# Research on the Application of Kohonen Neural Networks in Football Clustering Analysis

**Cenyue Wang, Youn Poong Oh***
*Department of Physical Education, Kunsan National University, Jeollabuk-do, Gunsan-si, Republic of Korea,*
**Corresponding Author.*

**Abstract: This study applies Kohonen neural networks for clustering analysis in sports. By iteratively optimizing the objective function, it effectively avoids numerous subjective factors, providing a novel and efficient approach for obtaining objective clustering results. The results demonstrate that clustering analysis using Kohonen neural networks offers a clear practical value in evaluating the comprehensive strength of soccer teams. It serves as an effective method for rational, effective, objective, and quantifiable assessment of team tactics and strategies. Furthermore, this method is readily applicable to other comprehensive evaluations in competitive sports.**

**Keywords: Sports; Clustering Analysis; Kohonen Neural Networks; Soccer**

## 1. Introduction

Artificial Intelligence (AI) is a field developed through the integration of multiple disciplines, including philosophy, computer science, systems science, and cybernetics. It aims to simulate, extend, and enhance human intelligence. AI possesses the capability to discover and solve problems using human knowledge to improve survival capabilities. It is commonly employed to address problems that are challenging for traditional computing methods. This stage represents the apex of human understanding and transformation of the objective world in the era of the information revolution. Research suggests that the short-term goal of AI is to mimic and execute human cognitive functions through machines, while the long-term goal is to emulate human intellectual functions and cognitive activities through automated systems. The Kohonen neural network, proposed by Professor Teuvo Kohonen of the University of Helsinki in 1981, is an artificial neural network characterized by its excellent classification performance and complete self-organizing feature map [1]. This network is based on AI technology. In recent years, various methods for clustering analysis in sports have emerged. However, many of these methods are affected by numerous subjective factors during analysis, which can impact the objectivity of clustering results. The Kohonen neural network, which simulates the mapping functions of the human brain's neural system and operates as a competitive learning network, is capable of unsupervised self-organizing learning. Consequently, employing the Kohonen neural network for clustering analysis can effectively minimize human subjectivity and rapidly produce clustering results, providing an effective method for comprehensive evaluation in soccer.

## 2. Introduction to Kohonen Neural Networks and Clustering Mechanism

### 2.1 Kohonen Neural Networks

The Kohonen neural network is characterized by the following features: (1) It has clustering functionality, where data inputs can be represented by cluster centers, thus enabling data compression; (2) Through competitive learning, it performs weight training to automatically determine different cluster centers, effectively avoiding subjective factors; (3) The network training converges relatively quickly [2].

Kohonen neural networks, initially introduced as Self-Organizing Maps (SOMs), utilize self-organizing neural networks (abbreviated as HKCNN) for clustering analysis. They employ iterative algorithms to optimize the objective function for effective classification of data sets. Figure 1 illustrates the structure of the Kohonen neural network.

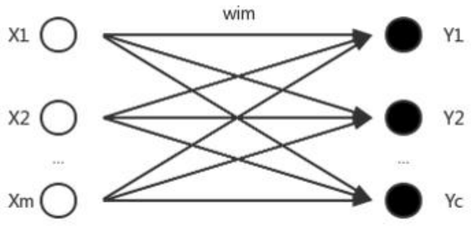


**Figure 1. Structure of the Kohonen Neural Network**

In the Kohonen neural network, the HKCNN network is a two-layer feedforward neural network. The first layer is the input layer, which contains m neurons, with the number of neurons equal to the dimensionality of the input sample vectors. The neurons in the first layer have linear response functions, while the second layer consists of competitive output neurons, with output values in the range {0, 1}. The total number of weight connections between the first and second layers is m×c. In the Kohonen neural network, the weight matrix is denoted as $W = [W1, W2, \ldots, Wc]^T$, where $W_i = [W_{i1}, W_{i2}, \ldots, W_{im}]^T$ and i=1,2,3,...,c. Here, $W_i$ represents the i-th cluster center. The HKCNN network in the Kohonen neural network employs a competitive learning algorithm. During competitive learning, only the winning neuron updates its weights for each input sample vector, while the weights of the losing neurons remain unchanged [3].

The competitive learning process is as follows: (1) Network Initialization; (2) Add Input Stimulus; (3) Distance Calculation; (4) Determine Minimum Distance; (5) Adjust Weights; (6) Return to Step (2).

