

УДК 621.311.238:004.032.26

ТОНКАЯ НАСТРОЙКА НЕЙРОННОЙ СЕТИ В ЗАДАЧАХ ПОЛУЧЕНИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ГАЗОТУРБИННОЙ ЭЛЕКТРОСТАНЦИИ**Килин Г.А., Кавалеров Б.В., Шулаков Н.В., Ждановский Е.О.***ФГБОУ ВО «Пермский национальный исследовательский политехнический университет», Пермь, e-mail: thisisforasm@rambler.ru, kbv@pstu.ru, shulakov@pstu.ru, zhdanovskiy.e@gmail.com*

Искусственные нейронные сети широко применяются в различных областях деятельности человека. Одним из наиболее распространенных алгоритмов обучения нейронной сети является алгоритм обратного распространения ошибки. Во время работы алгоритма обратного распространения ошибки происходит обновление весовых коэффициентов искусственной нейронной сети на каждой итерации обучения. Обновление весов происходит, начиная с выходного слоя по направлению к входному слою искусственной нейронной сети. В случае большого количества скрытых слоев может наблюдаться как эффект взрывного роста весов, так и, в некоторых случаях, эффект затухания градиента. В случае возникновения таких эффектов процесс обучения становится значительно более ресурсоемким и занимает больше времени. Одним из выходов возникающей ситуации является тонкая настройка весов нейронной сети. Тонкая настройка позволяет устанавливать весовые коэффициенты нейронной сети ближе к некоторому минимуму, чтобы процесс работы алгоритма обратного распространения ошибки занял меньше времени. Основная идея тонкой настройки заключается в последовательном и послойном обучении скрытых слоев и дальнейшем формировании из этих слоев результирующей нейронной сети. В статье приводится алгоритм тонкой настройки искусственной нейронной сети в виде его подробного описания и блок-схемы, а также сравнение обучения нейронной сети с использованием данного метода, а также без него.

Ключевые слова: искусственная нейронная сеть, газотурбинная электростанция, алгоритм обучения, перцептрон, тонкая настройка

NEURAL NETWORK FINE CONFIGURATION IN THE PROBLEMS OF RECEIVING A MATHEMATICAL MODEL OF A GAS-TURBINE ELECTRIC POWER STATION**Kilin G.A., Kavalеров B.V., Shulakov N.V., Zhdanovskiy E.O.***Federal State Educational Institution of Higher Education Perm National Research Polytechnic University, Perm, e-mail: thisisforasm@rambler.ru, kbv@pstu.ru, shulakov@pstu.ru, zhdanovskiy.e@gmail.com*

Artificial neural networks are widely used in various fields of human activity. One of the most common algorithms for learning a neural network is the back-propagation algorithm. During the operation of the back-propagation algorithm, the weights of the artificial neural network are updated at each training iteration. The weights are updated starting from the output layer towards the input layer of the artificial neural network. In the case of a large number of hidden layers, both the effect of explosive growth of weights, and, in some cases, the effect of attenuation of the gradient can be observed. In the event of such effects, the learning process becomes much more resource-intensive and takes more time. One of the ways out of the situation that arises is to fine-tune the weights of the neural network. Tweaking allows the neural network weights to be set closer to a certain minimum, so that the process of the operation of the back propagation error algorithm takes less time. The main idea of fine-tuning is the sequential and layer-by-layer learning of hidden layers and the further formation of the resulting neural network from these layers. The article provides an algorithm for fine-tuning an artificial neural network in the form of its detailed description and block diagram, as well as comparing the training of a neural network using this method, as well as without it.

Keywords: artificial neural network, gas turbine power station, learning algorithm, perceptron, fine tuning

Искусственные нейронные [1, 2] сети (ИНС) являются вычислительными структурами, состоящими из определенного количества однотипных элементов, соединённых между собой, названными искусственными нейронами. Каждый нейрон выполняет относительно простые функции, а именно получает и пересылает сигнал другим элементам данной сети. Часто данные процессы сравнивают с процессами, происходящими в нейронных сетях живых организмов, но искусственные нейронные сети имеют более простую структуру. В качестве примера для научных исследований был выбран многослойный перцептрон [1, 2] (рис. 1).

Нейронные сети находят своё применение в различных сферах деятельности

человека. В наше время нейронные сети используются для распознавания изображений [3, 4], получили применение в финансовой сфере [5]. Благодаря способности аппроксимировать нелинейность любого вида [1, 2] ИНС используют в различных отраслях промышленности. Например, ИНС используются в задачах идентификации и получения математических моделей газотурбинных установок (ГТУ) [6, 7], газотурбинных электростанций (ГТЭС) [8, 9]. Такие нейросетевые модели в дальнейшем используют в качестве объекта управления для настройки параметров регулятора системы автоматического управления (САУ). Кроме того, ИНС также используются в качестве системы диагностики газотурбинного двигателя [10].

