

УДК 004.932.75'1

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ФУНКЦИИ АКТИВАЦИИ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ ПРИ РЕШЕНИИ ЗАДАЧИ РАСПОЗНАВАНИЯ ОБРАЗОВ

Толмачев Сергей Владимирович, студент, направление подготовки 01.03.02 Прикладная математика и информатика, Оренбургский государственный университет, Оренбург
e-mail: tolmachev411@yandex.ru

Гришина Любовь Сергеевна, аспирант, направление подготовки 02.06.01 Компьютерные и информационные науки, Оренбургский государственный университет, Оренбург
e-mail: zabrodina97@inbox.ru

Научный руководитель: **Болодурина Ирина Павловна**, доктор технических наук, профессор, заведующий кафедрой прикладной математики, Оренбургский государственный университет, Оренбург
e-mail: prmat@mail.osu.ru

***Аннотация.** В настоящее время распознавание образов относится к одной из важнейших задач искусственного интеллекта. Одним из наиболее актуальных и эффективных методов её решения являются искусственные нейронные сети. Этот метод требует детальной настройки начальных параметров и алгоритмов обучения, от которых во многом зависит вероятность успешной классификации образов. Данное исследование направлено на построение искусственной нейронной сети для решения задачи распознавания рукописных цифр с целью анализа влияния на точность нейросети такого её структурного элемента, как функция активации. Обучение проводилось с использованием наиболее популярного метода обратного распространения ошибки. Результаты вычислительных экспериментов на наборе MNIST показали, что в качестве передаточной функции предпочтительно использование функции положительной срезки. Разработанное программное обеспечение может быть использовано для дальнейших исследований, направленных на поиск алгоритма выбора оптимальной конфигурации нейронных сетей для различных задач машинного обучения.*

***Ключевые слова:** распознавание образов, искусственные нейронные сети, персептрон, функция активации, метод обратного распространения ошибки.*

***Для цитирования:** Толмачев С. В., Гришина Л. С. Исследование влияния функции активации искусственной нейронной сети при решении задачи распознавания образов // Шаг в науку. – 2021. – № 4. – С. 77–82.*

RESEARCH OF THE INFLUENCE OF THE ACTIVATION FUNCTION OF AN ARTIFICIAL NEURAL NETWORK IN SOLVING THE PROBLEM OF PATTERN RECOGNITION

Tolmachev Sergey Vladimirovich, student, training program 01.03.02 Applied Mathematics and Informatics, Orenburg State University, Orenburg
e-mail: tolmachev411@yandex.ru

Grishina Lyubov Sergeevna, postgraduate student, training program 02.06.01 Computer and information sciences, Orenburg State University, Orenburg
e-mail: zabrodina97@inbox.ru

Research advisor: **Bolodurina Irina Pavlovna**, Doctor of Technical Sciences, Professor, Head of the Department of Applied Mathematics, Orenburg State University, Orenburg
e-mail: prmat@mail.osu.ru

***Abstract.** Currently, pattern recognition is one of the most important tasks of artificial intelligence. One of the most relevant and effective methods of its solution is artificial neural networks. This method requires detailed configuration of the initial parameters and learning algorithms, which largely determine the probability of successful image classification. This study is aimed at building an artificial neural network to solve the problem of recognizing*

handwritten numbers in order to analyze the impact on the accuracy of the neural network of such a structural element as the activation function. The training was carried out using the most popular backpropagation method. The results of computational experiments on the MNIST set have shown that it is preferable to use the positive cutoff function as the transfer function. The developed software can be used for further research aimed at finding an algorithm for choosing the optimal configuration of neural networks for various machine learning problems.

Key words: pattern recognition, artificial neural networks, perceptron, activation function, backpropagation method.

Cite as: Tolmachev, S. V., Grishina, L. S. (2021) [Research of the influence of the activation function of an artificial neural network in solving the problem of pattern recognition]. *Shag v nauku* [Step into science]. Vol. 4, pp. 77–82.

