

УДК: 519.2

ПРОГНОЗИРОВАНИЕ РЕЗУЛЬТАТОВ СОРЕВНОВАНИЙ ПО КОННОМУ СПОРТУ С ПОМОЩЬЮ НЕЙРОННЫХ СЕТЕЙ

Хамадиярова В.Е.¹

¹Тюменский индустриальный университет, Россия, г. Тюмень, ул. Володарского, 38, e-mail: general@tyuiu.ru

Рассматривается возможность использования технологии нейронных сетей как инструмента для прогнозирования результатов участия спортсменов в соревнованиях по конному спорту, в частности, по выездке; рассмотрена тема прогнозирования в спорте, её значимость, а также актуальность таких прогнозов в России; рассмотрена предметная область, а именно выездка, программы выездки и способ оценивания всадников; определены ключевые характеристики всадника, лошади и соревнования, которые могут влиять на результат выступления пары (всадника и лошади) на соревнованиях по выездке; проведён поиск, сбор и обработка результатов соревнований по выездке с 2011 по 2019 год, полученных из базы данных; рассмотрены различные программные комплексы с возможностью создания и обучения нейронных сетей и определён наиболее оптимальный вариант. Смоделирован симулятор нейронной сети в среде MATLAB с помощью пакета расширений Deep Learning Toolbox (ранее известный как Neural Network Toolbox). На основе подготовленного обучающего набора данных разработанный симулятор нейронной сети обучен прогнозировать результаты спортсменов по конному спорту, а именно, по выездке. Полученная сеть апробирована на тестовых данных. Полученные данные проанализированы и сделаны соответствующие выводы.

Ключевые слова: прогнозирование, нейронные сети, MATLAB, Neural Network Toolbox, конный спорт

PREDICTION THE RESULTS OF COMPETITIONS IN EQUESTRIAN SPORT BY USING NEURAL NETWORKS

Hamadijarova V.E.¹

¹Tyumen Industrial University, Russia, Tyumen, Volodarskogo, 38, e-mail: bykova_a@mirea.ru

The possibility of using neural network technology as a tool to predict the results of the participation of athletes in equestrian competitions, in particular dressage, is being considered; The topic of predicting in sports, its significance, as well as the relevance of such forecasts in Russia are considered; the subject area is examined, namely dressage, dressage programs and a method for assessing riders; defines the key characteristics of the

rider, horse and competition, which can affect the performance of a couple (rider and horse) in dressage competitions; Search, collection and processing of the results of dressage competitions from 2011 to 2019, obtained from the database; various software systems with the possibility of creating and training neural networks are considered and the most optimal option is determined. A neural network simulator was simulated in MATLAB using the Deep Learning Toolbox (formerly known as the Neural Network Toolbox) extension pack. Based on the prepared training data set, the developed neural network simulator is trained to predict the results of athletes in equestrian sports, namely dressage. The resulting network is tested on test data. The data obtained are analyzed and the corresponding conclusions are made.

Keywords: prediction, neural networks, MATLAB, Neural Network Toolbox, equestrian sport

Введение.

Прогнозированию результатов выступлений спортсменов на соревнованиях придаётся немалое значение. На достоверность таких прогнозов возлагается довольно большая ответственность, поскольку они могут играть ключевую роль в планируемом тренировочном процессе спортсмена и в оптимизации отбора кандидатов в сборную команду. Как следствие от результатов прогнозирования может зависеть успешность выступления спортсмена на предстоящих соревнованиях.

Данная тема довольно интересна для исследования, поскольку использование компьютерных технологий, в частности, нейронных сетей при прогнозировании в области физической культуры и спорта на данный момент ещё не вышло на высокий уровень.

Прогнозирование результатов спортсменов в конном спорте является достаточно актуальной темой в России, так как русские всадники не брали призовых мест на Олимпийских играх с 1980 года, а количество олимпийских медалей, как известно, всегда служило показателем развития того или иного вида спорта в стране.

Стоит отметить, что на Олимпийских играх успешно выступают представители тех видов спорта, которые наиболее полно используют возможности научного обеспечения процесса подготовки, в том числе и прогнозирования. [2, с. 6]

Теме прогнозирования в спорте посвящено немало работ. Например, в книге "Физкультура и спорт" Баландина В.И., Блудова Ю.М. и Плахтиенко В.А. [1] рассматриваются актуальные проблемы прогнозирования в спорте. Используя обширный исследовательский материал из различных видов спорта, авторы дают методические и методологические подходы к разработке методов экспертных оценок, экстраполяции и моделирования в спорте.