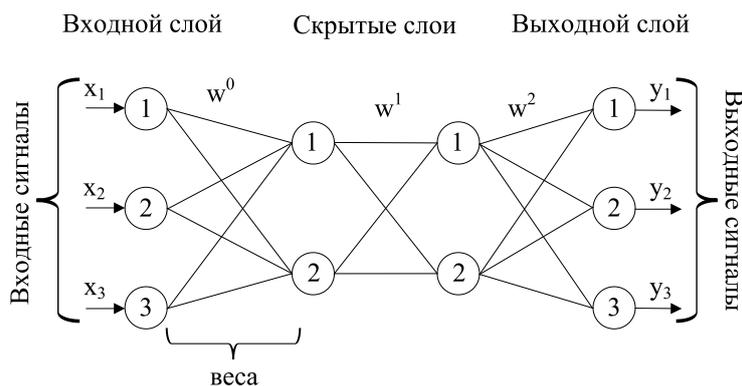


Рис. 1. Структура многослойного персептрона

Однако необходимо учесть, что чем сложнее поставлена задача перед исследователем, тем сложнее проходит процесс обучения нейронной сети. ИНС в основном разделяют на сети, которые обучаются с учителем, и те, которые обучаются без учителя [1, 2]. В данной статье будет рассматриваться метод обучения с учителем, использующий алгоритм обратного распространения ошибки [1, 2] для обучения полученной ИНС. Данный алгоритм является самым распространённым способом обучения нейронной сети, однако у данного метода есть ряд нюансов, как правило, этим алгоритмом обучают относительно не большие сети, с одним или несколькими скрытыми слоями, поскольку при увеличении структуры есть вероятность возникновения затухания, либо взрывного роста весов [11, 12]. Взрывной рост может произойти, если веса слишком большие, либо значение производной в точке слишком велико, а затухание происходит, если значение весов или производной в точке очень мало. Особенно такая ситуация характерна в случае рекуррентных искусственных нейронных сетей, так как такие сети содержат обратные связи [1, 12].

Тонкая настройка ИНС

В ходе работы алгоритма обратного распространения ошибки происходит обновление весовых коэффициентов на каждой итерации обучения, начиная с выходного слоя по направлению к входному слою ИНС. В случае большого количества скрытых слоев наблюдается эффект затухания градиента [11, 12], то есть чем ближе к входному слою, тем меньше изменение весового коэффициента в ходе алгоритма обучения. Тем самым процесс обучения замедляется. Одним из способов решения возникшей проблемы является тонкая настройка весов (предобучение) нейронной сети [13, 14].

Предположим, для получения модели ГТЭС мы выбрали сеть прямого распространения следующей структуры: $x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_n$. Каждый x_i – это количество нейронов в слое, x_n – выходной слой. Обозначим за x_0 размерность входного вектора, который подается на вход слоя x_1 . Так же у нас есть массив данных для обучения D_0 – это пары вида «вход, ожидаемый выход», и мы хотим обучить сеть, используя алгоритм обратного распространения ошибки. Но перед этим осуществим тонкую настройку весов каждого скрытого слоя по алгоритму, представленному на рис. 2.



Рис. 2. Алгоритм тонкой настройки весов скрытого слоя

Визуализация данного алгоритма (рис. 2) представлена на рис. 3.

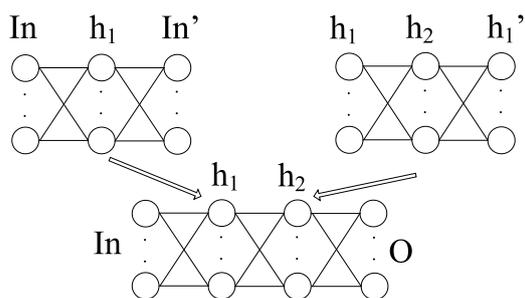


Рис. 3. Визуализация алгоритма тонкой настройки весов скрытого слоя

Где In – входной слой нейронной сети; h_1 – первый скрытый слой нейронной сети; h_2 – второй скрытый слой нейронной сети; O – выходной слой нейронной сети.

Для тестирования была взята архитектура нейронной сети без обратных связей, с двумя скрытыми слоями, по 10 нейронов в каждом скрытом слое.

Каждый из скрытых слоев ИНС был обучен по алгоритму (рис. 2), и из этих слоев была сформирована результирующая ИНС (рис. 3). В итоге были получены следующие результаты (рис. 4–5).

