

УДК 004.8

А. Н. Иванов, С. А. Мустафина, Н. Д. Морозкин

АНАЛИТИЧЕСКИЙ ОБЗОР ИСТОРИИ РАЗВИТИЯ НЕЙРОСЕТЕВЫХ АРХИТЕКТУР И АЛГОРИТМОВ

Ключевые слова: нейронные сети, нейросети, алгоритмы обучения, машинное обучение.

В статье рассматривается история возникновения искусственных нейронных сетей – математических моделей нервной системы живых организмов. Авторами рассматриваются биологические нейроны как прототипы для разработки моделей искусственных нейронов, их строение и функционирование. Описываются попытки реализации искусственных нейронных сетей и нейрокомпьютеров на их основе с начала XX века. Приводится математический аппарат простейшего искусственного нейрона: его сумматорные и активационные функции. Обозначены основные требования, предъявляемые к активационным функциям с целью дальнейшего эффективного обучения сети. В хронологическом порядке перечислены основные достижения исследователей, направленные на совершенствование старых и разработку новых архитектур нейронных сетей, среди которых: перцептрон Розенблатта, многослойный перцептрон, когнитрон и неокогнитрон, нейронная сеть Хопфилда и карта Кохонена, нейронная машина Больцмана, сети адаптивного резонанса (ART), рекуррентные, сверточные и развертывающиеся нейронные сети, генеративные состязательные сети (GAN), глубокие сверточные и остаточные сети. Также приводятся математические формулы и алгоритмы, лежащие в основе различных методов машинного обучения нейросетей: метода обратного распространения ошибки (backpropagation), встречного распространения ошибки, Больцмановского обучения, обучения Коши, алгоритма глубокого обучения. Описываются преимущества и области применения для каждой конфигурации искусственных нейронных сетей.

A. N. Ivanov, S. A. Mustafina, N. D. Morozkin

ANALYTICAL REVIEW OF THE HISTORY OF NEURAL NETWORK ARCHITECTURES AND ALGORITHMS

Keywords: neural networks, NNs, training algorithms, machine learning.

This paper describes the history of artificial neural networks, i. e., mathematical models of nervous system in living organisms. The authors consider biological neurons as prototypes for developing the models of artificial neurons, as well as their structure and functioning. We also describe the attempts to implement artificial neural networks and neural computers based on them, which have been made since the beginning of the 20th century. This paper also describes the mathematical tools for the simplest artificial neuron with its summatory and activation functions. It also outlines the core requirements set for activation functions in order to further efficiently train the network. We chronologically arranged the main achievements of researchers, aimed at improving the existing neural network architectures and developing the new ones, such as Rosenblatt's perceptron, multi-layer perceptron, cognitron and neocognitron, Hopfield's neural network and Kohonen map, Boltzmann's neural machine, adaptive resonance networks (ART), recurrent, convolutional and deconvolutional neural networks, generative adversarial networks (GAN), and deep convolutional and residual networks. The paper also provides mathematical formulas and algorithms that underlie various methods of neural network machine training, such as back-propagation, counter-propagation, Boltzmann training, Cauchy training, and deep training algorithms. Advantages and applications are described regarding each configuration of artificial neural networks.

История развития и создания нейронных сетей тесно связана с достижениями в области изучения их прообраза – биологических нейронов живых организмов [1]. Несмотря на то, что до сих пор не удалось подробно изучить работу головного мозга человека и нервной системы, нейрофизиология достигла значительных успехов в изучении принципов функционирования отдельных нейронов [2-4].

Биологический нейрон – это структурно-функциональная единица нервной системы живых организмов [5]. Он представляет собой клетку мозгового вещества, которая способна реагировать на оказываемые на неё воздействия. Эти воздействия передаются посредством биологических сигналов, которые нейрон способен трансформировать (обрабатывать, ослабевать, усиливать и передавать другим нейронам) [6]. Несмотря на то, что нейроны все же имеют незначительные различия в зависимости

от их специфики и месторасположения [7, 8], следует отметить их общее строение (рис.1).

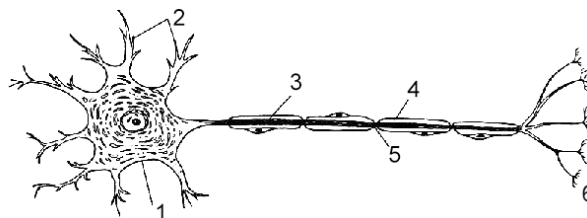


Рис. 1 – Строение нейрона: 1 – тело нейрона, 2 – дендриты, 3 – аксон, 4 – миелиновая оболочка, 5 – перехват Ранвье, 6 – аксонные терминалы

Клеточная мембрана нейрона имеет сложную форму и включает большое количество отходящих в разные стороны отростков, которые называются дендритами (2). Количество концевых участков дендритов может варьироваться в интервале от 5 до

10000 [9]. Через данные отростки нейрон способен получать электрические сигналы от других аналогичных нейронов. Таким образом, дендриты условно можно называть информационными входами нейрона. При этом на поверхности клетки присутствует выступ, называющийся аксонным холмиком, из которого выходит единственный и самый длинный отросток – аксон (3). Электрические возмущения по данному отростку передаются исключительно в направлении от тела клетки (1) к периферии. Таким образом, аксон является информационным выходом нейрона. Строгая направленность передачи сигналов обусловлена особенностью их распространения, а именно за счет последовательного возмущения мембраны клетки и образования с обеих сторон от неё разности электрического потенциала. Наличие инерционности в процессах, отвечающих за передачу биоэлектрических импульсов, приводит к невозможности срабатывания нейрона дважды за малый промежуток времени.

На конце аксон обычно разветвляется и соединяется с дендритами других нейронов. Такие соединения называются аксонными терминалями или синапсами (6). Передача электрического импульса от аксона одного нейрона к дендриту другого осуществляется за счет выброса специальных биологически активных веществ (интермедиатов) [10].

