Prediction of The Result in 400m Hurdle Races in Two Years Training Cycle

Krzysztof Przednowek[1], Janusz Iskra[2], Justyna Lenik[1], Stanisław Cieszkowski[1]

[1.] Faculty of Physical Education, University of Rzeszow, Rzeszow, Poland
[2.] Faculty of Physical Education and Physiotherapy, Technical University of Opole, Opole, Poland
krzprz@ur.edu.pl

Abstract: In this study, artificial neural networks conducting prediction of results in 400m hurdle races were presented. The determined models predict results that should be obtained by the given athlete for suggested training loads in a two-year training cycle. All models were determined based on training data of 21 athletes – Polish Athletics Association National Team members. The athletes featured a high training level (result in 400m hurdles: 51.26 ± 1.24 s). To assess the prediction ability of designed models, the leave-one-out cross-validation method was used. From the conducted analysis it follows, that the method generating the least prediction error is the artificial neural network with 6 neurons in the hidden layer and the hyperbolical tangent activation function. The optimal model generates prediction error at the level of $RMSE_{CV}$=0.75 s. The obtained model can be a tool supporting the coach in planning the training loads over a two-year training cycle.
[Przednowek K, Iskra J, Lenik J, Cieszkowski S. **Prediction of The Result in 400m Hurdle Races in Two Years Training Cycle.** *Life Sci J* 2014;11(11):794-797]. (ISSN:1097-8135). http://www.lifesciencesite.com. 143

Keywords: sport prediction, 400-metres hurdle races, artificial neural networks

1. Introduction

Requirements of sports competition involve a continuous search for new solutions that could increase the training process effectiveness. It entails both training and organizational measures. Solution of those problems should be approached in an innovative way. One of examples of such approach is the effort to use the predictive abilities of prediction models. From the coach's point of view, prediction of result is of great importance in sports training process.

Sport prediction includes both prediction of sports talent development (Papic, 2009; Roczniok, 2013) and sports results forecast (Maszczyk, 2011; Przednowek and Wiktorowicz, 2013). Prediction of results involves individual events; it uses the season statistical analysis of the progress of team sports competitions (Haghighat et al., 2013).

This work contains the analysis of result prediction, taking into account the training process. Maszczyk et al. (2011) implemented regression models, based on which it was possible to predict the result in javelin throw. Przednowek and Wiktorowicz (2011) presented in turn prediction model designed by means of ridge regression, used for prediction of results in race walking. On the other hand, Drake and James (2011) presented regression models including the level of selected physiological parameters and result in race walking over the distance of 5, 10, 20 and 50 km, and Chatterjee et al. (2009), designed a non-linear regression equation used for prediction of maximum oxygen efficiency of athletes practicing football. Moreover, the studies are aimed at selection of sport event that would be most appropriate for young professional training candidates. For that purpose, both knowledge of experts in the given kind of sports and results of various motor tests are applied (Roczniok et al., 2013; Papić et al., 2009). The research that had been conducted so far includes also a broad application of artificial neural networks in sport prediction (Haghighat et al., 2013). Artificial neural networks are also used for sport talent prediction, handball players' tactics identification, or for swimmers' training efficiency analysis (Pfeiffer et al., 2012). In numerous studies, application of neural networks in different aspects of sport training are presented (Ryguła, 2005; Silva et al., 2007; Maszczyk et al., 2012). Those models support among others sports selection, training control or training loads planning.

The aim of the research was verification of artificial neural networks that could be applied for prediction of results in 400-metres hurdle races in a two-years training cycle. The verification was carried out based on training data of athletes running the 400-meters hurdle races and featuring a very high level of sport abilities.

2. Material and Methods

For development of result prediction in a two-year training cycle, 27 variables were applied (Tab. 1). Prediction of result over a two-year training cycle involves generation of predicted result that should be obtained after both the first and second year of the same two-year training cycle. On the other hand, training loads are generated as sums of the given training means applied during the whole two-year cycle. For calculation of models, 29 collected templates were used. The small number of templates results from the duration of the analyzed period. It takes two years to collect one training (template).

