



УДК 519.237

DOI: 10.31799/2077-5687-2021-2-44-51

ВЛИЯНИЕ СТРУКТУРЫ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ НА ЕЁ ЭФФЕКТИВНОСТЬ НА ПРИМЕРЕ ПРОГНОЗИРОВАНИЯ МАКРОЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ

Б. С. Ермаков

Санкт-Петербургский государственный университет аэрокосмического приборостроения

В статье рассмотрена ситуация того, как структура искусственной нейронной сети может влиять на ее результаты. Исследуется многослойный перцептрон, перед которым стоит задача прогнозирования макроэкономических показателей. Проведены множественные эксперименты с сетями разной структуры – разным количеством скрытых слоев и разным количеством нейронов в этих слоях. По результатам экспериментов сделан вывод о том, что, в целом, до определенного предела увеличение размеров сети улучшает точность прогноза, но на некотором этапе дальнейшее увеличение становится нецелесообразным. Также эксперименты свидетельствуют о том, что проблема переобучения для перцептронов большого размера может быть не так существенна, как для других моделей машинного обучения, например, регрессионных.

Ключевые слова: искусственные нейронные сети, прогнозирование, многослойный перцептрон, переобучение, размер искусственной нейронной сети.

Для цитирования:

Ермаков Б. С. Влияние структуры искусственной нейронной сети на ее эффективность на примере прогнозирования макроэкономических показателей // Системный анализ и логистика: журнал.: выпуск №2(28), ISSN 2077-5687. – СПб.: ГУАП., 2021 – с. 44-51. РИНЦ. DOI: 10.31799/2077-5687-2021-2-44-51.

THE IMPACT OF ARTIFICIAL NEURAL NETWORK'S STRUCTURE ON ITS EFFICIENCY FOR FINANCIAL INDICATORS FORECASTING

B. S. Ermakov

State University of Aerospace Instrumentation

The article investigates the influence of artificial neural network's structure on the results, with example of multilayer perceptron for forecasting some of the financial indicators. Multiple tests were made with various networks structures: different numbers of hidden layers and different numbers of neurons in these layers. Based on tests results, the increase of network's size is effective to a certain extent, but at some point the further size increase is unreasonable. Also, the test results demonstrate that overfitting problem for multilayer perceptron is not as crucial as for the other machine learning models, such as regression.

Key words: artificial neural networks, forecasting, multilayer perceptron, overfitting, artificial neural network's size.

For citation:

Ermakov B. S. The impact of artificial neural network's structure on its efficiency for financial indicators forecasting // System analysis and logistics.: №2(28), ISSN 2077-5687. – Russia, Saint-Petersburg.: SUAI., 2021 – p. 44-51. DOI: 10.31799/2077-5687-2021-2-44-51.

Введение

В 21 веке искусственные нейронные сети (ИНС) получают все более широкое распространение. Это мощный и универсальный инструмент, который может применяться, в частности, для прогнозирования сложных явлений. Одной из популярных архитектур ИНС является многослойный перцептрон, представляющий собой набор слоев, состоящих из нейронов. Количество этих слоев и количество нейронов в каждом слое оказывает существенное влияние на процесс обучения и на конечный результат. На данный момент не существует известного способа однозначно выбрать размер сети, который наилучшим образом соответствовал бы решаемой задаче. В общем случае, чем больше ИНС по размеру, тем больше информации об исследуемом объекте она способна вместить и как следствие – точнее воспроизводить исследуемые закономерности. Однако увеличение размеров сети, во-первых, ведет к увеличению времени, затрачиваемого на её обучение, а во-вторых, чревато т.н. проблемой «переобучения» – ситуации, в которой ИНС слишком точно подстраивается под



обучающую выборку и теряет способность к обобщению для данных за пределами этой выборки.

Целью данной статьи является исследование того, как структура ИНС (количество скрытых слоев и нейронов в них) влияет на точность прогнозов, предоставляемых обученной нейросетью, на примере прогнозирования значений нескольких макроэкономических параметров.

Искусственные нейронные сети и переобучение

Искусственная нейронная сеть – математическая модель, имитирующая поведение биологических нейронных сетей. ИНС состоит из входов, выходов и слоев нейронов между ними. Искусственный нейрон представляет собой функцию от линейной комбинации выходов нейронов предыдущего слоя [1]:

$$\varphi = F(\sum_{i=1}^m w_i x_i + b),$$

где φ – выход сигнал нейрона, x_i – выход i -го нейрона предыдущего слоя, w_i – вес синаптической связи с i -м нейроном предыдущего слоя, b – пороговое значение, m – количество нейронов в предыдущем слое, F – функция активации нейрона, обычно используются сигмоида или гиперболический тангенс.