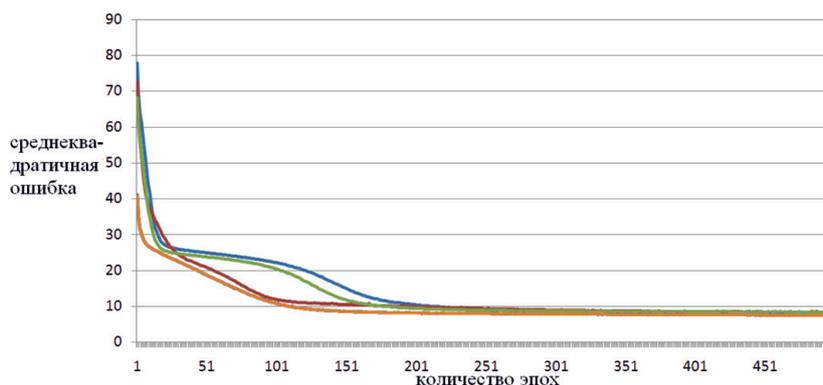


Рис. 4. Изменение ошибки по эпохам обучения для обучающей выборки (синяя, зеленая, красная – нет тонкой настройки; коричневая, голубая, фиолетовая – есть тонкая настройка)

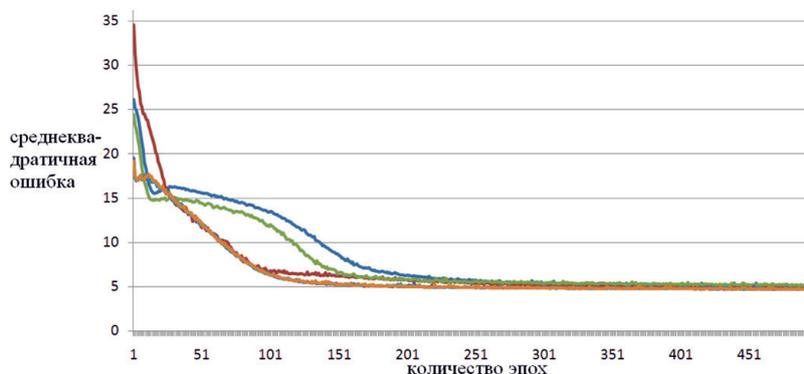


Рис. 5. Изменение ошибки по эпохам обучения для тестовой выборки (синяя, зеленая, красная – нет тонкой настройки; коричневая, голубая, фиолетовая – есть тонкая настройка)

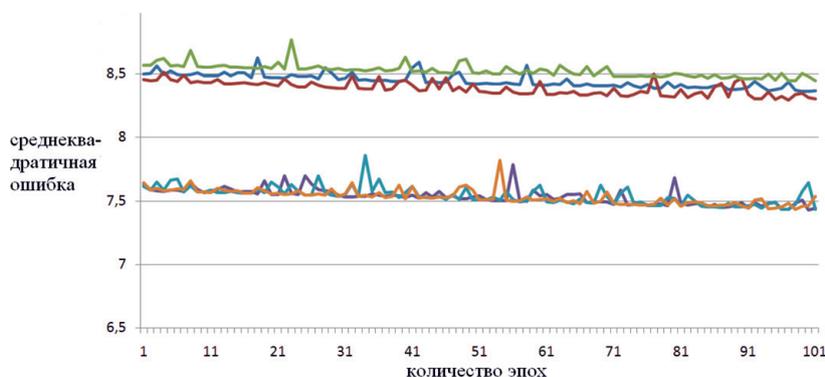


Рис. 6. Изменение ошибки последних 100 эпох обучения для обучающей выборки (синяя, зеленая, красная – нет тонкой настройки; коричневая, голубая, фиолетовая – есть тонкая настройка)

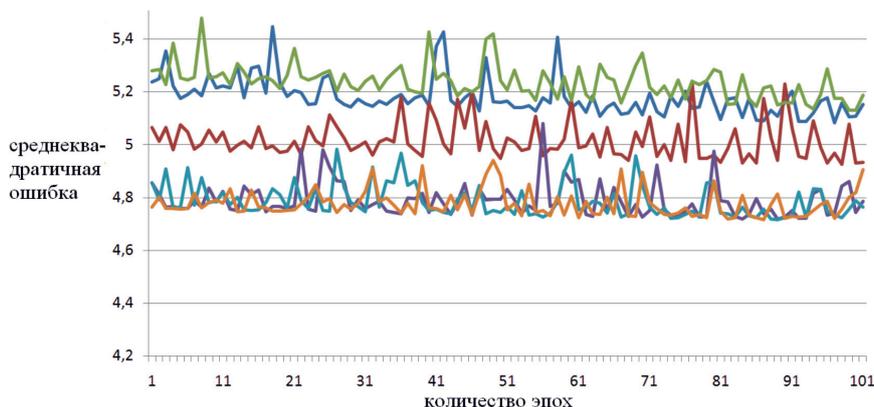


Рис. 7. Изменение ошибки последних 100 эпох обучения для тестовой выборки (синяя, зеленая, красная – нет тонкой настройки; коричневая, голубая, фиолетовая – есть тонкая настройка)

