

ВИКИПЕДИЯ

# Искусственный нейрон

Материал из Википедии — свободной энциклопедии

**Иску́сственный нейро́н** (*математический нейрон Маккаллока — Питтса*, *формальный нейрон*<sup>[1]</sup>) — узел искусственной нейронной сети, являющийся упрощённой моделью естественного нейрона. Математически, искусственный нейрон обычно представляют как некоторую нелинейную функцию от единственного аргумента — линейной комбинации всех входных сигналов. Данную функцию называют *функцией активации*<sup>[2]</sup> или *функцией срабатывания*, *передаточной функцией*. Полученный результат посылается на единственный выход. Такие искусственные нейроны объединяют в сети — соединяют выходы одних нейронов с входами других. Искусственные нейроны и сети являются основными элементами идеального нейрокомпьютера.<sup>[3]</sup>

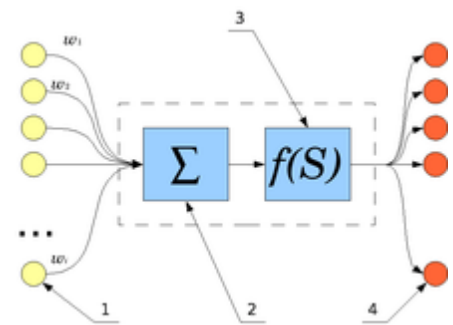


Схема искусственного нейрона

1. Нейроны, выходные сигналы которых поступают на вход данному
2. Сумматор входных сигналов
3. Вычислитель передаточной функции
4. Нейроны, на входы которых подаётся выходной сигнал данного
5.  $w_i$  — веса входных сигналов

## Содержание

### Биологический прототип

### История развития

### Связи между искусственными нейронами

### Математическая модель

### Передаточная функция нейрона

### Классификация нейронов

### Основные типы передаточных функций

Линейная передаточная функция

Пороговая передаточная функция

Сигмоидальная передаточная функция

Логистическая функция

Гиперболический тангенс

Модифицированный гиперболический тангенс

Радиально-базисная функция передачи

Другие функции передачи

### Стохастический нейрон

### Моделирование формальных логических функций

### См. также

## Примечания

## Литература

## Внешние ссылки

## Биологический прототип

Биологический нейрон состоит из тела диаметром от 3 до 100 мкм, содержащего ядро (с большим количеством ядерных пор) и другие органеллы (в том числе сильно развитый шероховатый ЭПР с активными рибосомами, аппарат Гольджи), и отростков. Выделяют два вида отростков. Аксон — обычно длинный отросток, приспособленный для проведения возбуждения от тела нейрона. Дендриты — как правило, короткие и сильно разветвлённые отростки, служащие главным местом образования влияющих на нейрон возбуждающих и тормозных синапсов (разные нейроны имеют различное соотношение длины аксона и дендритов). Нейрон может иметь несколько дендритов и обычно только один аксон. Один нейрон может иметь связи с 20 тысячами других нейронов. Кора головного мозга человека содержит около 80 миллиардов нейронов.

## История развития

Математическая модель искусственного нейрона была предложена У. Маккалоком и У. Питтсом вместе с моделью сети, состоящей из этих нейронов. Авторы показали, что сеть на таких элементах может выполнять числовые и логические операции<sup>[4]</sup>. Практически сеть была реализована Фрэнком Розенблаттом в 1958 году как компьютерная программа, а впоследствии как электронное устройство — перцептрон. Первоначально нейрон мог оперировать только с сигналами логического нуля и логической единицы<sup>[5]</sup>, поскольку был построен на основе биологического прототипа, который может пребывать только в двух состояниях — возбужденном или невозбужденном. Развитие нейронных сетей показало, что для расширения области их применения необходимо, чтобы нейрон мог работать не только с бинарными, но и с непрерывными (аналоговыми) сигналами. Такое обобщение модели нейрона было сделано Уидроу и Хоффом<sup>[6]</sup>, которые предложили в качестве функции срабатывания нейрона использовать логистическую кривую.