В ходе изучения функционирования нервной клетки было выявлено взаимодействие между сигналами, поступавшими на вход нейрона, а также активируется ли нейрон, посылая импульс следующим клеткам или нет [11]. Так, результат зависит от того, на какой именно из входов был подан сигнал. Единичный электрический импульс, поданный на одни входы, был достаточен для активации нейрона. При этом некоторые входы давали срабатывание только при одновременном воздействии на них. Таким образом, влияние разных дендритов оказывается разным. При этом воздействие на некоторые дендриты приводит, наоборот, к деактивации уже активированного с помощью других входов нейрона.

Следовательно, влияние входных импульсов можно классифицировать на две группы: активирующие (возбуждающие) и деактивирующие (подавляющие) [12]. Возможно также, что некоторые входы могут вообще не оказывать воздействие на функционирование нейрона. Конечный результат зависит только от комплексного (суммарного) воздействия всех входных импульсов [13].

На основе вышесказанного и была разработана математическая модель биологического нейрона – искусственный нейрон [14]. Создание и развитие нейронных сетей представляло собой длительный путь, который сопровождался этапами роста интереса к данной теме и этапами спада.

В 1943 году была издана публикация Мак-Келлога и Питтса [15], которые описывали работу нейрона как бинарное устройство, производящее логические операции над двоичными данными. Учитывая то, что для активации биологического нейрона требуется преодоление входными электрическими сигналами некоторого порога, предложенная

модель была названа «пороговой логикой». Была сконструирована электрическая схема, описывающая сеть из данных нейронов, способная подстраивать проводимость связей между ними. Таким образом, данная сеть имела возможность оптимизации и обучения, что позволяло её использовать для решения широкого круга задач.

В 1948 году благодаря работам Винера была заложена основа для появления новой научной дисциплины под названием «кибернетика». Данная наука изучает обработку, хранение и видоизменение информации в сложных системах, будь то биологический организм, общество и т.д.

В 1949 году для обучения искусственных нейронов нейрофизиолог Дональд Хебб предложил первый алгоритм обучения, направленный на целенаправленное изменение величин синаптических связей [16]. Данный алгоритм основывался на теории нейропластичности. Предложенная Хеббом нейросеть была способна к самообучению и легла в основу более эффективных современных алгоритмов обучения искусственных сетей.

Сеть Хебба была впервые воссоздана на вычислительной машине в 1954 году сотрудниками Массачусетского технологического института Кларком и Фарли. Вскоре, в 1956 году рядом ученых были программно реализованы различные вариации нейросетей.

В 1958 году был создан наиболее эффективный прототип нейронной сети – однослойный перцептрон (персептрон) [17]. Автором её стал американский ученый Фрэнк Розенблат, который использовал перцептрон для задачи классификации символов. Он же построил первый нейрокомпьютер «Марк-1», который обладал способностью к самообучению.

Математически однослойный перцептрон является аналогом линейного аппроксиматора [18]. На вход персептрона подается вектор входных активаций x . Для каждого входа ставится в соответствие определенный вес w . Веса можно представить в общем виде как один вектор весов. Тогда активация нейрона возникает, если сумма поэлементных произведений входов на соответствующие веса (скалярное произведение данных векторов) превосходит некоторое пороговое значение T . В этом случае нейрон возвращает значение 1, иначе – 0. В математическом виде это можно записать следующим образом.

$$y = \begin{cases} 1, & \sum_{i=1}^n x_i \cdot w_i > T, \\ 0, & \sum_{i=1}^n x_i \cdot w_i \leq T. \end{cases}$$

Переместив порог T в левую часть неравенств и обратив его знак, получаем смещение нейрона b .

$$y = \begin{cases} 1, & \sum_{i=1}^n x_i \cdot w_i + b > 0, \\ 0, & \sum_{i=1}^n x_i \cdot w_i + b \leq 0. \end{cases}$$

В целом же, искусственный нейрон имеет более широкий спектр свойств, чем перцептрон Розенблата. Нейрон может быть математически описан двумя функциями: сумматорной и активационной.

Сумматорная функция в большинстве случаев стандартна и не претерпевает изменений в нейронных сетях различной архитектуры [19]. Она основывается на том факте, что биологический нейрон реагирует только на суммарное значение входных активаций, на их суперпозицию. Данная функция выглядит следующим образом:

$$s(x) = \sum_{i=1}^n x_i \cdot w_i + b.$$

Результат сумматорной функции отправляется в качестве аргумента активационной функции и уже её результат становится выходным значением нейрона. В качестве активационных в нейросетях выступают самые разнообразные функции, а именно:

1) Линейная:

$$f(s) = s.$$

2) Пороговые:

- Модифицированная функция Хэвисайда:

$$f(s) = \begin{cases} 1, & s > 0, \\ 0, & s \leq 0, \end{cases}$$

или

$$f(s) = \begin{cases} 1, & s > 0, \\ -1, & s \leq 0. \end{cases}$$

3) Сигмоидальные:

- Логистическая функция (функция Ферми):

$$f(s) = \frac{1}{1 + e^{-s}}.$$

- Гиперболический тангенс:

$$f(s) = \tanh(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}}.$$

- Рациональная:

$$f(s) = \frac{s}{1 + |s|}.$$

4) Экспоненциальные [20]:

- Классическая экспоненциальная:

$$f(s) = e^{-s}.$$

- Радиальная:

$$f(s) = e^{-s^2}.$$

5) Полулинейные:

- ReLU (Rectified Linear Unit):

$$f(s) = \begin{cases} s, & s > 0, \\ 0, & s \leq 0. \end{cases}$$

- Полулинейная с насыщением:

$$f(s) = \begin{cases} 1, & s \geq 1, \\ s, & 0 < s < 1, \\ 0, & s \leq 0. \end{cases}$$

6) Треугольная, модульная, квадратичная, синусоида и т.д.

