A STUDY TO SELECT THE BEST METHOD FOR RANKINGS OF COUNTRIES

Shahram Shafiee[1], Mehrali Hematinezhad[1], Mohsen sharari[1], Arman Dehpanah[2], Misagh Hosseini[1]
[1] *Guilan University, Rasht-Iran*
[2] *Allameh Tabataba'i University, Tehran,Iran*

Original scientific paper

Abstract
The purpose of this study was to predict the success of participating countries in the Asian Games in 2010 through macro-economic variables, political, cultural and social. The statistical sample in this research was the participating countries in Asian Games from 1974 to 2010. Ranks of countries were predicted based on KNN model and using three methods number of gold medals, total of medals and weight of medals. The research results showed that among the three methods used, ranking method by using total of medals earned the highest correlation coefficient (0.9135) between original and predicted ranks compared with methods of based on the weight of medals (0.8941) and based on the numbers of gold medals (0.8501). In all three methods, predicted ranks of China and Iran were consistent with their original ranks. Also, the ranks of Japan, South Korea and Malaysia have been predicted only with 1 rank different from their original ranks.

Keywords: *Prediction, Asian Games, Macro variables, rankings of countries, KNN Model*

INTRODUCTION

In the modern and complex world of sport an accurate, scientific and timely decisions, has a very important and determining role in the failure or success. Meanwhile, factors such as number of criteria, the complexity of data and dynamic environment, has encountered making decision on sport with a serious challenge in recent decades. Nowadays new scientific methods have been chosen for predicting outcomes in sporting events, which by using these methods in addition to identifying Influential factors in the obtained results, also the events results and sport teams classification can be predicted (Lovalgia, M.J.; Lucus, J.W., 2005). Achieving international and especially Asian sports success has become increasingly important in most of the countries. Politicians, planners, media and press consider the medals won in the tournaments as an indicator of international success, despite the International Asian Games Committee's objections and serious comments that the Medals Table cannot be a sign of competence of one country to another country. As a result, success in elite sports has been regarded increasingly as a valuable source in achievement of non- sporting goals (Green, M., 2007).

Several studies have increased the knowledge of development method of elite sports and had a major role in better understanding the sport system and the factors that shape them (Bergsgard, N.A.; Houlihan, B.; Mangset, P.; Nodland, S.I.; Rommetveldt, H., 2007; Green, M., 2007; Oakley, B.; Green, M., 2001). This subject caused the increasement of government's desire and direct intervention for the development of these sports through substantial financial investment (Bergsgard, N.A.; Houlihan, B.; Mangset, P.; Nodland, S.I.; Rommetveldt, H., 2007). In particular, the question is why some countries in international sporting competitions and events are more successful than other countries? This issue is clearly related to the work of policymakers and planners, those who tend to improve their position in the games table (De Bosscher, V.; De knop, P.; Van Bottenburg, M.; Shibli, S.; Bingham, J., 2009). However, amount of countries that spend more expenditures in sport to compete with other countries are increasing, but there are little evidences to determine the factors affecting on sporting success in international levels (Heinila, K., 1982).

On the other hand, the question is whether predictions derived from statistical models are more accurate and reliable than predictions by experts based on subjective judgments? Statistical models may provide more accurate forecasts of human judgment because they are using objective criteria to avoid bias and inaccurate interpretation of data. However, sometimes statistical models cannot consider non-quantitative factors (Hematinezhad, M.; Gholizadeh, M.H.; Ramezaniyan, M.R.; Shafiee, Sh.; Ghazi Zahedi, A., 2011). In fact the subjective judgments by experts due to the use of qualitative criteria in the unknown and uncertain conditions may do better predictions compared to the statistical models (Song, Ch.; Boulier, B.L.; Stekle, H.O., 2007).

There have been many surveys for comparison of these two methods in different issues such as Medicine, college success, Business decision-making, weather forecasting, Macroeconomics prediction, Inflation rate, Political Elections and

etc that in most of them statistical models showed better predictions (Bunn, D.; Wright, G., 1991; Wright, G.; Lawrence, M.J.; Collopy, F., 1996; Grove, W.M.; Zald, D.H.; Lebow, B.S.; Snitz, B.E.; Nelson, C., 2000; Webby, R.; O'Connor, M., 1996; Grove, W.M.; Meehl, P.E., 1996).
Also, in relation with this subject in sport more research has been done (Abrahart, R.J.; Kneale, P.E.; See, L.M., 2004; Boulier, B.L.; Stekler, H.O., 1999; Forrest, D.; Simmons, R., 2000; Forrest, D.; Goddard, J.; Simmons, R., 2005) and all of them stated that the sporting forecasts based on data and information are quite different with prediction by random and chance. In fact, they provide forecasts more accurate than Lottery.

In general, most researchs in the field of prediction can be placed in a classification as bellow:
Researches that lack appropriate methods for predicting desired events and concentrated mostly on the speculation of experts, analysts, journalists and sport gamblers and data mining have considered rarely in their predictions (Strumbelj, E.; Sikonja, M.R., 2010; Easton, S.; Uylangco, K., 2010; Min, B.; Jinhyuck, K.; Chongyoun, Ch.; Hyeonsang, E.; McKay, R.I., 2008; Scheibehenne, B.; Bröder, A., 2007).

