

Forecasting the competitive performance of young athletes based on artificial intelligence technology

UDC 796.015.844.4



Dr. Hab., Associate Professor **R.S. Nagovitsyn**¹

Dr. Hab., Professor **I.G. Gibadullin**¹

PhD, Associate Professor **O.N. Batsina**¹

PhD, Associate Professor **I.A. Mokrushina**¹

¹Tchaikovsky State Physical Education and Sport Academy, Tchaikovsky

Corresponding author: gto18@mail.ru

Abstract

Objective of the study was to develop a program for predicting the competitive performance of young athletes based on artificial intelligence technology.

Methods and structure of the study. As part of the scientific work, the collection and processing of individual data of athletes (n=56) was carried out according to 38 characteristics, ranked into 2-4 categories in three key areas: heredity, environment and individual.

Results and conclusions. As a result of data processing using deep neural networks and machine learning algorithms, two categories of prediction were identified: athletes who achieved a sports title or the highest category, and athletes who did not reach this level. The control testing of the created program showed only 11% of the error probability in predicting the competitive performance of young athletes. The author's program made it possible to identify reliable patterns: if a young athlete's mother has a sports title, then with 79% probability he will be effective in future competitive activities, and if he is still trained by a mentor with experience from 16 to 30 years, then the probability of reaching the highest level or sports title rises to 86%.

Keywords: *young athletes, competitive performance, artificial intelligence, forecasting, intellectual program.*

Introduction. Today, information technologies are becoming more and more active and multifaceted part of the life of the younger generation, students and society as a whole [2]. Every year it is more and more difficult to imagine any of the spheres of human social activity without various achievements of scientific and technological progress in the field of digitalization [3]. The society has reached the stage of development at which programs, neural networks, various intelligent systems and digital technologies have the most significant impact on it [1]. The introduction of such technologies, including the most "top" of them, such as artificial intelligence programs, using machine learning, are beginning to make more and more important in the process of continuous modernization to improve the quality of human life [6].

One of the most progressive social spheres of a person, in which various intellectual technologies are actively used, is physical culture and sports activity [4]. In the last decade, various artificial intelligence systems with a continuous increase begin to have an increasingly significant impact on it [1]. In the field of physical culture and sports, these technologies are used to analyze large arrays of sports data, in terms of statistics on the actions of athletes and referees based on video recordings of a match or indicators of the physical activity of competitors [1, 4, 6]. Intelligent systems provide virtual sports assistance, assist in supporting the physical culture and sports potential of athletes, in the implementation of e-sports and sports betting activities [1, 5]. In the period of preparation for competitions, the coaching staff and the athletes themselves begin to actively use the results of intellectual



analysis to make the most rational decisions, and in some cases, artificial intelligence technologies make it possible to fully automate the training process [2, 4, 5].

Thus, the theoretical analysis of the existing possibilities of using artificial intelligence to improve the quality of physical culture and sports activities of the younger generation shows that today there is a special need and need for experimental research in this direction [1, 4].

Objective of the study was to develop a program for predicting the competitive performance of young athletes based on artificial intelligence technology.

Methods and structure of the study. At the preliminary stage of the study (May 2022), individual data of athletes (n=56) involved in various sports or who have already completed their sports careers were collected and processed. To collect information, an analysis of archival data of athletes from children's and youth sports schools in the cities of Chaikovsky (Perm Territory), Izhevsk and Glazov (Udmurt Republic) was carried out, as well as their personal and telephone surveys. For the study, data were collected from athletes during the time period when they were young

Table 1. Classification of the data of young athletes on the grounds and their categories

