

УДК 004.8

***АРХИТЕКТУРА, ФУНКЦИОНАЛ И НАЗНАЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ КАК ОДНОГО ИЗ ПРИМЕРОВ МАШИННОГО ОБУЧЕНИЯ***

***Ларин С.Э.***

*студент,*

*ФГБОУ ВО «Калужский государственный университет*

*им. К.Э. Циолковского»*

*Калуга, Россия*

***Белаш В.Ю.***

*к.пед.н., доцент,*

*ФГБОУ ВО «Калужский государственный университет*

*им. К.Э. Циолковского»*

*Калуга, Россия*

**Аннотация:** В представленной статье рассмотрены принципы функционирования нейронных сетей, их особенности и строение. В современном мире системы искусственного интеллекта стали популярны благодаря своим богатым возможностям и эффективности использования. В настоящее время существует множество задач, требующих практического применения интеллектуальных систем. Наиболее перспективными направлениями являются компьютерное зрение, диалоговые системы, анализ текста, робототехника, беспилотные автомобили и генерация контента: текста, изображений, музыки. Такое множество областей применения позволяет предположить, что интеллектуальные системы представляют собой уникальный набор для решения множества задач различной сложности при выполнении анализа и обработки большого объема данных. Тема данной научной работы обретает актуальность в свете современных возможностей использования нейронных сетей в образовании, что может привести к улучшению качества обучения, повышению эффективности учебных ресурсов

и созданию персонализированных образовательных маршрутов для обучающихся.

**Ключевые слова:** архитектура, информационные технологии, модель, нейронная сеть, перцептрон.

***ARCHITECTURE, FUNCTIONALITY AND PURPOSE OF NEURAL NETWORKS AS ONE OF THE EXAMPLES OF MACHINE LEARNING***

***Larin S.E.***

*student,*

*Kaluga State University named after K. E. Tsiolkovsky*

*Kaluga, Russia*

***Belash V.Yu.***

*Ph.D., Associate Professor,*

*Kaluga State University named after K. E. Tsiolkovsky*

*Kaluga, Russia*

**Abstract:** The article presents the principles of functioning of neural networks, their features and structure. In the modern world, artificial intelligence systems have become popular due to their rich capabilities and efficiency of use. Currently, many tasks require the practical application of intelligent systems. The most promising areas are computer vision, dialog systems, text analysis, robotics, self-driving cars and content generation: text, images, and music. Such a multitude of applications suggests that intelligent systems are a unique set for solving a variety of tasks of varying complexity when performing analysis and processing a large amount of data. The topic of this scientific work is becoming relevant in the light of the modern possibilities of using neural networks in education, which can lead to an improvement in the quality of education, increase the efficiency of educational resources and create personalized educational routes for students.

**Keywords:** architecture, information technology, model, neural network, perceptron.

С точки зрения биологии, центральная нервная система состоит из крайне обширной сети нейронов, приблизительно в количестве 10 миллиардов, которые взаимодействуют друг с другом через множество связей. Нейроны представляют собой особые клетки, способные формировать и передавать электрохимические импульсы.

Нейрон имеет бесчисленное количество ветвящихся проекций двух типов: дендрит и аксон. Дендриты выполняют роль входного канала нервного импульса, поступающих от других нейронов, которые вызывают определенное возбуждение клетки в ее теле размером от 3 до 100 микрон, а затем распространяются по аксону [4]. В итоге сотни миллиардов нейронов в мозге формируют сложную сеть, которая обеспечивает функционирование всех аспектов жизнедеятельности человека. Соединения между аксонами одной нейронной клетки и дендритами других клеток, осуществляемые через синапсы, служат для передачи сигналов. Когда нейрон активируется, происходит распространение электрического сигнала по его аксону. Данный сигнал может быть передан другим нейронам с помощью синапсов, что приводит к их активации.

Активация происходит в том случае, когда общий уровень сигналов, поступающих от дендритов к ядру, переходит на определенный уровень (порог активации) [6]. На рисунке 1 показано строение биологического нейрона.

