.25	7	1.25	9	12.8	53.95	19.75
Denmark	1.33	9	5.33	2.67	8.33	7	48.3	18
Ukraine	0.67	12.67	6	1.33	16	11	52.57	16.33
Sweden	1.67	13	2.33	2.33	11.67	11.3	46.3	17
Poland	0.67	15.67	4.67	1	16	11.3	45.2	23
Netherlands	0.67	20	7.33	1	11.67	16.7	56.73	22
Ireland	0.33	9.33	2.67	3.67	14.33	7.67	33.87	16.67

Table 2. Defensive Statistic Data of Euro 2012

Team	Fumble w_9	Shot w_{10}	Foul w_{11}
Germany	0.8	11.2	9.8
Spain	0.17	8.33	13.83
England	1	29.33	15
Czech	1.5	14.5	18.25
Portugal	0.8	10.2	18
Italy	1.17	13.5	14.83
Greek	1.75	20.75	12
Russia	1	13	14.33
Croatia	1	14.33	20.67
France	1.25	9.75	12.75
Denmark	1.67	19.67	12.67
Ukraine	1.33	12.67	10.33
Sweden	1.67	17.67	17
Poland	1	10.33	18.67
Netherlands	1.67	14.33	10
Ireland	3	22.67	17

3. Information Entropy Model Based on Genetic Algorithm

Normalization of original evaluating matrix is essential. Suppose there are m evaluating indicators, and n evaluating objects, then an original indicators value matrix $(a_{ij})_{m\times n}$ is formed, where a_{ij} is the data of the evaluating object on the indicator. After

[1] Data of table1 and talbe 2 comes from：http://euro2012.sina.com.cn/

normalization the original evaluating matrix, $(a'_{ij})_{m\times n}$ can be got, and $a_{ij}^{'} \in [0,1]$. Among these indicators, where the bigger is the better, $a_{ij}^{'} = a_{ij} \ / a_{j\max}^{*}$.While the smaller is the better, $a_{ij}^{'} = a_{j\min}^{*} / a_{ij}$ [3-5].

Information entropy has been widely used in system evaluation theory. Information entropy is defined for a set number of columns:

$X = \{X_1, X_1 \cdots, X_n\}(X_i \geq 0, \sum_{i=1}^{n} X_i = 1)$, function $H(X) = -\sum_{i=1}^{n} X_i \ln X_i$ is the information entropy of sequence X, X_i is the attribute information.

Suppose the number of teams is n, and the number of indicator is m, the benchmark value in the evaluation period is a'_{ij}, the weight of each indicator in team performance evaluation system is w_j, the method to construct the objective distinction degree of team performance using information entropy is:

STEP1: construct the comprehensive performance equation of the i team:

$$B_i = \sum_{j=1}^{m} w_j a'_{ij}$$

Where, B_i is the comprehensive performance of a team.

STEP2: construct the comprehensive performance sequence of normalized team:

$C_i = \dfrac{B_i}{\sum_{i=1}^{n} B_i}$ where, the normalized performance of the i team is C_i 。

STEP3: construct the distinction degree of information entropy.

$$H(C) = -\sum_{i=1}^{n} C_i \ln C_i$$

According to information entropy definition, combining with normalized performance, the information entropy function of comprehensive team performance:

$$H(C) = -\sum_{i=1}^{n} C_i \ln C_i$$

Here, $H(C)$ is the information entropy of evaluation results. Obviously, the larger the value of $H(C)$ is, the greater the distinction between the evaluation of the results, indicating that the greater the overall performance differences between the various teams. Therefore, $H(C)$ can be the goal of weight solution, getting the optimal weight vector and then evaluate each team. However, the goal should satisfy the subjective judgment of the experts, so as to enable quantitative evaluation requirements consistent with experts. Thus, the established subjective constraints can jointly construct the nonlinear programming model to optimize the weights.

$$\begin{cases} \max \quad H(w) = -\sum_{i=1}^{n} \dfrac{\sum_{j=1}^{m} w_j a'_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_j a'_{ij}} \ln \dfrac{\sum_{j=1}^{m} w_j a'_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_j a'_{ij}} \\ w_j^{k1} > w_j^{k2} \\ \sum_{j=1}^{m} w_j = 1 \end{cases}$$

$w_j^{k1} > w_j^{k2}$ indicates the indicator of j^{k1} is more important than that of j^{k2}. For such non-linear programming, planning software can be used to seek the solution.