## Связи между искусственными нейронами

Связи, по которым выходные сигналы одних нейронов поступают на входы других, часто называют *синапсами* по аналогии со связями между биологическими нейронами. Каждая связь характеризуется своим *весом*. Связи с положительным весом называются *возбуждающими*, а с отрицательным — *тормозящими*<sup>[7]</sup>. Нейрон имеет один выход, часто называемый *аксоном* по аналогии с биологическим прототипом. С единственного выхода нейрона сигнал может поступать на произвольное число входов других нейронов.

## Математическая модель

Математически нейрон представляет собой взвешенный сумматор, единственный выход которого определяется через его входы и матрицу весов следующим образом:

$$y = f(u), \text{ где } u = \sum_{i=1}^n w_i x_i + w_0 x_0$$

Здесь  $x_i$  и  $w_i$  — соответственно сигналы на входах нейрона и веса входов, функция  $u$  называется индуцированным локальным полем, а  $f(u)$  — передаточной функцией. Возможные значения сигналов на входах нейрона считают заданными в интервале  $[0, 1]$ . Они могут быть либо дискретными (0 или 1), либо аналоговыми. Дополнительный вход  $x_0$  и соответствующий ему вес  $w_0$  используются для *инициализации* нейрона<sup>[8]</sup>. Под инициализацией подразумевается смещение активационной функции нейрона по горизонтальной оси, то есть формирование порога чувствительности нейрона<sup>[5]</sup>. Кроме того, иногда к выходу нейрона специально добавляют некую случайную величину, называемую сдвигом. Сдвиг можно рассматривать как сигнал на дополнительном, всегда нагруженном, синапсе.

## Передаточная функция нейрона

Передаточная функция  $f(u)$  определяет зависимость сигнала на выходе нейрона от взвешенной суммы сигналов на его входах. В большинстве случаев она является монотонно возрастающей и имеет область значений  $[-1, 1]$  или  $[0, 1]$ , однако существуют исключения. Также для некоторых алгоритмов обучения сети необходимо, чтобы она была непрерывно дифференцируемой на всей числовой оси<sup>[8]</sup>. Искусственный нейрон полностью характеризуется своей передаточной функцией. Использование различных передаточных функций позволяет вносить нелинейность в работу нейрона и в целом нейронной сети.

## Классификация нейронов

В основном, нейроны классифицируют на основе их положения в топологии сети. Разделяют:

- *Входные нейроны* — принимают исходный вектор, кодирующий входной сигнал. Как правило, эти нейроны не выполняют вычислительных операций, а просто передают полученный входной сигнал на выход, возможно, усилив или ослабив его;
- *Выходные нейроны* — представляют из себя выходы сети. В выходных нейронах могут производиться какие-либо вычислительные операции;
- *Промежуточные нейроны* — выполняют основные вычислительные операции<sup>[9]</sup>.

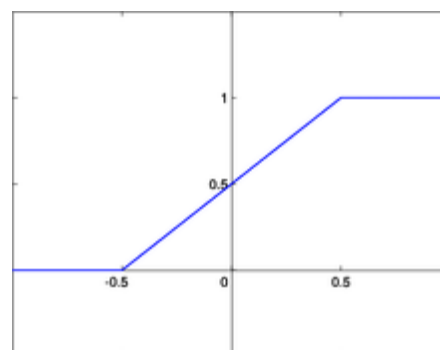
## Основные типы передаточных функций

### Линейная передаточная функция

Сигнал на выходе нейрона линейно связан со взвешенной суммой сигналов на его входе.

$$f(x) = tx,$$

где  $t$  — параметр функции. В искусственных нейронных сетях со слоистой структурой нейроны с передаточными функциями такого типа, как правило, составляют входной слой. Кроме простой линейной функции могут быть использованы её модификации. Например полулинейная функция (если её аргумент меньше нуля, то она равна нулю, а в остальных случаях, ведет себя



Линейная функция активации с насыщением

как линейная) или шаговая (линейная функция с насыщением), которую можно выразить формулой<sup>[10]</sup>:

$$f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq 1 \\ x & \text{else} \end{cases}$$

При этом возможен сдвиг функции по обеим осям (как изображено на рисунке).

