УДК 004.31

ОБЗОР НЕКОТОРЫХ АКТИВАЦИОННЫХ ФУНКЦИИ МАТЕМАТИЧЕСКОГО НЕЙРОНА

Малышев А.Н., Данилов Е.А.
Пензенский государственный технологический университет, Пенза, e-mail: EvgenyDplus@mail.ru

В данной статье был проведен обзор некоторых активационных функций математического нейрона. Отмечено, что искусственные нейронные сети важное направление в современной науке позволяющее найти решения для многих задач. Приведены важные моменты в развитии искусственных или математических нейронов. Рассмотрена математическая модель нейрона и описана ее работа. Указана необходимость применения активационной функции на выходе математического нейрона для преобразования выходного сигнала. Перечислены некоторые активационные функции. Наиболее простая это функция единичного скачка. Часто применяется логистическая функция из семейства сигмоидальных функций. Последней описывается активационная функция гиперболический тангенс применяемая биологами и относящейся к семейству сигмоидальных функций. Для каждой из активационных функций приводится графическое изображение и математическая формула. Указаны преимущества и недостатки разных активационных функций. Рассмотрены основные свойства логистической функции. Приведено ее преимущество перед функцией единичного скачка. Отмечено что гиперболический тангенс обладает всеми полезными свойствами логистической функции. Указано что искусственные нейронные сети способны решать задачи до последнего времени считавшиеся трудно решаемыми или вообще не решаемыми на вычислительных машинах с классической архитектурой.

Ключевые слова: нейронные сети, искусственные нейронные сети, искусственный нейрон, математический нейрон, перцептрон, активационная функция, функция активации, единичный скачок, сигмоидальная функция, логистическая функция, гиперболический тангенс

OVERVIEW OF SOME ACTIVATION FEATURES MATH NEURON

Malyshev A.N., Danilov E.A.
Penza State Technological University, Penza, e-mail: EvgenyDplus@mail.ru

This article has reviewed some activation functions of a mathematical neuron. It is noted that artificial neural networks are an important direction in modern science allowing to find solutions for many problems. The important points in the development of artificial or mathematical neurons are given. The mathematical model of a neuron is considered and its work is described. The necessity of using the activation function at the output of a mathematical neuron to convert the output signal is indicated. Some activation features are listed. The simplest is the binary step function. Often, a logistic function from the family of sigmoidal functions is used. The latter describes the activation function of the hyperbolic tangent used by biologists and belonging to the family of sigmoidal functions. For each of the activation functions, a graphic image and a mathematical formula are provided. The advantages and disadvantages of different activation functions are indicated. The main properties of the logistic function are considered. Its advantage over the binary step function is given. It is noted that the hyperbolic tangent has all the useful properties of the logistic function. It is indicated that artificial neural networks are capable of solving problems that until recently were considered difficult to be solved or not solved at all on computers with classical architecture.

Keywords: neural networks, artificial neural networks, artificial neuron, mathematical neuron, perceptron, activation function, binary step, sigmoidal function, logistic function, hyperbolic tangent

Искусственные нейронные сети – это одно из направлений исследований в области искусственного интеллекта, основанное на попытках воспроизвести техническими средствами свойства нервной системы человека.

Нейрон Маккаллока–Питтса был ранней моделью функционирования мозга. Эта линейная модель могла распознавать две категории выходов, проверяя, является значение $f(x, w)$ положительным или отрицательным. Конечно, чтобы модель соответствовала желаемому определению категорий, нужно было правильно подобрать веса. Веса задавал человек. В 1950-е годы был изобретен перцептрон – первая модель, которая могла в процессе обучения находить веса, определяющие категории, имея примеры входных данных из каждой категории. Модель адаптивного линейного элемента, относящаяся примерно к тому же времени, просто возвращала само значение $f(x)$ для предсказания вещественного числа и также могла обучаться предсказанию чисел на данных [1].

Искусственный нейрон является структурной единицей искусственной нейронной сети и представляет собой аналог биологического нейрона.

С математической точки зрения искусственный нейрон – это сумматор всех входящих сигналов, применяющий к полученной взвешенной сумме некоторую простую, в общем случае, нелинейную функцию, непрерывную на всей области определения [2] (рис. 1). Обычно, данная функция монотонно возрастает. Полученный результат посылается на единственный выход.

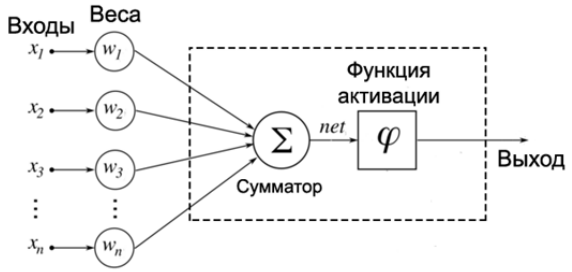

Рис. 1. Математический нейрон

Искусственные нейроны объединяются между собой определенным образом, образуя искусственную нейронную сеть. Каждый нейрон характеризуется своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Он обладает группой синапсов – однонаправленных входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал поступает на синапсы следующих нейронов.