Researches that have used data restricted to a short-term period, for example, data of one year, for their predictions. A forecasting is Scientific when determines its status in the future by review of past incidents and events about a phenomenon (McHale, I.; Morton, A., 2011; Klaassen, F.; Magnus, J., 2003; Del Corrala, J.; Rodriguez, J.P., 2010; Boulier, B.L.; Stekler, H.O., 2003; Rue, H.; Salvesen, O., 2000).

Researches that have used simple techniques for predicting. Most of the applied statistical models were regression techniques, which their ability was somewhat limited (Goddard, J., 2010; Grant, A.; Johnstone, D., 2010; Boulier, B.L.; Stekler, H.O., 2003; Rue, H.; Salvesen, O., 2000; Goddard, J.; Asimakopoulos, I., 2004; Boulier, B.L.; Stekler, H.O., 1999).

Of course, also there were some scientific researchs that have used more appropriate methods for predicting sporting events, which have been mentioned as follows:
Iyer and Sharda (2009) have used neural networks model to forecast cricket player's future performance based on their earlier performance from 1985 to 2006- 2007 season. In this study, players were divided into three categories of successful, middle and unsuccessful. The Neural network models using four sets of data were studied increasingly. Then, the neural models intended for prediction of the future performance of cricket players were used. The results showed that neural networks can provide accurate and valuable decision making in players selection process (Iyer, S.R.; Sharda, R., 2009).
Forrest et al (2010) in their paper attempted to predict the national team medal totals participating in the Beijing 2008 Summer Olympics. In this paper was used a statistical model based on a regression analysis of number of medals in the previous period and GDP. Final forecasts were successful about the fundamental changes in medal shares relative to the 2004 Games, namely the increase of medals for China and Great Britain and the substantial fall in medals for Russia (Forrest, D.; Sanz, I.; Tena, J.D., 2010).

Condon et al (1999) have predicted the success of countries in the 1996 Olympics through a neural network. In this study, Information of 271 Sporting Event for over 17 independent variables was collected from 195 countries. After examining 27 different models – three models using ordinary least squares (OLS), linear regression models and 24 neural network models, finally they concluded that the neural network model in comparison to regression and OLS model is more suitable tool for prediction the success of countries in the Olympics (Condon, E.M.; Golden, B.L.; Wasil, E.A., 1999).

In this study, by using the strengths of past research has been tried to overcome the limitations of previous research. Therefore, we have focused our research on multiple variables, long period of time (from 1974 to 2010) and using a strong quantitative model.

METHODS

2.1. This part consists of below steps:
1) Economic, political, social and cultural variables that are important contributors for international sporting success were identified by a comprehensive literature review. Then the variables were given to relevant experts in order to rank them according to their importance, and then each variable was given a specific point according to its rank. At the end, the first eleven variables were selected as effective variables to predict the ranking of participating countries in Asian Games. The proposed model for predicting the Asian Games in Figure 1 is given

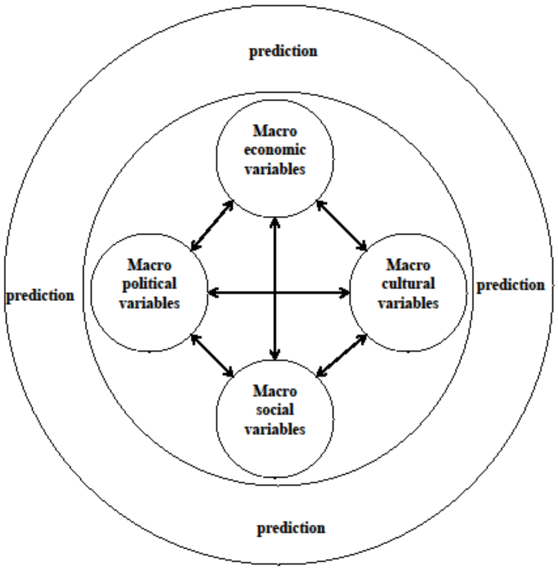

Fig 1: Model of Predicting the success of nations at the Asian Games

2) In second step, the information of selected variables for the participating countries was collected from 1974 to 2010. Additionally, the information of the countries includes Uzbekistan, Kazakhstan, Tajikistan and Kyrgyzstan was given from 1994 to2010. The countries include Afghanistan, North Korea and Iraq was removed from this research, because of the lack of their information. For example information of macro-economic, political, social and cultural variables in Japan was reported in Table 1.

Table 1: information of macro-economic, political, social and cultural variables in Japan

	2010	2006	2002	1998	1994	1990	1986	1982	1978	1974
urban population	66.6	66.2	65.5	65	64.3	63.1	61.1	60	58.48	56.08
education expenditures	3.7	3.17	3.17	3.19	3.6	4.08	4.73	3.83	3.93	3.4
age structure	64.3	65.91	67.44	68.81	69.71	69.71	68.44	67.52	67.41	68.04
GDP real growth rate	-5.369	2.039	0.262	-2.049	0.864	5.572	2.831	3.377	5.27	-1.23
GDP per capita	37555	34141	30736	30494	38047	24432	16428	9141	8421	4157
unemployment	5.145	4.1	5.4	4.1	2.9	2.1	2.8	2.3	1.71	1.12
population	127.471	127.676	127.4	126.349	125.116	123.438	121.446	118.451	115.259	111.631
inflation average	-1.407	-0.3	-0.87	0.584	0.599	3.067	0.615	2.739	1.5	1.7
Current account balance	2.84	3.907	2.847	3.087	2.732	1.45	4.299	0.629	0.407	-1.015
life expectancy at birth	82.25	82.32	81.56	80.5	79.69	78.82	78.06	76.92	76.04	74.39
merchandise trade	31.45	28.1	19.24	17.33	14.12	17.33	16.95	24.9	18.41	25.64

3) The ranking of participating countries in Asian Games takes place in 2 methods:

a) The ranking of participating countries is based on the number of golden medals acquired each country.

b) The ranking of participating countries is based on the number of total medals acquired each country.