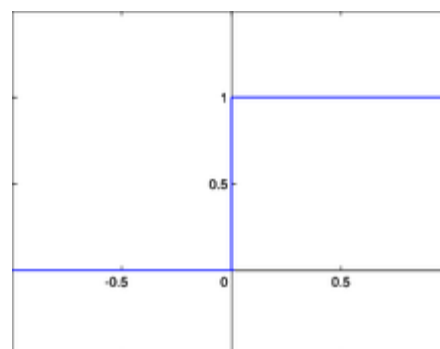
Недостатками шаговой и полулинейной активационных функций относительно линейной можно назвать то, что они не являются дифференцируемыми на всей числовой оси, а значит не могут быть использованы при обучении по некоторым алгоритмам.

## Пороговая передаточная функция

Другое название — функция Хевисайда. Представляет собой перепад. До тех пор пока взвешенный сигнал на входе нейрона не достигает некоторого уровня  $T$  — сигнал на выходе равен нулю. Как только сигнал на входе нейрона превышает указанный уровень — выходной сигнал скачкообразно изменяется на единицу. Самый первый представитель слоистых искусственных нейронных сетей — перцептрон<sup>[11]</sup> состоял исключительно из нейронов такого типа<sup>[5]</sup>. Математическая запись этой функции выглядит так:

$$f(x) = \begin{cases} 1 & \text{if } x \geq T \\ 0 & \text{else} \end{cases}$$

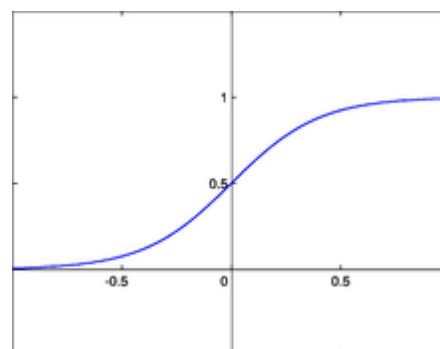
Здесь  $T = -w_0 x_0$  — сдвиг функции активации относительно горизонтальной оси, соответственно под  $x$  следует понимать взвешенную сумму сигналов на входах нейрона без учёта этого слагаемого. Ввиду того, что данная функция не является дифференцируемой на всей оси абсцисс, её нельзя использовать в сетях, обучающихся по алгоритму обратного распространения ошибки и другим алгоритмам, требующим дифференцируемости передаточной функции.



Пороговая функция активации

## Сигмоидальная передаточная функция

Один из самых часто используемых, на данный момент, типов передаточных функций. Введение функций сигмоидального типа было обусловлено ограниченностью нейронных сетей с пороговой функцией активации нейронов — при такой функции активации любой из выходов сети равен либо нулю, либо единице, что ограничивает использование сетей не в задачах классификации. Использование сигмоидальных функций позволило перейти от бинарных выходов нейрона к аналоговым<sup>[12]</sup>. Функции передачи такого типа, как правило, присущи нейронам, находящимся во внутренних слоях нейронной сети.