В частности, классический однослойный перцептрон Розенблата представляет собой единичный искусственный нейрон, в котором в качестве активационной функции выступает модифицированная функция Хэвисайда.

Следует отметить, что ввиду особенностей работы и требований алгоритмов обучения, предпочтение отдается активационным функциям, которые удовлетворяют следующим критериям:

- ограниченная область значений (насыщаемость);
- дифференцируемость (при этом, чем более простой вид имеет производная функции, тем лучше);
- должна быть неубывающей;
- быть нелинейной.

В связи с вышеприведенными качествами, наибольшее распространение имеют сигмоидальные активационные функции [19]. Своё название они получили благодаря графику данных функций, которые напоминают букву «s». Самой эффективной в большинстве случаев остается логистическая функция. Стоит отметить, что через эту функцию можно выразить и другую сигмоидальную функцию – гиперболический тангенс, который является лишь результатом масштабирования и смещения исходной функции. Учитывая, что чаще всего активационную функцию дополняют множителем перед переменной s , то можно считать обе функции идентичными. Их графики приведены на рисунках 2 и 3.

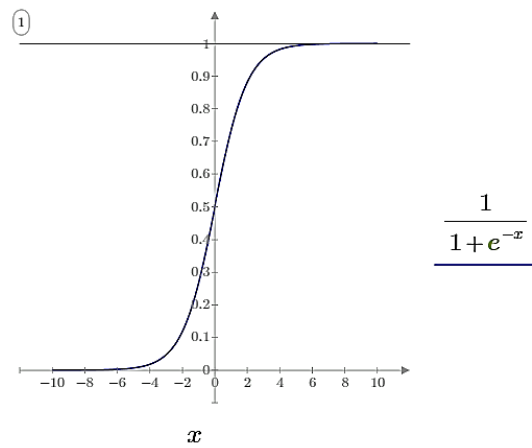


Рис. 2 – График логистической функции

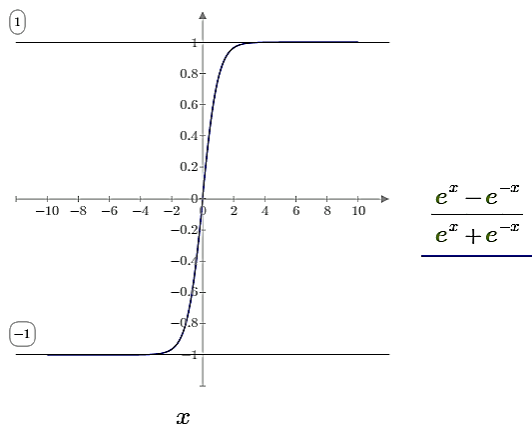


Рис. 3 – График гиперболического тангенса

Схематично однослойный перцептрон Розенблата представлен на рисунке 4.

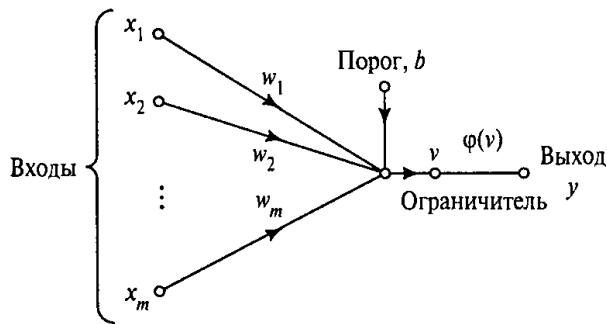


Рис. 4 – Структура однослойного перцептрона Розенблата

В 60-х годах в России также велись исследования в области изучения работы нейронных сетей. Так, например, биофизиками Александром Петровым и Михаилом Бонгардом были предприняты попытки поиска задач, которые невозможно решить при помощи перцептрона.

Как было выяснено, существующие на данный момент нейросети не были способны решать многие задачи, в том числе, легко реализуемые другими методами. Так, например, нерешенной являлась задача «XOR» или «исключающего или». Простейший логический вентиль XOR имеет два входа. На выходе значение 1 будет в том и только в том случае, если логическая единица будет строго на одном из входов. Тогда для перцептрона Розенблата выходит система неравенств, накладывающая ограничения на веса:

$$\begin{cases} w_1 + b > 0, \\ w_2 + b > 0, \\ w_1 + w_2 + b \leq 0, \\ b \leq 0. \end{cases}$$

Очевидно, что весов w_1 и w_2 , которые бы удовлетворяли данной системе, не существует.

В 1969 году специалист в области искусственного интеллекта Марвин ли Минский в соавторстве с Сеймуром Папертом, наряду с собственными доказательствами обучаемости перцептрона, обосновал ограниченность перцептрона и его принципиальную неспособность решать широкий спектр задач [21].

Вышеприведенные недостатки, которые скопились в конце 60-ых годов, а также смерть автора перцептрона привели к потере интереса относительно нейронных сетей у ученых. Данная тенденция продолжалась более 15-и лет. Несмотря на это, небольшой круг ученых продолжал заниматься совершенствованием нейросетей. Среди них были такие исследователи, как Кохонен, Гроссберг, Андерсон и др.

В частности, исследования велись для комбинаций отдельных перцептронов, организованных в слои – многослойных перцептронов. Они представляют собой сети, в которых нейроны одного слоя связаны со всеми нейронами предыдущего слоя. Первый слой обычно состоит из нейронов с заданными значениями на выходе, то есть имитируют входной сигнал. Промежуточные слои называют «скрытыми». Результат работы сети считается с последнего слоя. Структура такой сети представлена на рисунке 5.

