УДК 004.89

ПРОГНОЗИРОВАНИЕ РЕЗУЛЬТАТОВ XXXII ЛЕТНИХ ОЛИМПИЙСКИХ ИГР 2020 С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Крутиков А.К., Клюкин В.Л., Подковырин В.Д.

ФГБОУ ВО «Вятский государственный университет», Киров, e-mail: yadrodisk@yandex.ru

В статье описывается эксперимент по прогнозированию количества золотых медалей и мест в неофициальном командном зачете Олимпийских Игр 2020 г. в городе Токио, с использованием обобщенно-регрессионной искусственной нейронной сети и искусственной нейронной сети с радиально базисными функциями. В качестве национальных сборных команд, для которых будет осуществлен прогноз, выбраны национальные команды Российской Федерации (РФ), Соединенных Штатов Америки (США), Китайской Народной Республики (КНР) и Японии, которая является хозяйкой проводимых игр. Для произведения прогноза разработаны модули в программной среде MATLAB. Целью исследования является оптимизация, повышение точности прогнозирования распределения медального плана мультиспортивных событий, с использованием модульной системы на основе искусственных нейронных сетей. Нейронные сети обладать высокой способностью к обобщению, анализируя объемные наборы данных, предоставляемых для прогноза. Способность к обобщению позволяет решать задачи аппроксимации и регрессии, в виде которых может быть представлена задача прогнозирования. Для проведения эксперимента разработаны обучающие выборки, фрагмент одной из них приведен в виде графика. Результаты экспериментов приведены в таблице. Произведен анализ полученных данных, а также анализ целесообразности дальнейшего использования разработанных модулей.

Ключевые слова: прогнозирование, искусственная нейронная сеть, спортивный результат, неофициальный командный зачет, обобщенно-регрессионная нейронная сеть, нейронная сеть с радиально-базисными функциями

FORECASTING THE RESULTS OF THE XXXII SUMMER OLYMPIC GAMES 2020 USING ARTIFICIAL NEURAL NETWORKS

Krutikov A.K., KluykinV.L., Podkovyrin V.D.

Vyatka State University, Kirov, e-mail: yadrodisk@yandex.ru

The article describes an experiment to predict the number of gold medals and places in the unofficial team competition of the 2020 Olympic Games in Tokyo, using a generalized regression artificial neural network and an artificial neural network with radially basic functions. The national teams for which the forecast will be made are the national teams of the Russian Federation (RF), the United States of America (USA), the people's Republic of China (PRC) and Japan, which is the host of the games. Modules in the MATLAB software environment have been developed for making a forecast. The aim of the study is to optimize and improve the accuracy of the process of planning sports training and forecasting the distribution of the medal plan for multi-sport events, using a modular system based on artificial neural networks. Neural networks have a high ability to generalize by analyzing the large data sets provided for the forecast. The ability to generalize allows you to solve approximation and regression problems, which can be represented as a prediction problem. For the experiment, developed the training sample, a fragment of one of them is given in the form of a graph. The results of the experiments are shown in the table, the analysis of the data obtained, as well as the analysis of the feasibility of further use of the developed modules.

Keywords: forecasting, artificial neural network, sports results, unofficial team standings, generalized regression neural network, a neural network with radial basis functions

Олимпийские игры с 1896 г. являются самым крупным мультиспортивным мероприятием в мире. XXXII Олимпийские игры пройдут в Японии в городе Токио в 2020 г. Для большинства ведущих стран мира, данное событие является ключевым в мире спорта, и используется во внешнеполитических и внутриполитических целях.

Национальные олимпийские комитеты и государственные спортивные системы выстраивают подготовку спортивных сборных команд, с учетом идущих олимпийских циклов, а в некоторых случаях, ориентируясь на конкретные Олимпийские игры (ОИ). Например, спортивные реформы КНР к ОИ-2008 в г. Пекине (в том числе, закон о спорте в КНР [1]), планомерная подготовка и усиление национальных сборных РФ к ОИ-2014 в г. Сочи [2]. Неофициальный командный зачет (НКЗ) – распространенная система оценки выступления команд на играх, которая не признается международным олимпийским комитетом, но активно тиражируется национальными средствами массовой информации, национальными олимпийскими комитетами и спортивными федерациями.