Таким образом, целью данного исследования является разработка и научное обоснование методики прогнозирования результатов соревнований по конному спорту с помощью нейронных сетей.

Задачи исследования:

1. Определить ключевые факторы, влияющие на выступления всадников в соревнованиях по выездке.
2. Смоделировать симулятор нейронной сети в среде MATLAB с помощью пакета расширения Deep Learning Toolbox (ранее Neural Network Toolbox).
3. На основе подготовленного обучающего набора данных обучить разработанный нейросимулятор прогнозировать результаты спортсменов.
4. Проанализировать полученные данные и сделать соответствующие выводы.

Материалы и методы.

Результаты соревнований берутся с сайта <https://www.equestrian.ru/> [5]. Данный сайт предоставляет обширную базу данных, которая содержит такую информацию как: положения и результаты соревнований, информация о конноспортивных клубах и школах, данные о спортсменах, данные о лошадях.

Методы исследования:

1. Теоретические методы научного познания: дедукция, аналогия, анализ и синтез, обобщение, специализация, абстрагирование и моделирование.
2. Эмпирические методы познания: сравнение и эксперимент.

Основная часть.

Выездка - это олимпийский вид конного спорта, в котором всадник демонстрирует мастерство управления лошадью, а лошадь, в свою очередь, демонстрирует свою гибкость, аккуратность, динамичность и послушание. Участникам соревнований необходимо строго по памяти проехать езду, которая состоит из различных элементов выездки.

Сложность езды зависит от программы. В данном случае будут рассматриваться следующие программы: предварительный приз, командный приз, личный приз, малый приз, средний приз, большой приз

За каждый элемент всадник получает от 0 до 10 баллов, далее баллы складываются, и вычисляется процент положительных баллов. Побеждает та пара, которая сумела проехать езду на наибольшее количество процентов.

На результат выступления спортсмена может повлиять множество факторов, начиная с места проведения соревнований и заканчивая настроением лошади. Для данного же исследования взяты следующие параметры: пол всадника, возраст всадника, разряд или звание всадника, пол лошади, возраст лошади и программа соревнования.

Результаты выступлений всадников на соревнованиях с 2011 по 2019 год, взятые из базы данных, представлены в виде таблицы Excel (Рис. 1).

Год	Данные всадника				Данные лошади			Данные соревнования			Результат	
	Пол всадника (0 - м, 1 - ж)	Год рождения всадника	Возраст всадника	Разряд (0 - б/р, 1 - 1 разряд, 2 - 2 разряд, 3 - 3 разряд, 4 - КМС, 5 - МС, 6 - МСМК)	Пол лошади (0 - жер., 1 - коб., 2 - мер.)	Год рождения лошади	Возраст лошади	Количество участников	Уровень (0 - предварительный, 1 - командный, 2 - личный, 3 - малый, 4 - средний, 5 - большой)	Место	% Успеха	
2011	1	1974	37	4	0	2000	11	4	3	1	100,0	
2011	0	1983	28	5	1	2000	11	4	3	2	75,0	
2011	1	1988	23	1	0	1999	12	4	3	3	50,0	
2011	1	1983	28	0	2	2001	10	4	3	4	25,0	
2012	1	1998	14	0	2	2005	7	18	0	1	100,0	
2012	0	1999	13	3	1	2000	12	18	0	2	94,4	
2012	1	2001	11	0	2	1994	18	18	0	3	88,9	
2012	1	2000	12	0	2	2005	7	18	0	4	83,3	
2012	1	1998	14	0	2	2000	12	18	0	5	77,8	
2012	0	2000	12	0	1	2000	12	18	0	6	72,2	
2012	1	2000	12	0	0	1998	14	18	0	7	66,7	
2012	1	1999	13	0	2	2001	11	18	0	8	61,1	
2012	1	1999	13	0	0	1994	18	18	0	9	55,6	
2012	1	1998	14	0	2	1994	18	18	0	10	50,0	
2012	1	1999	13	0	2	1999	13	18	0	10	50,0	
2012	1	1999	13	0	0	1993	19	18	0	12	38,9	
2012	1	1999	13	0	0	2004	8	18	0	13	33,3	
2012	1	1998	14	0	2	2005	7	18	0	14	27,8	

Рис. 1 Таблица подготовленных результатов соревнований в формате Excel

В данной таблице присутствуют следующие столбцы:

