

УДК 628.336

## МАТЕМАТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ НЕЙРОННЫХ СЕТЕЙ

Колодей Г.А., студент гр.133702

Белорусский государственный университет информатики и радиоэлектроники

г. Минск, Республика Беларусь

Чепикова В.В. – магистр технических наук

**Аннотация.** В докладе приводится пример математического представления нейронных сетей и их виды.

**Ключевые слова.** Нейронная сеть, математика, виды нейронных сетей, принцип работы нейронных сетей.

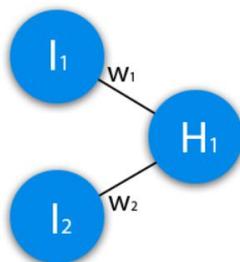
### Введение

Нейронные сети, также известные как искусственные нейронные сети (ИНС) или смоделированные нейронные сети (СНС), представляют собой подмножество алгоритмов машинного обучения и служат основой алгоритмов глубокого обучения. Концепция «нейронных сетей» возникла в попытке смоделировать процессы, происходящие в человеческом мозгу при передаче сигналов между биологическими нейронами.

Искусственные нейронные сети (ANN) состоят из образующих слоев узлов: слой входных данных, один или несколько скрытых слоев и слой выходных данных. Каждый узел (искусственный нейрон) связан с другими узлами с определенным весом и пороговым значением. Если вывод какого-либо узла превышает пороговое значение, то этот узел активируется и отправляет данные на следующий уровень сети. В противном случае данные на следующий уровень сети не передаются.

Для обучения и постепенного повышения точности нейронных сетей применяются обучающие данные. При достижении требуемой точности алгоритмы обучения превращаются в мощные инструменты для вычислений и [искусственного интеллекта](#), что позволяет использовать их для классификации и кластеризации данных с высокой скоростью. Задачи из области распознавания речи или изображений можно выполнить за несколько минут, а не за несколько часов, как при распознавании вручную. Одной из наиболее известных нейронных сетей является алгоритм поиска Google.

### Принцип работы нейронных сетей



$$1) H_{1_{input}} = (I_1 * w_1) + (I_2 * w_2)$$

$$2) H_{1_{output}} = f_{activation}(H_{1_{input}})$$

Рис.1

В этом примере (рис.1) изображена часть нейронной сети, где буквами I обозначены входные нейроны, буквой H — скрытый нейрон, а буквой w — веса. Из формулы видно, что входная информация — это сумма всевозможных входных данных, умноженных на которые соответствуют им веса. В это время дадим на вход 1 и 0. Пускай w<sub>1</sub>=0.4 и w<sub>2</sub> = 0.7. Входные данные нейрона H<sub>1</sub> станут следующими: 1\*0.4+0\*0.7=0.4. Сейчас когда у нас имеется входные

данные, мы можем получить выходные данные, подставив входное значение в функцию активации (подробнее о ней далее). Сейчас, когда у нас имеются выходные данные, мы передаем их дальше. И так, мы воспроизводим для всевозможных слоев, пока не дойдем до выходного нейрона. Запустив такую сеть в первый раз мы увидим, что ответ далек от верно, потому что сеть не натренирована. Дабы усовершенствовать итоги мы будем ее тренировать. Но прежде чем узнать как это делать, давайте введем немного терминов и свойств нейронной сети.

### **Виды нейронных сетей**

Нейронные сети можно разделить на несколько типов в зависимости от целевого назначения. Вот список наиболее распространенных типов нейронных сетей, имеющих практическое применение:

Перцептрон — первая нейронная сеть, созданная Фрэнком Розентблаттом в 1958 году.

Эта статья посвящена в основном нейронным сетям с прямой связью или многослойным перцептронам (MLP). Они состоят из следующих слоев: входной, один или несколько скрытых слоев и выходной. Хотя такие нейронные сети формально классифицируются как MLP, на самом деле они состоят из сигмовидных нейронов, а не перцептронов, поскольку большинство реальных проблем не линейны. Данные, поступающие в эти модели, используются для обучения; они лежат в основе алгоритмов компьютерного зрения, обработки естественного языка и других нейронных сетей.

