recruitment process, selection process, neural networks

Igor RYGUŁA*, Robert ROCZNIOK*

APPLICATION OF NEURAL NETWORKS IN OPTIMIZATION OF THE RECRUITMENT PROCESS FOR SPORT SWIMMING

The essence of the recruitment and selection in sports depends on determining an aptitude vector of the candidate to sport training. For this reason, the recruitment process may be optimized by determining possibly large amount of information on the sport level of the candidate with as small as possible number of tested characteristics, using a mathematical model based on neural networks. The main aim of this work was verification of the usefulness of neural models in optimization of the process of recruitment process, both at sprint distance of 50 m and typically endurance distance of 800 m. The material for the investigation was a group of 80 young swimmers in the youngster and junior category from the Silesian macro-region. For the purpose of verification of the usefulness of neural models, statistical analysis were made of the measurement results of young swimmers and two neural models were developed - for sprint distance (50 m) and endurance distance (800 m). The developed models, based on the architecture of perception networks, has shown capability of generalization and prediction, which has enabled to reach conclusions on practical possibility of using neural networks in optimization of the recruitment and selection process for sport swimming.

1. INTRODUCTION

The scientific approach to problem of sports (naturally competitive sports are meant here), recently obtains a unique dimension. A dimension where traditional branches of science, such as pedagogy Carl, psychology Dorsch, sports medicine Hollman or physiology Stegman in the individual approach are insufficient. The development of cybernetics, with its theory of models and modelling has won significant methodological meaning, has contributed to the increased attraction of application of this tool in numerous fields of practical and scientific activity of man. This also applies to the area of sport training and process of recruitment, where they infiltrating through ideas developed in such disciplines as statistics and forecasting. Using self-learning neural networks [8, 20] for optimization of the recruitment process and predicting results, we are able to make simulation of the sporting score on the basis of input data of the athlete alone. Naturally, the reliability of information will rely on the correctness of developed mathematical model, on the construction and kind of neural network (its ability to generalize) and the volume of materials available for its testing. If these conditions are fulfilled, we will be able to optimize the recruitment process and predict the score of competition swimmers with high dose of reliability.

* Dep. of Theory of Sport, University School of Physical Education 40-065 Katowice, 72a Mikołowska Str. Poland

Hitherto applications of neural models in the field of sport sciences are very limited and related rather to object recognition (discriminative analysis) and their grouping (taxonomic analysis) [19]. These are relatively simple tasks, easy to solve with neural networks. In the sport task hitherto solved [21], it was enough to group data into two sets: learning and test set. The innovative application of neural networks in problems concerning competition swimming has been presented by Edelmann-Nusser, Hohrmann and Hennenberg. They have shown possibilities of application of nonlinear mathematical models based on artificial neural networks for prediction of scores of the elite woman swimmers in 200 m backstroke finals in Sydney 2000 Olympic Games. The data available to the investigators included the scores of the swimmers from nineteen earlier competitions in 200 m backstroke swimming and data from the last four weeks preceding the Games. The investigators have used perceptron networks with ten neurons in input, two in the hidden layer and one output neuron, which were subjected to teaching. The output neuron was the score in 200 m backstroke swimming. The results obtained were very accurate. The error of prediction was only 0.05 s at total time of swimming of 2:12.64 (min:s) (time predicted by model was 2:12.59) [7].

These results indicate that artificial neural networks are excellent tools for modelling and prediction of nonlinear phenomena. Very few similar works were published and the one discussed concerned expert competition swimmers. Availability of simple application, enabling coaches to predict future scores of young athletes at the initial stage of training, would give the coaches a powerful tool to be used in all stages of selection.

2. RESEARCH QUESTIONS AND HYPOTHESES

1. Is the application of artificial neural networks to optimization of the recruitment to competition swimming possible?
2. What will be prediction quality of developed models?
3. What will be structure and quality of developed neural models?

The hypothesis is assumed that on the basis of experience and collected and precisely described data, artificial neural networks may be used to support the process of recruitment and prediction in competition swimming.

3. MATERIAL AND METHODS

The material for the investigation was a group of 80 young swimmers in the youngster and junior category from the Silesian macro-region. The objects for the investigation were assigned in accordance with the principles of purposeful selection. The basis criterion was the completeness of collected measurement data for a given person. The investigation may be classified as empirical and explorative of cut-set character, therefore the basic investigation method was direct-participating observation, which was complemented[1] by diagnostic polling.

For the purpose of verification of the research hypothesis and questions formulated earlier, statistical analyses were made of measurement results for young swimmers. Neural models were developed for quality index 50 m and for quality index 800 m.