In the first method because of using only the number of earned gold medals in ranking of countries, considers identical two countries which are equal in the number of gold, but are unequal in number of silver and bronze medals. This limitation is corrected in the second method, but the limitation of this model is that it assumes weight of gold, silver and bronze medals equal. For example, if the total number of earned medals of two countries be equal, the country that has more number of gold or silver medals has no superiority over the country which has more bronze medals. We have proposed a third way to resolve this problem, namely based on the following method in which to each medal was given the weight of its own and then we ranked the countries according to:

Gold medals (3) + Silver medals (2) + Bronze medals

At the next stage, we examined the performance of these three methods in order to identify the best method for countries ranking.

In this research we used WEKA (Waikato Environment for Knowledge Analysis) software that is a popular suite of machine learning software written in Java, developed at the University Of Waikato, New Zealand. WEKA is free software available under the GNU General Public License. The WEKA workbench contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to this functionality. WEKA, an open source collection of data mining algorithms written in java, is a solid exploratory tool for those interested in mining their collected data (Witten & frank, 2005). There are various algorithms for predicting the rank of countries and we used the k-nearest neighbor algorithm.

2.2. K-Nearest Neighbor Algorithm

The K-Nearest Neighbor (KNN) algorithm is one of the oldest and simplest techniques for general, non-parametric classification and based on supervised learning (Bay, S.D., 1999). The purpose is to gain nearest k sample from the available training data when a new sample appears and classify the appeared sample according to most similar class (Mitchell, T.M., 1997).

An ordering function is defined by giving a point x' of the d-dimensional input feature space $fx': R^d \to R$. The typical ordering function is based on the Euclidean metrics: $f_{x'}(x) = \|x - x'\|$. It becomes possible to order the entire set of training samples X with respect to x' by means of an ordering function.

This corresponds to define a function $R_{x'}: \{1, \ldots, N\} \to \{1, \ldots, N\}$ that reorders the indexes of the N training points. Blanzieri and Melgani had defined this function recursively, as follows:

$$\begin{cases} R_{x'}(1) = \arg\min_i f_{x'}(x_i) & \text{With } i \in \{1, \ldots, N\} \\ R_{x'}(j) = \arg\min_i f_{x'}(x_i) & \text{With } i \in \{1, \ldots, N\} \text{ and } i \neq R_{x'}(1), \ldots, i \neq R_{x'}(j-1) \text{ For } j = 2, \ldots, N. \end{cases} \quad (1)$$

Therefore, $x_{r_{x'}(j)}$ is the point of the set X in the *j*th position in terms of distance from *x*′, namely the *j*th nearest neighbor, and $f_{x'}\left(x_{r_{x'}(j)}\right) = \left\|x_{r_{x'}(j)} - x'\right\|$ is its distance from *x*′. According to the above definition, the decision rule of the KNN classifier for binary classification problems is defined by the following majority rule:

$$KNN\ (x) = sign\left(\sum_{i=1}^{n} y_{r_x(i)}\right) \quad (2)$$

In which, $y_{r_x(i)} \in \{-1, +1\}$ is the class label of the *i*th nearest training sample (Blanzieri, E.; Melgani, F., 2008).

Generally proximity is defined with Euclidean distance. Mitchell (1997) had described Euclidean distance exactly with a formula.

An arbitrary example *x* be described by the feature vector $\langle a_1(x), a_2(x), \ldots, a_n(x)\rangle$ where $a_r(x)$ indicates the value of *r*th feature of example *x*. Then the distance between two examples x_i and x_j is defined to $d(x_i, x_j)$ as follows:

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n} (a_r(x_i) - a_r(x_j))^2} \quad (3)$$

Thereafter, unknown sample is assigned to most similar class from KNN. Also, KNN technique is used to guess an actual value for an unknown sample (Aci, M.; Avci, M., 2011).

At first, selecting suitable *k* value and distance measurement determines the performance of a KNN classifier. Uneven distribution of data points makes the process of determining the k value difficult. Generally, for making the borders smooth among the classes larger k values are selected in the event of noised data sets. Different heuristic methods such as cross-validation can chose a good k. Nearest Neighbor Algorithm is a particular case where the class is forecasted to be the class of the closest training sample. It is impossible to select same k value for all various applications (Song, Y.; Huang, J.; Zhou, D.; Zha, H.; Giles, C.L., 2007).

Different efforts have been done for offering new methods in order to enhance the performance of KNN technique by using previous knowledge like the distribution of the data and feature choice. Discriminant Adaptive NN (DANN), Adaptive Metric NN (ADAMENN), Weight Adjusted KNN (WAKNN), Large Margin NN (LMNN) are some of these methods (Song, Y.; Huang, J.; Zhou, D.; Zha, H.; Giles, C.L., 2007).