Подготовка к ОИ рождает конкуренцию спортивных государственных систем, не только в, непосредственно, спортивной ее части, но и в методической, материально-ресурсной, управленческой и кадровой областях. Оптимизация расходов на подготовку национальных команд и команд олимпийского резерва, оценка турнирного графика и ряда многих других факторов требует

тщательного планирования. Именно спортивное планирование такого рода относится к задачам долгосрочного или, в отдельных случаях, сверхдолгосрочного спортивного прогнозирования [3]. Спортивное прогнозирование является составляющей процесса спортивной подготовки спортсменов как индивидуальных, так и командных видов спорта. Результатами прогнозов пользуются спортивные функционеры, чиновники, тренерские группы, сами спортсмены, а также букмекеры. Временной промежуток зависит от целей и задач прогнозирования, и помимо перечисленного, можно выделить среднесрочные, краткосрочные и сверхкраткосрочные прогнозы [3].

Использование для достижения определенных целей современных технических разработок является залогом успешной конкуренции. К современным быстро развивающимся прикладным научным областям относятся информационные технологии, технологии искусственного интеллекта [4], методы машинного обучения и т.д. Все они активно применяются при решении задач прогнозирования, в том числе спортивного прогнозирования. Для планирования траекторий развития ситуаций и моделирования этих процессов с помощью ЭВМ разработаны системы поддержки принятия решений, экспертные системы, системы нечеткой логики. И это лишь малая часть тех технических и информационных средств, которые применяются для осуществления прогнозов. Как инструмент решения задач прогнозирования [5], в том числе спортивного [6, 7], успешно зарекомендовали себя и искусственные нейронные сети (ИНС). В данной работе речь пойдет об использовании таких ИНС, как обобщенно-регрессионные нейронные сети (GRNN) и нейронные сети с радиально-базисными функциями (RBF).

Целью проводимых исследований являются оптимизация и повышение точности процесса планирования спортивной подготовки как индивидуальных спортсменов, так и команд в целом, и прогнозирование возможного медального зачёта в мультиспортивных состязаниях с использованием системы на основе ИНС.

Материалы и методы исследования

В качестве главного объекта исследований, в данной работе, выступает прогнозирование места в неофициальном командном зачёте на ОИ-2020 сборных команд четырех стран – РФ, США, КНР и Японии. Также прогнозируется количество золотых медалей ОИ-2020, завоёванных этими командами.

Научными сотрудниками лаборатории «Интеллектуальные системы» им. В.А. Байкова ФГБОУ ВО Вятского государственного университета разработаны два модуля программной системы прогнозирования: на основе обобщенно-регрессионной нейронной сети и радиально-базисной нейронной сети. Модули разработаны в среде MATLAB. Нейронная сеть, по сравнению с рядом других средств прогнозирования, имеет преимущество в способности принимать и обрабатывать наборы параметров, которые могут не учитываться другими приложениями. В спортивном прогнозировании подбор и анализ как основных, так и множества дополнительных параметров, влияющих на итоговые результаты, играет ведущую роль в достижении точности прогноза.

В зависимости от типов и особенностей спортивных событий необходимо менять структуру нейронных сетей, математические модели, используемые в каждом отдельном модуле ИНС и, соответственно, варьировать обучающие выборки (ОВ). Обучающая выборка, в зависимости от типа прогнозирования и сложности лингвистической переменной, являющейся непосредственным результатом прогноза, формируется из заранее подготовленных и обработанных наборов параметров. Например, для прогнозирования индивидуального результата легкоатлета, в обучающую выборку необходимо включить не только его контрольно-тестовые показатели и профильные результаты в избранной дисциплине, но и климатические показатели мест проведений сборов и самих турниров, фармакологические, антропометрические параметры и т.п.

Для эксперимента, описанного в данной работе взят набор параметров, рекомендуемых профессорами Эндрю Бернарда и Меган Буссе [8]. Согласно результатам их последнего исследования, а также результатам, изложенных в работе [9], ключевую роль в распределении мест в неофициальном командном зачёте играют: ВВП на душу населения государства, итоги выступления сборной на предыдущих ОИ, численность населения, фактор выступления в родной стране, и ряд специализированных спортивных параметров. Отдельно в обучающую выборку добавлено количество комплектов наград, разыгрываемых на ОИ. Обучающая выборка представляет собой массив векторов. Каждый вектор – это указанные выше параметры национальной сборной перед определенными ОИ, и каждому исходному вектору соответствует определённое число в результирующем векторе. Соответственно, в зависимости от цели прогноза, этим

числом может быть место национальной команды в НКЗ, либо число золотых медалей, завоёванных спортсменами этой сборной. Фрагмент результирующих векторов обучающей выборки для первого варианта прогноза приведен на рис. 1.