Сигмоидальная функция активации

## Логистическая функция

Математически логистическую функцию можно выразить так:

$$\sigma(x) = \frac{1}{(1 + \exp(-tx))}$$

Здесь  $t$  — это параметр функции, определяющий её *крутизну*. Когда  $t$  стремится к бесконечности, функция вырождается в пороговую. При  $t = 0$  сигмоида вырождается в постоянную функцию со значением 0,5. Область значений данной функции находится в интервале (0,1). Важным достоинством этой функции является простота её производной:

$$\frac{d\sigma(x)}{dx} = t\sigma(x)(1 - \sigma(x))$$

То, что производная этой функции может быть выражена через её значение, облегчает использование этой функции при обучении сети по алгоритму обратного распространения<sup>[13]</sup>. Особенностью нейронов с такой передаточной характеристикой является то, что они усиливают сильные сигналы существенно меньше, чем слабые, поскольку области сильных сигналов соответствуют пологим участкам характеристики. Это позволяет предотвратить насыщение от больших сигналов<sup>[14]</sup>.

## Гиперболический тангенс

Использование функции гиперболического тангенса

$$\operatorname{th}(Ax) = \frac{e^{Ax} - e^{-Ax}}{e^{Ax} + e^{-Ax}}$$

отличается от рассмотренной выше логистической кривой тем, что его область значений лежит в интервале (-1;1). Так как верно соотношение

$$\operatorname{th}\left(\frac{t}{2}x\right) = 2\sigma(x) - 1,$$

то оба графика отличаются лишь масштабом осей. Производная гиперболического тангенса, разумеется, тоже выражается квадратичной функцией значения; свойство противостоять насыщению имеет место точно также.

## Модифицированный гиперболический тангенс

Использование функции модифицированного гиперболического тангенса

$$\operatorname{mth} x = \frac{e^x - e^{-x}}{e^{ax} + e^{-bx}}, (a, b > 1),$$

масштабированной по оси ординат до интервала [-1;1] позволяет получить семейство сигмоидальных функций.

## Радиально-базисная функция передачи

Радиально-базисная функция передачи (RBF) принимает в качестве аргумента расстояние между входным вектором и некоторым наперед заданным центром активационной функции. Значение этой функции тем выше, чем ближе входной вектор к центру<sup>[15]</sup>. В качестве радиально-базисной можно, например, использовать функцию Гаусса:

$$y = \exp\left(-\frac{(S - R)^2}{2\sigma^2}\right).$$

Здесь  $S = ||\mathbf{X} - \mathbf{C}||$  — расстояние между центром  $\mathbf{C}$  и вектором входных сигналов  $\mathbf{X}$ . Скалярный параметр  $\sigma$  определяет скорость спада функции при удалении вектора от центра и называется *шириной окна*, параметр  $R$  определяет сдвиг активационной функции по оси абсцисс. Сети с нейронами, использующими такие функции, называются RBF-сетями. В качестве расстояния между векторами могут быть использованы различные метрики<sup>[16]</sup>, обычно используется евклидово расстояние:

$$S = \sqrt{\sum_{j=1}^N (x_j - c_j)^2}.$$

Здесь  $x_j$  —  $j$ -я компонента вектора, поданного на вход нейрона, а  $c_j$  —  $j$ -я компонента вектора, определяющего положение центра передаточной функции. Соответственно, сети с такими нейронами называются вероятностными и регрессионными<sup>[17]</sup>.

В реальных сетях активационная функция этих нейронов может отражать распределение вероятности какой-либо случайной величины, либо обозначать какие-либо эвристические зависимости между величинами.

## Другие функции передачи

Перечисленные выше функции составляют лишь часть от множества передаточных функций, используемых на данный момент. В число других передаточных функций входят такие как<sup>[18]</sup>:

- Экспонента  $f(x) = \exp(-Ax)$ ;
- Тригонометрический синус;
- Модульная:  $f(x) = |x|$ ;
- Квадратичная.