Generally, KNN technique requires the following Steps (Aci, M.; Avci, M., 2011):
1. Selection of k value: k value is completely up to user. Generally after some tests a k value is selected according to results.
2. Distance calculation: Any distance measurement can be used for this step. In general, Euclidean and Manhattan distances which are the most known distance measurements are preferred.
3. Distance sort in ascending order: Selection of k value is also important in this step. The obtained distances are sorted in ascending order and k of minimum distances is taken.
4. Classification of nearest neighbors: Classes of *k* nearest neighbor are identified.
5. Finding dominant class: In the final step, required data is classified by using maximum proportion according to class of identified k nearest neighbor. This proportion is calculated for each class of k nearest neighbor with the number of data belonging to that class over k.

Let $\mathrm{P} = \{\mathrm{p1}, \mathrm{p2}, \mathrm{p3}, \dots, \mathrm{pc}\}$ is the set of k nearest neighbor probabilities for each class where c is the number of class. Maximum proportion is calculated as in bellow Equation:

$$P_{max} = \max({}^{P_i}/_{k})$$

(4)

RESULTS

Table (1) contains output obtained from the KNN model using three methods (number of gold medals, total of medals and weight of medals) to predict the ranks of participating countries in the Asian Games in 2010.

Table 1. The output of KNN model for Asian Games in 2010

	Output by gold of medals	Output by total of medals	Output by weight of medals
Correlation coefficient	0.8501	0.9135	0.8941
Mean absolute error	3.5	3.0357	3
Root mean squared error	4.8624	4.2636	4.766
Relative absolute error	46.8764 %	35.9201 %	35.5491 %
Root relative squared error	54.6964 %	41.6565 %	46.6864 %
Total Number of Instances	28	28	28

In the above table, correlation coefficient in ranking methods using total of medals between original and predicted ranks for 28 participating countries in 2010 has been reported 0.9135. This high correlation coefficient indicates high ability of this method for ranking of countries compared with other methods. Then, correlation coefficients for methods based on the weight of medals and numbers of gold medals have been placed respectively 0.8941 & 0.8501. The amount of mean absolute error for the method based on the weight of medals was 3 that were lower than other methods.

The information of 28 participating countries in Asian Games in 2010 with their Macro-economic, political, social and cultural variables was reported in table 2.

Table 2. Information about Macro-economic, political, social and cultural variables of participating countries in Asian Games in 2010

	urban population	education expenditures	age structure (15-64)	GDP real growth rate	GDP per capita	unemployment	population	inflation average	Current account balance	life expectancy at birth	merchandise trade
China	44	1.9	72.1	8.504	2202	4.1	1341.41	3.12	6.239	74.68	59.2
Japan	66.6	3.7	64.3	-5.369	37555	5.145	127.471	-1.407	2.84	82.25	31.45
Korea South	81.7	4.2	72.3	-0.987	20329	3.3	48.91	2.9	1.607	79.05	92.27
Iran	69	4.8	72.9	1.484	3411	12.57	75.35	8.5	2.3	70.06	46.73
Thailand	33.7	4.9	70.5	-3.436	3577	1.39	67.653	3.245	2.496	73.6	130.86
Kazakhstan	58.2	4.41	70.2	1.465	6346	7.8	15.584	7.303	0.715	68.51	81.74
Uzbekistan	36.9	9.4	67	6.978	764	0.2	28.246	9.151	5.055	72.51	55.92
India	29.8	3.2	64.3	5.355	871	10.7	1215.94	13.162	-2.172	66.8	40.6
Qatar	95.7	3.3	76.8	11.467	34449	0.5	1.352	1.033	25.111	75.94	90.12
Malaysia	71.3	4.5	63.6	-3.631	6347	3.5	28.233	2	15.379	74.12	160.71
Singapore	100	3.2	76.7	-3.328	33174	2.078	4.832	2.097	21.986	80.74	361.62
Saudi Arabia	82.3	5.7	59.5	-0.8	15886	10.476	26.106	5.2	9.1	73.12	94.03
Bahrain	83.9	2.9	70.1	3.04	14908	15	1.06	2.393	5.486	75.91	143.34
Hong Kong	100	3.3	74.6	-3.623	29273	4.387	7.122	2	12.052	82.34	354.39
Kuwait	98.4	3.8	70.7	-1.51	18756	1.639	3.606	4.463	31.62	77.97	79.92
Philippines	65.7	2.6	60.6	0.994	1499	7.2	94.013	4.953	3.492	71.83	64.82
Vietnam	28.3	5.3	68.3	4.606	778	5	88.257	12	-6.907	74.37	158.11
Mongolia	57.3	5.1	67.9	0.5	1322	3	2.734	7.341	-6.641	66.57	117.07
Indonesia	52.6	3.5	66	3.99	1753	7.5	234.557	4.724	1.414	70.79	51.98
Syria	54.6	4.9	59.9	3.019	1893	8.5	21.762	5.037	1.13	74.23	59.09
Tajikistan	26.5	3.5	62.1	2	465	2.2	6.536	7.039	-7.27	66.75	91.08
Jordan	78.5	2	59.4	3	2707	13	6.126	5.278	-8.911	72.71	116.2
Lebanon	87.1	2	67.1	7	6163	9.2	3.908	5.003	-12.79	72.05	72.47
Kyrgyzstan	36.4	6.6	64.5	1.465	589	5.57	5.431	8.428	-12.47	67.37	112.66
Pakistan	36.6	2.9	59.1	1.966	866	6.195	169.38	11.5	-3.824	66.53	38.11
Laos	32	2.3	56.2	4.584	656	2.5	6.497	6.865	-10.131	64.97	44.56
Nepal	17.7	3.8	59.2	3.995	355	46	28.285	11.762	-2.77	66.39	37.02
Bangladesh	27.6	2.4	61.4	5.419	436	5.1	167.671	7.385	2.088	66.15	49.31

Checking Information about country's Macro-economic, political, social and cultural variables in 2010 indicated that among the 28 participated countries in the games, 12 countries were remarkable in terms of having the highest and lowest amount in variables.