### 2.2 Clustering Mechanism of Kohonen Neural Networks

Biological research has revealed several characteristics of neural cells in the human cerebral cortex: (1) Neural cells are organized in an orderly manner, generally arranged in a two-dimensional spatial pattern; (2) Neural cells in different regions of the space effectively control various body movements, with cells in specific regions responding sensitively to particular stimuli. Some neural cells exhibit strong responses due to high sensitivity, while others may show minimal to no response due to lower sensitivity. This organization allows for the orderly control of body movements; (3) Neural cells in different spatial regions develop varying sensitivities to different stimuli, which are closely related to an individual's postnatal motor habits and training, such as the phenomenon of "left-handedness."

The Kohonen neural network simulates the neural cells of the human brain. The input layer represents various stimulus signals, and the output layer nodes simulate neural cells. Neural excitation is primarily triggered by signal stimuli. After receiving the sample data stimulus at the input points, the signals are transmitted through network connections to the output nodes. The output nodes exhibit varying sensitivities to different inputs and influence neighboring nodes through lateral connections. Clustering in Kohonen neural networks involves continuously learning from samples to capture the intrinsic structural characteristics of the data. The network reflects these structural features through the Kohonen neural network. The number of input nodes is determined by the number of clustering variables, and the output nodes represent the number of clusters. The learning objective is to ensure that output nodes provide consistent outputs for samples with similar structural characteristics. Learning involves an iterative process of adjusting weights to approach consistent outputs [4].

### 2.3 Topological Structure of Kohonen Neural Networks

Fundamentally, the topological structure of Kohonen neural networks is typically composed of two layers with full connectivity and a feedforward design, as illustrated in Figure 2. The characteristics of this network topology include: (1) The network comprises two layers, namely the input layer and the output layer, with the latter also referred to as the competitive layer; (2) Each input node in the input layer is connected to nodes in the output layer; (3) The output nodes are distributed in a two-dimensional structure, with lateral connections between nodes. For output nodes, a specific neighborhood range results in neighboring cells, or adjacent nodes.

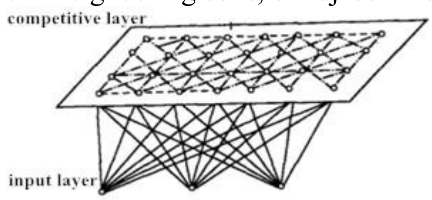


**Figure 2. Topological Structure of the Kohonen Neural Network**

## 3. Application of Kohonen Neural Networks in Football Clustering Analysis

Currently, in China, football defense and attack metrics are rarely used to comprehensively analyze various factors in matches. This limitation prevents a holistic assessment of tactical effectiveness during competitions, resulting in a lack of generalizability and comparability of the data, and significantly constrains its reference value [5]. Therefore, starting from the characteristics of football matches, this study uses key offensive and defensive metrics such as average goals scored, corner kicks, assists, goals conceded, free kicks, turnovers, fouls, and shot accuracy as evaluation criteria. The Kohonen neural network is employed to analyze and assess the comprehensive strength of football teams participating in the FIFA World Cup.

### 3.1 Research Data

The study focuses on the players of the 16 teams that reached the knockout stage of the 22nd FIFA World Cup held in Qatar, with a total of 48 matches played. The performance and tactical statistics of the teams are presented in Table 1.

