

# Neural Network of Vector Quantization as a Tool for Predicting Sports Achievements in Individual and Team Sports

Krutikov A.K.

Department of electronic computers  
Vyatka State University  
Kirov, Russia  
yadrodisk@yandex.ru

Meltsov Vasily Yurevich

Department of electronic computers  
Vyatka State University  
Kirov, Russia  
meltsov69@mail.ru

Podkovyrin Vitaliy Dmitrievich

Department of physical education  
Vyatka State University  
Kirov, Russia  
yadrodisk@yandex.ru

**Abstract** — Planning and forecasting in sport is an integral element of the sphere and system of physical culture and sport functioning, and many related areas of activity. The development of information technologies, methods and means of artificial intelligence gives wide opportunities for their application in the field of sports forecasting. The article describes the features of design and application of the system that predicts sports results in team and individual sports.

The functional structure of the system is given, the basic principles of its operation are considered. A feature of the structure is the introduction of a data clustering module based on a neural network of vector quantization of signals. The algorithm of this module functioning and the possibility of its expansion for detailing the formed forecast is described. The recommendations to users on the formation and detailing of the overall sample, evaluation of the generated sample effectiveness and analysis of the results are defined. The results of experiments on synthetic samples and real methods are presented and analyzed. Convenient and intuitive interface allows the system to be used by specialists of different profiles in the field of physical culture and sports, starting with the athletes themselves and ending with the heads of sports clubs, organizations and federations.

**Keywords**—*sports forecasting; software system; neural network; training sample; vector quantization; LVQ-network*

## I. INTRODUCTION

Forecasting in sports is used by experts everywhere. The outcomes of individual matches and the results of the teams 'performance throughout the championship, the results of the athletes' performance at individual stages of the world Cup and the final places of the participants are predicted. At the same time, customers of the forecast are often interested not only in individual achievements of athletes and numerical values of numerous individual indicators of the matches (the

number of corners, violations of the rules in each half, yellow and red cards, etc.), but also more global indicators. For example, prospects of spectator interest growth to sports discipline taking into account introduction of essential changes in rules, the approximate medal plan of the United team at the forthcoming championship, growth rates of world records, long-term prospects of certain sports and all Olympic movement development. Specialists of various sports fields use such forecasts to achieve their goals. Bookmakers, coaching staffs, sports functionaries, breeders are somehow forced to form sports forecasts and analyze them [1, 2, 3].

## II. LITERATURE REVIEW

Part of forecasts in sport is made in the format of empirical assessment, when an expert in a particular sport determines, evaluates and fixes his opinion on the future event. Some forecasts are made using ready-made algorithms and formulas, for example, using time series, the Wingate equation, various methods of regression analysis and other approaches based on mathematical statistics. A number of specialists use ready-made software products based on machine learning technologies and applied artificial intelligence to make predictions [4, 5, 6, 7, 8].

In the latter case, we can distinguish several basic approaches and elements of the mathematical apparatus, implemented with the help of modern technologies for designing software systems [9, 10, 11, 12]:

- use of time series analysis apparatus, including two main methods-extrapolation method and regression analysis method [13, 14];

- the use of analytical modeling apparatus used in the construction of various models of the predicted object, and then transformed into computer models for experiments [15];
- the use of machine learning, in particular, various clustering methods, involving the correlation of the predicted result to a particular class (cluster) of reliable events [16, 17, 18];
- the use of expert assessments that form the basis of the knowledge base in the design and implementation of software expert systems [5].

Software for the implementation of extrapolation methods, regression and cluster analysis is very diverse and is most often based on specialized mathematical packages. However, very few of these applications are adapted for sports forecasting.

In recent years, the apparatus of artificial neural networks has been used as one of the tools for implementing regression and cluster analysis [16, 17, 19]. In this work the specialized program module on the basis of a neural network of vector quantization for implementation of sports achievements forecasts in individual and team sports, for the purpose of various qualification athletes preparation plans improvement is considered.