Сверточные нейронные сети (CNN) похожи на сети с прямой связью, но обычно используются для распознавания изображений, обнаружения образов и/или компьютерного зрения. Для обнаружения закономерностей на изображениях с помощью таких сетей применяются законы линейной алгебры, в частности правила перемножения матриц.

Рекуррентные нейронные сети (RNN) включают обратную связь. Эти алгоритмы обучения в основном используются на данных временных рядов для прогнозирования будущих событий, таких как курсы акций на фондовых биржах или объем продаж.

### **Математическое представление нейронных сетей**

Под катом простое и лаконичное введение в математическое представление нейронных сетей для практиков, интересующихся теорией: от перцептрона до сети с двумя скрытыми слоями.

1. Однослойная нейронная сеть (перцептрон, рис.2)

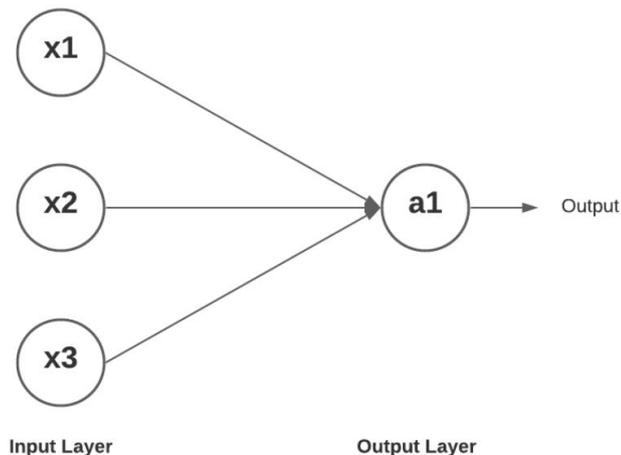


Рис.2

Вот как будет выглядеть математическое уравнение, чтобы получить значение a1 (выходной узел) как функцию входных данных x1, x2, x3.

$$a_1^{(2)} = g(\theta_{10}^{(1)} x_0 + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2 + \theta_{13}^{(1)} x_3)$$

В приведенном выше уравнении верхний индекс веса представляет слой, а нижний индекс веса представляет собой вес соединения между входными и выходными узлами. Таким образом, вес первого уровня представляет собой между узлом 1 в следующем слое и узлом 2 в текущей смене.

2. Нейронная сеть с одним открытым слоем (рис.3).

Ниже представлена нейронная сеть со скрытым слоем с тремя нейронами, входным слоем с тремя входными нейронами и выходным слоем с одним нейроном.

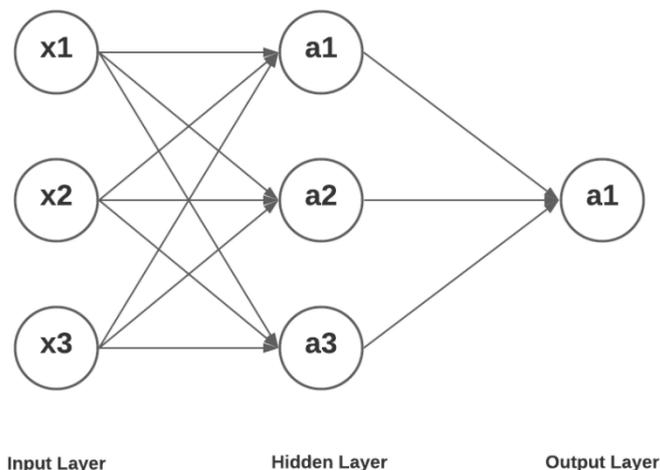


Рис.3

Вот как может выглядеть математическое уравнение (рис.4) для получения значений a1, a2 и a3 в слое 2 в зависимости от входных данных x1, x2, x3. Кроме того, значение a1 в слое 3 отображается как функция значений a1, a2 и a3 в слое 2.