Table 1. The variables used to build the models of prediction (N=29)

Var.	Description	$\bar{x}$	sd	V[%]
y_1	Expected result of 400m hurdles after one year (s)	51,36	1,23	2,4
y_2	Expected result of 400m hurdles after two year (s)	50,86	1,14	2,2
x_1	Age (years)	21,8	1,6	7,1
x_2	Body Mass Index	22,1	0,9	4,1
x_3	Current result of 400m hurdles (s)	51,98	1,37	2,6
x_4	Maximal speed (m)	7339,0	2496,5	34,0
x_5	Technical speed (m)	8952,6	5450,0	60,9
x_6	Technical and speed exercises (m)	8067,9	3580,9	44,4
x_7	Speed endurance (m)	27755,2	24145,4	87,0
x_8	Specific hurdle endurance (m)	24993,4	9927,4	39,7
x_9	Pace runs (m)	306225,9	81862,2	26,7
x_{10}	Aerobic endurance (m)	758346,6	112923,7	14,9
x_{11}	Strength endurance I (m)	59212,4	25013,1	42,2
x_{12}	Strength endurance II (n)	13333,9	9436,9	70,8
x_{13}	General strength of lower limbs (kg)	269710,3	138652,6	51,4
x_{14}	Directed strength of lower limbs (kg)	120181,0	47606,3	39,6
x_{15}	Specific strength of lower limbs (kg)	84147,6	81961,0	97,4
x_{16}	Trunk strength (amount)	106725,0	80801,5	75,7
x_{17}	Upper body strength (kg)	5870,7	6055,4	103,1
x_{18}	Explosive strength of lower limbs (amount)	1726,9	630,0	36,5
x_{19}	Explosive strength of upper limbs (amount)	879,3	465,3	52,9
x_{20}	Technical exercises – walking pace (min)	923,6	376,6	40,8
x_{21}	Technical exercises running pace (min)	1081,5	471,6	43,6
x_{22}	Runs over 1-3 hurdles (amount)	182,7	56,6	31,0
x_{23}	Runs over 4-7 hurdles(amount)	360,5	111,8	31,0
x_{24}	Runs over 8-12 hurdles (amount)	325,4	168,1	51,7
x_{25}	Hurdle runs in varied rhythm (amount)	1615,8	663,7	41,1

Prediction of results over a two-year training cycle was conducted by means of artificial neural networks (Bishop, 1995); it is due to the fact, that there are two results at the model output (after the first and the second year). Networks carrying out that task have 25 neurons in the input layer and 2 in the output layer (Fig. 1). All networks were designed using STATISTICA 10 software (StatSoft, 2011). The number of neurons in the hidden layer was established in a leave-one-out cross-validation (LOOCV) process (Arlot, 2010).

In this paper, due to the small amount of data, LOOCV was selected, in which a testing set consists of one pair of selected data $(\mathbf{x}_i, y_i)$, whereas the number of tests is equal to the number of data n. As the performance criterion of the model the square root of the mean square error was calculated on the basis of the formulas:

$$MSE_{CV} = \frac{1}{rn}\sum_{i=1}^{n}(y_i - \hat{y}_{-i})^2,$$

$$RMSE_{CV} = \sqrt{MSE_{CV}},$$

where: $\hat{y}_{-i}$ – the output value of the model constructed in the i-th step of cross-validation based on a data set containing no testing pair $(\mathbf{x}_i, y_i)$, r – number of inputs of model, MSE_{CV} – mean square error, $RMSE_{CV}$ – root mean square error.

In addition to the cross-validation error, which allows evaluation of the predictive ability of the model, a training error – describing the measure to which it matches the data – will be considered as well. A training error is defined as

$$MSE_T = \frac{1}{rn}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2,$$

$$RMSE_T = \sqrt{MSE_T},$$

where $\hat{y}_i$ – the output value of the model built in the i-th step, based on the full data set, MSE_T – mean square error of training, $RMSE_T$ – root mean square error of training.