Каждый синапс характеризуется величиной синаптической связи или ее весом w_i, который является эквивалентом электрической проводимости биологических нейронов [3].

Текущее состояние нейрона определяется, как взвешенная сумма (*net*) его входов x_i.

$$net = \sum_{i=1}^{n} x_i w_i$$

Просто так подавать взвешенную сумму на выход достаточно бессмысленно. Нейрон должен как-то обработать ее и сформировать адекватный выходной сигнал. Именно для этих целей и используют функцию активации.

Она преобразует взвешенную сумму в число, которое и является выходом нейрона (выход нейрона обозначим переменной *out*).

Для разных типов искусственных нейронов используют самые разные функции активации. В общем случае их обозначают символом ϕ(*net*). Указание взвешенного сигнала в скобках означает, что функция активации принимает взвешенную сумму как параметр [3].

Активационная функция (Activation function) (ϕ(*net*)) – функция, принимающая взвешенную сумму как аргумент. Значение этой функции и является выходом нейрона (*out*).

$$out = \phi(net).$$

Самыми известными функциями активации являются:

– функция единичного скачка;

– семейство сигмоидальных или логистических функций;

– гиперболический тангенс.

Самый простой вид функции активации. Выход нейрона может быть равен только 0 или 1. Если взвешенная сумма больше определенного порога *b*, то выход нейрона равен 1. Если ниже, то 0.

Запишем эту функцию активации математически:

$$out(net) = \begin{cases} 0, \ net < b \\ 1, \ net \geq b \end{cases}.$$

Выход нейрона (*out*) зависит от взвешенной суммы (*net*) следующим образом: если *net* меньше порога (*b*), то *out* равен 0. А если *net* больше или равен порогу *b*, то *out* равен 1.

Графически эту функцию активации можно изобразить следующим образом (рис. 2).

На горизонтальной оси расположены величины взвешенной суммы. На вертикальной оси – значения выходного сигнала. Как легко видеть, возможны только два значения выходного сигнала: 0 или 1. Причем 0 будет выдаваться всегда от минус бесконечности и вплоть до некоторого значения взвешенной суммы, называемого порогом. Если взвешенная сумма равна порогу или больше него, то функция выдает 1.

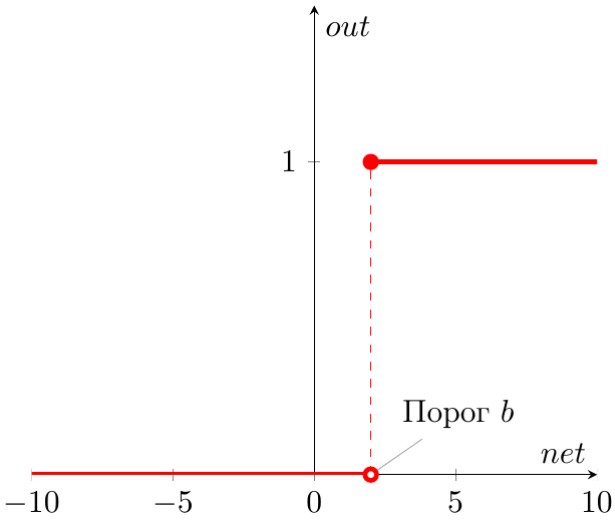

Рис. 2. Функция единичного скачка

Существует целое семейство сигмоидальных функций, некоторые из которых применяют в качестве функции активации в искусственных нейронах [4]. Все эти функции обладают некоторыми очень полезными свойствами, ради которых их и применяют в нейронных сетях. Самая часто используемая в нейронных сетях сигмоида – логистическая функция (рис. 3).

График этой функции представляет некоторое подобие латинской буквы S, откуда и пошло название семейства этих функций.

Математически активационная функция описывается:

$$out(net) = \frac{1}{1 + \exp(-a \cdot net)}.$$

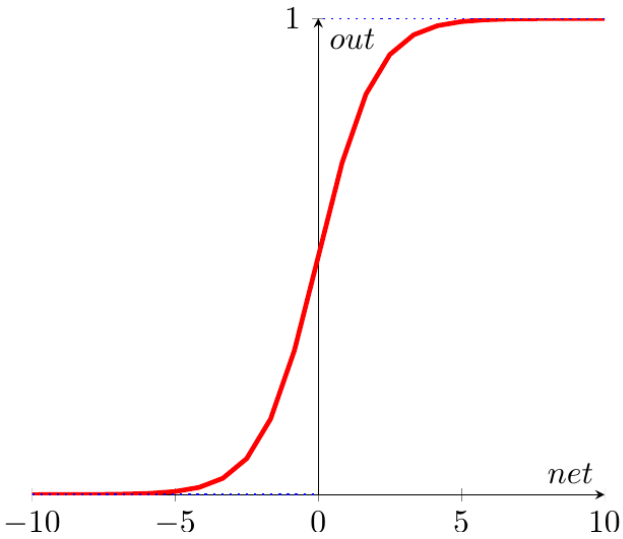

Рис. 3. Логистическая функция

Параметр *a* – это число, которое характеризует степень крутизны функции. На рис. 4 представлены логистические функции с разными значениями *a*.