Если функция активации нелинейная, а сигнал распространяется по направлению от входов к выходам, ИНС называется многослойный перцептрон.

Для обучения используется N обучающих примеров, которые представляют собой множество пар входных данных и соответствующих им выходных данных: $(x_n, y_n), n = 1 \dots N$. Функция ошибки вычисляется как средняя по набору примеров обучающей выборки среднеквадратичная ошибка:

$$E = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^k (o_{ni} - y_{ni})^2,$$

где E – функция ошибки, o_{ni} – фактический i -й выход ИНС для данного входа x_n , y_{ni} – требуемое значение i -го выхода ИНС, k – количество выходов ИНС.

Целью в ходе обучения ИНС является минимизация функции ошибки (обычно с помощью градиентного спуска). Параметры сети – веса связей между нейронами – на каждой итерации изменяются таким образом, чтобы уменьшалось расхождение между выходами ИНС и реальными данными. В результате нейросеть подстраивается под воспроизведение закономерностей, присутствующих в обучающей выборке.

Однако, в случае если объем обучающей выборки недостаточно велик или размер нейросети слишком большой для данной выборки, существует риск «переобучения» сети [2]. Такое явление свойственно многим моделям. Например, в [3] переобучение наглядно проиллюстрировано на примере аппроксимации синусоидальной функции полиномом:

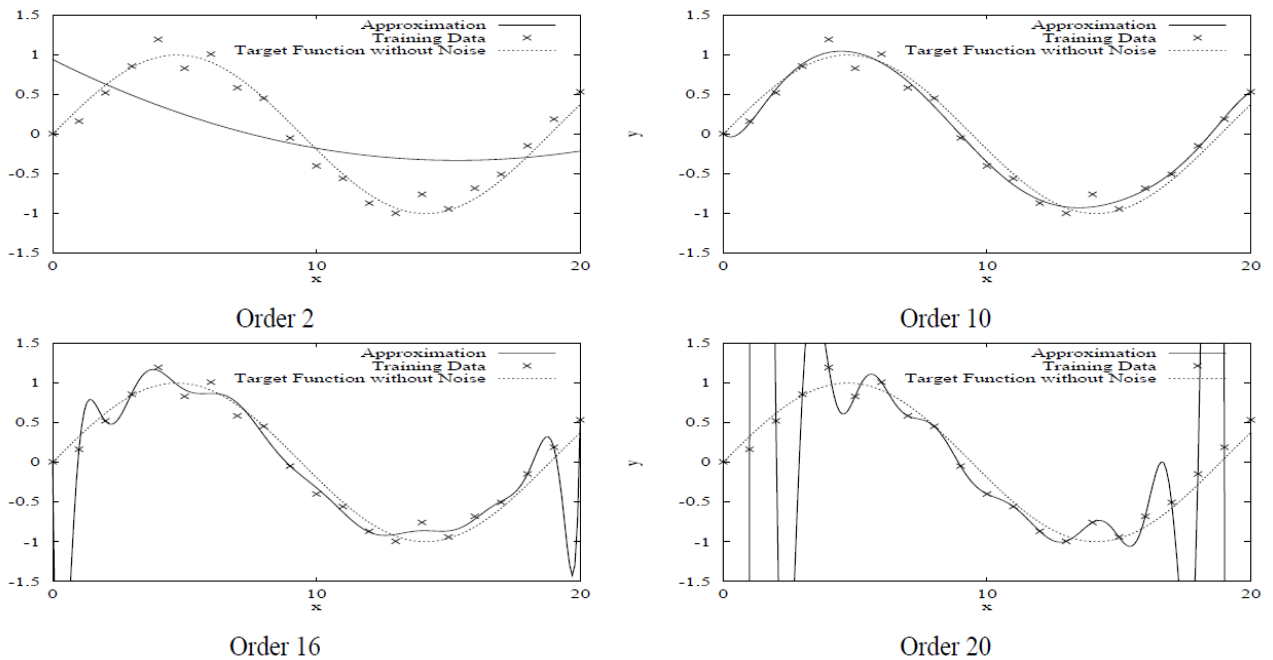


Рис.1. Аппроксимация сигмоидальной функции полиномами различной степени

Полином десятой степени аппроксимирует исходную функцию достаточно хорошо, полином 16-й степени уже имеет существенные отклонения. Полином 20-й степени почти безупречно проходит через все точки исходной выборки, но между ними его поведение не имеет ничего общего с исходной функцией. Такая модель слишком сложна для решаемой задачи, в ней слишком много свободных параметров, как следствие – модель в состоянии с высокой точностью воспроизвести обучающую выборку, но за ее пределами избыточная сложность модели приводит к паттернам, отсутствовавшим в исходных данных – теряется способность к обобщению.