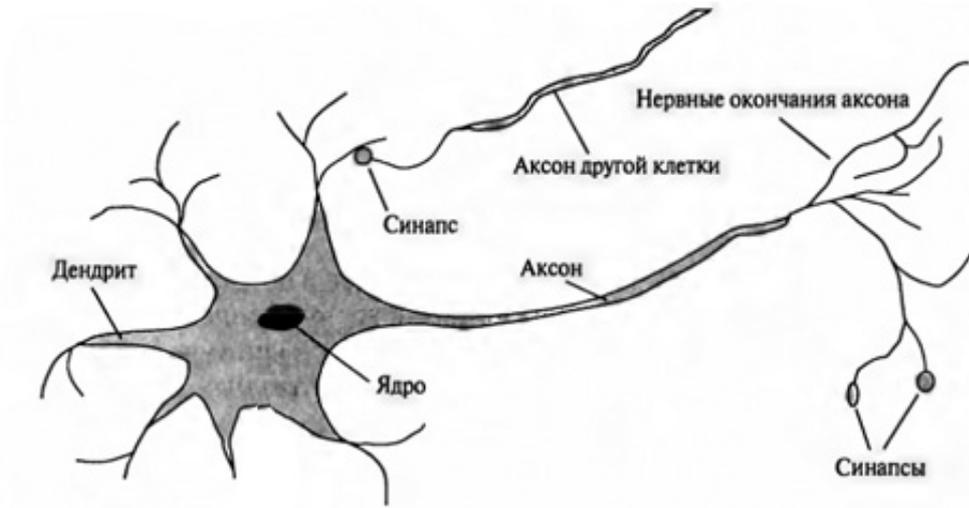


Рис. 1. Строение нейрона<sup>1</sup>

Модель нейрона представляет собой упрощенную схему биологического нейрона. Происходит получение входных сигналов от других нейронов, которые умножают на весовые коэффициенты. Суммарный выход этих умножений оценивает активность нейрона [1].

На рисунке 2 представлена модель искусственного нейрона.

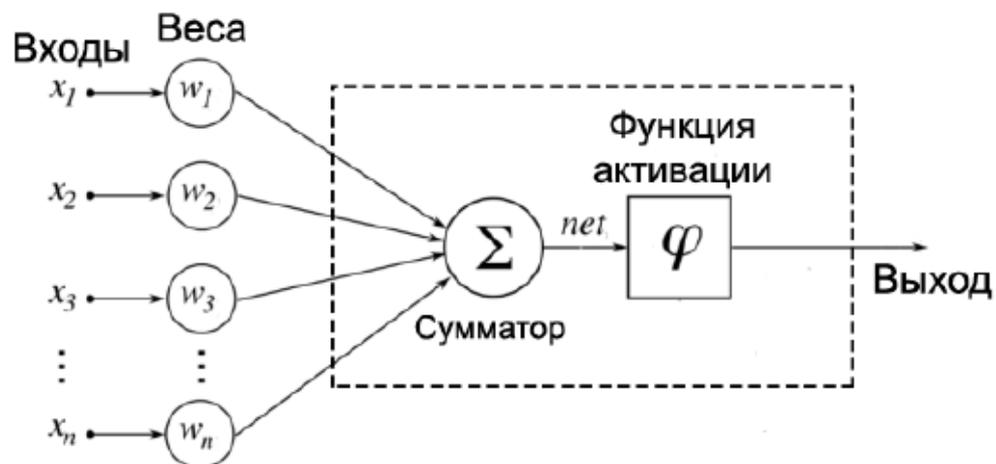


Рис. 2. Модель искусственного нейрона<sup>2</sup>

<sup>1</sup> составлено авторами

<sup>2</sup> составлено авторами

Входные сигналы умножаются на веса, а затем результаты суммируются. Полученное значение поступает на входные данные функции активации, которая трансформирует его в конечное значение.

После рассмотрения модели искусственного нейрона необходимо изучить одну из ключевых составляющих нейронных сетей, а именно их архитектуру. Архитектура нейронных сетей определяет, какие виды операций выполняются в сети, а также как эти операции выполняются.

Перейдем к рассмотрению основных моделей нейронных сетей.