Во-первых, давайте представим выходные значения, обработанные в трех скрытых нейронах скрытого слоя. Входной слой представлен как слой 1, скрытый слой — это слой 2, а выходной слой — это слой 3 (рис.5).

$$a_1^{(2)} = g(\theta_{10}^{(1)} x_0 + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2 + \theta_{13}^{(1)} x_3)$$

$$a_2^{(2)} = g(\theta_{20}^{(1)} x_0 + \theta_{21}^{(1)} x_1 + \theta_{22}^{(1)} x_2 + \theta_{23}^{(1)} x_3)$$

$$a_3^{(2)} = g(\theta_{30}^{(1)} x_0 + \theta_{31}^{(1)} x_1 + \theta_{32}^{(1)} x_2 + \theta_{33}^{(1)} x_3)$$

Рис.5

Давайте определим выходное значение узла в выходном слое. Значение представлено как функция  $a_1$ ,  $a_2$  и  $a_3$  в предыдущих узлах, которые могут быть представлены как значения  $x_1$ ,  $x_2$  и  $x_3$  во входном слое (рис.6).

$$a_1^{(3)} = g(\theta_{10}^{(2)} a_0^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)} + \theta_{13}^{(2)} a_3^{(2)})$$

Рис.6

3. Нейронная сеть с одним скрытым слоем (3 нейрона) и выходным слоем (2 нейрона)  
 Ниже представлена нейронная сеть со скрытым слоем с тремя нейронами, входным слоем с двумя входными нейронами и выходным слоем с двумя нейронами (рис.7).

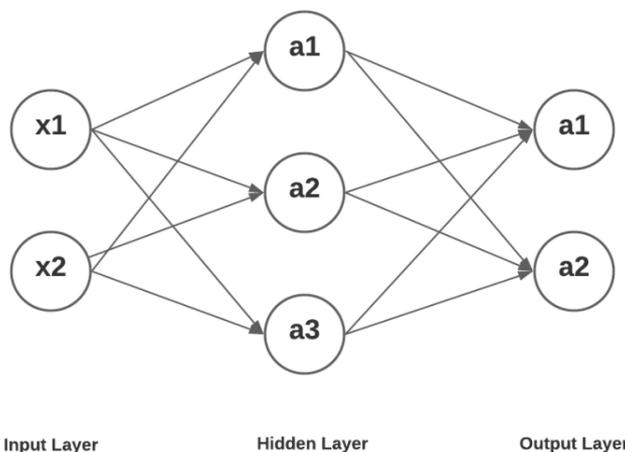


Рис.7

Вот как может выглядеть математическое уравнение для получения значения  $a_1$ ,  $a_2$ ,  $a_3$  в слое 2 как функции входных значений  $x_1$ ,  $x_2$ . Также значение  $a_1$ ,  $a_2$  в слое 3 представлено как функция значений  $a_1$ ,  $a_2$ ,  $a_3$  в слое 2.

Во-первых, давайте представим выходные значения, обработанные тремя скрытыми нейронами в скрытом слое. Входной слой представлен как слой 1, скрытый слой — как слой 2, а выходной слой — как слой 3.

$$a_1^{(2)} = g(\theta_{10}^{(1)} x_0 + \theta_{11}^{(1)} x_1 + \theta_{12}^{(1)} x_2)$$

$$a_2^{(2)} = g(\theta_{20}^{(1)} x_0 + \theta_{21}^{(1)} x_1 + \theta_{22}^{(1)} x_2)$$

$$a_3^{(2)} = g(\theta_{30}^{(1)} x_0 + \theta_{31}^{(1)} x_1 + \theta_{32}^{(1)} x_2)$$

Давайте определим выходное значение узлов в выходном слое. Значение представлено в предыдущих узлах как функция  $a_1$ ,  $a_2$ ,  $a_3$ , которые могут быть представлены во входном слое как значения  $x_1$ ,  $x_2$ ,  $x_3$ .

$$a_1^{(3)} = g(\theta_{10}^{(2)} a_0^{(2)} + \theta_{11}^{(2)} a_1^{(2)} + \theta_{12}^{(2)} a_2^{(2)} + \theta_{13}^{(2)} a_3^{(2)})$$

$$a_2^{(3)} = g(\theta_{20}^{(2)} a_0^{(2)} + \theta_{21}^{(2)} a_1^{(2)} + \theta_{22}^{(2)} a_2^{(2)} + \theta_{23}^{(2)} a_3^{(2)})$$

#### 4. Сети глубокого обучения с двумя скрытыми слоями.

Наконец, давайте посмотрим, как выходные значения узлов  $a_1$  в выходном слое могут быть математически выражены как функция входных данных  $x_1$ ,  $x_2$ . Вот схема сети глубокого обучения, которая имеет два скрытых слоя, один с тремя узлами, а другой с двумя узлами. Затем есть входной слой с двумя входными узлами и выходной слой с одним выходным узлом. Вот схема упрощенной сети глубокого обучения.