| Signs: categories |
|---|
| Heredity |
| Gender: young man, girl |
| Zodiac sign by elements: fire, earth, air, water |
| The decade of the zodiac sign: 1-10, 11-19, 20-31 |
| Maximum oxygen consumption: <25, 25-35, >35 |
| Blood type: 1, 2, 3, 4 |
| Hemoglobin: <115, 115-150, >150 |
| Hematocrit: <35, 35-45, >45 |
| Minerals (%): <7, 7-10, >10 |
| Total amount of liquid (%): <40, 40-50, >50 |
| Protein (%): <10, 10-15, >15 |
| Muscle fibers are white/red: <b, b=k, <k |
| Vital capacity of the lungs: <1500, 1500-2000, >2000 |
| Respiratory volume: <180, 180-250, >250 |
| Environment |
| Age at the time of «enrollment» in the section: 7-8, 9-10, 11-12 |
| The marital status of a young athlete: a complete family, an incomplete family, an orphan |
| Place of residence: city of more than 100 thousand, city of less than 100 thousand, village |
| Elder brother / sister: no or not involved in sports, has a rank, title |
| Father: no or not involved in sports, has a rank, has a title |
| Mother: no or not involved in sports, has a rank, has a title |
| Coach qualification: has a rank, Candidate Master of Sports / Master of Sports title, International Master of Sports / Honored Master of Sports title |
| The performance of the trainees of the coach over the past 5 years: categories, CMS title, MS/IMS title |
| Trainer experience: up to 5 years, from 6 to 15, from 16 to 30, from 31 and above |
| Father or mother work in the field of FC and sports: no, yes, part-time |
| Education: school, gymnasium, lyceum |
| Who enrolled in the section: himself/friend, grandmother/grandfather, brother/sister, father/mother |
| Individual |
| The presence of the All-Russian sign «Ready for work and defense»(GTO): no, there is |
| Average grade in school for the previous year: <4,2, 4,2-4,7, >4,7 |
| Skipping classes for no good reason / all workouts(%): >10,3-10,<3 |
| Skipping classes for a good reason / all workouts (%): >15, 5-15, <5 |
| Physical Development according to GTO testing: unsigned, bronze / silver, gold |
| Average performance in the first two/three competitions: <3, 2-3, >2 |
| Cooper's Strength Endurance Test (min): <1,2, 1,2-2, >2 |
| Muscle mass index to height and weight: <15, 15-20, >20 |
| Fat content (%): <25, 25-35, >35 |
| Fat-free mass (%): <75, 75-65, >65 |
| Muscle mass in relation to the lower to upper extremities:<1,2,1,2-1,5,>1,5 |
| Muscle mass in relation to right and left: <1,05, 1,05-1,1, >1,1 |
| Muscle mass in relation to the limbs to the body: <1,2, 1,3-1,5, >1,5 |



athletes and were just starting their sports career. The experimental sample included athletes who currently have the highest sports category “Candidate for Master of Sports of Russia” or the sports title “Master of Sports of Russia” (n=14), athletes of III-I sports categories (n=17) and athletes with youth sports ranks or no ranks (n=25).

At the main stage of the study (June 2022), a program was developed to predict the competitive performance of young athletes based on artificial intelligence technology using the Orange analytical system. In the interface of the Orange software, a workflow was created to analyze data using process intelligent models. To implement the experimental work, deep neural networks and machine learning algorithms for categorical classification were used: “Logistic regression” and “Random forest” [5]. In the process of analyzing big data, the predictive and classification functionality of the Orange intellectual platform was used. For the “training” of the author’s program on the intellectual platform Orange, the data of not all athletes were used, but only those selected by random sampling (n=32) for training testing. An important condition for testing the program was the uniformity of the sample, namely, the tested sample should include both athletes of the highest category and titles, and those with sports categories or not having sports qualifications.

At the final stage of the study, a system analysis of the program for predicting the competitive performance of young athletes in the Orange intellectual system created at the previous stage made it possible to determine the main classification for predicting sports performance. The program was finally tested with a validation sample (n=15), which was also selected randomly, but with a mandatory condition for sample uniformity. With the help of the data of the validation set, the program already independently re-

vealed the forecast based on machine learning, taking into account the classification of patterns [5]. When determining the error, additional “training” of the intellectual system was carried out until the program found the correct and optimal patterns for predicting the correct answer [4]. At the end of this stage, the program was tested by a control sample (n=9), which also included athletes with titles, categories and no qualifications. The data of the control sample entered into the program made it possible to identify the final percentage of the reliability of forecasting of the author’s intellectual development.

Results of the study and their discussion.

Based on a preliminary analysis of the special scientific literature and the features of the versatile monitoring of young athletes, a system of signs and their categories was compiled to implement the process of intellectual forecasting [1, 4, 6]. The system was classified according to 38 features, each of which was ranked into 2-4 categories, in three key areas: heredity, environment and individual (Table 1):

After defining the features of comparison by category, the individual data of the athletes participating in the experiment, obtained at a time when they were still young and novice athletes, were loaded into the Orange analytical system. It should be noted that not all indicators were collected for each athlete, since some of the indicators were lost over the years or were not previously included in the monitoring of the athlete, especially the analysis of body composition. As a result of data processing using deep neural networks and machine learning algorithms “Logistic Regression” and “Random Forest”, two categories of prediction were identified: athletes with a sports title or the highest rank, and athletes who have not reached this level. After additional “training” of the intelligent system with the help of the validation sample, the pro-