For example Japan in GDP real growth, Inflation, Merchandise trade variables had the lowest and in GDP per capita the highest amount. Bahrain, despite of having the lowest population, obtained the first rank in the Unemployment rate. In terms of Population and Education expenditures China respectively had the highest and lowest amount. Kuwait in terms of Current account balance was higher than other countries. India obtained the highest Inflation rate among the other countries. Qatar in terms of Age structure and GDP real growth and Hong Kong in Urban population and Life expectancy at birth obtained the highest amount among the countries.

In table (3) original and predicted ranks of participating countries in the Asian Games in 2010 has been reported. As you see in the Table (3), three methods for predicting the ranking of countries have been used. Checking information showed that in all three methods, predicted ranks of China and Iran was consistent with their original ranks. Also, the rank of Japan, South Korea and Malaysia countries in all three methods have been predicted only with 1 rank different from their original rank.

Table 3. The information about original and predicted rank of participating countries in Asian Games in 2010

	original rank by gold medals	predicted rank by gold medals	original rank by total medals	predicted rank by total medals	original rank by weight medals	predicted rank by weight medals
China	1	1	1	1	1	1
Japan	3	2	3	2	3	2
Korea South	2	3	2	3	2	3
Iran	5	5	4	4	5	5
Thailand	8	4	16	5	8	4
Kazakhstan	4	6	8	6	4	6
Uzbekistan	6	8	7	7	6	7
India	7	7	5	8	7	8
Qatar	14	9	6	9	14	9
Malaysia	9	10	9	10	9	10
Singapore	13	11	14	11	13	11
Saudi Arabia	17	12	21	12	17	12
Bahrain	19	13	10	13	19	13
Hong Kong	10	14	13	14	10	14
Kuwait	18	15	11	15	18	15
Philippines	15	17	12	16	15	16
Vietnam	11	18	15	17	11	17
Mongolia	16	16	17	19	16	18
Indonesia	12	19	18	18	12	19
Syria	26	20	19	20	26	20
Tajikistan	23	21	22	22	23	21
Jordan	21	22	23	21	21	22
Lebanon	24	24	26	23	24	23
Kyrgyzstan	22	23	20	24	22	24
Pakistan	20	26	28	25	20	25
Laos	27	26	24	26	27	26
Nepal	28	28	25	28	28	27
Bangladesh	25	25	27	27	25	28

In checking each of the three methods separately we see that the ranking based on the number of gold medals, predicted ranks of countries China, Iran, India, Mongolia, Lebanon, Nepal and Bangladesh were consistent with their original ranks. In method based on total of medals, countries China, Iran, Uzbekistan, Indonesia, Bangladesh and Tajikistan were similar in predicted rank and original rank. Finally, in method based on the weight of medals, predicted rank of China and Iran was consistent with their original rank.