Введение

Область распознавания образов связана с автоматическим обнаружением закономерностей в данных с помощью компьютерных алгоритмов и с использованием этих закономерностей для принятия таких действий, как классификация данных по различным категориям. На данный момент задача распознавания образов является основной во многих интеллектуальных системах, таких как: распознавание лиц [5], отпечатков пальцев, речи [2], жестов, машинное зрение [6], диагностика в медицине [3], геология. Идея проектирования интеллектуальных вычислительных устройств по образу и подобию биологических систем привела к созданию теории нейронных сетей, ставшей одним из самых мощных и полезных подходов к разработке искусственного интеллекта. Нейронные сети уже являются ключевым структурным элементом многих технологий в повседневной жизни, таких как системы автоматического распознавания автомобильных номеров или системы считывания почтовых индексов, написанных от руки.

Заметим, что производительность нейронных сетей зависит от различных факторов, таких как количество скрытых слоев, количество скрытых узлов, алгоритм обучения и функция активации каждого узла. Тем не менее, основной упор в исследованиях нейронных сетей делается на изучение алгоритмов и архитектур, а важность функций активации игнорируется. Выбор функций активации может сильно повлиять на сложность и производительность нейронных сетей и играет существенную роль в сходимости алгоритмов обучения.

1. Обзор исследований

Вопрос зависимости качества нейронной сети от её параметров для различных задач машинного обучения изучается непрерывно.

Авторский коллектив из Федерального университета Пернамбуку в публикации [10] рассматривает вопрос повышения производительности моделей искусственных нейронных сетей для прогнозирования временных рядов. В частности, в работе предлагаются три новые функции активации и оценивается их эффективность с использованием наборов данных финансового рынка.

В экспериментах использовались два алгоритма обучения: алгоритмы обратного распространения ошибки с модификацией Флетчера-Ривса и Левенберга-Марквардта.

В публикации [9] С. И. Хашина производится оценка различных функций активации для задач регрессии в компьютерной графике. Автор осуществляет попарное сравнение на трех обучающих выборках при неизменном количестве скрытых слоев и узлов в них. В ходе испытаний лучший результат получается при использовании функции ReLU во всех внутренних слоях.

Автор исследования [7] анализирует влияние параметров нейронной сети на качество прогнозирования и скорость её обучения. В работе тестируются различные варианты архитектур нейронных сетей без обратной связи на синтетических и реальных данных с различными алгоритмами обучения. В рамках исследования выполнен детальный обзор методов обучения, а также отмечены их преимущества и недостатки.

Таким образом, обзор показал, что поиск алгоритма настройки наиболее эффективной конфигурации нейронной сети для различных её приложений на текущий момент является актуальной темой исследований.

Целью данной работы является построение искусственной нейронной сети для решения задачи распознавания цифр, в процессе проектирования и тестирования которой будет продемонстрирована значимость выбора функции активации на точность и скорость обучения.

2. Постановка задачи

Рассмотрим базу данных образцов рукописного написания цифр MNIST, предложенную Национальным институтом стандартов и технологий США с целью сравнения методов распознавания изображений с помощью машинного обучения. База данных содержит тренировочный набор из 60000 размеченных изображений и тестовый набор объемом в 10000 прецедентов.

Пусть X – множество изображений цифр, Y – множество цифр. Тогда обучающая выборка X^l представляет собой множество пар объект-ответ $X^l = (x_i, y_i)_{i=1}^l$,

где

$x_i \in X$ – 8-битное изображение 28×28 пикселей, заданное матрицей значений цветов,
 $y_i \in Y$ – известная цифра на объекте.

Пусть существует $y: X \rightarrow Y$ – некоторая зависимость, значения которой известны только на объектах обучающей выборки. Требуется построить алгоритм $a: X \rightarrow Y$, способный классифицировать произвольный объект $x \in X$. Алгоритм a является приближением неизвестной зависимости y , осуществляющей отображение зависимости между изображениями цифр и их классами.