**Table 1. Performance and Tactical Statistics of the Top 16 Teams in the 22nd FIFA World Cup**

| Rk | Team | G | A | FK | CK | GA | Tkl | F | OFF | SoT | LPS |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Argentina | 2.58 | 1.28 | 2.44 | 4.87 | 0.58 | 42.72 | 15.26 | 3.26 | 19.26 | 45.69 |
| 2 | France | 2.01 | 1.72 | 2.87 | 7.01 | 0.44 | 48.26 | 19.26 | 2.64 | 14.62 | 33.54 |
| 3 | Croatia | 1.44 | 1 | 1.3 | 5.58 | 0.85 | 50.72 | 17.62 | 3.18 | 16.92 | 44.58 |
| 4 | Morocco | 1.15 | 0.44 | 1.15 | 7.58 | 0.85 | 45.58 | 19.26 | 1.26 | 8.26 | 50.64 |
| 5 | Netherlands | 2.01 | 1 | 1.8 | 5.61 | 1 | 42.81 | 14.16 | 5.29 | 14.26 | 54.69 |
| 6 | England | 1.21 | 0.81 | 1.01 | 4 | 0.6 | 37.41 | 14.62 | 4.621 | 11.84 | 38.95 |
| 7 | Brazil | 1.41 | 0.61 | 0.61 | 4.41 | 1.2 | 47.52 | 20.26 | 2.16 | 13.28 | 43.28 |
| 8 | Portugal | 1.26 | 1.01 | 1.41 | 4.21 | 1.4 | 43.26 | 18.64 | 1.48 | 13.39 | 40.95 |
| 9 | Japan | 1 | 0.76 | 2 | 5.01 | 0.75 | 47.35 | 24.62 | 26.58 | 17.15 | 44.69 |
| 10 | Senegal | 1.51 | 1 | 0.26 | 4.75 | 1.25 | 35.26 | 16.95 | 3.25 | 10.64 | 42.95 |
| 11 | Australia | 1.24 | 0.51 | 2 | 4 | 1 | 46.26 | 15.45 | 1.36 | 13.25 | 56.21 |
| 12 | Switzerland | 1.51 | 1 | 2 | 5.25 | 0.75 | 34.26 | 16.24 | 2.54 | 12.95 | 43.28 |
| 13 | Spain | 1.24 | 0.26 | 1.51 | 4.25 | 1.25 | 50.26 | 17.15 | 3.61 | 10.48 | 56.18 |
| 14 | United States | 1.51 | 1 | 3.51 | 6.7 | 1.75 | 43.16 | 21.26 | 5.62 | 12.95 | 43.26 |
| 15 | Poland | 1.66 | 1 | 4.01 | 6.2 | 1.66 | 59.26 | 21.26 | 4.62 | 15.84 | 43.58 |
| 16 | South Korea | 0.66 | 0.61 | 2.76 | 4.25 | 0.66 | 35.68 | 20.65 | 1.36 | 4.45 | 40.36 |

### 3.2 Kohonen Neural Network Clustering Design

In this study, key offensive and defensive metrics from the World Cup, including assists, goals, corner kicks, goals conceded, free kicks, turnovers, tackles, offside occurrences, fouls, and shot accuracy, are used as sample input vectors during the training of the Kohonen neural network. Consequently, the network has 10 input layer nodes, m=10 [6]. The goal is to cluster the 16 football teams into categories of Excellent, Good, Average, and Poor based on these metrics, thus the network has 4 output layer nodes, c=4. Figure 3 illustrates the structure of the neural network.

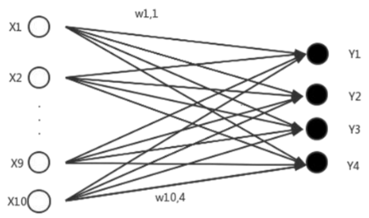


**Figure 3. Neural Network Structure for Football Analysis**

### 3.3 Football Learning Training Based on Neural Networks

In the Kohonen neural network, different cluster centers are represented by weights, and the number of output layer nodes corresponds to the number of clusters. During network training, initial weight values are chosen randomly within the range of 0 to 1, as they have minimal impact on the final training results. The network is trained using 16 samples, corresponding to the 16 national football teams, with each sample consisting of 10-dimensional vectors, representing 10 metrics. The network training is conducted for a total of 2000 iterations [7].

For the selection of the learning coefficient $\eta(t)$ , the goal is to ensure a rapid decrease during the first quarter of the training and then a slower decrease towards zero. During the first 500 iterations, $\eta(t)$ decreases from 1 to 0.1 quickly, with the change function defined as: $\eta(t) = (4999-9*t)/4990$. From the 501st to the 2000th iteration, $\eta(t)$ decreases from 0.1 to 0, with the change function defined as: $\eta(t)=1/$ (t-

491). This behavior is intended to ensure that η(t) decreases rapidly at first to quickly capture the clustering structure of the input vectors [8]. Subsequently, η(t) is adjusted based on a smaller base value to gradually approach zero, allowing for more precise adjustment of weights and ensuring a more comprehensive and objective clustering result [9].