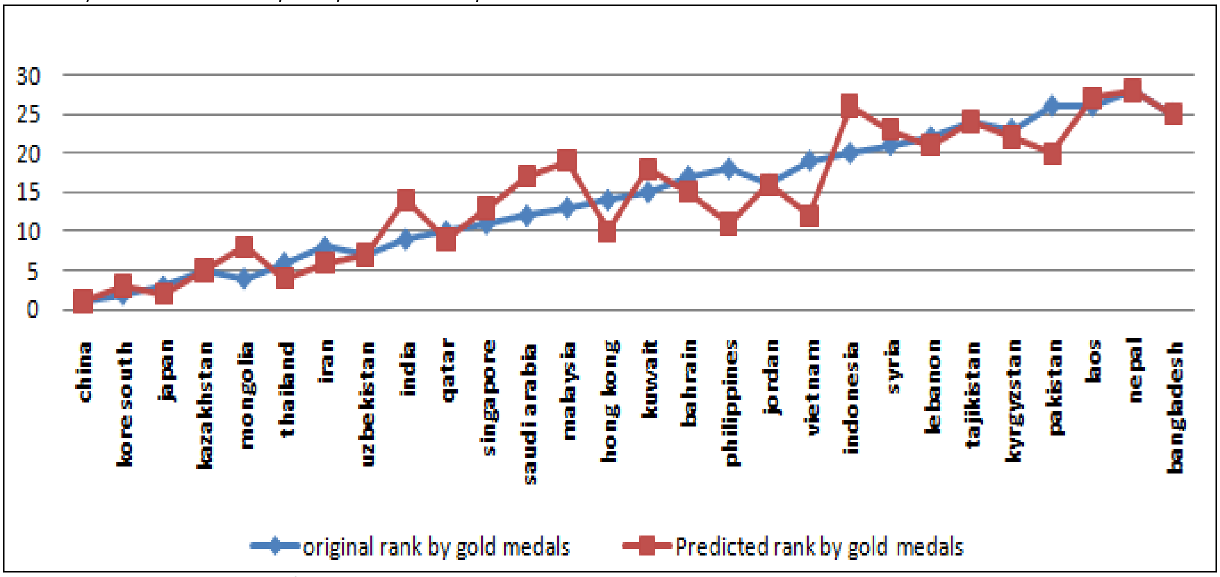

Figure 1: The diagram of original and predicted ranks based on gold medals

Figure (1) indicates a comparison between predicted ranks and original ranks of participating countries in the Asian Games 2010, based on the number of gold medals. KNN model of 28 countries mentioned predicted the rank of 19 countries (68%) with maximum 3 differences, 7 countries (25%) between maximum 4 to 6 difference and 2 countries (7%) more than 6 differences in comparison with their original ranks.

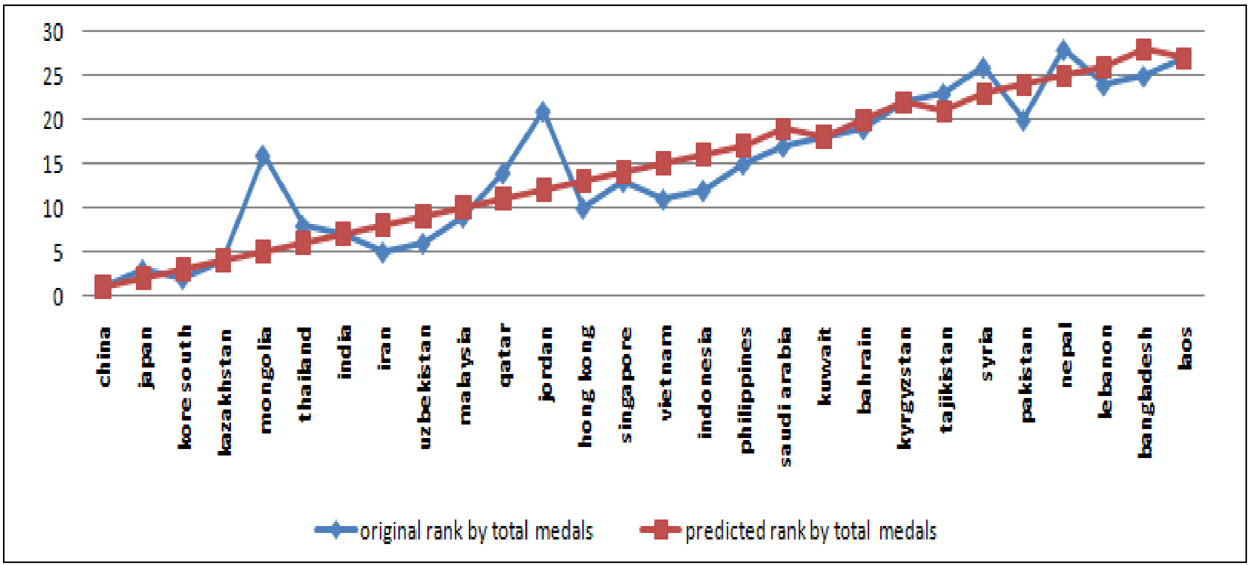

Figure 2: The diagram of original and predicted ranks based on total of medals

Figure (2) indicates a comparison between predicted ranks and original ranks of participating countries in the Asian Games 2010, based on the total of medals. KNN model of 28 countries mentioned predicted the rank of 23 countries (82%) with maximum 3 differences, 3 countries (11%) between maximum 4 to 6 difference and 2 countries (7%) more than 6 differences in comparison with their original ranks.