### III. RESEARCH METHODOLOGY

The scientific staff of the Vyatka State University laboratory "Intelligent systems" together with the coaching and teaching staff of the physical education Department developed a prototype of a software system for predicting sports results based on the neural network approach [20]. The generalized structure of the system is shown in Fig. 1. The system consists of several modules, each of which implements a certain model of an artificial neural network. The primary processing unit receives the initial data, converts it to the required form and transmits it for processing to all (or only the user-specified) modules of the next level. The collector produced by the modules answers and, depending on inherent in it the algorithm either outputs all the obtained results in tabular form for analysis by experts, or selects the "best" of them according to the chosen criterion, or makes their aggregation subject to a comprehensive co-factor for "optimal" and set the system settings. The latter option requires a complex and time-consuming configuration of the hybridization unit and is available at the development and testing stage.

A number of experiments were carried out with the implementation of various models of neural networks. The modules were used as a regression analysis tool to predict numerical individual results of an athlete [20], and as a cluster analysis tool to predict the outcomes of sporting events.

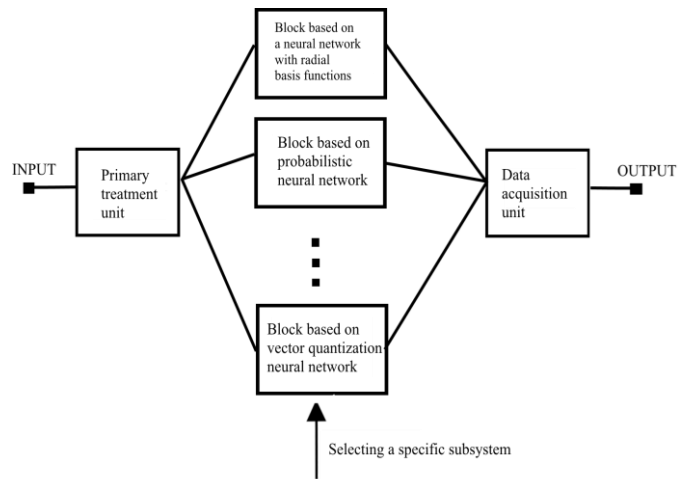


Fig. 1. Structure of forecasting software system

One of the models that successfully implements the algorithm of cluster analysis is a neural network of vector quantization (LVQ) [20]. Therefore, the system was supplemented with a software module containing a two-layer LVQ network. The competing layer performs vector clustering, and the line layer maps clusters to user-defined target classes. In the process of training LVQ-networks, the weights of neurons are adjusted taking into account the suitability of training examples and clusters to one class. Thus, the trained LVQ-network produces clustering of input vectors describing a specific applied research problem.

After the design and software implementation of this module, experiments were conducted to predict sports results. The prediction was to determine the class that most adequately corresponds to the vector coming to the inputs of the network. Each class is associated with a specific sporting event (for example, class\_1 – victory, class\_2 – defeat, class\_3 – draw, etc.).

### IV. RESULTS

For carrying out valid experiments the training sample taking into account features neural network of vector quantization functioning was formed. The set of vector elements is a set of anthropometric, physiological, technical, psychological parameters of an athlete (team). The resulting vector element is the number of the class to which the training vector belongs. When predicting a particular sporting event, the inputs of the neural network receive the vector of the most important statistical indicators of the athlete (team) for a certain period of training and competitive training. The output of the neural network is the value of the resulting element, that is, the number of the event class. Several configurations of the LVQ neural network with different structure, different number of neurons and different number of learning epochs were trained during the experiments. The fight between Russian boxer Sergei Kovalev (1.0) and Briton Anthony Yard (0.1) for the WBO world title was chosen as one of the main predicted events. The match ended with the victory of the Russian [21].

A prototype of a system for predicting sports results on the basis of neural networks, including a module containing LVQ

network, was developed in MATLAB and occupies 3.4 MB on the hard disk. Recommended system requirements for PC: Intel Core i3, 1.7 GHz, RAM from 1 GB, 200 MB of free hard disk space, Windows XP/7/8/10.