Table 2. The results of the reliability of the categories of features that significantly affect the prediction of the competitive performance of young athletes (%)

| Categories of features on the result «will be performed by a candidate for master of sports of Russia or the title of MS Russia» | | 1 | 1-2 | 1-3 | 1-4 |
|---|---|----|-----|-----|-----|
| 1 | Sports mother: has a sports title | 79 | 86 | 93 | 100 |
| 2 | Trainer experience: 16 to 30 | | | | |
| 3 | Father or mother work in the field of FC and sports: part-time | | | | |
| 4 | The performance of the trainees of the coach over the past 5 years: the title of Master of Sports (MS) / International Master of Sports (IMS) | | | | |
| Categories of signs according to the result «Candidate for Master of Sports of Russia or the title of MS of Russia will not fulfill the category» | | 1 | 1-2 | 1-3 | 1-4 |
| 1 | Who «recorded» in the section: grandmother or grandfather | 83 | 88 | 93 | 100 |
| 2 | Training of a young athlete: lyceum | | | | |
| 3 | Skipping classes without a valid reason from all training sessions (%): >10 | | | | |
| 4 | Average grade in school for the previous year: <4.2 | | | | |



gram was verified by the data of the control sample. As a result, the program showed only one error out of nine, which corresponds to only 11% of the probability of an error created by the author's program based on artificial intelligence. The program showed the only error, only for a group of 6 athletes who did not reach the highest sports category and the title of candidate master or master of sports of Russia.

Thus, the program for predicting the competitive performance of young athletes based on artificial intelligence technologies was tested and finally developed. Nevertheless, when using the program in practice to improve the training process, data were needed on which features and categories had the greatest impact on performance or, conversely, on their insignificance for further success in competitive activity. In this regard, by changing the functionality of the intellectual program, categories of features were identified that most significantly affect the result of predicting the competitive performance of young athletes (Table 2)

After identifying the categories of features that most reliably affect the result of predicting the competitive effectiveness of young athletes by analyzing the data of the entire experimental sample (N = 56), the percentage ratio was obtained in various combinations of categories of features (Table 2). The effectiveness of forecasting in both categories is 100% recorded if the young athlete manifests itself in monitoring jointly 4 categories of features. If the young athlete is fixed in the 3rd category of features (1-3), then he will be performed in the future in the future in the future athletic, the highest category or title or, conversely, as shown in table 2, in 93% of the probability will not reach this level. In turn, if a young athlete has a sports title (1 sign with a category for execution), then with a 79% probability it will be effective in future competitive activities. And if, it will still be trained by a mentor with experience from 16 to 30 years old (1-2 signs), then the probability of achieving the highest sports category and the title of master of sports of Russia increases to 86%. On the other hand, if the athlete was "recorded" in the section Grandpa or grandfather (1 sign with a category for non-implementation), then he will not reach a high level in his career with 83%. And if he still studies at the Lyceum (1-2 signs), then the probability of not reaching the sporting highest category and the title of Russia in him increases to 88%.

Conclusions. Thus, the study attempts to look into the future and create an intelligent big data technology that is fully consistent with the development of modern sports. The developed author's program based on artificial intelligence makes it possible to predict the com-

petitive performance of a novice athlete and identifies key categories of features that positively or negatively affect the possibility of a novice athlete achieving the highest sports category or title in the future. The use of neural networks and machine learning algorithms, as shown by the results of the study, improves the quality of sports selection, which will allow timely individualization and improvement of the training process. In this direction, it is artificial intelligence technologies with a properly selected algorithm for clustering, classification and forecasting that have a strong and stable ability for intellectual analysis, choosing the main data that are significant for improving the effectiveness of a young athlete's sports trajectory.

References

1. Morhat P.M. Iskusstvennyy intellekt v sfere sporta: vozmozhnosti, napravleniya i sposoby zadaystvovaniya [Artificial intelligence in the field of sports: opportunities, directions and methods of involvement]. *Teoriya i praktika fizicheskoy kultury*. 2018. No. 10. pp. 95-97.
2. Nagovitsyn R.S., Vladykina I.V., Senator S.Yu. Programma podgotovki k sdache norm vserossiyskogo kompleksa «GTO» na osnove mobilnogo obucheniya [The program of preparation for passing the norms of the All-Russian complex "GTO" based on mobile learning]. *Teoriya i praktika fizicheskoy kultury*. 2015. No. 1. pp. 46-48.
3. Nagovitsyn R.S., Rassolova E.A., Senator S.Yu., Torbina I.I. Razrabotka veb-portala dlya podgotovki studentov k testirovaniyu po normam GTO [Development of a web portal for preparing students for testing according to the GTO norms]. *Teoriya i praktika fizicheskoy kultury*. 2016. No. 1. pp. 39-42.
4. Yue Q. Dynamic database design of sports quality based on genetic data algorithm and artificial intelligence. *Comput Intell Neurosci*. 2022. No. 16. 7473109.
5. Wang T., Park J. Design and implementation of intelligent sports training system for college students' mental health education. *Front Psychol*. 2021. No. 12. 634978.
6. Zhang B., Jin H., Duan X. Physical education movement and comprehensive health quality intervention under the background of artificial intelligence. *Front Public Health*. 2022. No. 10. 947731.