## Стохастический нейрон

Выше описана модель детерминистического искусственного нейрона, то есть состояние на выходе нейрона однозначно определено результатом работы сумматора входных сигналов. Рассматривают также стохастические нейроны, где переключение нейрона происходит с вероятностью, зависящей от индуцированного локального поля, то есть передаточная функция определена как:

$$f(u) = \begin{cases} 1 & \text{с вероятностью } P(u) \\ 0 & \text{с вероятностью } 1 - P(u) \end{cases},$$

где распределение вероятности  $P(u)$  обычно имеет вид сигмоида:

$$\sigma(u) = \frac{A(T)}{1 + \exp(-u/T)},$$

а нормировочная константа  $A(T)$  вводится для условия нормализации распределения вероятности  $\int_0^1 \sigma(u) du = 1$ . Таким образом, нейрон активируется с вероятностью  $P(u)$ .

Параметр  $T$  — аналог температуры (но не температуры нейрона) и определяет беспорядок в нейронной сети. Если  $T$  устремить к 0, стохастический нейрон перейдет в обычный нейрон с передаточной функцией Хевисайда (пороговой функцией).

## Моделирование формальных логических функций

Нейрон с пороговой передаточной функцией может моделировать различные логические функции.

$$f(u) = \begin{cases} 1 & \text{if } u \geq T \\ 0 & \text{else} \end{cases}, \text{ где } u = \sum_{i=1}^n w_i x_i + 0$$

Таблицы иллюстрируют, каким образом можно, задав веса входных сигналов и порог чувствительности, заставить нейрон выполнять конъюнкцию (логическое «И») и дизъюнкцию (логическое «ИЛИ») над входными сигналами, а также логическое отрицание входного сигнала<sup>[19]</sup>. Этих трех операций достаточно, чтобы смоделировать абсолютно любую логическую функцию любого числа аргументов.

<b>НЕ</b>	<b>T</b>	-1,0	
	<b>w</b>	-1,5	
	<b>x</b>	0	1
	<b>f</b>	1	0

<b>И</b>	<b>T</b>	1,5			
	<b>w<sub>1</sub></b>	1,0			
	<b>w<sub>2</sub></b>	1,0			
	<b>x<sub>1</sub></b>	0	0	1	1
	<b>x<sub>2</sub></b>	0	1	0	1
	<b>f</b>	0	0	0	1

или	<b><i>T</i></b>	0,5			
	<b><i>w<sub>1</sub></i></b>	1,0			
	<b><i>w<sub>2</sub></i></b>	1,0			
	<b><i>x<sub>1</sub></i></b>	0	0	1	1
	<b><i>x<sub>2</sub></i></b>	0	1	0	1
	<b><i>f</i></b>	0	1	1	1

## См. также

- Мемистор
- Мемристор
- Нейрон
- Связывающий нейрон
- Модель биологического нейрона

## Примечания

- Л. Г. Комарцова, А. В. Максимов «Нейрокомпьютеры», МГТУ им. Н. Э. Баумана, 2004 г., ISBN 5-7038-2554-7
- По аналогии с *нейронами активации*
- Миркес Е. М.*, Нейрокомпьютер. Проект стандарта. (<http://pca.narod.ru/MirkesNeurocomputer.htm>) — Новосибирск: Наука, 1999. — 337 с. ISBN 5-02-031409-9
- В статье *McCulloch W.S., Pitts W. A logical Calculus of Ideas Immanent in Nervous Activity* — Bull. Mathematical Biophysics, 1943 online (<http://www.springerlink.com/content/61446605110620kg>) (недоступная ссылка)
- Ясницкий, 2005, с. 29.
- В работе *Widrow B., Hoff M.E. Adaptive switching circuits*. 1960 IRE WESTCON Conferencion Record. — New York, 1960
- В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.11
- В. А. Терехов — Нейросетевые системы управления — с.12-13
- В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.14
- В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.12
- Очень часто в литературе можно встретить название *персептрон*
- Ясницкий, 2005, с. 34.
- CIT forum — Нейрокомпьютеры — архитектура и реализация ([http://www.citforum.ru/hardware/neurocomp/neurocomp\\_02.shtml](http://www.citforum.ru/hardware/neurocomp/neurocomp_02.shtml))
- В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.13
- Ясницкий, 2005, с. 77.
- В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.349



- .7. В. В. Круглов, В. В. Борисов — Искусственные нейронные сети. Теория и практика — с.348
- .8. Text
- .9. Ясницкий, 2005, с. 30.