- Год – год проведения соревнования;
- Пол всадника – мужской (0), женский (1);
- Год рождения всадника;
- Возраст всадника – рассчитывается по формуле:

$$\text{Возраст всадника} = \text{Год} - \text{Год рождения всадника}$$
- Разряд – без разряда (0), 1 разряд (1), 2 разряд (2), 3 разряд (3), кандидат в мастера спорта (4), мастер спорта (5), мастер спорта мирового класса (6);
- Пол лошади – жеребец (0), кобыла (1), мерин (2);
- Год рождения лошади;
- Возраст лошади – рассчитывается по формуле:

$$\text{Возраст лошади} = \text{Год} - \text{Год рождения лошади}$$
- Количество участников – количество участников соревнования;
- Уровень – программа соревнований (предварительный приз (0), командный приз (1), личный приз (2), малый приз (3), средний приз (4), большой приз (5));
- Место – место, занятое парой;
- % успеха – рассчитывается по формуле:

$$\% \text{ успеха} = 100 * \frac{(\text{Количество участников} - \text{Место} + 1)}{\text{Количество участников}}$$

Всего за 2011-2019 год получилось 807 записей. Бралась результаты следующих соревнований: Чемпионат и первенство Уральского федерального округа, Чемпионат России, Кубок Победы и "Гордости России. Кубок Отечественного конного завода".

Для реализации нейронных сетей к настоящему времени разработано множество программных пакетов. Наиболее известными являются: MATLAB Deep Learning Toolbox, RapidMiner, NeuroShell 2, Process Advisor, NeuroSolutions и др.

Исходя из функциональных возможностей наиболее оптимальным вариантом является MATLAB Deep Learning Toolbox. Более того, среда MATLAB имеет дружелюбный графический пользовательский интерфейс разработки, исследования и управления процессами в режимах обучения и работы НС.

В процессе обучения сети на её входы подаются входные данные и производится сопоставление значения, полученного на выходе, с целевым (желаемым). На основании результата сравнения (отклонения полученного значения от желаемого) вычисляются величины изменения весов и смещения, уменьшающие это отклонения.

Создаём сеть, состоящую из 6 входов, 1 выхода и одним скрытым слоем. Число нейронов скрытого слоя равно десяти.

Для начала импортируем данные из таблицы Excel в виде двух матриц: `input_data` – входящие данные сети, `output_data` – выходные данные сети. Входными данными будут являться столбцы "Пол всадника", "Возраст всадника", "Разряд", "Пол лошади", "Возраст лошади", "Уровень", выходными данными – "% Успеха".

Создавать сеть будем с помощью "Fitting app" (приложение приближения функций), а тренировать сеть с помощью алгоритма Левенберга-Марквардта. Данный алгоритм предназначен для оптимизации параметров нелинейных регрессионных моделей. Предполагается, что в качестве критерия оптимизации используется среднеквадратичная ошибка модели на обучающей выборке. Алгоритм заключается в последовательном приближении заданных начальных значений параметров к искомому локальному оптимуму.

[3]

В окне процесса обучения сети "Neural Network Training" (Рис. 2) можно наблюдать всю необходимую информацию.

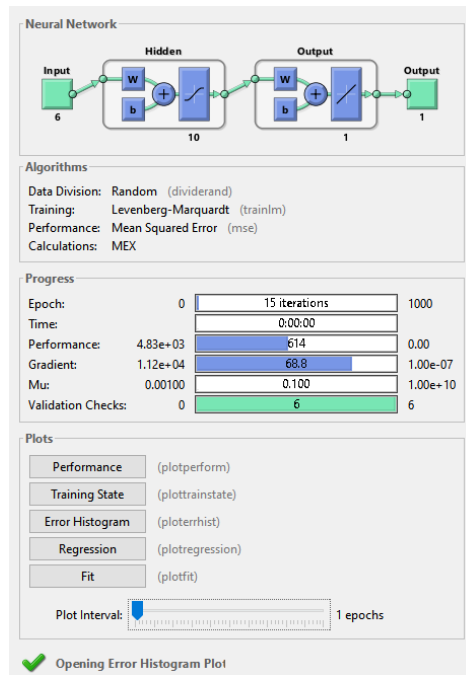


Рис. 2 Окно процесса обучения сети

По кнопке "Performance" можно увидеть график обучения сети (Рис. 3), показывающий поведение ошибки обучения.