Таким образом, получена формальная постановка задачи многоклассовой классификации, которую будем решать с помощью искусственных нейронных сетей с использованием алгоритма обратного распространения ошибки.

3. Нейросетевое решение задачи распознавания цифр

3.1 Архитектура

Существует определенная классификация ИНС [4, 1]. Их можно разделить по нескольким критериям: количеству слоев, типу связей, алгоритму обучения, типу решаемой задачи.

Для распознавания образов и видео, обработки естественного языка используются сверточные нейронные сети, получившие в последнее время широкое распространение. Они обеспечивают частичную устойчивость к изменениям масштаба, смещениям, смене ракурса и искажениям. Однако такая архитектура требует сложной настройки параметров и её применение лучше подходит для решения более сложных задач распознавания. Поэтому, в качестве архитектуры для решения поставленных задач используем многослойный персептрон (MLP).

Данная модель сети относится к искусственным нейронным сетям прямого распространения, в которых связи между узлами не цикличны. В такой сети распространение сигнала происходит строго в одном направлении, от входных узлов к выходным узлам через скрытые (внутренние), если они существуют. Многослойный персептрон имеет три или более последовательно соединенных слоя. Каждый нейрон в одном слое посылает сигнал всем нейронам в последующем слое. Во многих приложениях сети с данной структурой в качестве функции активации используют сигмоидную функцию.

В зависимости от типа структур нейронов ИНС подразделяются на гомогенные (однородные) и гетерогенные. Сети, в которых все нейроны имеют единую функцию активации, называются гомогенными, в противном случае – гетерогенными. В соответствии с целью исследования и со структурой прецедентов в данном исследовании будет использоваться однородный трехслойный персептрон с 28×28 входными и 10 выходными нейронами.

3.2. Метод обучения

Наиболее популярным методом обучения многослойного персептрона является метод обратного распространения ошибки [8], основанный на вычислении градиента функции потерь. Основным принципом метода заключается в распространении сигнала ошибки в направлении, обратном прямому распространению сигнала. Преимуществом данного метода является быстрое вычисление градиента, возможность обобщения на произвольные передаточные функции и функции потерь, а также возможность динамического и распределенного обучения. Следует учитывать, что данный метод имеет медленную сходимость и для его использования необходимо, чтобы функция активации нейрона была дифференцируемой.

3.3. Функции активации

Функция активации рассчитывает выходное значение нейрона на основании взвешенной суммы входов и некоторого порогового значения. Чаще всего на практике применяются следующие функции активации: ReLU, Leaky ReLU, сигмоида и гиперболический тангенс.

Сигмоида – это гладкая монотонная нелинейная функция, которая применяется для сглаживания значений некоторой величины. Областью определения данной функции является вся числовая прямая $(-\infty, +\infty)$ а область значений представляет собой интервал $(0; 1)$.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

Гиперболический тангенс – это скорректированная сигмоидная функция, равная отношению полуразности и полусуммы двух экспоненциальных функций в точках x и $-x$. Данная функция отображает множество всех действительных чисел во множество значений диапазона от -1 до 1 . Выход функции центрирован относительно нуля.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = 2\sigma(2x) - 1 \quad (2)$$

Производные $\sigma(x)$ и $\tanh(x)$ могут быть выражены через их значения, что облегчает использование данных функций при обучении сети по алгоритму обратного распространения. Однако их применение может приводить к загибанию градиента или «параличу сети» из-за горизонтальных асимптот.

Функция положительной срезки (ReLU) – кусочно-непрерывная функция, реализующая простой пороговый переход в нуле. Заменяет все отрицательные значения на нуль. Преимуществом этой функции является простота математического преобразования, а недостатком – выключение ча-

сти нейронов сети из-за равенства нулю градиента («проблема умирающего ReLU»).

$$f(x) = \max(0, x) \quad (3)$$

Leaky ReLU – это модификация функции положительной срезки, предложенная для решения проблемы умирающих нейронов. График данной функции при отрицательных значениях образует не горизонтальную прямую, а наклонную с небольшим коэффициентом, что позволяет добиться ненулевого градиента.

$$f(x) = \begin{cases} 0,01x & x < 0 \\ x & x \geq 0 \end{cases} \quad (4)$$

В связи с существенными различиями между функциями активации: их областью допустимых значений, свойствами дифференцируемости и ограниченности, промежутками знакопостоянства и т.д., нейросетевое решение будет изменяться и иметь разные точность и скорость сходимости.