3. Results

In the search for the optimal models, MLP networks with hidden neurons from 1 to 10 were analyzed. Considering the number of templates (29), the application of a greater number of hidden neurons would result in the phenomenon of oversizing (Bishop, 1995). The obtained results show that only perceptron with tanh function feature a relatively small error (Fig. 1) at the level of 0.75 s. The best network contains 6 neurons in the hidden layer. The MLP (exp) network generates very large errors what disqualifies it as a prediction tool over the analyzed period of time (Fig. 2). It is significant, that both MLP networks feature a very good ability to project learning data, what is confirmed by zero learning error. Similarly to the MLP networks, for RBF, structures from 1 to 10 neurons in the hidden layer were searched (Fig. 3). The presented errors clearly illustrate poor prediction abilities in execution of the considered task. Apart from that, that network features a very large learning error, what is also an indication that that method should be rejected.

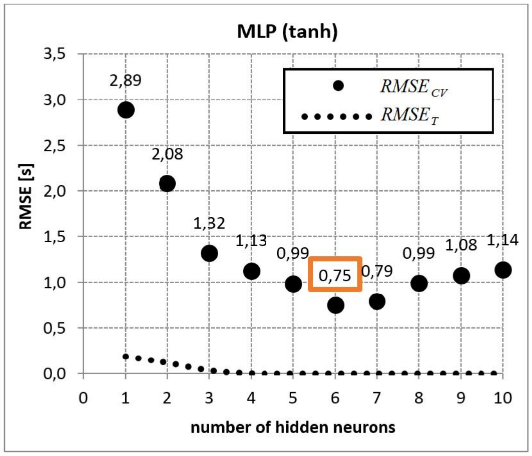

Figure 1. Prediction error for MLP with *tanh* function

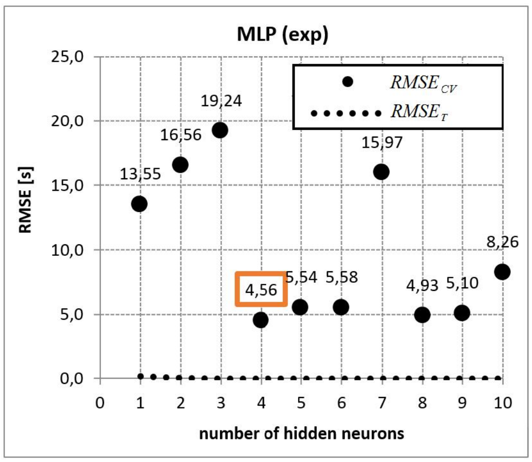

Figure 2. Prediction error for MLP with *exp* function

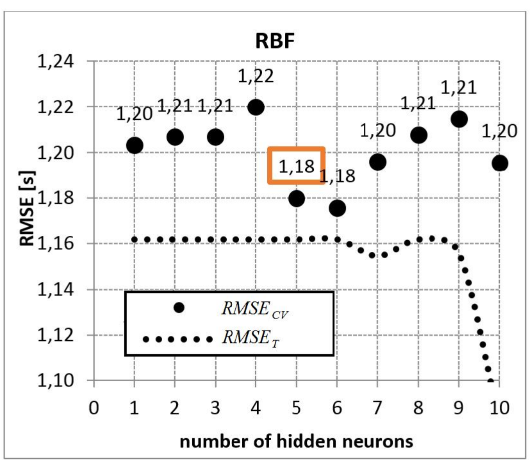

Figure 3. Prediction errors for RBF network

The conducted analysis showed that the best network implementing the task of result prediction over a two-year training cycle features an error at the level of 0.75 s. Taking into account the fact that neural models act according to the "black box" principle, the comparative analysis with other studies is difficult or even impossible. In this case, it is also impossible to determine the optimal set of input variables. The values of errors in the determined model show also that network features a zero learning error, and in consequence, it will perfectly project the learning data.