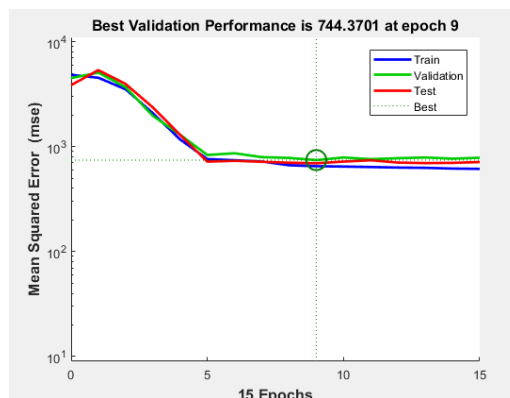


Рис. 3 График обучения сети

Из графика видно, что за 9 эпох достигнуто значение среднеквадратической ошибки равное 744,3701. Эпоха – это однократное представление всех обучающих входных данных на входы сети. Функция обучения использует обучение с ранним остановом (early stopping) в качестве средства борьбы с переобучением. Из графика видно, что обучение остановлено, когда ошибка на проверочном множестве перестала уменьшаться.

Графики состояния обучения – "Training State" показаны на Рис. 4.

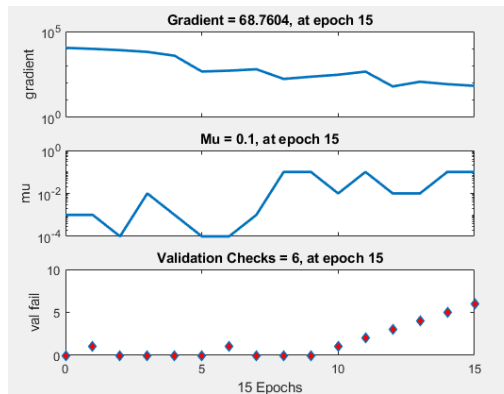


Рис. 4 Графики состояния обучения

На графике "val fail" показано изменение ошибки на контрольном множестве. Видно, что после 9 эпохи ошибка начинает расти. График "gradient" показывает изменение градиента функционала ошибки обучения по весам сети. График "mu" отражает изменение параметра обучения метода Левенбер-га-Марквардта (в других методах отражаются другие параметры). Гистограмма ошибок "Error Histogram" (Рис. 5) показывает, на каком числе примеров (Instances) сеть дает ту или иную погрешность. Погрешность вычисляется как разность между целевым значением и выходом сети. На графике отображаются погрешности для обучающего, проверочного и тестового множеств.

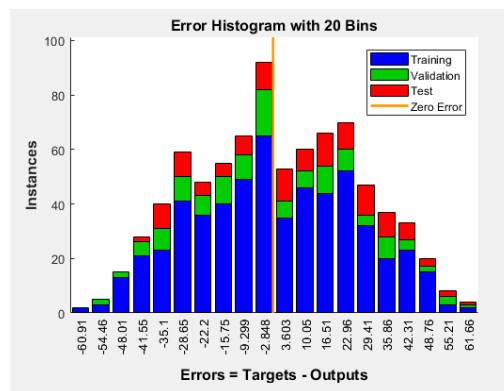


Рис. 5 Гистограмма ошибок

Из гистограммы видно, что большинство ошибок лежит между -2,848 и 3,603.

График "Regression" (Рис. 6) показывает линейную регрессию результатов обучения сети на трех рассмотренных подмножествах и на всех множествах. Для каждого результата рассчитывается коэффициент корреляции, строится график и выводится уравнение регрессии. [4]

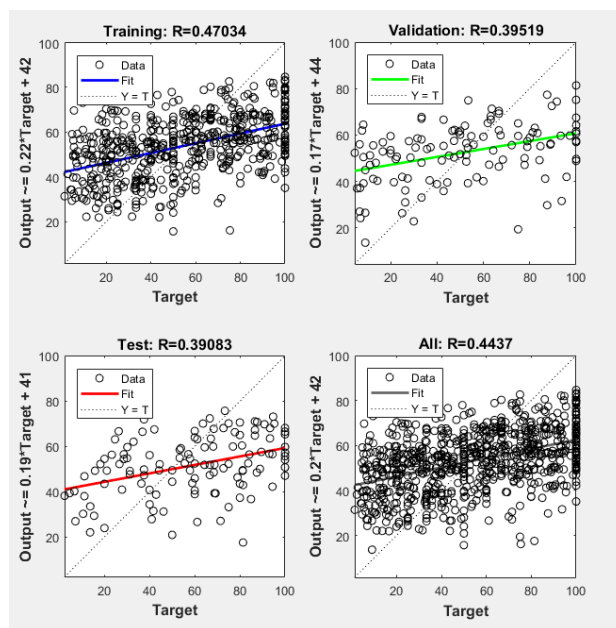


Рис. 6 Линейная регрессия результатов

Построим графики ожидаемых и полученных результатов (Рис. 7).