Pay attention to offending and defending the mainstream that the balance has already become sport development of current football,so

$|w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 - w_9 - w_{10} - w_{11}|$ is a very little positive number. Therefore, we establish the model as bellows:

$$\begin{cases} \max \quad H(w) = -\sum_{i=1}^{n} \dfrac{\sum_{j=1}^{m} w_j a'_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_j a'_{ij}} \ln \dfrac{\sum_{j=1}^{m} w_j a'_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} w_j a'_{ij}} \\ |w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 + w_8 - w_9 - w_{10} - w_{11}| < 0.15 \\ \sum_{j=1}^{11} w_j = 1 \end{cases}$$

This is a nonlinear programming that can't be solved in a traditional way.A genetic algorithm (GA) is a search technique used in computing to find solutions to optimization problems[6]. Genetic algorithms can be categorized as meta heuristics with global perspective. Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover. Genetic algorithms are implemented as a computer simulation in which a population of abstract representations of candidate solutions to an optimization problem evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. We will use genetic algorithm to solve it [7, 8, 17, 18].

4. Evaluation

The weight of the 11 indicators calculated by information entropy is shown in Table 3.

Table 3. The Weight of the 11 Indicators

indicator	w_1	w_2	w_3	w_4	w_5	w_6
weight	0.08744	0.067934	0.026577	0.051155	0.04672	0.083112
indicator	w_7	w_8	w_9	w_{10}	w_{11}	
weight	0.066657	0.078749	0.126959	0.150261	0.214928	

Table 2 shows that in terms of indicators, the gap of possession percentage among teams is the widest. So weight of possession percentage is the greatest, followed by weight of shooting and foul. The score of the technical and tactical ability of every team, as shown in Table 4.

Table 4. Score and Competition Results

Team	offensive score	ranking	Defensive score	ranking	Comprehensive socore	ranking	result
Germany	0.43355	2	0.353673	2	0.787265	2	top four
Spain	0.443131	1	0.429579	1	0.872676	1	champion
England	0.360405	7	0.204678	15	0.565143	10	top eight
Czech	0.293678	14	0.216136	12	0.509818	15	top eight
Portugal	0.383239	4	0.266734	6	0.649996	5	top four
Italy	0.375543	6	0.253204	9	0.628764	8	second place
Greek	0.288366	15	0.248176	10	0.536567	13	top eight
Russia	0.402836	3	0.264863	7	0.667734	4	third in their group
Croatia	0.317642	11	0.210847	13	0.528537	14	third in their group
France	0.332688	9	0.31086	5	0.64356	6	top eight
Denmark	0.305931	13	0.242799	11	0.548743	12	third in their group
Ukraine	0.312911	12	0.318923	3	0.631855	7	third in their group
Sweden	0.345223	8	0.207666	14	0.552927	11	fourth in their group
Poland	0.329022	10	0.255593	8	0.584647	9	fourth in their group
Netherlands	0.376022	5	0.310899	4	0.686967	3	fourth in their group
Ireland	0.26953	16	0.186307	16	0.455839	16	fourth in their group

From Table 4, we see that Spanish team makes the highest score of technical and tactical ability, and wins the champion; Germany and Netherlands teams rank the second and third respectively, thus entering the semi-finals; while in terms of technical and tactical ability, Ireland and Czech team rank backward, and their competition results are not desirable. Overall, score in technical and tactical ability stands at the same level with competition results.

5. SOM Neural Networks Introduction

There are four common clustering methods-k-means, hierarchical clustering, SOM and FCM, among them, as for k-means, the initial point is unstable and randomly selected. Consequently, although the unstable hierarchy clustering of clustering results does not need to determine the number of category, the clustering quality is restricted and it cannot be corrected once a split or merger is executed [9-11]. FCM is sensitive to the initial cluster centers and it needs to artificially determine the number of clusters, so it is easy to fall into local optimal solution. SOM clustering algorithm simulates the Self Organizing Feature Map function of the brain systems, and it is a competitive learning network, and the learning is a kind of self-organizing learning without supervision. As shown in Figure 1, SOM network has a total of two layers: the input and output layer. The input layer: it gathers the outside information together and brings them to each neuron in the output layer. The form of the input layer is similar with that of the BP network, and their number of nodes and sample dimension are the same. For the output layer, it is also a competitive layer, whose neuron arrangement patterns include one-dimensional linear array, two-dimensional plane array and three-dimensional grid array. The most typical one is the two-dimensional form, for it is more similar with the cerebral cortex image, as shown in Figure 1.