During the investigation, following features were measured:

State variables - quantities characterizing body build of the candidate and the results of efficiency tests (measurements were made during recruitment for sport swimming):

- X1 - Body height [cm].
- X2 - Body mass [kg].
- X3 - Static wide jump [cm].
- X4 - Swimmer's step [number of full cycles at 25 m] [3].
- X5 - Vital lung capacity [cm^3] [14].

Quality indices - quantities characterizing sport development (measurements were after 12 month of swimming training):

- $VAR6_{50m}$ - 50m free-style score [pts.]
- $VAR6_{800m}$ - 800m free-style score [pts.]

4. QUALITY INDICES

To build a model, a variable or two must be chosen, called quality indices, the value of which at the end of the training cycle should be maximal (number of point in the multi-event table).

5. ASSUMPTIONS CONCERNING METHOD OF TRAINING OF THE SUBJECTS

In construction of the model, it should be remembered that the model designed take not into account the influence of the sport training on obtained level of development. Evaluating potential of the candidate, we are assuming peculiarity of the sport training of the swimmers, implemented with the use of the method described in papers of Bartkowiak [2], Costill [5] and Costill et al. [6].

[1] Complementing is a feature of a research method or technique bases on its supporting and complementing action. [19]

Table 1. Basic descriptive statistics of state variables and quality indices of investigated athletes [10].

	Mean	Median	Standard deviation	Skew	Kurtosis $_{-3}$
Wide jump	147,03	147,50	18,85	0,08	-0,67
Technique	29,63	27,50	11,19	0,42	-0,92
Height	136,84	138,75	5,89	-0,70	-0,53
Mass	30,55	30,40	4,24	0,87	1,28
Lung cap.	2610,00	2550,00	286,22	1,09	1,31
50 m	151,13	142,50	68,33	0,27	-0,90
800 m	73,83	71,50	39,35	0,41	-1,04

6. CONSTRUCTION OF NEURAL MODEL

In the process of modelling optimization of recruitment for sport swimming, unidirectional multilayer networks were used (multilayer perceptrons), trained with the use of the algorithm of backward error propagation. The training has multi-stage character. During each step, consecutive approximations of the optimal parameter values are determined. The implemented steps are called teaching epochs. The epoch contains a single-time presentation of all teaching cases and implemented modification of the network parameters based on it [12, 15]. The purpose of teaching is reducing the value of error to the level accepted by the creator of the network [18].

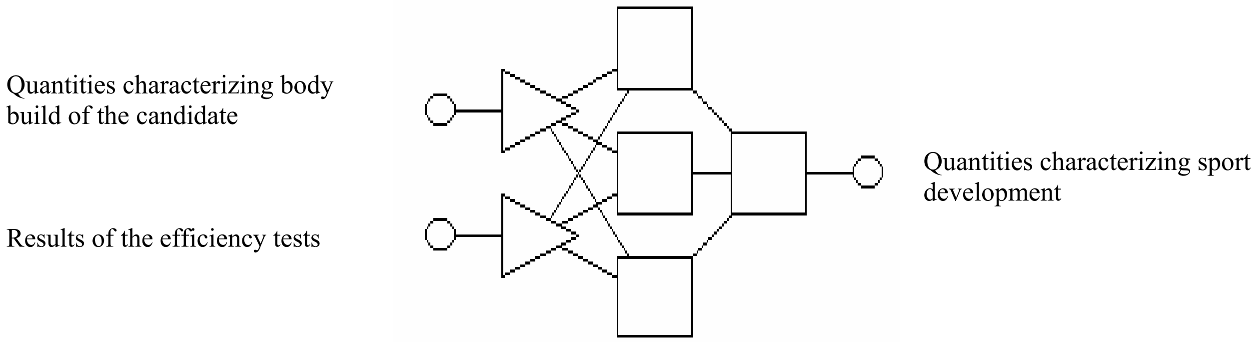


Fig. 1. Diagram of the neural model of the interdependence of variables of sport training [17].

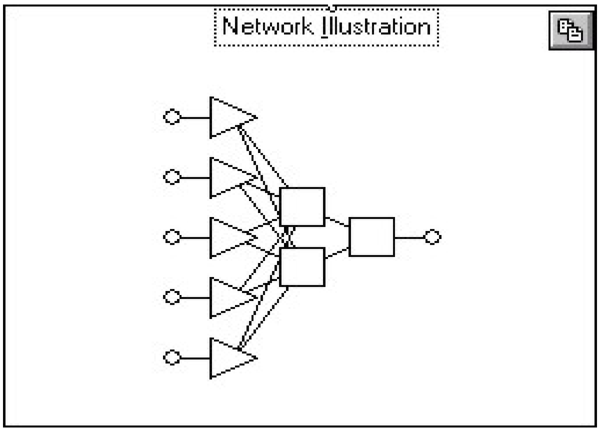
The model so designed has prognostic character, because on the basis of initial quantities characterizing the candidate, the level of his or her future development may be predicted. Model obtained in this way may have wide application potential with basic function of predicting future level of sport development of the candidate for training.