Касательно оптимального размеров нейронной сети нет единого мнения. Традиционный подход заключается в том, что количество свободных параметров (весов связей между нейронами) должно быть в несколько раз меньше количества обучающих примеров. Однако есть и альтернативная точка зрения, согласно которой количество параметров нейросети должно быть примерно таким же по размеру как обучающая выборка, как, например, показано в [4].

Для решения проблемы переобучения в контексте нейросетей были предложены различные методы, среди них следующие [2]:

- ранняя остановка – в ходе обучения нейросети и по мере убывания ошибки на обучающей выборке, ее эффективность также оценивается по тестовой выборке, и как только ошибка на ней начнет расти обучение останавливается;
- сокращение сети – во время или после обучения из модели удаляются избыточные элементы;
- расширение обучающих данных – помимо очевидного добавления дополнительных обучающих примеров (которое не всегда является возможным), в некоторых случаях также добавляют новые данные сгенерированные на основе имеющихся;
- регуляризация – уменьшение значений параметров модели: даже если их слишком много, при отсутствии параметров со слишком большими (по модулю) значениями уменьшается вероятность существенных отклонений от реальных данных.

Переобучение является значительной проблемой, однако в данной статье экспериментально будет показано, что в случае нейронных сетей, в отличие от других моделей машинного обучения, переобучение может не оказывать существенного негативного влияния на результат.



Эксперименты по прогнозированию

Чтобы оценить влияние структуры и размеров нейросети на ее эффективность, была проведена серия экспериментов по прогнозированию макроэкономических параметров. Прогнозировались значения курса доллара к рублю, стоимость барреля нефти марки Brent и индекс Московской биржи. На вход сети подавались средние значения по каждому из этих параметров за прошлые 2, 4, 6 и 10 дней. На выходе сеть предоставляла прогноз по этим трем параметрам на день вперед. Использовались дневные значения показателей за 10 лет: с января 2010 года по декабрь 2019 года (около 2.5 тысяч значений).

В ходе экспериментов проводились серии по 100 испытаний для каждой из конфигураций ИНС: от 1 до 3 скрытых слоев, от 10 до 1500 нейронов в слое. Нейросеть была реализована на языке Python с использованием библиотеки Scikit-learn.

Нейросеть имела архитектуру многослойный перцептрон, функцией активации был гиперболический тангенс, метод обучения – модификация градиентного спуска предложенная в [5], размер набора для стохастического градиентного спуска – 200 обучающих примеров. В нейросети использовалась L2 регуляризация [6] со стандартным значением параметра 0.0001. Точность в каждом серии экспериментов рассчитывалась как средняя точность по каждому из 200 примеров тестовой выборки (остальные примеры использовались для обучения).

Результаты испытаний приведены в таблице 1: второй столбец представляет собой количества нейронов в скрытых слоях ИНС, третий – среднюю точность прогноза по 100 испытаниям для данной структуры ИНС, четвертый – среднее время ушедшее на обучение ИНС.

Таблица 1 – Результаты экспериментов

№	ИНС	Точность, %	Время, с
1	[10]	92.57	0.2
2	[20]	94.62	0.2
3	[30]	95.47	0.2
4	[50]	96.58	0.2
5	[100]	97.86	0.3
6	[200]	98.56	0.5
7	[500]	98.89	1.0
8	[1000]	98.99	2.6
9	[1500]	99.00	4.7
10	[10, 5]	91.59	0.4
11	[20, 10]	94.44	0.4
12	[50, 20]	97.01	0.4
13	[50, 50]	97.98	0.7
14	[100, 50]	98.32	0.9
15	[100, 100]	98.57	1.0
16	[200, 50]	98.67	1.1
17	[200, 100]	98.77	1.4
18	[200, 200]	98.82	1.8
19	[500, 100]	98.93	2.6
20	[500, 200]	99.03	3.7
21	[500, 500]	99.07	8.5
22	[1000, 100]	99.02	4.9
23	[1000, 500]	99.11	18.7
24	[10, 10, 10]	92.56	0.3



25	[50, 30, 10]	96.85	0.6
26	[50, 50, 50]	98.13	0.8
27	[100, 50, 20]	98.19	0.8
28	[100, 100, 100]	98.55	1.1

Далее представлены графики зависимости точности прогноза от структуры ИНС.