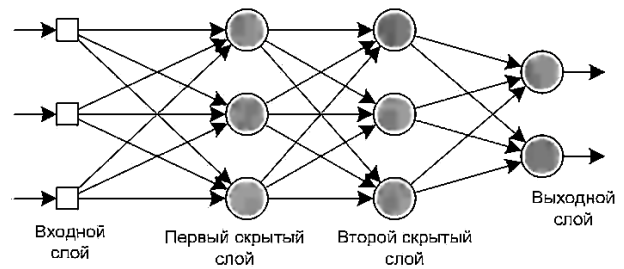


Рис. 5 – Структура многослойного перцептрона

Так как значения импульсов передаются последовательно от первого слоя к последнему, многослойный перцептрон называют ещё сетью с прямым распространением активаций (feed forward neural networks).

В 1974 году П. Вербосом был описан новый метод обучения – алгоритм обратного распространения ошибки (backpropagation) [22]. К аналогичным результатам впоследствии пришли Д. Паркер (1982 г.), Дэвид И. Румельхарт, Дж. Е. Хинтон и Рональд Дж. Вильямс (1986 г.), а также советские учёные С.И. Барцев и В.А. Охонин (1986г.).

Алгоритм обратного распространения ошибки основан на минимизации функционала ошибки нейросети методом градиентного спуска. Изменение весов нейронов при этом происходит в направлении, противоположном направлению градиента. Данный факт требует, чтобы активационная функция нейрона была дифференцируемой. Функционал ошибки, вычисляемый для обучения нейронов выходного слоя, передается нейронам нижележащих слоёв в соответствии с их вкладом в итоговую невязку.

Функцию ошибки можно представить, как:

$$E = \hat{y} - y,$$

где \hat{y} и y – соответственно, истинное и расчётное значения выходного параметра.

Тогда величина изменения весов будет записана как:

$$\Delta b = \Delta,$$

$$\Delta w_i = \Delta \cdot a_i^{-1}(s).$$

- Для выходного слоя:

$$\Delta = E \cdot a'(s).$$

- Для скрытых слоев:

$$\Delta = \Delta^{+1} \cdot w_i^{+1} \cdot a'(s),$$

где Δ – величина ошибки для данного нейрона;

Δ^{+1} – величина ошибки для нейрона следующего слоя;

$a'(s)$ – значение производной активационной функции от результата сумматорной функции;

$a_i^{-1}(s)$ – значение активации нейрона предыдущего слоя (входа для данного нейрона);

w_i^{+1} – вес синаптической связи между данным нейроном и нейроном следующего слоя.

Данный метод позволил ускорить процесс обучения и расширить круг успешно решаемых нейросетью задач. Однако, и он имел определенные недостатки, связанные в основном с методом градиентного спуска, а именно замедление обучения при выходе значения целевой функции на «плато», а также незащищенность от схождения в локальном

минимуме. Несмотря на это, он послужил мощным импульсом для развития нейронных сетей.

В 1975 году японский исследователь Фукусима разработал нейросеть новой архитектуры – когнитрон [23]. Она моделировала принцип работы зрительной коры головного мозга, в которой нейроны подразделяются на тормозящие и возбуждающие. Благодаря слоистой структуре наблюдается уменьшение размерности входного изображения, что в конечном итоге сводится к решению задачи классификации.

В 1980 году Фукусимой была представлена более совершенная модель нейронной сети, названная неокогнитроном [24]. Её основное отличие состоит в наличии чередующихся «простых» (US) и «сложных» (UC) слоёв. Такая структура сети позволяет ей быть устойчивой к помехам, шуму, повороту и искажению изображения.

В 1982 году Джон Хопфилд стал автором полносвязной нейронной сети с реализацией обратных связей [25]. В данной сети нейроны могли выполнять все три функции, а именно быть входными до обучения, скрытыми в процессе обучения и выходными после него. При этом каждый искусственный нейрон был соединён со всеми остальными. Для стабилизации значений нейронов в системе был введен дополнительный гиперпараметр – температура, понижение которой способствовало сходимости значений.

Сеть Хопфилда, часто называемая ещё ассоциативной памятью (благодаря стабилизации в ней происходит запоминание образов), стала прообразом для прочих рекуррентных нейронных сетей, в которых результат активации нейрона напрямую или же через промежуточные нейроны, зависит от собственного значения на предыдущих итерациях обучения.

Также в 1982 году были открыты самоорганизующиеся карты Кохонена (Kohonen networks) [26].

В процессе обучения данных сетей происходит соревновательное обновление весов отдельных нейронов. При этом на результат данных нейронов также оказывают значения результаты работы соседних нейронов. Данное свойство позволяет применять карты Кохонена в задачах классификации и кластеризации.

В 1986 году была опубликована нейронная машина Больцмана (BM) [27]. Она аналогична по строению сетям Хопфилда. Отличием является то, что нейроны в данной полносвязной сети строго классифицированы по признаку на входные/выходные и скрытые. Обучают такие сети алгоритмом обратного распространения ошибки или алгоритмом *contrastive divergence*, в котором градиент вычисляется на основе цепей Маркова.

Также были предложены ограниченные машины Больцмана [28]. Отличие их состоит в том, что данные сети не являются полносвязными, за счет чего происходит упрощение сети и ускорение процесса её обучения.

В 1987 году Роберт Хехт-Нильсон путём оптимизации действующих алгоритмов обучения разработал алгоритм встречного распространения ошиб-

ки [29]. Сети, обучаемые данным способом, содержат в своем составе элементы карты Кохонена и слой Гроссберга. В ходе обучения определяется нейрон, запоминающий входные активации, а уже затем по нему формируется значение выходных нейронов. Данный алгоритм позволяет сети обучаться в десятки и сотни раз быстрее алгоритма обратного распространения ошибки. Однако, он менее точен и не способен сводить функционал ошибки к нулю. В связи с этим данный алгоритм эффективен на начальных этапах обучения нейронных сетей.

В 1987 году были разработаны нейронные сети, обладающие свойством пластичности. Это свойство характеризует способность сетей обучаться новым данным без потери накопленной информации в предыдущих процессах обучения [30].