Как описывалось ранее, для эксперимента, разработаны два модуля, в основе каждого из которых лежит определённая нейронная сеть. Первый модуль содержит обобщенно-регрессионную нейронную сеть (GRNN) [10]. GRNN-сеть предназначена для решения задач регрессии, и в данном случае, прогнозирование будет представлено как механизм регрессии. GRNN-сеть будет копировать все имеющиеся обучающие векторы и использовать их для оценки так называемого отклика в произвольной точке сети. Конечный результат работы GRNN-модуля – это взвешенное среднее значение выходов по всем обучающим векторам, где величины весов отражают расстояние от этих векторов до той точки, в которой производится текущее оценивание. Первый промежуточный слой нейронов GRNN-сети состоит из радиальных элементов. Второй промежуточный слой нейронов содержит элементы, которые помогают оценить взвешенное среднее значение [10].

Второй модуль разработан на основе нейронной сети с радиально базисными функциями (RBF). RBF-сеть является нейронной сетью с прямыми связями, и используется преимущественно для решения задач аппроксимации [11]. Поэтому прогнозирование, с использованием разработанного модуля, будет представлять собой решение задачи аппроксимации. Фактически это задача поиска значения функции от несколькими переменных, где в качестве параметров будут выступать описанные ранее значения обучающей выборки. Структура RBF-сети приведена на рис. 2.

Первый слой нейронов является радиальным слоем, второй слой – линейный. Недостаток такой модели – серьезная зависимость размерности от обучающей выборки. То есть размерность радиального слоя должна соответствовать количеству векторов в обучающей выборке, соответственно, крупная выборка спровоцирует громоздкую RBF-сеть. В данной работе подобрана оптимальная обучающая выборка, и объём, который занимает компьютерная модель «оптимальной» RBF-сети, не превысил 1,22 Мб.

Результаты исследования и их обсуждение

Применяемая для экспериментов обучающая выборка построена на основе открытых источников в сети Интернет [12]. Нейронные сети обучены в соответствии с правилами, относящимися к соответствующей модели ИНС. Стоит отметить, что для GRNN-сети составлено две вариации обучающей выборки. Одна из выборок разделена на четыре набора, и фактически представляет четыре различных выборки, в соответствии с указанной страной. В другой выборке данные по всем рассматриваемым странам сведены в один массив, с отдельными числовыми элементами, являющимися кодами стран. Для RBF-сети подобного образца получить не удалось, и сеть обучалась на основе отдельных выборок по рассматриваемым странам.

Целевая ошибка обучения во всех экспериментах варьировалась от 0,1 до 0,0001, при этом изменение целевой ошибки не давало изменения результата прогноза. Минимальная среднеквадратичная ошибка (MSE) не превысила значения 0,00542.

Образец фрагмента обучающей выборки, составленной на основе параметров национальной сборной КНР приведен на рис. 3.