Именно из-за этих свойств логистическая функция чаще всего используются в качестве функции активации в искусственных нейронах.

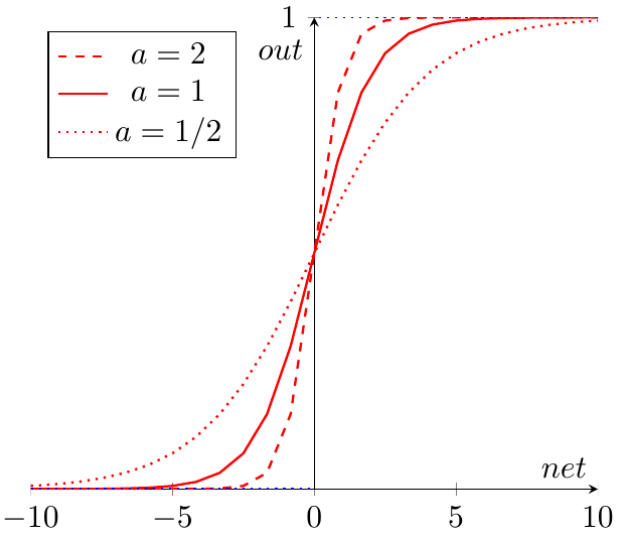

Рис. 4. Логистическая функция с разными значениями параметра a

Значение, получаемое при использование логистической функции с разными значениями параметра *a* в качестве функции активации, будет между 0 и 1. Причем чем больше взвешенная сумма, тем ближе выход будет к 1 (но никогда не будет точно ей равен). И наоборот, чем меньше взвешенная сумма, тем ближе выход нейрона будет к 0.

Свойства логистической функции:

– она является «сжимающей» функцией, то есть вне зависимости от аргумента (взвешенной суммы), выходной сигнал всегда будет в пределах от 0 до 1;

– она более гибкая, чем функция единичного скачка – ее результатом может быть не только 0 и 1, но и любое число между ними;

– во всех точках она имеет производную, и эта производная может быть выражена через эту же функцию.

Есть еще одна сигмоида – гиперболический тангенс изображена на рисунке 5. Он применяется в качестве функции активации биологами для более реалистичной модели нервной клетки [5].

Такая функция позволяет получить на выходе значения разных знаков (например, от –1 до 1), что может быть полезным для ряда сетей.

Функция записывается следующим образом:

$$out(net) = \mathrm{th}\left(\frac{net}{a}\right).$$

Параметр *a* также определяет степень крутизны графика этой функции.

Гиперболический тангенс обладает всеми полезными свойствами, которые имеет и логистическая функция.

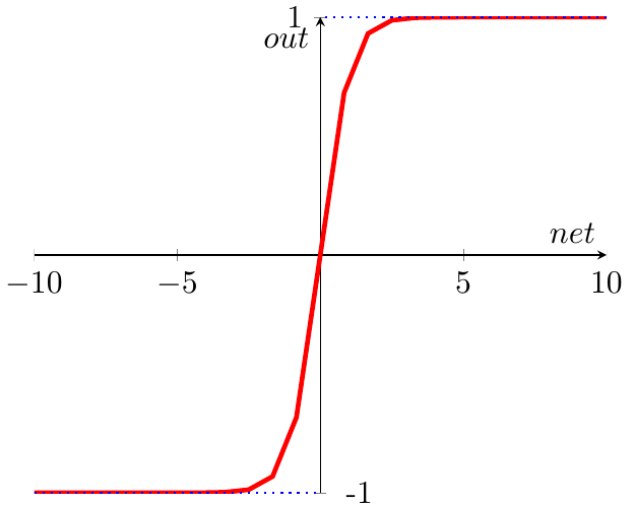

Рис. 5. Гиперболический тангенс

Математическая модель искусственного нейрона с *n* входами:

$$out = \phi\left(\sum_{i=1}^{n} x_i w_i\right),$$

где ϕ – функция активации, взвешенная сумма, как сумма *n* произведений входных сигналов на соответствующие веса:

$$\sum_{i=1}^{n} x_i w_i .$$

Перспективы развития искусственных нейронных сетей обусловлены тем, что даже на существующем уровне внутренней организации искусственные нейронные сети способны справляться с рядом задач, считавшихся до последнего времени трудными или вообще не поддающимися решению при помощи компьютеров традиционной архитектуры.

Список литературы

1. Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с анг. А.А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.

2. Искусственный нейрон. Структура искусственного нейрона [Электронный ресурс]. – https://neuronus.com/theory/nn/150–iskusstvennyj-nejron.html (дата обращения 11.02.2019).

3. Радько П. Электронный учебник [Электронный ресурс]. – https://neuralnet.info/book/ (дата обращения 11.02.2019).

4. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение. – СПб.: Питер, 2018. – 480 с.

5. Аксенов С.В., Новосельцев В.Б. Организация и использование нейронных сетей (методы и технологии) / под общ. ред. В.Б. Новосельцева. – Томск: Изд-во НТЛ, 2006. – 128 с.