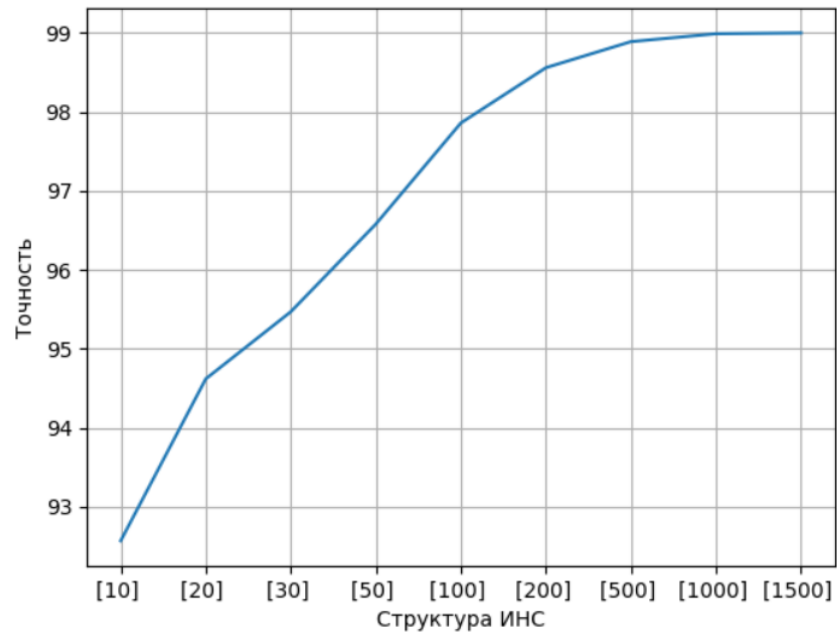


Рис. 2. Точность прогноза ИНС с одним скрытым слоем

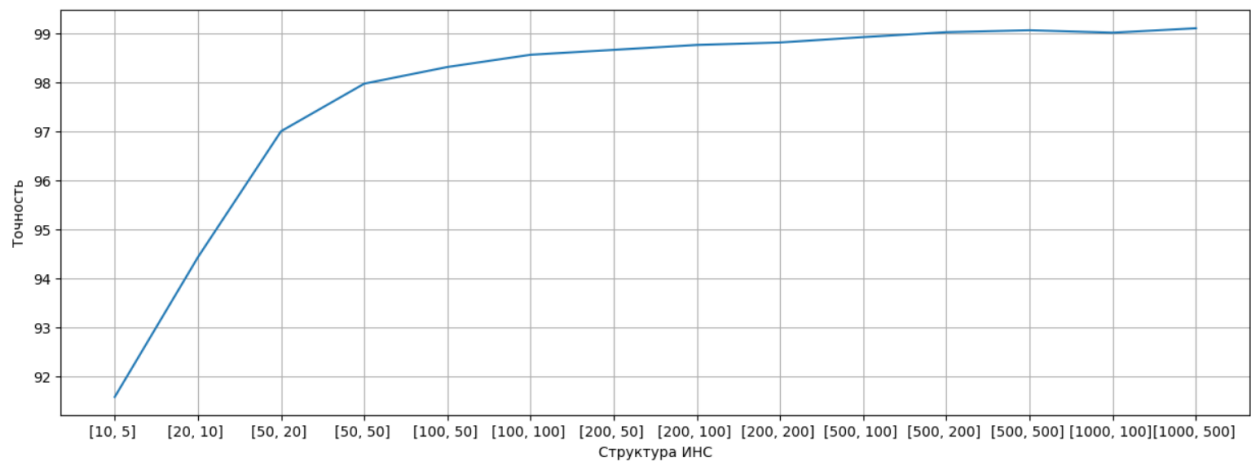


Рис. 3. Точность прогноза ИНС с двумя скрытыми слоями

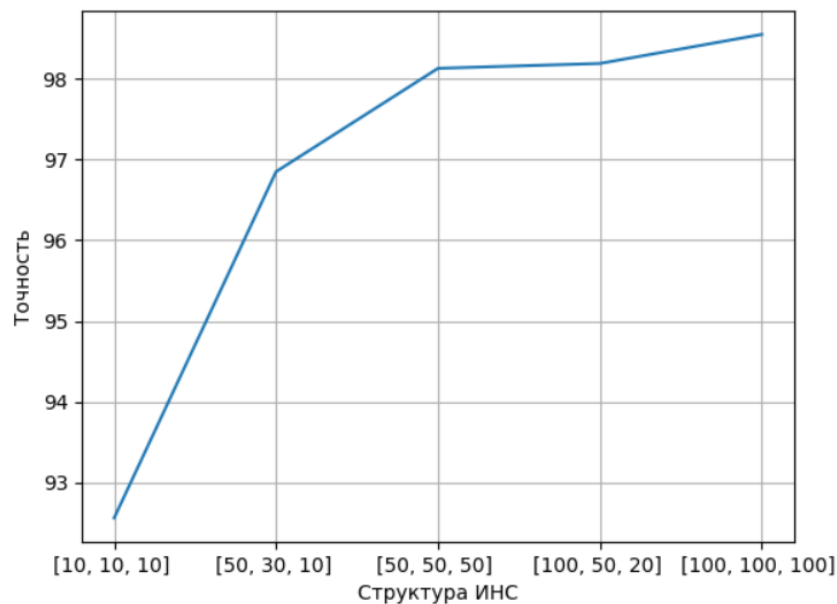


Рис. 4. Точность прогноза ИНС с тремя скрытыми слоями

По таблице и графикам видно, что по мере увеличения числа нейронов в слоях точность прогноза растет, но растет с замедлением. Время, требуемое на обучение ИНС, также увеличивается с увеличением размерности, но также нелинейно. Это связано с особенностями библиотеки NumPy (используемой в библиотеке Scikit-learn), которая использует векторные и матричные операции для ускорения вычислений [7].

Оценка результатов

Хотя точность прогноза и растёт с увеличением размеров ИНС, после определенного предела прирост точности незначителен. Например, если сравнить испытания 18 и 19, то видно, что в данном случае увеличение количества нейронов в 1,5 раза дало прирост точности всего в 0,11%, аналогичная ситуация наблюдается для остальных нейросетей большого размера.

Также увеличение количества скрытых слоев не всегда ведет к сколько-нибудь заметному улучшению точности прогноза. Например, сеть с одним скрытым слоем в 200 нейронов дает точность всего на 0.1% меньшую, чем сеть с двумя слоями в 200 и 50 нейронов. Еще более показательное сравнение испытаний 15 и 28: сеть с двумя слоями по 100 нейронов дает точность на 0.02% большую чем сеть с тремя слоями по 100 нейронов! Отсюда следует вывод, что в зависимости от поставленной задачи на определенном этапе дальнейшее усложнение нейросети неоправданно в