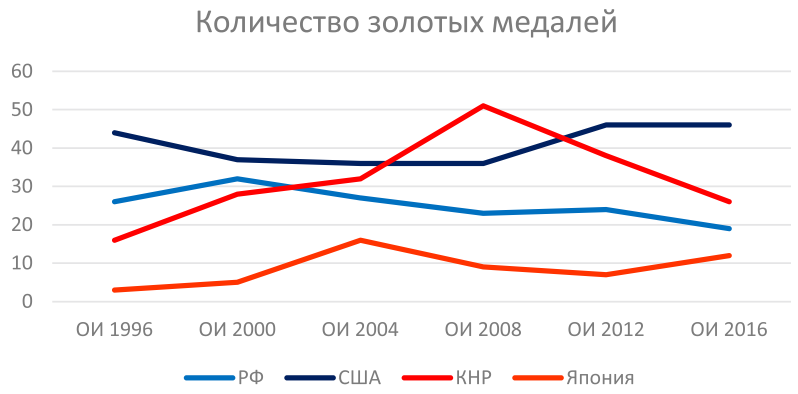

Рис. 1. Фрагмент обучающей выборки

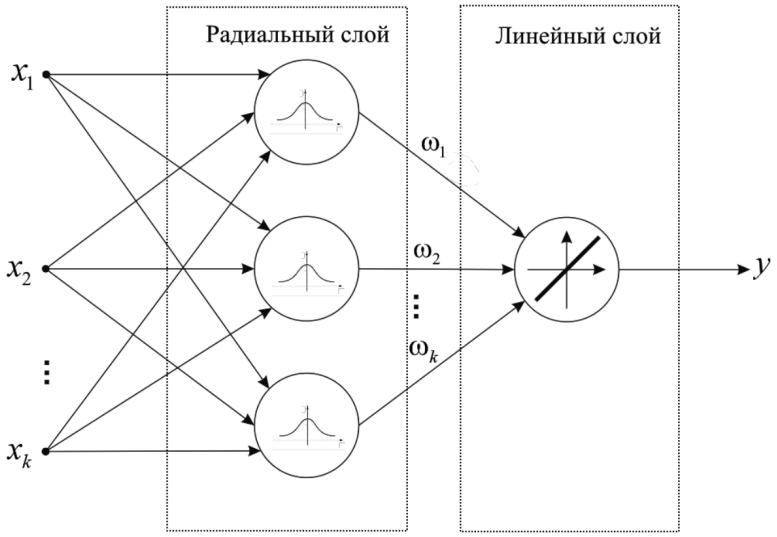

Рис. 2. Структура RBF-сети

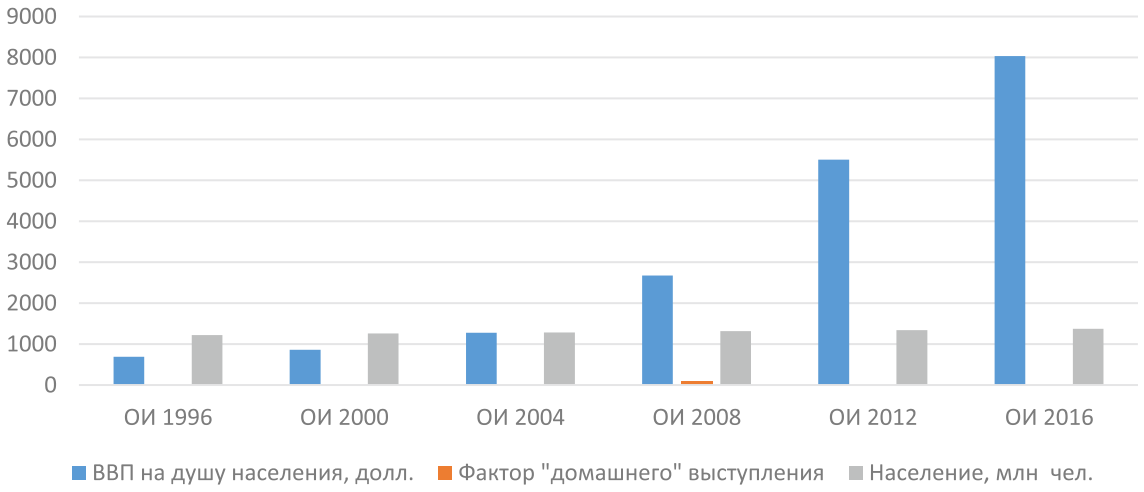

Рис. 3. Фрагмент обучающей выборки на основе параметров национальной сборной КНР

Результаты прогноза по золотым медалям на ОИ-2020 приведены в табл. 1. Важно отметить, что количество золотых медалей в первой таблице не является отдельным элементом обучения для экспериментов, в ходе которых получена вторая таблица. Таким образом, прогнозируемое число золотых медалей не связано с прогнозируемым местом.

Результаты прогноза по местам НКЗ ОИ-2020 оформлены и приведены в табл. 2. Прогноз будет актуален, в случае, если олимпийский турнир состоится в указанные Международным олимпийским комитетом сроки. В результате прогноза, по данным большинства экспериментов, лидерство сохранит сборная США. В свою очередь, РФ и КНР сохранят места в неофициальном командном зачёте, аналогично Играм 2016 г. Определенный рывок может совершить Япония, с учетом фактора домашнего выступления, что показывают отдельные эксперименты с GRNN-сетью. Точность прогноза можно будет оценить, только при условии участия всех национальных сборных в полном составе, под флагом своего государства.

Таблица 1

Результаты прогноза

	GRNN (ОВ № 1), золотых медалей	GRNN (ОВ № 2), золотых медалей	RBF (ОВ № 1), золотых медалей
Олимпийская сборная РФ	24	24	19
Олимпийская сборная США	46	46	36
Олимпийская сборная КНР	26	26	12
Олимпийская сборная Японии	3	36	5

Таблица 2

Результаты прогноза

	GRNN (ОВ № 1), место	GRNN (ОВ № 2), место	RBF (ОВ № 1), место
Олимпийская сборная РФ	4	4	3
Олимпийская сборная США	1	1	1
Олимпийская сборная КНР	3	3	2
Олимпийская сборная Японии	15	3	5

Время обучения нейронных сетей, во всех экспериментах не превысило 45 секунд. С учетом временного промежутка прогноза (долгосрочный), отдельной таблицы с временем обучения не приводится, а фактическое время обучения во всех экспериментах считается незначительным, и не влияющим на результат прогноза.