7. RESULTS OF MODELING

The computations made were intended for evaluation of the usefulness of unidirectional neural models for the optimization of recruitment and for predicting scores obtained by future competition swimmers. The computations were made in two stages, attempting the description and prediction of variables $VAR6_{50m}$ and $VAR6_{800m}$.

8. RESULTS FOR $VAR6_{50M}$

We have chosen three-layer perceptron network with two neurons in the hidden layer.



dd

Fig. 2. Perceptron network

Table 2. Results of computation

	Tr. $VAR6_{50}$	Ve. $VAR6_{50}$	Te. $VAR6_{50}$
Data Mean	151,1	157,2	139,2
Data S.D.	67,57	76,96	58,21
Error mean	-0,24	3,93	14,55
Error S.D.	20,16	22,75	17,84
Abs E. Mean	16,07	17,84	19,78
S.D. Ratio	0,30	0,29	0,31
Corelation	0,95	0,95	0,95

Data Mean - mean value determined on the basis of the real values of clarified variable;
Data S.D. - standard deviation determined on the basis of the real values of clarified variable;
Error Mean - mean value of error (calculated as a difference between the real and calculated value of clarified variable;

Abs E. Mean - mean value of absolute values of errors (calculated as a difference between the real and calculated value of clarified variable;
Error S.D. - standard deviation of errors;
S.D. Ratio - ratio of the standard deviation of errors and standard deviation of clarified variable;
Correlation - Person correlation coefficient computed between the real values of clarified variable and values calculated by the model;
Tr. Var6 (Ve. Var6, Te. Var6) - teaching (validating, testing) set;
VAR6 - freestyle score in points on 50 m or 800 m distance.

Presented results indicate the possibility of description of shaping of the VAR_{50m} variable by perceptron models. The value of S.D. Ratio for the teaching set is 0.29, which perhaps is not excellent result, but looking at the results of validating and testing set of 0.29 and 0.30, we may decide that the network is able to generalize and predict results. The practical value of obtained model seems to be corroborated by high values of correlation coefficient 0.95. The results obtained convince on proper functioning of the network.

9. RESULTS FOR $VAR6_{800M}$

We have chosen three-layer perceptron network with two neurons in the hidden layer.

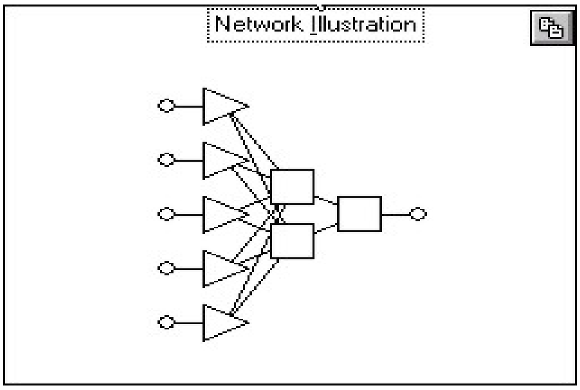


Fig. 3. Perceptron network

Table 3. Results of computation

	Tr. $VAR6_{50}$	Ve. $VAR6_{50}$	Te. $VAR6_{50}$
Data Mean	73,01	71,20	81,30
Data S.D.	38,35	47,29	40,49
Error mean	-0,72	1,39	3,83
Error S.D.	13,79	17,89	11,45
Abs E. Mean	11,69	14,03	8,99
S.D. Ratio	0,36	0,38	0,28
Corelation	0,93	0,93	0,96

Presented results for the VAR_{800m} output variable are slightly worse compared with VAR_{50m}.In spite of the fact that the S/D Ratio values are higher and equal 0.35 for the reaching set, which is far worse result than in the case of neural model for the VAR_{50m} variable, looking at the results of validating and testing set of 0.37 and 0.28, not significantly differing from the teaching set, it may be concluded that the network poorer describes the investigated phenomenon, but anyway has the ability to generalize and predict the results. The value of the obtained model is confirmed by the high values of the correlation coefficient 0.93 - 0.95. The results obtained guarantee correctness of the conclusion on the existing possibility of description of shaping the VAR_{800m} variable by perceptron models.