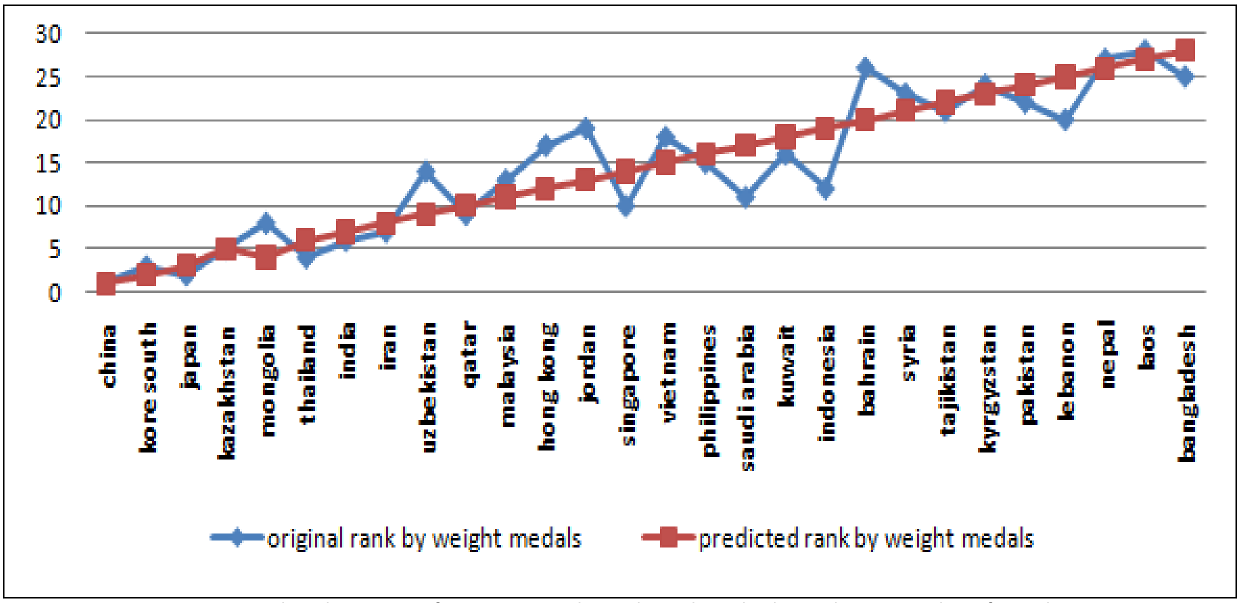

Figure 3: The diagram of original and predicted ranks based on weight of medals

Figure (3) indicates a comparison between predicted ranks and original ranks of participating countries in the Asian Games 2010, based on the weight of medals. KNN model of 28 countries mentioned predicted the rank of 19 countries (68%) with maximum 3 differences, 8 countries (25.5%) between maximum 4 to 6 difference and 1 country (3.5%) more than 6 differences in comparison with their original ranks.

DISCUSSION

The purpose of this study was to predict the success of participating countries in the Asian Games in 2010, by using KNN model because of its simplicity, efficiency and high ability. In this study we have predicted ranks of participating countries in the Asian Games 2010 with a conceptual model based on macro-economic, political, social and cultural variables which can lead to sporting success in the international levels and also by using three methods (the number of gold medals, total of medals and weight of medals).

Research results showed that among the three methods used, ranking method by using total of medals earned the highest correlation coefficient (0.9135) between original and predicted ranks for 28 participating countries in the Asian Games 2010 that indicates the high ability of this method for ranking of countries. Then, correlation coefficients for methods based on the weight of medals and numbers of gold medals have been placed respectively 0.8941 & 0.850.
Another result of this study is that in all three methods, predicted ranks of China and Iran was consistent with their original ranks. Also, the rank of Japan, South Korea and Malaysia countries in all three methods have been predicted only with 1 rank different from their original rank.

The reason that the predicted rank of a country is not exactly consistent with its original rank could be due to the two factors:
1) In this research only Macro variables have been studied. But to obtain better results, micro and medium variables should also be placed under investigation and thus, create an optimized pattern by the consensus of all of the variables.
2) In this research had been used the macro-economic, political, social and cultural variables. In fact, predicting of political, social and cultural phenomena because of their complexities and the existence of the parameters influencing on they is very difficult if not impossible.
Based on each three methods China earned the first rank that indicates the high ability of China among Asian countries, but the remarkable thing is about the country of Japan. Despite gaining third rank in games, Japan's rank is predicted the second in this model. This indicates that Japan's position, based on macro variables has the potential to earn second rank. Thus, can be conclude that Japan is the only country which is capable of taking the place of China in coming years in terms of the used macro variables, although it is not possible in the near future.

This model has a very high efficiency for successful countries or the same 4 top countries. For example, according to the model's findings the possibility of comparing Iran with China, South Korea and Japan, which has ranks of first to third, has been provided well. Therefore, based on macro variables, we can calculate the rank distance of Iran with these three countries. For example, the inflation rate in 2010 in China, South Korea and Japan, was respectively 3.12%, 2.9% and -1.4%, while inflation rate of Iran was 8.5% in this year. Comparing this information we conclude that the 3 Asian top countries have less inflation than Iran. Macro politicians and planners of Iran should estimate inflation rate in 2014 with regard to the difference between the three countries in different years based on their inflation rate, and then try to achieve those levels. Such an interpretation for all the variables affecting the ranking of countries in the Asian Games is possible.

REFERENCE

1. Abrahart, R.J.; Kneale, P.E.; See, L.M. (2004). *Neural Networks for Hydrological Modeling.* Rotterdam: A.A. Balkema Publishers (Leiden, The Netherlands), a number of Taylor & Francis Group plc.
2. Aci, M.; Avci, M. (2011). K nearest neighbor reinforced expectation maximization method. *Expert Systems with Applications , 38*, 12585-12591.
3. Bay, S.D. (1999). Nearest neighbor classification from multiple feature subsets. *Intelligent Data Analysis , 3* (3), 191.
4. Bergsgard, N.A.; Houlihan, B.; Mangset, P.; Nodland, S.I.; Rommetveldt, H. (2007). *Sport policy. A comparative analysis of stability and change.* London: Elsevier.
5. Blanzieri, E.; Melgani, F. (2008). Nearest neighbor classification of remote sensing images with the maximal margin principle. *IEEE Transactions on Geoscience and Remote Sensing , 46* (6), 1804-1811.
6. Boulier, B.L.; Stekler, H.O. (1999). Are sports seedings good predictors?: an evaluation. *International Journal of Forecasting , 15*, 83-91.