В рамках данного исследования рассмотрены все представленные функции активации и проведе-

на оценка их эффективности при решении задачи распознавания цифр.

4. Результаты исследования

В ходе выполнения работы была реализована нейронная сеть на языке программирования Python и проведена серия вычислительных экспериментов при различных конфигурациях сети.

Эксперимент 1. Анализ влияния количества эпох обучения

В данном эксперименте выполнялось последовательное обучение нейронной сети при 10 эпохах. Скорость обучения была принята равной 1. После прохождения каждой из 10 эпох производилось тестирование нейронной сети на тестовом наборе данных. В результате эксперимента нейронная сеть с функцией активации ReLU показала наилучшие результаты, достигнув за 10 эпох точности 96%. Чуть менее эффективной получилась нейронная сеть с функцией активации Leaky ReLU (рисунок 1), которая по итогам эксперимента могла правильно распознавать цифры с вероятностью 94.7%.

Зависимость точности от количества эпох

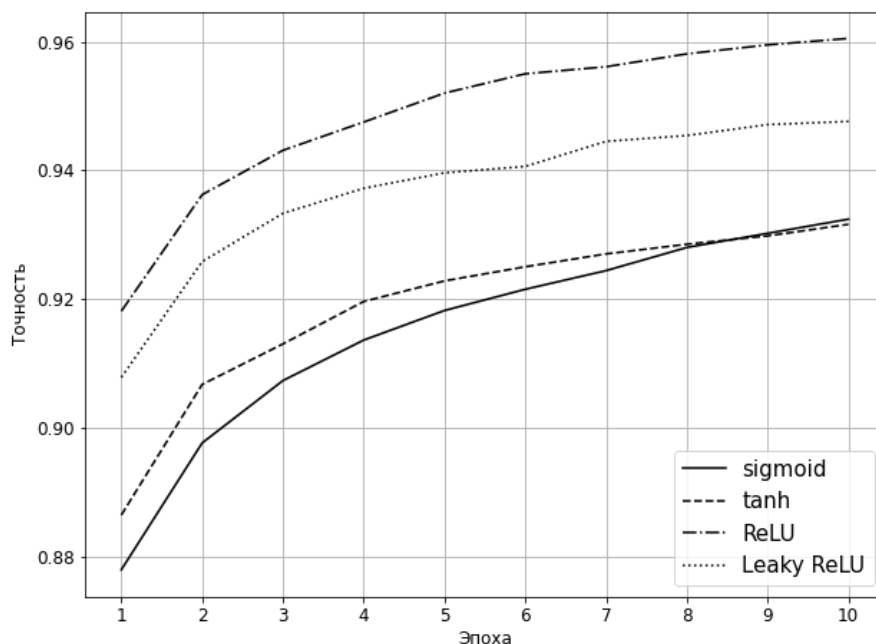


Рисунок 1. Результаты первого эксперимента

Эксперимент 2. Анализ влияния коэффициента обучения

В ходе второго эксперимента производилось варьирование скорости обучения в диапазоне значений от 0,1 до 1,0 с шагом 0,1. Количество эпох было зафиксировано на 10, так как при данном их числе в предыдущем эксперименте были получе-

ны наилучшие показатели. Точность нейронной сети с функциями активации ReLU и Leaky ReLU возрастала, пока скорость обучения не превышала 0,9. При использовании в структуре нейронной сети остальных функций активаций точность распознавания увеличивалась вместе с ростом коэффициента обучения постоянно (рисунок 2).