Two different algorithms were used to train the network. Table I shows the results of training the neural network with the use of specialized learning algorithm LVQ2.1. The LVQ2.1 algorithm stops after a fixed number of training epochs. It is specified by the program. The training time, in this case, directly depends on the number of epochs. The experiments were carried out at different values of the initial learning rate.

It should be noted that the table shows the results of forecasting for the learning rate of 0.001. With the increase in the initial rate of education, the accuracy of forecasts decreases. Thus, at the initial learning rate of more than 0.5 significantly increases the root mean square error of learning, and, consequently, the accuracy of the forecast falls and the network gives erroneous conclusions.

When conducting experiments with predicting the results of this sporting event, in 84% of cases outcome is predicted correctly. Having more detailed (insider) information about the boxer, the coaching staff can form forecasts, defining a specific event, or a number of events, in respect of which it is necessary to remove some uncertainty. For example, knowing about a minor injury of the opponent, it is possible to predict the number and strength of a certain type blows, and to prepare the corresponding defense or counterattack.

TABLE I. THE RESULTS OF THE FORECAST USING THE LEARNING ALGORITHM LVQ2.1

<i>Number of neurons</i>	<i>The number of training epochs</i>	<i>Time of learning</i>	<i>The result of the forecast</i>
1	500	22 sec.	(0,1)
7	500	22 sec.	(0,1)
24	500	24 sec.	(1,0)
36	500	29 sec.	(1,0)
77	500	40 sec.	(1,0)
171	500	58 sec.	(1,0)
223	500	1 min. 26 sec.	(1,0)
345	500	1 min. 43 sec.	(1,0)
475	500	2 min. 18 sec.	(1,0)
551	500	2 min. 27 sec.	(1,0)
668	500	3 min. 05 sec.	(1,0)
851	500	3 min. 52 sec.	(1,0)
1390	500	6 min. 12 sec.	(1,0)
2390	500	14 min. 47 sec.	(1,0)

To improve the accuracy of the forecast (with a constant prediction time) and its detail, it is possible to combine several

neural networks of vector quantization into a separate subsystem. Detailing is necessary if the total forecast (one of the total forecasts) may depend on another predicted event in this experiment. For example, statistics the number of strikes (percentage of strikes hit the target), forecasting of percent slopes, predictable physical form (fatigue) opponent in the round, the other predicted units (at the discretion of the relevant expert on the sport). At the same time, for the final neural network, which gives the final forecast, it is necessary to prepare a special training sample, based on the obtained prognostic results.

An example of combining modules using artificial neural networks of vector quantization into a separate subsystem is shown in Fig. 2.

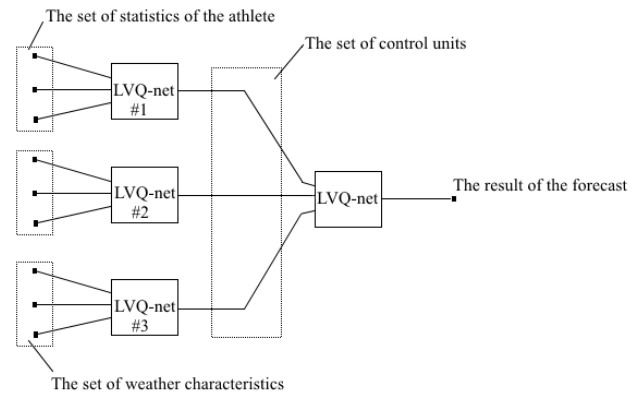


Fig. 2. The results of the forecast using the learning algorithm LVQ2.1

In this example, LVQ-network #1 is trained on the basis of the athlete statistical indicators. As a result of the forecast it gives some control unit - event\_1. LVQ-network #2 gives, respectively, the forecast for event\_2 and LVQ network #3 - on event\_3. The second and third networks can be trained on their own unique data sets, but they can also be trained on the same vector, producing different control results as a forecast. If it is necessary to obtain a specific numerical parameter for the intermediate forecast, a generalized regression neural network, a neural network with radio-basis functions or a cascade neural network [20] can act instead of one of the LVQ networks. From the obtained control units (predicted events), a vector is formed, which enters the input of the final LVQ-network trained earlier. The result of this module is the final forecast (event) required by specialists.