Свойством пластичности обладают сети адаптивного резонанса (AR), открытые Стивеном Гроссбергом и Гейлом Карпендером [31, 32]. Данные сети обладают специфичным гиперпараметром – порогом «бдительности». При подаче на вход вектора входных активаций происходит активация определенных нейронов в слое сети. При этом активируясь, нейрон снижает степень активации других нейронов сети. Если активация нейрона с максимальным выходным значением превосходит порог «бдительности», то происходит обучение нейросети с обновлением весов. В противном случае данный нейрон исключается и процесс повторяется снова. Значение гиперпараметра определяет точность детализации, с которой искусственная нейронная сеть будет запоминать данные из обучающей выборки.

Нейронные сети адаптивного резонанса достаточно разнообразны. Наибольшее распространение получили:

- ART1 – сеть, оперирующая двоичными данными;
- ART2 и ART3 – сети, оперирующие непрерывными значениями;
- FuzzyART – сеть, основанная на законах нечёткой логики.

В 1988 году Брумхедом и Лоу были представлены сети на основе радиально-базисных функций (Radial Basis Function – RBF) [33]. Данные сети отличаются от многослойного перцептрона радиально-базисной функцией в качестве функции активации нейронов единственного скрытого слоя.

В 1990 году Джеффри Элман разработал рекуррентные нейронные сети (Recurrent Neural Networks – RNN) [34]. Данные искусственные нейронные сети являются аналогами сетей прямого распространения за тем лишь исключением, что каждый нейрон имеет синаптическую связь с самим собой, то есть образует петлю. За счет этого достигается влияние на результат работы сети не только текущих, но и предыдущих входных активаций.

Такие сети высоко эффективны при обработке потоковой информации: изображений, видео, аудио, текстовой информации.

Стоит отметить, что для данных сетей процесс обучения необходимо осуществлять на последовательной обучающей выборке – перемешивание данных является недопустимым. Также несмотря на то,

что рекуррентные сети анализируют временные зависимости, они обладают кратковременной памятью вследствие затухания предыдущих сигналов.

В 1998 году Ян Лекун предложил для распознавания и классификации изображений использовать свёрточные нейронные сети (Convolutional Neural Networks – CNN) [35].

Данные искусственные нейронные сети содержат большое число скрытых слоев, число нейронов в которых при движении вглубь уменьшается в геометрической прогрессии. Такие сети способны анализировать входное изображение, осуществляя его последовательное сканирование.

В 2000-х годах исследования специалистов по нейросетям были в основном посвящены решению проблемы нахождения глобального минимума в процессах обучения. Данная задача решалась путем применения различных приложений стохастического обучения (Больцмана, Коши).

Больцмановское обучение аналогично такому алгоритму оптимизации как «имитация отжига». Для этого изначально задаётся «температура» T с достаточно высоким значением [36]. Далее начинают последовательно изменять веса сети в ту или иную сторону на величину:

$$P(w) = e^{-\frac{w^2}{T^2}}.$$

Если изменение веса способствовало уменьшению ошибки на выходе из сети, то данное изменение сохраняется. В противном случае изменение происходит с вероятностью, зависящей от температуры по формуле:

$$P(c) = e^{-\frac{c}{kT}}.$$

После каждого шага значение температуры постепенно уменьшают. По завершении данного процесса большинство весов связей приобретают значения, соответствующие глобальному минимуму ошибки.

Обучение Коши отличается от Больцмановского заменой функции распределения Больцмана на функцию Коши:

$$T(t) = \frac{T_0}{1+t},$$

$$P(c) = \frac{T(t)}{T(t)^2 + x^2}.$$

Так как распределение Коши имеет неопределённую величину дисперсии, то в ходе обучения веса меняются значительно сильнее. Это приводит к ещё большему увеличению скорости обучения, однако возрастает риск возникновения «паралича» сети.

В 2007 году Джеффри Хинтон разработал алгоритм глубокого обучения [37]. Данный метод эффективен для обучения многослойных нейронных сетей. Особенностью алгоритма является использование ограниченной машины Больцмана для обучения нейронов в начальных слоях сети.

В 2010 году были разработаны развёртывающиеся нейронные сети (Deconvolutional Networks – DN) [38]. Данные сети также называют обратными графическими сетями. Они похожи на свёрточные сети,

однако количество нейронов в них, напротив, увеличивается от слоя к слою. Развёртывающиеся нейронные сети эффективно справляются с задачами восстановления образа из малого набора исходных данных.

В 2014 году Ян Гудфеллоу предложил генеративные состязательные сети (GAN) [39]. Такие сети являются симбиозом двух нейросетей, которые могут быть различной архитектуры (наиболее распространены свёрточные сети и сети прямого распространения). Выходы первой нейросети являются входами для второй нейросети.

Первая половина сети, называемая генератором, на основе входных данных формирует некоторые новые данные. Эти данные обрабатываются уже второй частью сети, называемой дискриминатором. Дискриминатор определяет источник данных, а именно являются ли эти данные исходными или же они синтезированы генератором.

Такие сети работают в режиме состязания, поскольку генератор обучается формировать наиболее непривычные данные для дискриминатора. В свою очередь последний обучается каждый раз их выявлять.

В 2015 году сотрудниками Массачусетского технологического института была опубликована работа [40] по созданию глубоких свёрточных обратных глубинных сетей (Deep Convolutional Inverse Graphics Networks – DCIGN). Данные сети представляют собой комбинацию свёрточных и развёртывающихся сетей.

Такие сети обладают широким спектром возможностей обработки изображений. В частности, они способны достраивать недостающие фрагменты фотографий, а также осуществлять трёхмерный поворот изображений, представленных в двумерном формате.

Также в 2015 году вышла работа [41], описывающая глубокие остаточные сети (Deep Residual Networks – DRN). Данные сети являются модифицированными сетями прямого распространения с большим количеством скрытых слоёв, между которыми имеются связи, соединяющие удалённые слои. В таких сетях исходные активации распространяются как последовательно по слоям, так и пропуская некоторые слои.