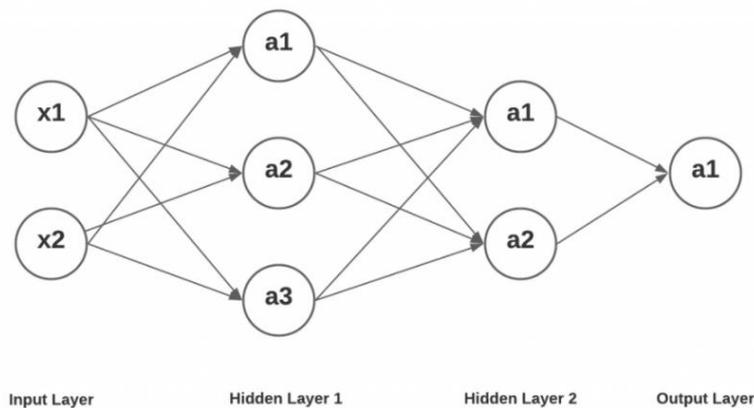


Рис.8

Значения в слое 2 ( $a_1$ ,  $a_2$ ,  $a_3$ ) и слое 3 ( $a_1$ ,  $a_2$ ) остаются такими же, как и в предыдущем разделе. Давайте построим значение 1 в выходном слое как функцию значений  $a_1$  и  $a_2$  в предыдущем слое (слой 3) (рис.8).

$$a_1^{(4)} = g(\theta_{10}^{(3)} a_0^{(3)} + \theta_{11}^{(3)} a_1^{(3)} + \theta_{12}^{(3)} a_2^{(3)})$$

## **Выводы**

Современные искусственные нейронные сети — это устройства, использующие огромное количество искусственных нейронов и связей между ними. Несмотря на то, что конечная цель развития нейронных сетей — полная имитация мыслительного процесса человека — не достигнута, они уже используются для решения многих задач обработки изображений, управления роботами и непрерывного производства, для понимания речи и синтеза, для диагностики заболеваний человека и технических неисправностей в машинах и устройствах, для прогнозирования курсов валют и др. Если мы перейдем к более прозаическому уровню, нейронные сети — это просто сети, состоящие из простых частей взаимосвязанных — формальных нейронов. В основе используемых концепций лежит идея о том, что нейроны можно моделировать довольно простыми автоматами и что вся сложность, гибкость и другие важные свойства мозга определяются связями между нейронами.

Краткое изложение о представлениях нейронных сетей в виде математических моделей:

1. Важно понимать обозначения, в которых вы будете представлять нейронную сеть как уравнение.
2. Первому или входному слою можно назначить номер 1, скрытому номер 2, а выходному
3. Весам между входным узлом в одном слое и узлом в следующем слое назначается верхний индекс — значение слоя, состоящего из входного узла. Нижний индекс веса состоит из двух чисел — числа, представляющего узел в следующем слое и номера входного узла.

## **Материал был взят на ресурсах:**

1. <https://habr.com/ru/post/254773/>
2. Книга «Введение в системы баз данных» Криса Дж. Дейта

UDC 628.336

## MATHEMATICAL REPRESENTATION OF NEURAL NETWORKS

Kolodey G.A.

*Belarusian State University of Informatics and Radioelectronics<sup>1</sup>, Minsk, Republic of Belarus*

*Chepikova V.V – Master of Technical Sciences*

**Annotation.** The report provides an example of the mathematical representation of neural networks and their types.

**Keywords.** Neural network, mathematics, types of neural networks, the principle of neural networks.