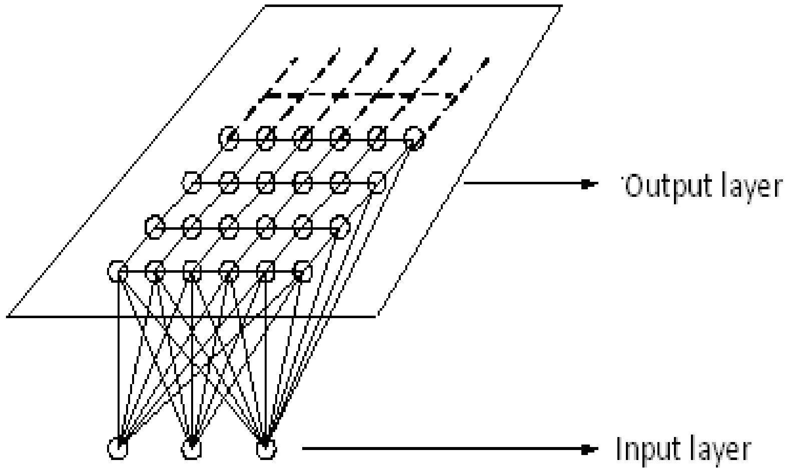

Figure 1 Mapping Network Structure of Self-organizing Feature

Working principle [12, 13]: SOM network operation can be divided into two phases-training and operation. In the training phase, the network randomly inputs the samples in training set, and for a particular input pattern, one of the nodes in the output layer will generate maximum response and win out, while at the beginning of the training phase, it is difficult to identify the specific node in the output layer that generates maximum response. When the input mode type is changed, winning node in the two-dimensional plane is also changed. Due to lateral mutual excitement, the nodes nearby the winning node also have a greater impact, so the input direction of weight vector connecting the winning node and all the nodes in the winning area shall be adjusted to different degrees, the different degrees of adjustment will gradually decreases according to the distance between the winning node and all the nodes in the surrounding area.

Through a self-organized way, the network adjusts network weights with a large number of training samples, and finally makes all the nodes in the output layer become neurons that are sensitive to a particular pattern type and the corresponding inner star weight vector become the central vector of each input mode. And when the features of two modes are similar, locations of their corresponding nodes are also closer. Consequently, orderly characteristic pattern that can reflect the sample mode classification is formed in the output layer [14, 15, 19, 20].

6. Clustering Analysis Based on SOM Clustering Method

The sample dimension in Table 1 is 11, so the input layer has 11 nodes. We classify the 16 teams into strong, general and weak teams, so the input layer has 3 neurons. After iteration of 10000 times using tool in Matlab2008a, reasonable study and evaluation of 16 teams' technical and tactical ability can be completed by the self-organizing feature mapping network. The result of the program operation and clustering results are shown in Figure 1 and Table 5.

Table 5. Program Operation Result

Numbers in Training	Clustering Result															
1000	4	3	1	2	1	3	2	4	1	4	4	3	2	1	4	2
5000	4	3	1	2	1	3	2	4	1	4	4	3	2	1	4	2
10000	4	3	1	2	1	3	2	4	1	4	4	3	2	1	4	2

Table 2 shows that in SOM clustering operation tends to be stable, and produce good clustering results.

7. Clustering Results Evaluation

Judging from the clustering results in Table 5, we can get the specific concrete results, as shown in Table 6.

Table 6. SOM Clustering Results

Category	Teams
1	England, Portugal, Croatia, Poland
2	Czech, Sweden, Greek, Ireland
3	Spain, Italy, Ukraine
4	Germany, France, Netherlands, Russia, Denmark

Then, we calculate the average value of the category of each index, and the results are shown in Table 7 and Table 8 [12, 13].

Table 7. Average of Each Classification

Category	Goal	Shoot	Corner Kick	Offside	Free Kick	Shoot Kick	Possession Percentage	Crossing
1	1.22	14.02	5.72	1.42	14.65	9.82	43.03	24.38
2	1.06	10.02	3.19	2.75	14.06	8.00	41.70	15.92
3	1.11	16.00	6.11	2.39	16.67	13.39	56.69	16.50
4	1.28	16.18	6.73	1.73	10.59	12.94	55.60	20.32

Table 8 Average of Each Classification

Category	Fumble	Shot	Foul
1	0.95	16.05	18.09
2	1.98	18.90	16.06
3	0.89	11.50	13.00
4	1.28	13.59	11.91

Category 1 includes England, Portugal, Croatia and Poland, whose technical and tactical abilities are average, and their Possession Percentage and Shoot Kick are 43.03% and 9.82 respectively, the ratio between goal and fumble (goal / fumble) is 1.28:1, their attacking and defending balances are good, but their fouls number 18.09 times, significantly more than that of the other groups. Especially the Croatia team fouls an average of 20.67 times and gets 3 yellow cards per game with more barbaric style.