### 3.4 Program Design

The system development environment used is Visual C++. Visual C++ provides a visual and integrated user interface that supports the development of Windows applications. The project is developed using Visual C++ 6.0 [10]. Figure 4 illustrates the program flow diagram.

### 3.5 Experimental Results

After training the Kohonen neural network, the clustering results are obtained as shown in Table 2. In this table, 1 denotes Excellent, 2 denotes Good, 3 denotes Average, and 4 denotes Poor. The specific results are detailed in Table 2.

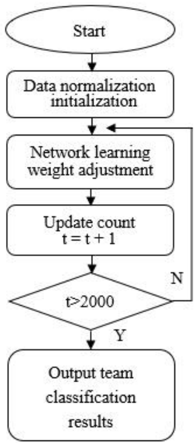


**Figure 4. Program Flow Diagram**

**Table 2. Clustering Results from Kohonen Neural Network Training**

| Rk | Team | Grade | Clusters |
|---|---|---|---|
| 1 | Argentina | Excellent | 1 |
| 2 | France | Excellent | 1 |
| 3 | Croatia | Excellent | 1 |
| 4 | Morocco | Average | 3 |
| 5 | Netherlands | Excellent | 1 |
| 6 | England | Excellent | 1 |
| 7 | Brazil | Good | 2 |
| 8 | Portugal | Excellent | 1 |
| 9 | Japan | Good | 2 |
| 10 | Senegal | Excellent | 1 |
| 11 | Australia | Average | 3 |
| 12 | Switzerland | Excellent | 1 |
| 13 | Spain | Good | 2 |
| 14 | United States | Good | 2 |
| 15 | Poland | Excellent | 1 |
| 16 | South Korea | Excellent | 1 |

## 4. Conclusion

Based on the Kohonen neural network, clustering analysis of the different metrics for 16 football teams aligns well with the overall match results. This indicates that the Kohonen neural network provides a practical and effective method for assessing the comprehensive strength of sports teams. It offers a rational, effective, objective, and quantitative evaluation of team tactics and strategies. This method is easily adaptable for application in other competitive assessments.

## Reference

[1] Luo Heng. Research on the Construction of a Risk Warning Model for Sports Event Broadcast Rights Operation Based on BP Neural Networks. Sports Science and Technology Literature Bulletin, 2023, 31(05):184-187+221.

[2] Chen Deping, Feng Yan. Application of Kohonen Neural Networks in Clustering Analysis of Sports Activities. Sports Science, 2005, 25(3):143-145.

[3] Zhao Xiying, Jiang Yu, Liu Peng. Research on a Comprehensive Evaluation Model of Middle School Students' Physical Fitness Based on Clustering Analysis and BP Artificial Neural Networks. Sports Science and Technology, 2022, 43(06):33-36.

[4] Huang Yaming, Li Xinwei, Cui Lei, et al.

 

Research on Mutual Citation Relationships of Scientific Journals - A Comparison of the Application of Kohonen Neural Networks and System Clustering Methods. Journal of Medical Informatics, 2007, 28(5):440-444.
[5] Luo Zheng, Zhang Xueqian. Detection Model of Malicious Domain Names Based on Evolutionary Thinking Algorithm Optimized S-Kohonen Neural Networks. Information Network Security, 2020, 20(06):82-89.
[6] Wen Nan. Bibliometric Report on Research Literature of Ethnic Traditional Sports in Domestic Universities - Visualization Analysis Based on Citespace. Martial Arts Research, 2018, 3(05):116-119+123.
[7] Su Yanyan, Qiu Zhiliang, Li Ge, Lu Shenglian, & Chen Ming. A Review of 2D Single-Person Pose Estimation Based on Deep Learning. Computer Engineering and Applications, 2024(07), 1-21.
[8] Zhu Shanshan. Research on Machine Learning Evaluation Methods for Knee Joint Force in Sports. Journal of Lanzhou University of Arts and Science (Natural Science Edition), 2024(03), 100-104.
[9] Sun Lin, Ma Zhongli, Chen Hao. Quantitative Analysis of Russia Sports Research in China Based on CNKI. Journal of Hebei Sports University, 2014, 28(03):30-33+63.
[10] Mao Jie, Mei Yan. Application of Gray ART Clustering Analysis Method in Monitoring Biochemical Indicators of Competitive Sports. Journal of Wuhan Sports University, 2005(10):56-58.