силу того, что получаемый прирост эффективности слишком мал по сравнению с увеличением временных затрат на вычисления: в рассматриваемой задаче даже относительно большие конфигурации ИНС требуют лишь несколько секунд на обучение, но в некоторых случаях размер обучающей выборки и количество нейронов может измеряться миллионами – время, требуемое на обучение нейросети в таком случае становится критически важным. По-видимому, в контексте данной задачи структура сети с двумя скрытыми слоями является наиболее подходящей.

Тем не менее, в целом, прослеживается тенденция к увеличению точности прогноза нейросети по мере увеличения ее размеров. В связи с этим встает вопрос о потенциальном переобучении нейросети. Считается, что в результате переобучения точность результатов растет на обучающей выборке, но после определенного этапа начинает падать на тестовой выборке (примерах, которые не использовались в обучении). В данном случае точность различных нейросетей оценивалась именно на тестовой выборке, поэтому проблема переобучения не должна была повлиять на результат. Можно предположить, что выборка из 200 тестовых примеров недостаточно велика и не позволяет



обнаружить неспособность нейросети к обобщению, вызванную переобучением. Однако в результате увеличения тестовой выборки в два раза (и соответствующего уменьшения обучающей выборки на 200 примеров) средняя точность прогноза упала всего на 0,7%, что вполне оправданно с учетом того, что количество материала для обучения нейросети стало меньше, а количество неизвестных ей примеров, используемых для тестирования, стало больше.