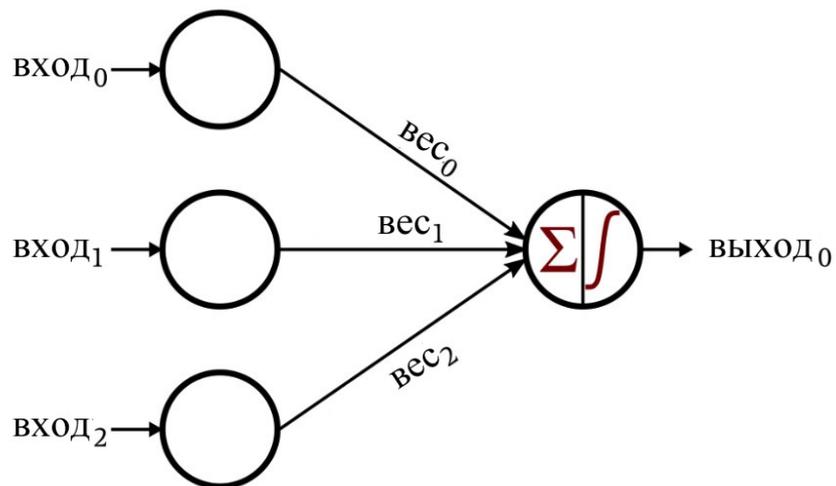
Перцептрон – это модель машинного обучения, которая может научиться разделять данные на два класса. Она была разработана Фрэнком Розенблаттом в 1957 году и является одной из первых нейронных сетей [2].

В перцептроне нейроны организованы в виде слоев. В простейшей форме слоистой сети присутствует входной слой исходных узлов, который проецируется непосредственно на выходной слой нейронов (вычислительные узлы), но не наоборот. Простыми словами, данная сеть является строго слоистой.

В случае, когда во входном и выходном слоях по четыре узла, сеть называется однослойной, причем обозначение "однослойная" относится к выходному слою, состоящему из вычислительных узлов (нейронов).

Входной слой исходных узлов не учитывается, так как в нем вычисления не производятся.

На рисунке 3 представлена архитектура простого перцептрона.

Рис.3. Архитектура перцептрона<sup>3</sup>

Перцептрон – линейный классификатор, это говорит о том, что он создает гиперплоскость для разделения данных на различные классы. Он может обучаться методом стохастического градиентного спуска, который позволяет ему улучшать свои предсказания на основе обратной связи, получаемой из примеров данных. Одним из преимуществ перцептрона является его простота и скорость обучения. Его можно использовать для таких задач как классификация текста, изображений и аудиосигналов.

Однако у перцептрона есть некоторые ограничения, такие как его неспособность решать линейно неразделимые задачи и склонность к переобучению большого набора данных.

Перцептрон, несмотря на свои ограничения, является важным алгоритмом машинного обучения, который используется в качестве основы для более сложных нейронных сетей. После изучения простого перцептрона рассмотрим многослойный перцептрон, который использует метод обратного распространения.

Многослойный перцептрон – это нейронная сеть, которая учится распознавать сложные взаимосвязи в данных. Она состоит из нескольких

---

<sup>3</sup> составлено авторами

слоев нейронов, каждый из которых вносит свой вклад в формирование конечного результата обработки информации. Веса, соединяющие нейроны, определяют, как информация будет передаваться от одного слоя к другому [3].

На рисунке 4 представлена архитектура многослойного перцептрона.