По рис. 4 и 5 наглядно видно, что в случае тонкой настройки весовых коэффициентов (предобучения) ИНС уменьшение ошибки происходит значительно интенсивнее. Рассмотрим последние 100 эпох обучения (рис. 6 и 7).

Заключение

Объяснение полученным результатам можно дать следующее: при обучении первого скрытого слоя ИНС создается модель, которая по экспериментальным данным, подаваемым на вход ИНС, генерирует некоторые скрытые признаки, то есть весовые коэффициенты ИНС сразу помещаются в некоторый минимум, необходимый для вычисления этих скрытых признаков. В дальнейшем, с каждым последующим обучением скрытых слоев ИНС, вычисляются признаки признаков, а весовые коэффициенты ИНС всегда помещаются в состояние, достаточное для вычисления этих иерархических признаков. Уже когда дело доходит до алгоритма обучения с учителем, по сути, эффективно обучаться будут только 2–3 слоя от выхода, на основании тех гиперпризнаков, что были вычислены раньше, а те, в свою очередь, будут незначительно меняться в угоду решаемой задачи.

Стоит отметить, что для проверки использовалась очень простая архитектура нейронной сети и небольшое количество экспериментальных данных, а также эпох обучения. В случае решения задачи с большим количеством экспериментальных и большого количества скрытых слоев ИНС, а также большого количества эпох обучения, разница между предварительно настроенной нейронной сетью и обычным способом инициализированной сетью должна оказаться значительно больше.

Исследование выполнено при финансовой поддержке РФФИ и Пермского края в рамках научного проекта № 19-48-590012.

Список литературы

1. Хайкин С. Нейронные сети: полный курс, 2-е изд. М.: Издательский дом Вильямс, 2008. 1104 с.
2. Друки А.А. Применение сверточных нейронных сетей для выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном // Известия Томского политехнического университета. 2014. Т. 324. № 5. С. 85–92.
3. Солдатова О.П., Гаршин А.А. Применение сверточной нейронной сети для распознавания рукописных цифр // Компьютерная оптика. 2010. Т. 34. № 2. С. 251–259.
4. Дорогой Я.Ю. Архитектура обобщенных сверточных нейронных сетей // Вестник Национального технического университета Украины. 2011. № 54. С. 229–234.
5. Ефремова Е.А., Дунаев Е.В. Применение нейронных сетей для прогнозирования финансовых временных рядов // Автоматизированные системы обработки информации, управления и проектирования // Доклады ТУСУРа. 2004. С. 192–196.
6. Asgari H. et al. Artificial neural network-based system identification for a single-shaft gas turbine. *Journal of Engineering for Gas Turbines and Power*. 2013. Т. 135. № 9. P. 092601–092607.
7. Килин Г.А., Бахирев И.В., Кавалеров Б.В. Получение нелинейной модели ГТУ на основе нейронной сети // Автоматизация в электроэнергетике и электротехнике. 2015. Т. 1. С. 72–77.
8. Ждановский Е.О., Кавалеров Б.В., Килин Г.А. Разработка нейросетевой модели газотурбинной электростанции для настройки регуляторов газотурбинной установки // Фундаментальные исследования. 2017. № 12–3. С. 479–485.
9. Килин Г.А., Кавалеров Б.В. Разработка математической модели газотурбинной электростанции на основе технологии нейронных сетей // Климовские чтения – 2016: перспективные направления развития двигателестроения. 2016. С. 229–233.
10. Легконогих Д.С. Применение нейросетевых технологий в системах диагностики авиационных силовых установок // Известия Самарского научного центра Российской академии наук. 2012. Т. 14. № 4(2). С. 639–643.
11. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks. *Intern. conference on artificial intelligence and statistics*. 2010. P. 249–256.
12. Pascanu R., Mikolov T., Bengio Y. Understanding the exploding gradient problem. *Tech. Rep. Montreal, Universite de Montreal Publ.* 2012. P. 11.
13. Erhan D. et al. Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research*. 2010. Т. 11. № Feb. P. 625–660.
14. Hinton G.E., Salakhutdinov R.R. Reducing the dimensionality of data with neural networks. *Science*. 2006. Т. 313. № 5786. P. 504–507.