The designed multi-module LVQ network was implemented in MATLAB (Fig. 3).

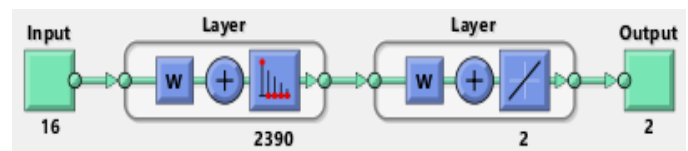


Fig. 3. LVQ network implementation in MATLAB

In General, when using LVQ-network as a forecasting tool, there is a tendency to increase the accuracy of forecasts when the structure of the neural network is enlarged. However, this usually increases the training time.

The developed neural network of vector quantization, successfully coped with the task of forecasting and in most experiments correctly determined the winner (table 1). The statistics used for the training sample were taken from open sources on the Internet: sports statistics websites, ratings of international professional Boxing federations, etc. Having a large number of statistical data, it is possible to enlarge and Refine the training sample, and, consequently, there will be new hidden dependencies determined by the neural network. It can also affect the result. Individual coaching staffs, leading specialists, personal trainers, working directly with specific athletes, having their own statistics can process it and include it in the training sample, while providing a more accurate forecast of the expected event.

## V. CONCLUSION

The use of the presented software system will allow to carry out forecasts not only to the trained specialists-analysts, for example, applied mathematicians and IT-professionals, but also to all workers of various level in the sphere of physical culture and sports. Depending on the tasks, the system can be configured for short-term, medium-term and long-term forecasts. With proper work with sports statistics and improvement of the training samples formation, the modules of the system can become a serious tool in predicting the outcomes of sports events, team results and individual achievements of athletes.

The modular architecture of the system makes it easy to add new nodes and components, taking into account the rapidly developing scientific industry of applied intelligent systems and artificial neural networks [23,24]. Experiments with various learning algorithms, algorithms for preprocessing the training sample, procedures for determining the most important parameters of the training sample will also be continued. A number of methodological recommendations for working with the system have been developed and are constantly being finalized. After analyzing the functioning of all modules, the transition from the prototype of the system in the MATLAB environment to its implementation in one of the high-level languages will be made for more effective use by all specialists of the sports sphere.