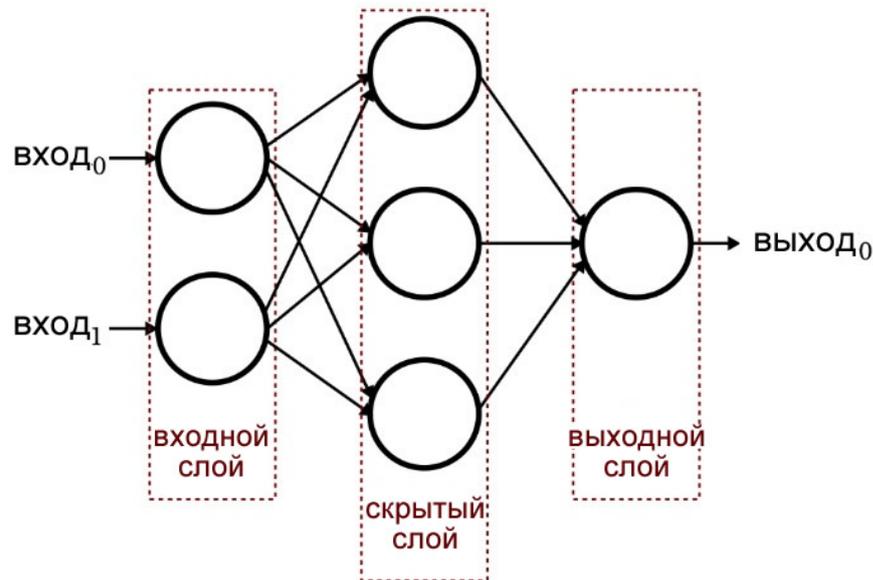


Рис. 4. Архитектура многослойного перцептрона<sup>4</sup>

Многослойный перцептрон – мощный инструмент, который может использоваться для решения сложных задач, требующих обучения и адаптации: распознавание образов, классификация, прогнозирование.

Свёрточные нейронные сети (CNN), благодаря использованию свертки, позволяют обрабатывать изображения с пространственной структурой, выявляя повторяющиеся узоры в них, позволяющую эффективно использовать структуру входных данных.

Свёрточные нейронные сети представляют собой инструмент компьютерного зрения, который используется для решения широкого круга задач, используемых для идентификации и категоризации изображений. Они

<sup>4</sup> составлено авторами

основаны на принципах работы зрительной системы человека и способны выявлять сложные закономерности в изображениях.

Основным мотивом для разработки нейронной сети послужило понимание Хьюбелом и Визелем работы зрительной коры кошки [5].

В архитектуре CNN каждый слой сети является трехмерным, который имеет пространственную протяженность и глубину, соответствующую количеству признаков.

Понятие глубины одного слоя в CNN отличается от понятия глубины по количеству слоев. Входной слой представляет собой изображение, которое разбивается на отдельные каналы, соответствующие различным цветам. Скрытые каналы представляют собой наборы признаков, которые кодируют различные типы объектов на изображении. Операция свертки используется для объединения признаков из соседних пикселей и создания новых признаков. Фильтр свертки представляет собой трехмерную матрицу весов, которая определяет, как будут объединены признаки [8].

Свёртка фильтра с пространственной областью входного изображения определяет значение скрытого состояния в следующем слое сверточной нейронной сети. Операция между фильтром и пространственными областями в слое выполняется во всех возможных позициях для определения следующего слоя (в котором активации сохраняют свои пространственные связи из предыдущего слоя). Все слои, кроме конечного набора из двух или трех слоев, сохраняют свою пространственную структуру. Поэтому можно представить, какие участки изображения влияют на те или иные активации в слое. Признаки в слоях нижнего уровня отражают линии или другие базовые формы, в то время как признаки в слоях более высокого уровня отражают более сложные формы

Конволюционные сети стали золотым стандартом для обработки изображений, поскольку они могут эффективно выявлять пространственные закономерности в данных. Они часто используются для распознавания

Дневник науки | [www.dnevniknauki.ru](http://www.dnevniknauki.ru) | СМИ Эл № ФС 77-68405 ISSN 2541-8327

иллюстраций и фотографий, обнаружении объектов и для обработки простого текста. Конволюционные нейронные сети представляют собой яркий пример того, что выбор архитектуры нейронной сети должен осуществляться с учетом семантических представлений о рассматриваемой области данных.

На рисунке 5 представлена архитектура свёрточной нейронной сети.

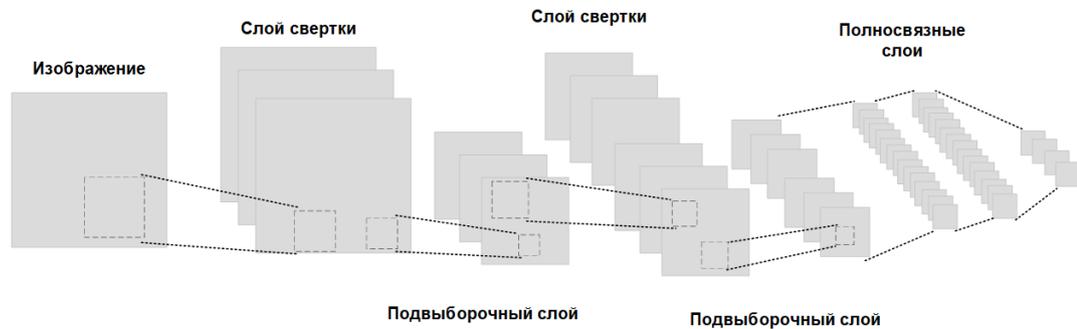


Рис. 5. Архитектура свёрточной нейронной сети<sup>5</sup>

Рекуррентные нейронные сети (RNN) – это тип нейронных сетей, которые могут осуществлять обработку данных, поступающих в виде последовательности. Они реализуют алгоритм, создавая петли, которые позволяют им запоминать информацию из предыдущих шагов и использовать ее для принятия решений на следующих шагах. Это делает их идеальными для задач, связанных с анализом данных, которые имеют временную зависимость, таких как распознавание речи, машинный перевод и прогнозирование временных рядов [7].