Зависимость точности от скорости обучения

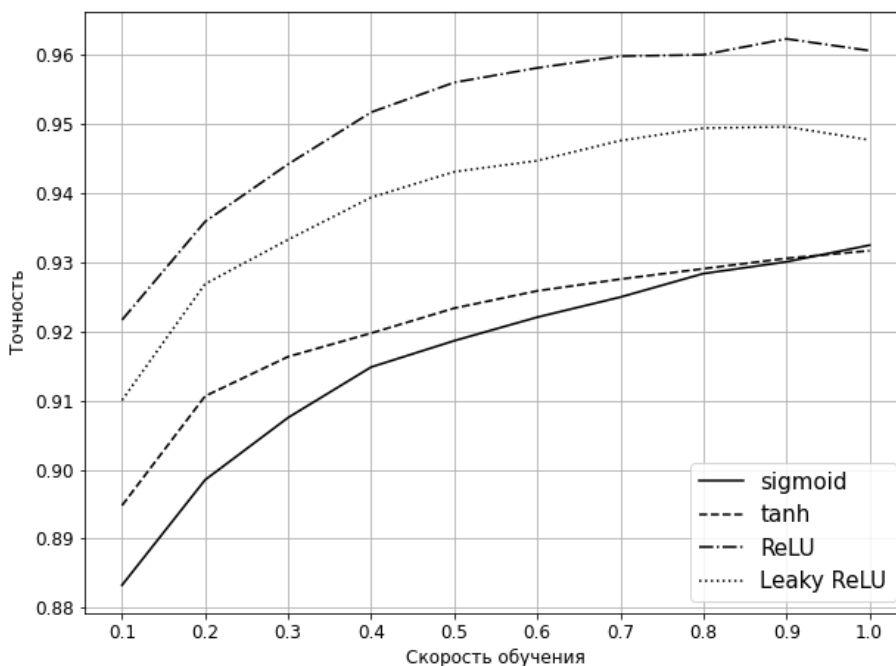


Рисунок 2. Результаты второго эксперимента

Эксперимент 3. Анализ влияния количества скрытых узлов

В последнем эксперименте изменялось количество скрытых узлов нейронной сети при постоянной скорости обучения, равной 0,9. По результатам эксперимента можно наблюдать, что точность нейронной сети с функцией активации гиперболичес-

кий тангенс является монотонно возрастающей функцией от мощности скрытого слоя, чего нельзя сказать при использовании других передаточных функций (рисунок 3). Наилучшие точность и скорость обучения по-прежнему достигаются при использовании кусочно-непрерывных ReLU и Leaky ReLU.