10. DISCUSSION OF RESULTS

The developed models have prognostic character, on the basis of the quantities characterizing the candidate they should estimate the level of his or her future development. The process of network construction was implemented using the teaching set, containing data from different periods: earlier initial data, obtained during recruiting and later output quantities after a year of training. The obtained models should be complemented with new learning cases, which will lead to the improvement of the quality of these models, expressed by better generalization and prediction. The quality of obtained models enable conclusion on the feasibility of using neural models in optimization of recruiting process. The results of research confirmed the papers of Edelmann-Nusser, Hohrmann and Hennenberg, indicating possibility of using perceptron networks to predict the score of swimmers. In most cases it is found that the simplest networks give best generalization [4, 12].

Summing up, it may be said that on the basis of experience and collected and precisely described data, the artificial neural networks may be applied to assist in the process of recruiting and prediction in sport swimming. This enables to accept the research hypothesis formulated in the paper.

11. CONCLUSIONS

Collected data and the results of investigation enable to formulate the following conclusions:

On the basis of computations resulting from neural models it may be unequivocally said that there is a possibility to use artificial neural networks in optimizing of the recruitment and prediction of results in sport swimming.

On the basis of results obtained for both models it may be guessed that the quality of predictions for new cases for the models will be satisfactory, because the models have shown no tendency to over-learning (loss of ability to generalize).

The developed neural models have three-layer structure, with five neurons in the input layer, two in hidden layer and one in the output layer. Considering the small number of data, the quality of developed models is satisfactory. The value of the ratio of standard deviation of errors and standard deviation of clarified variable is on decent level. The quality of models is confirmed by high values of correlation coefficients ca. 0.95. The results obtained enable to ascertain the proper functioning of developed perceptron networks.

BIBLIOGRAPHY

[1] BARTKOWIAK E.: Sport swimming technique. COS, RCMSzKFiS, Warszawa 1995.
[2] BARTKOWIAK E.: Swimming. Program of training of children and young people. COS, RCMSzKFiS, Warszawa 1997.
[3] BARTKOWIAK E.: Collection of investigations on training loads of Polish swimmers. Sport Wyczynowy nr 8 – 9, 1986.
[4] CHLEWIŃSKI Z. (red): Models of mind. PWN. Warszawa 1999.
[5] COSTILL D.L. et al. 1991. Adaptations to swimming training: influence of training volume. Medicine and Science in Sports and Exercise, vol. 23, 3.
[6] COSTILL D.L., MAGLISCHO E.W., RICHARDSON A.B. 1995. Swimming. Blackwell Scientific Publications, Oxford.
[7] EDELMANN-NUSSER J., HOFMAN A., HENNEBERG B.: Modeling and Prediction of Competitive Performance in Swimming Upon Neural Networks. European Journal Sport Science, vol.2, issue 2. 2002 by Human Kinetics Publishers and the European College of Sport Science.
[8] KORBICZ J., OBUCHOWICZ A., UĆIŃSKI D.: Artificial neural networks. Basics and applications. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1994.
[9] LULA P.: Unidirectional neural networks in modeling of economic phenomena. AE, Kraków 1999.
[10] OSTASIEWICZ W. (red): Statistical methods of data analysis. Wydawnictwo Akademii Ekonomicznej im. Oskara Langego. Wrocław 1999.
[11] ROSENBLATT F.: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington D.C. 1961.
[12] RUMELHART D.E., MINTIN G.E., WILIAMS R.J.: Learning Representations by Back-Propagating Errors. Nature, vol.323.
[13] RYGUŁA I.: Elements of the optimization of the process of choice of children for training on the example of high jump. Uniwersytet Śląski, Katowice 1988.
[14] RYGUŁA I.(red): Selected elements of sport training. Katowice 1999.
[15] RYGUŁA I., LULA P.: The application of neuro-genetic networks for developing mathematical model in training of women short distance runners. Katowice 1999.
[16] RYGUŁA I.(red): Elements of theory, method, diagnosis and optimization of sport training). Katowice 2000.
[17] RYGUŁA I.: Tools of the system analysis of sport training. Katowice 2000.
[18] RYGUŁA I., LULA P., PALIWODA-PĘKOSZ G.: Comparative investigation of neural models used in the design of sport training. Antropomotoryka nr 24, 2002.
[19] RYGUŁA I.: Research process in sports sciences. AWF Katowice, 2003.
[20] TADEUSIEWICZ R.: Neural networks. Akademicka Oficyna Wydawnicza RM. Warszawa 1993.
[21] WITT A., BARTON G.: Artificial neural networks. Sport wyczynowy nr 9. 1996r.