На рисунке 6 представлена архитектура рекуррентной нейронной сети.

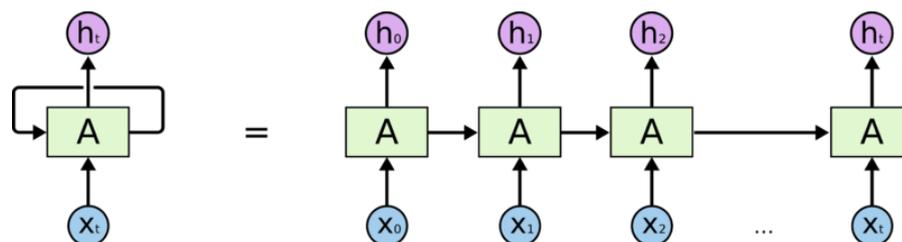


Рис. 6. Архитектура рекуррентной нейронной сети<sup>6</sup>

<sup>5</sup> составлено авторами

<sup>6</sup> составлено авторами

Важным моментом последовательности является то, что последовательные слова зависят друг от друга.

Весовые матрицы связей разделяются между несколькими связями во временной сети, чтобы обеспечить использование одной и той же функции на каждом временном интервале. Это совместное использование является ключом к специфическим знаниям, которые получает сеть.

Алгоритм обратного распространения (BP) – это метод обучения нейронных сетей, который работает путем корректировки весов сети в соответствии с ошибкой на выходе. BPTT – это особый тип BP, который учитывает совместное использование и временную протяженность при обновлении весов.

В области анализа временных рядов RNN также используются для таких задач, как предсказание цен на акции, температуру или давления.

Нейронные сети способны к обучению сложных закономерностей в данных без необходимости их явной формализации. Это позволяет им решать задачи, для которых сложно или невозможно разработать алгоритмы, основанные на знаниях о конкретной предметной области. Примерами таких задач являются распознавание образов и обработка естественного языка.

Однако, нейронные сети обладают и некоторыми недостатками. Одним из них является высокая вычислительная стоимость обучения, особенно при работе с большими массивами данных. Другим недостатком является сложность интерпретации, что затрудняет понимание того, как нейронные сети принимают решения.

### **Библиографический список**

1. Geron A. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. – O'Reilly Media, Inc., 2019. – 856 p. – P. 334.

2. Беспалов А. Обзор нейронных сетей // Habr. – URL: <https://habr.com/ru/articles/456186/> (дата обращения: 01.09.2023).
3. Биологический нейрон и его кибернетическая модель: сайт. – URL: <https://pandia.ru/text/78/046/86405.php> (дата обращения: 14.09.2023).
4. Горчаков А.В. Компьютерная интеллектуальная технология. Моделирование интеллектуальных процессов. – Томск: Изд-во Томского гос. пед. ун-та, 2005. – 460 с.
5. Лекция 3. Искусственные нейронные сети // Искусственный интеллект: Технологии и методы, Лекция 3, Интуит, 2018. – URL: <https://intuit.ru/studies/courses/88/88/lecture/20527?page=2> (дата обращения: 10.09.2023)
6. Нейронные сети, перцептрон // Викиконспекты, Информатика, НИИ ИТМО. – URL: <https://kurl.ru/gowuI> (дата обращения: 15.09.2023)
7. Рекуррентные нейронные сети // Викиконспекты, Информатика, НИИ ИТМО. – URL: <https://kurl.ru/rTQbE> (дата обращения: 12.09.2023).
8. Сверточные нейронные сети // Викиконспекты, Информатика, НИИ ИТМО. – URL: <https://kurl.ru/ZkFHI> (дата обращения: 06.09.2023)

*Оригинальность 87%*