Заключение

Эксперименты показывают, что тенденции спортивного прогнозирования, вне зависимости от его типа и временного промежутка, схожи. Также, вне зависимости от формулировки задачи на прогнозирования (аппроксимация, регрессия, кластеризация), нейронные сети весьма эффективно осуществляют прогноз, и дают на выходе конкретные численные значения с незначительными погрешностями [13–15].

С учетом высокой точности прогнозов, и дальнейшего совершенствования технологий прогнозирования, возможно использовать данные модули как аналитический инструмент при оценке медальных планов и планировании подготовки национального спортивного резерва.

Тестирование модулей на данный момент продолжается. Моделируются прогнозы различного типа, различных временных промежутков, в различных видах спорта. Планируется провести переход от прототипов в среде MATLAB, к программным реализациям на языках высокого уровня. Ведется работа по созданию механизма определения ряда ключевых параметров в обучающей выборке, т.е. набора прогнозоформирующих предикторов.

Список литературы

1. Закон о спорте КНР. Материалы VIII съезда Компартии КНР 29.08. 1995 VIII съезд партии. Пекин, 1995. 182 с.

2. Олимпийские Игры 2014 в г. Сочи [Электронный ресурс]. URL: https:// ruxpert.ru/ Олимпиада_в_Сочи_2014 (дата обращения: 01.02.2020).

3. Прогнозирование в спорте [Электронный ресурс]. URL: https://studopedia.ru/4_174852_lektsiya--prognozirovanie-v-sporte.html (дата обращения: 01.02.2020).

4. Мельцов В.Ю., Страбыкин Д.А. Вывод следствий с построением схемы логического вывода // Фундаментальные исследования. 2013. № 11–8. С. 1588–1593.

5. Галушкин А.И. Нейронные сети: основы теории. М.: ГЛТ, 2012. 496 с.

6. Perl J. Neural network-based process analysis in sport. In J. R.R. Dopico, J. Dorado de la Calle & A.P. Sierra (Eds.). Encyclopedia of artificial intelligence, III. 2008. P. 1212–1218.

7. Жуков С.В., Зеленский А.В. Использование нейронных сетей в построении оптимальной тренировочной траектории в биатлоне // Учёные записки университета им. П.Ф. Лесгафта. 2017. № 5 (147). С. 81–88.

8. Не счесть золота? [Электронный ресурс]. URL: https://www.forbes.ru/forbes/issue/2004-08/20514-ne-schest-zolota (дата обращения: 01.02.2020).

9. Ясницкий Л.Н., Павлов И.В., Черепанов Ф.М. Прогнозирование результатов Олимпийских игр 2014 года в неофициальном командном зачете методами искусственного интеллекта. // Современные проблемы науки и образования. 2013. № 6. [Электронный ресурс]. URL: https://science-education.ru/ru/article/view?id=11206 (дата обращения: 01.02.2020).

10. Крутиков А.К. Прогнозирование спортивных результатов в индивидуальных видах спорта с помощью обобщенно-регрессионной нейронной сети // Молодой ученый. 2018. № 12. С. 22–26. URL: https://moluch.ru/archive/198/48884/ (дата обращения: 09.03.2020).

11. Каллан Р. Нейронные сети: Краткий справочник. М.: Вильямс И.Д., 2017. 288 с.

12. IOC [Электронный ресурс]. URL: https://www.olympic.org/the-ioc (дата обращения: 09.03.2020).

13. Крутиков А.К. Особенности обучения нейронной сети векторного квантования при разработке специализированного программного модуля // Южно-Сибирский научный вестник. 2019. № 2. С. 150–154.

14. Мельцов В.Ю., Крутиков А.К., Долженкова М.Л., Чистяков Г.А., Подковырин В.Д. Применение нейросетевых алгоритмов кластерного анализа при решении задач прогнозирования спортивных событий // Научно-технический вестник Поволжья. 2018. № 12. С. 250–252.

15. Крутиков А.К. Совершенствование подходов к разработке программных систем прогнозирования в спорте: материалы 57-й Междунар. науч.студ. конф. 14–19 апреля 2019 г. Новосиб. гос. ун-т. Новосибирск: ИПЦ НГУ, 2019. 141 с.