За счёт указанных свойств данные сети способны обучаться сложноорганизованным данным как обычные глубокие нейросети, при этом процесс обучения осуществляется с более высокой скоростью.

Заключение

Последние исследования показывают, что за период 2016-2018 годов произошло резкое снижение численности стартапов, в основе которых лежат машинное обучение и, в частности, искусственные нейронные сети [42]. Это может свидетельствовать о завершении стадии завышенных ожиданий к данной технологии. Найденные архитектурные решения, а также предложенные алгоритмы обучения показали высокую, а зачастую, наибольшую из всех доступных методов эффективность для решения широкого

спектра прикладных задач. Однако, исследователи столкнулись с насыщением возможностей существующих нейросетей, в то время как остались области, в которых нейронные сети так и не достигли приемлемого успеха.

Преодоление возникших трудностей видится научным сообществом путём сведения подобных задач к классам, успешно решаемых нейросетевыми методами, а также через детальную предобработку исходных данных. Таким образом, на данном этапе развития науки возникла необходимость в тесной интеграции искусственных нейронных сетей с технологиями аналитики данных и BigData как единственно возможном дальнейшем векторе развития нейросетей.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-47-020068.

Литература

- 1.R.J. Baron, *The cerebral computer: An introduction to the computational structure of the human brain*. Psychology Press, Hove, UK, 2013. Pp. 552.
- 2.G.M. Shepherd, *Foundations of the neuron doctrine*. Oxford University Press, Oxford, UK, 2015. Pp. 384.
- 3.G. Indiveri, B. Linares-Barranco, T.J. Hamilton, A.V. Schaik, R. Etienne-Cummings, T. Delbruck, S. Liu, P. Dudek, P. Häfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, K. Hynna, F. Folowosele, S. Saighi, T. Serrano-Gotarredona, J. Wijekoon, Y. Wang, K. Boahen, Neuromorphic silicon neuron circuits, *Frontiers in neuroscience*, **5**, 73 (2011).
- 4.J.A. Bailey, R. Wilcock, P.R. Wilson, J.E. Chad, Behavioral simulation and synthesis of biological neuron systems using synthesizable VHDL, *Neurocomputing*, **74**, 14-15, 2392-2406 (2011).
- 5.E.R. Kandel, J.H. Schwartz, T.M. Jessell, *Principles of Neural Science*, McGraw-hill, New York, **4**, 1227-1246 (2000).
- 6.T. Zhang, Y. Zeng, B. Xu, Biological neural network structure and spike activity prediction based on multi-neuron spike train data, *International Journal of Intelligence Science*, **5**, 2, 102-111 (2015).
- 7.S. Fujita, Analysis of neuron differentiation in the central nervous system by tritiated thymidine autoradiography, *Journal of Comparative Neurology*, **122**, 3, 311-327 (1964).
- 8.S. Basu, A. Aksel, B. Condrion, S.T. Acton, Tree2Tree: Neuron segmentation for generation of neuronal morphology, *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, IEEE, Rotterdam, Netherlands, 2010. Pp. 548-551.
- 9.A. Gidon, I. Segev, Principles governing the operation of synaptic inhibition in dendrites, *Neuron*, **75**, 2, 330-341 (2012).
10. O. Hobert, I. Carrera, N. Stefanakis, The molecular and gene regulatory signature of a neuron, *Trends in neurosciences*, **33**, 10, 435-445 (2010).
11. N. Kasabov, To spike or not to spike: A probabilistic spiking neuron model, *Neural Networks*, **23**, 1, 16-19 (2010).
12. S. Yaghini Bonabi, H. Asgharian, S. Safari, M. Nili Ahmadi, FPGA implementation of a biological neural network based on the Hodgkin-Huxley neuron model, *Frontiers in neuroscience*, **8**, 379 (2014).
13. R. Yuste, From the neuron doctrine to neural networks, *Nat. Rev. Neurosci.*, **16**, 8, 487-497 (2015).
14. A. Basu, P.E. Hasler, Nullcline-based design of a silicon neuron, *IEEE Trans. CAS-I*, **57**, 11, 2938-2947 (2010).
15. W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, **5**, 4, 115-133 (1943).
16. D.O. Hebb, *The organization of behavior: A neuropsychological theory*. Psychology Press, Hove, UK, 2005. Pp. 335.
17. F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, *GPSR*, **65**, 6, 386-408 (1958).
18. T.M. McKenna, J.L. Davis, S.F. Zornetzer, *Single neuron computation*. Academic Press, San Diego, USA, 2014. Pp. 644.
19. А. Старилов, Нейронные сети – математический аппарат, *BaseGroupLabs*, (2011).
20. B. Karlik, A.V. Olgac, Performance analysis of various activation functions in generalized MLP architectures of neural networks, *Int. J. Artif. Intell. Exp. Syst.*, **1**, 4, 111-122 (2011).
21. M. Minsky, S.A. Papert, *Perceptrons: an introduction to computational geometry*. MIT press, Cambridge, USA, 2017. Pp. 258.
22. P.J. Werbos. Ph.D. thesis, Beyond regression: New tools for prediction and analysis in the behavioral sciences, Harvard University, Cambridge, USA, 1974. Pp. 453.
23. K. Fukushima, Cognitron: A self-organizing multilayered neural network, *Biological cybernetics*, **20**, 3-4, 121-136 (1975).
24. K. Fukushima, S. Miyake, Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition, *Competition and cooperation in neural nets*, Springer, Berlin, Heidelberg, 1982. Pp. 267-285.
25. J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, *Proc. Natl. Acad. Sci.*, **79**, 8, 2554-2558 (1982).
26. T. Kohonen, Self-organized formation of topologically correct feature maps, *Biological cybernetics*, **43**, 1, 59-69 (1982).
27. G.E. Hinton, T.J. Sejnowski, Learning and relearning in Boltzmann machines, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1*. MIT Press, Cambridge, USA, 1986. Pp. 282-317.
28. P.I. Smolensky, Information processing in dynamical systems: Foundations of harmony theory, *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1*. MIT Press, Cambridge, USA, 1986. Pp. 194-281.
29. R. Hecht-Nielsen, Counter propagation networks, *Applied optics*, **26**, 23, 4979-4984 (1987).
30. Z. Warraich, J.A. Kleim, Neural plasticity: the biological substrate for neurorehabilitation, *PM&R*, **2**, 12, S208-S219 (2010).
31. G.A. Carpenter, S. Grossberg, ART 2: Self-organization of stable category recognition codes for analog input patterns, *Applied optics*, **26**, 23, 4919-4930 (1987).
32. S. Grossberg, Competitive learning: From interactive activation to adaptive resonance, *ICGS*, **11**, 1, 23-63 (1987).
33. D.S. Broomhead, D. Lowe, Radial basis functions, multi-variable functional interpolation and adaptive networks, *Complex Systems*, **2**, 3, 321-355 (1988).
34. J.L. Elman, Finding structure in time, *ICGS*, **14**, 2, 179-211 (1990).
35. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. of the IEEE*, **86**, 11, 2278-2324 (1998).
36. С. Хайкин, *Нейронные сети: полный курс*. Вильямс, Москва, 2006. 1103 с.
37. G.E. Hinton, S. Osindero, Y.W. The, A fast learning algorithm for deep belief nets, *Neural Comput.*, **18**, 7, 1527-1554 (2006).
38. M.D. Zeiler, D. Krishnan, G.W. Taylor, R. Fergus, Deconvolutional networks, In: *IEEE Conf. on Computer Vision*