7. Boulier, B.L.; Stekler, H.O. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting , 19*, 257-270.
8. Bunn, D.; Wright, G. (1991). Interaction of judgmental and statistical forecasting methods: Issues and analysis. *Management Science , 37*, 501-518.
9. Condon, E.M.; Golden, B.L.; Wasil, E.A. (1999). Predicting the success of nations at the summer Olympics using neural networks. *Computers and Operations Research , 26* (13), 1243-1265.
10. De Bosscher, V.; De knop, P.; Van Bottenburg, M.; Shibli, S.; Bingham, J. (2009). Explaining international sporting success: An international comparison of elite sport systems and policies in six countries. *Sport Management Review , 12*, 113-136.
11. Del Corrala, J.; Rodriguez, J.P. (2010). Are differences in ranks good predictors for Grand Slam tennis matches? *International Journal of Forecasting , 26*, 551-563.
12. Easton, S.; Uylangco, K. (2010). Forecasting outcomes in tennis matches using within-match betting markets. *International Journal of Forecasting , 26*, 564-575.
13. Forrest, D.; Goddard, J.; Simmons, R. (2005). Odds-setters as forecasters: The case of English football. *International Journal of Forecasting , 21*, 551-564.
14. Forrest, D.; Sanz, I.; Tena, J.D. (2010). Forecasting national team medal totals at the Summer Olympic Games. *International Journal of Forecasting , 26*, 576-588.
15. Forrest, D.; Simmons, R. (2000). Forecasting sport: the behaviour and performance of football tipsters. *International Journal of Forecasting , 16*, 317-331.
16. Goddard, J. (2010). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting , 21*, 331-340.
17. Goddard, J.; Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixedodds betting. *International Journal of Forecasting , 23*, 51-66.
18. Grant, A.; Johnstone, D. (2010). Finding profitable forecast combinations using probability scoring rules. *International Journal of Forecasting , 26*, 498-510.
19. Green, M. (2007). Olympic glory or grassroots development? Sport policy priorities in Australia, Canada and the UK 1960-2006. *International Journal of the History of Sport , 24* (7), 921-953.
20. Grove, W.M.; Meehl, P.E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical statistical controversy. *Psychology, Public Policy, and Law , 2*, 293-323.
21. Grove, W.M.; Zald, D.H.; Lebow, B.S.; Snitz, B.E.; Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment , 12*, 19-30.
22. Heinila, K. (1982). The totalisation process in international sport. Toward a theory of the totalization of competition in top-level sport. *Sportwissenschaft , 3*, 235-253.
23. Hematinezhad, M.; Gholizadeh, M.H.; Ramezaniyan, M.R.; Shafiee, Sh.; Ghazi Zahedi, A. (2011). Predicting the success of nations in Asian games using Neural Network. *Sport SPA , 8* (1), 33-42.
24. Iyer, S.R.; Sharda, R. (2009). Prediction of athletes performance using neural networks: An application in cricket team selection. *Expert Systems with Applications , 36*, 5510-5522.
25. Klaassen, F.; Magnus, J. (2003). Forecasting the winner of a tennis match. *European Journal of Operational Research , 148* (2), 257-267.
26. Lovalgia, M.J.; Lucus, J.W. (2005). High visibility athletic programs in the Football. *The Sport Journal , 8* (2), 1-5.
27. McHale, I.; Morton, A. (2011). A Bradley-Terry type model for forecasting tennis match results. *International Journal of Forecasting , 27* (2), 619-630.
28. Min, B.; Jinhyuck, K.; Chongyoun, Ch.; Hyeonsang, E.; McKay, R.I. (2008). A compound framework for sports results prediction: A football case study. *Knowledge-Based Systems , 21*, 551-562.
29. Mitchell, T.M. (1997). *Machine learning.* New York: McGraw-Hill.
30. Oakley, B.; Green, M. (2001). The Production of Olympic Champions: International Perspectives on Elite Sport Development System. *European Journal for Sport Management , 8*, 83-105.
31. Rue, H.; Salvesen, O. (2000). Prediction and retrospective analysis of soccer matches in a league. *The statistician , 49* (3), 399-418.
32. Scheibehenne, B.; Bröder, A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting , 23*, 415-426.
33. Song, Ch.; Boulier, B.L.; Stekle, H.O. (2007). The comparative accuracy of judgmental and model forecasts of American football games. *International Journal of Forecasting , 23*, 405-413.
34. Song, Y.; Huang, J.; Zhou, D.; Zha, H.; Giles, C.L. (2007). *IKNN: Informative. K-Nearest Neighbor Pattern Classification* (Vol. 4702). in: PKDD, LNAI.
35. Strumbelj, E.; Sikonja, M.R. (2010). Online bookmakers' odds as forecasts: The case of European soccer leagues. *International Journal of Forecasting , 26*, 482-488.
36. Webby, R.; O'Connor, M. (1996). Judgmental and statistical time series forecasting: A review of the literature. *International Journal of Forecasting , 12*, 91-118.

37. Witten, I.H.; Franke, E. (2005). *Data mining: Partical mashine learning tools and techniques* (2nd edition ed.). sanfrancisco, CA: Morgan kaufmann.
38. Wright, G.; Lawrence, M.J.; Collopy, F. (1996). The role and validity of judgment in forecasting. *International Journal of Forecasting , 12* (1), 1-8.

Corresponding author:
Shahram Shafiee
Department of Physical Education & Sport Sciences,
Faculty of Sport Science and Physical Education,
GuilanUniversity, Rasht, Iran
Phone: +981316690685
Cell: 0098-9119176240
Fax: 0098-131-6690815
P.O. Box: 1438

E mail: Shafieeshahram@gmail.com

Received: 11 June 2012
Accepted: 01 October 2012