Category 2 includes Czech, Sweden, Greece and Ireland, who have relatively weak technical and tactical ability, and their possession percentage and shoot kick are 41.70.47

and 8.00 respectively, goal-fumble ratio is 0.53, these data are basically the lowest level among the groups and fouls number is larger. It is worth mentioning that Greece, although with weak technical and tactical ability, focuses on teamwork with positive struggling spirit, and enters the final eights, which should inspire the Chinese football team.

Category 3 includes Spain, Italy and Ukraine, who have the strongest technical and tactical ability. Spain is the champion in this European Cup, for it has top star midfielders like Xavi, Iniesta, Silva, Fabregas and Mata. Italy team is the runner-up for its players have strong offensive strength with exceptional defensive ability. Although not a traditionally strong teams, Ukraine takes the advantage of home field and gives full play to their technical and tactical ability. The most significant characteristics in Category 1 is higher possession percentage, reaching 56.69%, goal kick reaching 16.67 per game, while attack and defend is more balanced , the goal-fumble ratio is 1.25:1. Category 2 includes Germany, the Netherlands, France, Russia, and Denmark, who have relatively strong technical and tactical abilities. Their possession percentage and shoot kick are 52.77% and13.14 respectively, not as good as Category 3.

8. Conclusion

The results are generally agreed with the actual results of the competition, indicating that the application of entropy method and SOM clustering analysis in comprehensive evaluation of the overall attacking and defending strength of each team has practical significance. By the above analysis, we can see that the most important indicator reflecting a team's technical and tactical ability is possession percentage and goal kick. In modern football game, a team not only needs to have a good attacking and defensive capabilities, technical level, but also needs to have a positive struggling spirit and attitude. Of course, we also find that a team with higher technical and tactical skills does not necessarily get higher scores, for football game has contingencies and there more factors influencing the results of the games.

References

[1] F. Tang, X. K. Chen, H. Liu and X. J. Xu, "Journal of Beijing University of Physical Education", vol. 2, **(2001)**.
[2] T. Quan, X. H. Sun, "Journal of Applied Statistics and Management", vol. 30, no. 2, **(2011)**.
[3] Z. H. Zou, Y. Yun and J. N. Sun, "Journal of Environmental Sciences", vol. 18, no. 5, **(2006)**.
[4] B. D. Fan, "China Rural Water and Hydropower", vol. 9, **(1998)**.
[5] D. C. Fang, G. L. Liu, L. P. Zhang, *et al.*, "ValueEngineering", vol. 2, **(2004)**.
[6] R. Cheng, M. Gen, "Tsujimura , Computers＆Industrial Engineering", vol. 4, no. 30, **(1996)**.
[7] S. Niu, L.Chen, "Journal o f Northwest Normal University (Natural Science)", vol. 1, no. 47, **(2011)**.
[8] R. Storn and K. Price, "International Computer Science Institute", **(1995)**.
[9] Q. G. Hu, H. X. Chen and X. D. Xu, "Water Resources and Water Engineering", vol. 2, no. 112, **(2000)**.
[10] D. H. Liang and H. H. Jiang, "Environmental Monitoring in China", vol. 2, no. 18, **(2002)**.
[11] T. S. Wang and P. Zhong, 'Journal of Guangzhou Physical Education Institute, vol. 25, **(2005)**.
[12] V. Faber, "Los Alamos Science", vol. 22, **(1994)**.
[13] M. Erisoglu, N. Calis and S. Sakallioglu, "Pattern Recognition Letter", vol. 32, **(2011)**.
[14] M. C. Su and H. T. Chang, "IEEE Transactions on Neural Networks", vol. 12, no. 1, **(2001)**.
[15] R. Singh, V. Cherkassky and N. Papanikolopoulos, "IEEE Transactions on Neural Networks", vol. 1, no. 11, **(2000)**.
[16] H. Yin and N. MAllinson, "IEEE Transactions on Neural Networks", vol. 2, no. 12, **(2001)**.
[17] T. Back, "Evolutionary algorithms in theory and practice", Oxford University Press, **(1995)**.
[18] D. E. Goldberg, "Genetic algorithms in search optimization and machine learning", **(1989)**; Addison Wesley, NewYork.
9.2. Conference Proceedings
[19] G. Lin, "Proceedings of the 29th Chinese Control Conference", **(2010)**; Beijing, China.
[20] L. Congdong, Z. Quan and T. Yongli, "Proceedings of 2010 IEEE the 17th International Conference on Industrial Engineering and Engineering Management", **(2010)**.

Author

Minjian Wang: physical education teacher of Jiaxing Vocational Technical College in China.Title: lecturer,Degree: master of physical education ,key research area: football Teaching.