## Литература

---

- *Терехов В.А., Ефимов Д.В., Тюкин И.Ю.* Нейросетевые системы управления. — 1-е. — Высшая школа, 2002. — С. 184. — ISBN 5-06-004094-1.
- *Круглов В. В., Борисов В. В.* Искусственные нейронные сети. Теория и практика. — 1-е. — М.: Горячая линия - Телеком, 2001. — С. 382. — ISBN 5-93517-031-0.
- *Каллан Р.* Основные концепции нейронных сетей = The Essence of Neural Networks First Edition. — 1-е. — «Вильямс», 2001. — С. 288. — ISBN 5-8459-0210-X.
- *Ясницкий Л. Н.* Введение в искусственный интеллект. — 1-е. — Издательский центр «Академия», 2005. — С. 176. — ISBN 5-7695-1958-4.
- *Комарцова Л. Г., Максимов А. В.* Нейрокомпьютеры (<http://www.books.ru/shop/search/advanced?as%5Bisbn%5D=5703819083&as%5Bsub%5D=%E8%F1%EA%E0%F2%FC>). — 1-е. — Изд-во МГТУ им. Н.Э. Баумана, 2002. — С. 320. — ISBN 5-7038-1908-3.
- *Савельев А. В.* Концепция многопроцессорного нейрона // Искусственный интеллект. Интеллектуальные и многопроцессорные системы. — Донецк-Таганрог- Минск, 2006. — С. 293-300.
- *Савельев А. В.* Нейрокомпьютеры в изобретениях // журнал «Нейрокомпьютеры: разработка, применение». (<http://www.radiotec.ru/catalog.php?cat=jr7>), Издательство "Радиотехника". — Москва, 2004. — № 2-3. — С. 33-49.
- *Lakhmi C. Jain; N.M. Martin* Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications. — CRC Press, CRC Press LLC, 1998 ([http://science-library.at.ua/load/sistemy\\_iskusstvennogo\\_intellekta/nechetkie\\_sistemy\\_i\\_nechetkaja\\_logika/fusion\\_of\\_neural\\_networks\\_fuzzy\\_systems\\_and\\_genetic\\_algorithms\\_industrial\\_applications/4-1-0-5](http://science-library.at.ua/load/sistemy_iskusstvennogo_intellekta/nechetkie_sistemy_i_nechetkaja_logika/fusion_of_neural_networks_fuzzy_systems_and_genetic_algorithms_industrial_applications/4-1-0-5))
- *Емельянов В. В., Курейчик В. В., Курейчик В. Н.* Теория и практика эволюционного моделирования. — М: Физматлит, 2003. (<http://science-library.at.ua/load/5-1-0-2>)

## Внешние ссылки

---

- Нейромодель RF-PSTH (симулирующая структуру рецептивного поля (РП) и выходной нейронный сигнал PSTH) ([http://neuroclusterbrain.com/ru/neuron\\_model\\_ru.html](http://neuroclusterbrain.com/ru/neuron_model_ru.html))

---

Источник — [https://ru.wikipedia.org/w/index.php?title=Искусственный\\_нейрон&oldid=103500579](https://ru.wikipedia.org/w/index.php?title=Искусственный_нейрон&oldid=103500579)

---

**Эта страница в последний раз была отредактирована 24 ноября 2019 в 08:40.**

Текст доступен по лицензии [Creative Commons Attribution-ShareAlike](#); в отдельных случаях могут действовать дополнительные условия.

Wikipedia® — зарегистрированный товарный знак некоммерческой организации [Wikimedia Foundation, Inc.](#)