- and Pattern Recognition (CVPR), IEEE, San Francisco, USA, 2010. Pp. 2528-2535.
39. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, *NIPS*, **27**, 2672-2680 (2014).
40. T.D. Kulkarni, W. Whitney, P. Kohli, J.B. Tenenbaum, Deep convolutional inverse graphics network, *NIPS*, **28**, 2539-2547 (2015).
41. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *CVPR*, 770-778 (2016).
42. P. Zhegin, *AI investment activity – trends of 2018*. Towards Data Science, 2019.

© **А. Н. Иванов** - аспирант кафедры информационных технологий и компьютерной математики факультета математики и информационных технологий ФГБОУ ВО «Башкирский государственный университет», sanekclubstr@mail.ru; **С. А. Мустафина** - д.ф.-м.н., профессор, зав. кафедрой математического моделирования Стерлитамакского филиала ФГБОУ ВО «Башкирский государственный университет», mustafina_sa@mail.ru; **Н. Д. Морозкин** - д.ф.-м.н., профессор, ректор ФГБОУ ВО «Башкирский государственный университет», rector@bsunet.ru.

© **A.N. Ivanov** - postgraduate student of the department of information technologies and computer mathematics, faculty of mathematics and information technologies, Bashkir State University, 450076, Russia, Republic Bashkortostan, Ufa, Validy Str., 32, Tel/Fax: +7 (347) 272 63 70, E-mail: sanekclubstr@mail.ru; **S.A. Mustafina** - doctor of Science, Professor, head of chair of mathematical modeling, Sterlitamak Branch of Bashkir State University, 453103, Russia, Republic Bashkortostan, Sterlitamak, Lenin Ave, 37, Tel/Fax: +7 (347) 343 50 02, E-mail: mustafina_sa@mail.ru; **N.D. Morozkin** - doctor of Science, Professor, rector, Bashkir State University, 450076, Russia, Republic Bashkortostan, Ufa, Validy Str., 32, Tel/Fax: +7 (347) 229 96 16, E-mail: rector@bsunet.ru.

СОДЕРЖАНИЕ

ХИМИЯ

<i>Рыбакова А.И., Саутина Н.В., Галяметдинов Ю.Г., Головешкина Д.Ю.</i> Исследование влияния размера капель на структурные переходы в системе вода/аот/изопропилмирилат	5
<i>Гуревич П.А., Галяутдинова Г.Г., Босяков В.И., Егоров В.И., Сайфутдинов А.М., Балымова М.В.</i> Флуориметрическое определение методом ВЭЖХ антибиотика хлорамфеникола с двустадийной предколоночной дериватизацией	10
<i>Саутина Н.В., Захарова А.О., Галяметдинов Ю.Г.</i> Исследование изменений в самоорганизации бис-(2-этилгексил) сульфосукцината натрия в жидких средах методом флуоресцентного анализа	14