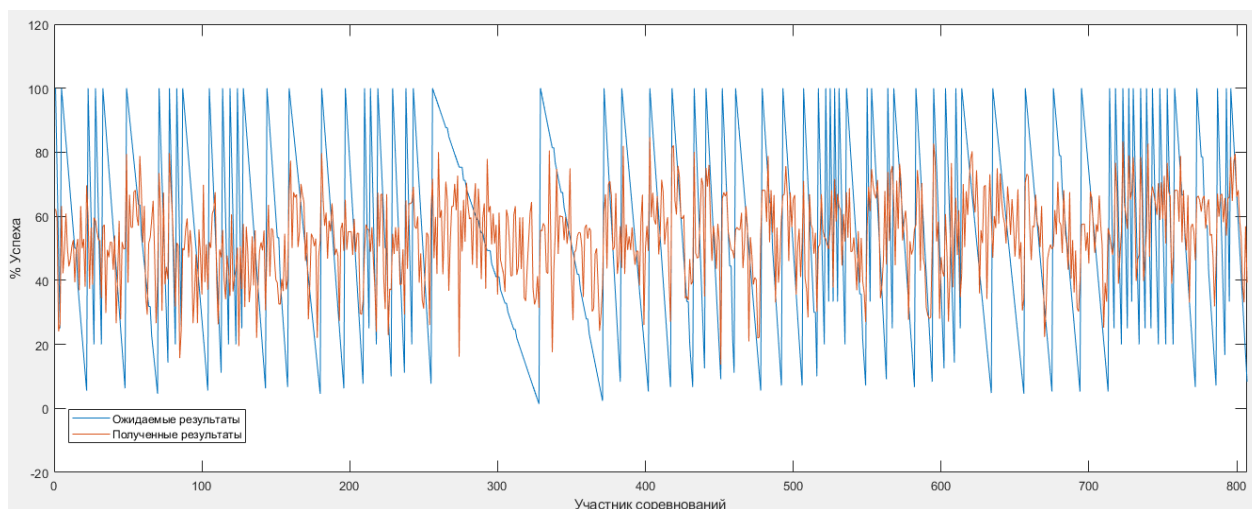


Рис. 7 Графики результатов

Проверим корректность работы сети с помощью команды `sim(<название_сети>, [массив входных данных])`, где название сети – имя созданной сети, а массив входных данных:

- Всадник: пол – женский; возраст – 41 год; разряд – мастер спорта;
- Лошадь: пол – мерин; возраст – 15 лет;
- Уровень соревнований: Средний приз

Результат команды: 79.2810 % успеха.

Теперь введём несколько другой массив данных:

- Всадник: пол – женский; возраст – 20 лет; разряд – 1;
- Лошадь: пол – мерин; возраст – 15 лет;
- Уровень соревнований: Средний приз

Результат команды: 49.4991 % успеха.

Первый результат больше следующего, так как первый всадник старше по возрасту и имеет звание мастера спорта, а следовательно является гораздо более опытным.

Результаты.

Результатами данной работы является обученная нейронная сеть, способная делать прогноз результатов спортсменов по выезде.

Заключение и выводы.

Исходя из полученных и выше изложенных данных можно сделать вывод, что данная сеть работает корректно, но слишком неточно. Следует либо пересмотреть ключевые факторы, влияющие на результат спортсмена, либо пересмотреть настройки сети.

Конный спорт – это очень непредсказуемый вид спорта, так как результат зависит не только от всадника, но и от лошади, поведение которой сложно предсказать.

Список литературы:

1. Баландин, В.И. Прогнозирование в спорте / В.И. Баландин, Ю.М. Блудов, В.А. Плахтиенко – М.: Физкультура и спорт, 1986. – 192 с.
2. Ерохина, О.А. Проблемы спортивного прогнозирования / О.А. Ерохина, Т.В. Сеницына, Макаренко Т. М. – Волжский – 6 с.
3. Алгоритм Левенберга-Марквардта: [Электронный ресурс]. URL: http://www.machinelearning.ru/wiki/index.php?title=Алгоритм_Левенберга-Марквардта (Дата обращения: 25.01.2020)
4. Использование Neural Network Fitting Tool: [Электронный ресурс]. URL: <https://megaobuchalka.ru/5/14821.html> (Дата обращения: 25.01.2020)
5. Equisterian.ru, конное обозрение / Лошади и конный спорт: [Электронный ресурс]. URL: <https://www.equestrian.ru/> (Дата обращения: 20.01.2020)