The analysis of similar studies proves that neural networks are very often used in execution of that kind of tasks. In most studies, artificial neural networks feature a satisfactory prediction error (Rygula, 2005; Silva et al., 2007; Maszczyk et al., 2011; Maszczyk et al., 2012; Pfeiffer and Hohmann, 2012). For comparison, in the study of Przednowek et al. (2012), prediction models for the result in 400 m hurdle race over the selected three-month training cycle was presented. The obtained model featured an error at the level of 0.72 s, what is a very good result.

4. Conclusion

In this study, application of artificial neural networks as tools of result prediction in 400m hurdle race over a two-year training cycle was analyzed. The best model, verified by means of LOOCV, was the neural network with 6 neurons in the hidden layer and the hyperbolic tangent activation function. The obtained result generates prediction error at the level of $RMSE_{CV}$=0.75 s, what confirms the validity of that method in execution of the task under consideration. However, it should be emphasized that model was developed using data of athletes featuring high training level. Therefore it will generate much more serious errors if applied for athletes at a lower training level.

Corresponding Author:
Krzysztof Przednowek, PhD
Department of Methodology and Informatics,
Faculty of Physical Education,
University of Rzeszow
Rzeszów, Poland
E-mail: krzprz@ur.edu.pl

References

1. Arlot S., Celisse A.: A survey of cross-validation procedures for model selection. *Statistics Surveys*, 2010; 4: 40–79.
2. Bishop C.: *Neural Networks for Pattern Recognition.* 1995, Oxford University Press.
3. Chatterjee P., Banerjee A. K., Das P., Debnath P.: A regression equation to predict VO_2max of

young football players of Nepal. *International Journal of Applied Sports Sciences,* 2009; 22(2): 113–121.
4. Drake A., James R.: Prediction of race walking performance via laboratory and field tests. W: *World race walking research* (red. Pupiš M.), 2011; 71–79.
5. Haghighat, M., Rastegari, H., Nourafza, N., Branch, N., Esfahan, I. (2013). A review of data mining techniques for result prediction in sports. *Advances in Computer Science: an International Journal*, 2(5), 7-12.
6. Maszczyk A., Roczniok R., Waśkiewicz Z., Czuba M., Mikołajec K., Zając A., Stanula A.: Application of regression and neural models to predict competitive swimming performance. *Peceptual and Motor Skills*, 2012; 114(2): 610–626.
7. Maszczyk A., Zając A., Ryguła I.: A neural network model approach to athlete selection. *Sports Engineering*, 2011; 13: 83–93.
8. Papić V., Rogulj N., Pleština V.: Identification of sport talents using a web-oriented expert system with a fuzzy module. *Expert Systems with Applications*, 2009; 36: 8830–8838.
9. Papić V., Rogulj N., Pleština V.: Identification of sport talents using a web-oriented expert system with a fuzzy module. *Expert Systems with Applications*, 2009; 36: 8830–8838.
10. Pfeiffer M, Hohmann A. Application of neural networks in training science. *Human Movement Science,* 2013; 31:344–359.
11. Przednowek K., Iskra J., Cieszkowski S.: The use of selected linear models in predicting the results of 400-metre hurdles races. W: *Current Research in Motor Control IV*, AWF Katowice, 2012.
12. Przednowek K., Wiktorowicz K.: Neuronowy system optymalizacji wyniku sportowego zawodników uprawiających chód sportowy. *Metody Informatyki Stosowanej*, 2011; 4:189–200.
13. Roczniok R., Maszczyk A., Stanula A., Czuba M., Pietraszewski P., Kantyka J., Starzyński M.: Physiological and physical profiles and on-ice performance approach to predict talent in male youth ice hockey players during draft to hockey team. *Isokinetics and Exercise Science,* 2013; 21(2): 121-127.
14. Ryguła I.: Artificial Neural Networks As a Tool of Modeling of Training Loads. *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, Shanghai, 2005; 2985–2988.
15. Silva A., Costa A., Olivveira P., Reis V., Saavedra J., Perl J., Rouboa A., Marinho D.: The use of neural network technology to model swimming performance. *Journal of Sports Science and Medicine*, 2007; 6: 117-125.
16. StatSoft, Inc. (2011). STATISTICA (data analysis software system), version 10. www.statsoft.com.

10/21/2014