ХИМИЧЕСКАЯ ТЕХНОЛОГИЯ

<i>Анашкин И.П., Клинов А.В., Бекетов Н.А.</i> Адсорбция Леннард-Джонсовского флюида на мембране кристаллического типа	20
<i>Андреева Г.В., Гильманов Р.З., Юсупова Л.М., Баранова Ю.Б., Петров Е.С., Мухаметзянова А.А., Корнилов А.В., Шиченков Р.А.</i> Механизм реакции окисления тринитротолуола в щелочной среде	23
<i>Бадретдинова Л.Х., Вахидов Р.М., Базотов В.Я.</i> Влияние размера частиц на период релаксации электрического заряда диэлектриков	26
<i>Евграфова А.А., Зайнуллина А.Э., Булидорова Г.В.</i> Влияние гидрофобизаторов на электрокинетический потенциал частиц пигментов	29
<i>Зинуров В.Э., Дмитриев А.В., Латыпов Д.Н., Соловьева О.В.</i> Влияние загрязнения пылеочистительного сепаратора мелкодисперсной пылью на энергетические затраты в ходе его эксплуатации	33
<i>Зайнуллина А.Э., Евграфова А.А., Сидоренко Д.А., Булидорова Г.В.</i> Исследование пресуемости композиций пигментов, применяемых в косметической промышленности	38
<i>Зинуров В.Э., Дмитриев А.В., Соловьева О.В., Латыпов Д.Н.</i> Исследование изменения эффективности очистки газового потока от мелкодисперсных частиц прямоугольным сепаратором при разной степени забивки дугообразных элементов пылью	42
<i>Калугин Л.Е., Дресвянников А.Ф.</i> Особенности анодного поведения титана в хлоридсодержащих растворах в присутствии фтористоводородной кислоты	47
<i>Клоков В.В., Морозова Н.Н., Кузнецова Г.В., Фазлыева А.Ф.</i> Силикатный кирпич в производстве газобетона автоклавного твердения	53
<i>Макарова Л.Н., Поветкин В.В.</i> Электрохимическое осаждение сплавов на основе цинка с легкоплавкими металлами	57
<i>Мануйко Г.В., Бронская В.В., Аминова Г.А., Антонов А.С., Назарова М.А., Плющев В.В.</i> Влияние водорода и температуры полимеризации пропилена на кинетическую неоднородность титан-магниевого катализатора	61
<i>Момзяков А.А., Киченин С.М., Дебердеев Р.Я.</i> Исследование влияния технологических параметров экструзионного аппарата на синтез стеарата цинка в твердой фазе	67
<i>Муродиён А., Сафаров А.Г., Кабутов К., Ботуров К., Сафиев Х.</i> Физико-химические и термодинамические характеристики антрацитов месторождения Назарайлок Республики Таджикистан	71
<i>Мусина Ф.А., Бронская В.В., Игнашина Т.В., Нурғалиева А.А., Харитонова О.С.</i> Экологически чистые химические технологии для устойчивого развития химической промышленности	79
<i>Нгуен Зюи Туан, Денисюк А.П.</i> Полнота горения композиций на основе нитратов различных металлов при атмосферном давлении	84
<i>Петров Е.С., Хазиева И.И., Гильманов Р.З., Петрова С.С., Филиппов Ю.В., Сафронова О.Е., Мухаметзянова А.А., Шиченков Р.А.</i> Разработка спрея - лекарственной формы для лечения сердечно-сосудистых заболеваний на основе производных 5-бромникотиновой кислоты	90

<i>Романова С.М., Мадякина А.М., Завьялова Н.Б.</i> Изучение возможности применения модифицированных нитратов целлюлозы в составе эмали марки НЦ-132П	95
<i>Чернышева А.В., Шкуро А.Е., Кривоногов П.С., Артемов А.В.</i> Исследование возможности химической сшивки древесно-полимерных композитов	99
<i>Хусаинов А.Д., Сеитова С.А., Шайхиев И.Г., Свергузова С.В.</i> Исследование продукта инертизации нефтешлама отходами ПЭНД в качестве ингредиента эластомерных композиций	102
<i>Юсупов Р.А., Бахтеев С.А., Ситникова Е.Ю., Вильданов А.Ф., Романова Р.Г., Кириллова Т.В.</i> Оптимизация метрологического обеспечения контроля технической воды	106
<i>Черезова Е.Н., Яшина Л.Ю., Хлебников М.К., Медведев А.Н.</i> Влияние деструктата силоксановых резин на термостабильность и деформационно-прочностные свойства ПВХ-композитов	113
<i>Ямалтдинова Э.Ф., Савельев С.Н., Савельева А.В., Фридланд С.В.</i> Очистка сернисто-щёлочных сточных вод окислительными методами	117

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

<i>Александров А.А., Кирпичников А.П., Ляшева С.А., Шлеймович М.П.</i> Анализ эмоционального состояния человека на изображении	120
<i>Горюнова С.М., Сайгин Н.А., Николаева Н.Г.</i> Использование V-образной модели для анализа рисков ситуаций испытательной лаборатории	124
<i>Белаишова Е.С., Хабибуллин Р.Р., Шарипов Б.Р., Канбекова К.Ю.</i> Алгоритмы для оценки ходов в играх антагонистического типа	130
<i>Терентьев С.А.</i> Система управления тренажером, обеспечивающая динамическое подобие технической системы с учетом действия возмущений	135
<i>Тутубалин П.И., Кирпичников А.П., Асадов Р.С., Ахметзянов И.И.</i> О платформе для реализации проекта искусственная поджелудочная железа	139
<i>Якимов И.М., Кирпичников А.П., Валова К.Д., Анишкина В.Н.</i> Имитационное моделирование компьютерной сети в системе Cisco Packet Tracer	145
<i>Якупов Д.Т., Мокишин В.В., Кирпичников А.П.</i> Сравнение методов спектральной кластеризации графовых моделей трубопроводных систем	150
<i>Иванов А.Н., Мустафина С.А., Морозкин Н.Д.</i> Аналитический обзор истории развития нейросетевых архитектур и алгоритмов	156
<i>Кирпичников А.П., Ляшева С.А., Шакирзянова А.А., Шакирзянов Р.М., Шлеймович М.П.</i> Обнаружение окружностей на изображениях с помощью метода Хафа и преобразования быстрой радиальной симметрии	164
<i>Ку Дык Тоан, До Шон Ха, Нуриев Н.К., Печеный Е.А.</i> Модели оптимизации местоположений точек доступа и назначение каналов в проектировании сети WLAN	168
<i>Кирпичников А.П., Буйнова Е.Л., Гараева Э.Р., Ризаев И.С.</i> Формирования маршрута движения транспорта при доставке товаров потребителям	175
<i>Назаров А.А., Поникаров С.И., Петлин К.А.</i> Математическая модель и численные исследования процессов лучистого теплообмена трехатомного газа в реакционной зоне проточного реактора дегидрирования углеводородов	181
<i>Кирпичников А.П., Ляшева С.А., Трегубов В.М., Шильников К.В., Шлеймович М.П.</i> Обнаружение трещин дорожного покрытия на изображениях	186
Правила для авторов	191