## References

- [1] A.O. Aleksina, D.V. Chernova, and A.Y. Aleksin, "Electronic interaction in the sphere of physical culture and sports services in Russia", *Advances in Intelligent Systems and Computing, Conference on Digital Transformation of the Economy: Challenges, Trends and New Opportunities*. 2018, vol. 908, pp 703-713.
- [2] S. Vos, M. Janssen, J. Goudsmit, C. Lauwerijssen, and A. Brombacher, "From Problem to Solution: Developing a Personalized Smartphone Application for Recreational Runners following a Three-step Design Approach", 11th conference of the International Sports Engineering Association, ISEA 2016, *Procedia Engineering*. 2016, vol. 147, pp. 799-805.
- [3] Q. Wang, "Application of Computer Technology in Sports", *Advances in Intelligent Systems and Computing, International Conference on Applications and Techniques in Cyber Intelligence*. 2019, vol. 1017, pp. 923-928.
- [4] L.N. Yasnitsky, A.V. Kirovasi, A.V. Rategov, and Cherepanov F.M, "Predicting outcomes of the 2015 world Championships in athletics methods of artificial intelligence", *Modern problems of science and education*. 2014, vol. 4, pp. 68-77. URL: <http://www.science-education.ru/ru/article/view?id=14423>.
- [5] D.A. Strabykin, *Intelligent systems logical forecasting of situations on the basis of deductive inference*, Kirov: Vyatka State University. 2014.
- [6] Y. Wang, V. Verba, D. Bezumnov, and E. Tarasenko, "Neural network using for prediction spinal diseases", 24th conference of FRUCT association. 2019, pp. 790-796.
- [7] "How neural networks help bookmakers". *Research of Bookmaker's Rating and expert's opinions*. URL: <https://bookmaker-ratings.ru/>.
- [8] L. Dimitrova, and K. Petkova, "Bayesian network-based causal analysis of injury risk in elite rhythmic gymnastics", *International Journal of Computer Science and Software Engineering (IJCSSE)*. 2014, vol. 2(1), pp. 50-61.
- [9] I.V. Bestuzhev-Lada, *Workbook on forecasting*, Moscow: Science. 1982.
- [10] M.L. Dolzhenkova, V.Yu. Meltsov, and D.A. Strabykin, "Method of consequences inference from new facts in case of an incomplete knowledge base", *Indian Journal Of Science And Technology*. 2016, vol. 9, 39, pp. 100413. DOI: 10.17485/ijst/2016/v9i39/100413.
- [11] V. Meltsov, G. Chistyakov, D. Strabykin, and M. Dolgenkova, "Accelerated logical inference in the intelligent control systems", *International Conference on Artificial Intelligence: Techniques and Applications, AITA 2016, DEStech Publications Inc*. 2016, pp. 1-7.
- [12] V.Yu. Meltsov, *High-performance systems of deductive inference: Monograph*, Yelm, WA, USA: Science Book Publishing House. 2014.
- [13] S. Russell, and P. Norvig, "Artificial intelligence: a modern approach", 3rd ed., Upper Saddle River, New Jersey: Prentice Hall. 2010.
- [14] R. Brachman, and H. Levesque, *Knowledge representation and reasoning*, Morgan Kaufmann. 2004.
- [15] A. Fischer, M. Do, T. Stein, T. Asfour, R. Dillmann, and H. Schwameder, "Recognition of individual kinematic patterns during walking and running — a comparison of artificial neural networks and support vector machines", *International Journal of Computer Science in Sport*. 2011, vol. 10(1), pp. 63-67.
- [16] S. Haykin. *Neural networks: full course*, 2nd ed. , Moscow: Publishing house "Williams". 2006.
- [17] T. Kohonen, *Self-organizing maps*, 3rd ed., Moscow: Binom. Laboratory knowledge. 2008.
- [18] S.H. Liao, J.L. Chen, and T.Y. Hsu, "Ontology-based data mining approach implemented for sport marketing", *Expert Systems with Applications*. 2009, vol. 36(8), pp.11045-11056.
- [19] R. MacGregor, "Using a description classifier to enhance knowledge representation". *IEEE Expert*. 2013, vol. 6(3), pp. 41-46. DOI:10.1109/64.87683.
- [20] V.Y. Meltsov, V.D. Podkovyrin, V.L. Klyukin, and A.K. Krutikov, "The use of cascade neural network direct transfer to predict athletic achievement in the shot put", *Scientific and Technical Bulletin of the Volga region*. 2018, vol. 4, pp. 136-139.
- [21] Wikipedia: Sergey Kovalev vs. Antony Yarde, URL: [https://ru.wikipedia.org/wiki/Сергей\\_Ковалев\\_-\\_Энтони\\_Ярд](https://ru.wikipedia.org/wiki/Сергей_Ковалев_-_Энтони_Ярд).
- [22] C. Pickerin, and J. Kiely, "The development of a personalised training framework: Implementation of emerging technologies for performance", *Journal of Functional Morphology and Kinesiology*. 2019, vol. 4, 2, C. 25.
- [23] V.Yu. Meltsov, V.A. Lesnikov, and M.L. Dolzhenkova, "Intelligent system of knowledge control with the natural language user interface", *International Conference "Quality Management, Transport and Information Security, Information Technologies"*. 2017, pp. 671-675. DOI: 10.1109/ITMQIS.2017.8085913.
- [24] Alexandr S. Kuznetsov. Russian Professor's meeting. *Russian Journal of Physical Education and Sport*. 2019, 14(1), pp. 17-22. DOI: 10.14526/2070-4798-2019-14-1-18-24