Зависимость точности от скрытых узлов

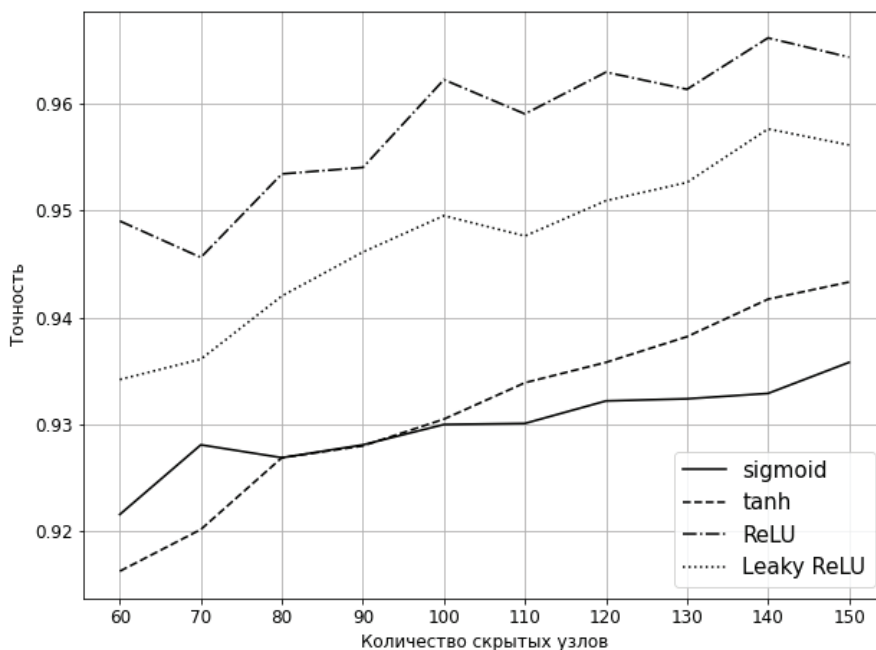


Рисунок 3. Результаты третьего эксперимента

Заключение

В рамках данного исследования была изучена теория искусственных нейронных сетей, на основании которой были выбраны структура и параметры проектируемой сети для решения задачи распознавания цифр. Был осуществлен выбор алгоритма обучения нейронной сети и её программная реализация.

Построенная нейронная сеть продемонстрировала хорошие результаты при классификации изображений: при различных параметрах сети вероятность правильного отнесения к определенному классу была не меньше 87%.

Проведенный при обучении нейросети сравнительный анализ функций активации показал, что функция положительной срезки позволяет быстрее достичь высокой точности распознавания при различных параметрах сети даже несмотря на свой главный недостаток – проблему умирающего ReLU. Также по итогам вычислительных экспериментов установлено, что при использовании в качестве передаточной функции гиперболического тангенса увеличение количества скрытых слоев в диапазоне от 60 до 150 непременно приводит к увеличению точности распознавания.

Литература

1. Воронов И. В., Политов Е. А., Ефременко В. М. Обзор типов искусственных нейронных сетей и методов их обучения // Вестник Кузбасского государственного технического университета. – 2007. – № 3 (61). – С. 38–42.
2. Гапочкин А. В. Нейронные сети в системах распознавания речи // Science Time. – 2014. – № 1. – С. 29–36.
3. Горбачевская Е. Н. Классификация нейронных сетей // Вестник Волжского университета им. В. Н. Тагичева. – 2012. – № 2 (19). – С. 128–134.
4. Дороничева А. В., Савин С. З. Методы распознавания медицинских изображений для задач компьютерной автоматизированной диагностики // Современные проблемы науки и образования. – 2014. – № 4. – С. 623.
5. Макаренко А. А. Калайда В. Т. Применение нейросетевых технологий для идентификации изображения лица человека // Доклады Томского государственного университета систем управления и радиоэлектроники. 2005. – № 3 (11). – С. 41–48.
6. Макаркин Д. А., Федорина Н. А., Горшкова К. В. Машинное зрение в робототехнических системах // Фундаментальные и прикладные научные исследования: актуальные вопросы, достижения и инновации: сборник статей XXX Международной научно-практической конференции. В 2 ч. Ч. 1. (Пенза, 15 янв. 2020 г.). – Пенза: 2020. – С. 141–143.
7. Рудой Г. И. Выбор функции активации при прогнозировании нейронными сетями // Машинное обучение и анализ данных. – 2011. – № 1. – С. 16–39.
8. Савкин Л. В., Дмитриев В. Г., Федоров Е. А. Многослойные перцептроны в бортовых системах космической техники: аппаратные подходы на базе ПЛИС FPGA // Фундаментальные и прикладные космические исследования: сборник трудов 13-ой конференции молодых учёных (Москва, 13–15 апр. 2016 г.). – Москва, 2016. – С. 96–106.
9. Хашин С. И. Сравнение активаторных функций нейросети // Вестник Ивановского государственного университета. – 2020. – № 1. – С. 101–105.
10. G. S. S. Gomes, T. B. Ludermir, L. M. M. R. Lima Comparison of new activation functions in neural network for forecasting financial time series // Neural Computing and Applications. 2011. – № 20. – С. 417–439.

Статья поступила в редакцию: 19.05.2021; принята в печать: 08.11.2021.

Авторы прочитали и одобрили окончательный вариант рукописи.