На основании этих наблюдений можно сделать вывод о том, что в рассматриваемой задаче эффект переобучения если и присутствовал в больших конфигурациях ИНС, то был незначителен и заметного негативного эффекта не оказал. Это хорошо согласуется с выводами статьи [3], в которой показано, что нейронные сети, обучаемые методом обратного распространения ошибки, аппроксимируют сложные функции намного лучше, чем другие методы, при этом даже при значительном увеличении размеров сети не наблюдается ухудшения качества, вызванного переобучением. Примечательно, что в этой статье также изучалась работа нейросетей, в обучении которых использовался метод сопряженных градиентов, и в этом случае переобучение проявлялось в полной мере. По-видимому, это в т.ч. специфика метода обратного распространения ошибки – обученные им нейросети даже при достаточно большом размере сохраняют способность к обобщению и не подвержены переобучению.

В статье [8] также показано, что малые веса нейронных связей намного важнее для решения проблемы переобучения, чем сам размер сети. В связи с этим использование регуляризации при обучении нейросети становится критически важным. Как было указано ранее, в проведенных экспериментах регуляризация применялась. Тот факт, что эффекты переобучения не проявились, свидетельствует в пользу озвученной точки зрения: регуляризация весов помогает избежать переобучения даже для сетей большого размера.

Можно сделать вывод, что нейросетям при использовании регуляризации и корректного алгоритма обучения свойственно избегать переобучения, в отличие от классических моделей типа регрессии. Также результаты, которые дают нейронные сети, во многом превосходят результаты, получаемые классическими методами, как, например, показано в [9]. И если в статье [4] показано, что нейронные сети большого размера хорошо справляются с задачами классификации, а в статье [3] – с задачами интерполяции, в данной статье нейронные сети показали высокую точность в задаче экстраполяции, без проявления эффектов переобучения, что говорит о том, что применение ИНС для прогнозирования уместно по крайней мере в некоторых случаях.

Заключение

По результатам экспериментов видно, что увеличение размеров нейронной сети в целом улучшает точность прогноза. Однако на определенном этапе дальнейшее увеличение может стать нецелесообразным, т.к. прирост эффективности незначителен в сравнении с увеличением временных затрат. Эффектов переобучения не выявлено даже для сетей большого размера. По-видимому, использование корректного метода обучения и регуляризации весов помогает избежать переобучения.

СПИСОК ЛИТЕРАТУРЫ

1. *Хайкин С.* Нейронные сети: полный курс, 2-е издание / Хайкин С. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
2. Xue Ying, An Overview of Overfitting and its Solutions / Xue Ying // Journal of Physics: Conference Series. – 2019.
3. *Lawrence S.* Overfitting and Neural Networks: Conjugate Gradient and Backpropagation / Lawrence S. // International Joint Conference on Neural Networks. – 2000. – pp. 114–119
4. *Sartori M.A.* A simple method to derive bounds on the size and to train multilayer neural networks



- / Sartori M.A., Antsaklis P.J. // IEEE Transactions on Neural Networks. – 1991. – Vol. 2. – P. 467 – 471.
5. *Kingma D.P.* Adam: A Method for Stochastic Optimization / Kingma D.P., Jimmy Ba // International Conference on Learning Representations. – 2015.
 6. Документация библиотеки Scikit-learn [Электронный ресурс] – URL: <https://bit.ly/2S22vXq> (дата обращения: 14.05.2021).
 7. Документация библиотеки NumPy [Электронный ресурс] – URL: <https://bit.ly/2SSrIt3> (дата обращения: 14.05.2021).
 8. *Bartlett P.L.* For valid generalization, the size of the weights is more important than the size of the network / Bartlett P.L. // Conference on Neural Information Processing Systems. – 1996. – p. 134-140
 9. *Menhazul A.* Comparative Study of Artificial Neural Network Regression and Classical Regression / Menhazul A., Maniruzzaman, Shaykhul I., Benojir A., Mohammad A., Faisal A. // Proceedings in International Conference on Bioinformatics and Biostatistics for Agriculture, Health and Environment. – 2017.

ИНФОРМАЦИЯ ОБ АВТОРЕ

Ермаков Булат Сергеевич —
аспирант

Санкт-Петербургский государственный университет аэрокосмического приборостроения
190000, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А
E-mail: ermakovb@yandex.ru

INFORMATION ABOUT THE AUTHOR

Ermakov Bulat Sergeevich —
postgraduate

Saint-Petersburg State University of Aerospace Instrumentation
67, Bolshaya Morskaya str., Saint-Petersburg, 190000, Russia
E-mail: ermakovb@yandex.ru