В. С. РОСТОВЦЕВ

ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

Учебник

Издание второе, стереотипное

ЛАНЬ

САНКТ-ПЕТЕРБУРГ
МОСКВА
КРАСНОДАР
2021

УДК 004.7
ББК 32.973я73

Р 78 **Ростовцев В. С.** Искусственные нейронные сети : учебник для вузов / В. С. Ростовцев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 216 с. : ил. — Текст : непосредственный.

ISBN 978-5-8114-7462-2

В учебнике приведены основные теоретические и практические сведения по разработке, обучению и применению искусственных нейронных сетей с использованием среды MatLab.

Учебник предназначен для студентов магистратуры направления «Информатика и вычислительная техника» и может быть полезен студентам других специальностей при изучении нейросетевых технологий, а также для слушателей курсов повышения квалификации и профессиональной переподготовки.

УДК 004.7
ББК 32.973я73

Обложка
Е. А. ВЛАСОВА

© Издательство «Лань», 2021
© В. С. Ростовцев, 2021
© Издательство «Лань»,
художественное оформление, 2021

2. ОДНОСЛОЙНАЯ НЕЙРОННАЯ СЕТЬ

2.1. Формальная модель нейрона

Искусственный нейрон имитирует в первом приближении свойства биологического нейрона. На вход искусственного нейрона поступает некоторое множество сигналов, каждый из которых является выходом другого нейрона. Каждый вход умножается на соответствующий вес, аналогичный синаптической силе, и все произведения суммируются, определяя уровень активации нейрона.

На рисунке 2.1 представлена модель, реализующая эту идею. Хотя сетевые парадигмы весьма разнообразны, в основе почти всех их лежит эта конфигурация.

Здесь множество входных сигналов, обозначенных x_1, x_2, …, x_n, поступает на искусственный нейрон. Эти входные сигналы, в совокупности обозначаемые вектором **X**, соответствуют сигналам, приходящим в синапсы биологического нейрона. Каждый сигнал умножается на соответствующий вес w_1, w_2, …, w_n, и поступает на суммирующий блок, обозначенный символом $\sum$.

Каждый вес соответствует «силе» одной биологической синаптической связи. Множество весов в совокупности обозначается вектором **W**. Суммирующий блок, соответствующий телу биологического элемента, складывает взвешенные входы алгебраически, создавая выход, который мы будем называть S. В векторных обозначениях это может быть компактно записано в виде произведения векторов **S = X W**.

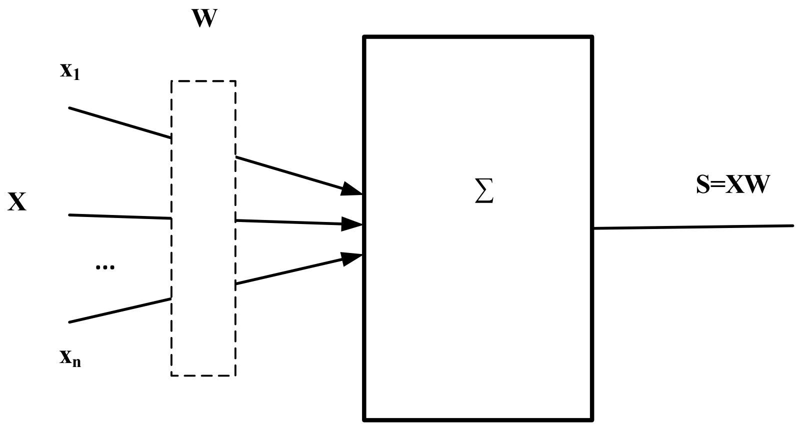

Рис. 2.1
Искусственный нейрон

2.2. Активационные функции

Сигнал S далее, как правило, преобразуется активационной функцией F и дает выходной нейронный сигнал OUT. Активационная функция может быть обычной линейной функцией

$$OUT = K(S), \qquad (2.1)$$

где K — постоянная, пороговой функции; $OUT = 1$, если $S > T$, $OUT = 0$ в остальных случаях; T — некоторая постоянная пороговая величина, или же функция, более точно моделирующая нелинейную передаточную характеристику биологического нейрона.

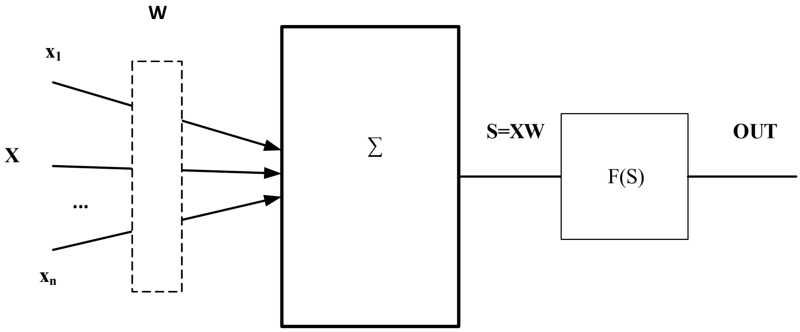

Рис. 2.2
Архитектура искусственного нейрона

На рисунке 2.2 блок, обозначенный $F(S)$, принимает сигнал S и выдает сигнал OUT. Если блок F сужает диапазон изменения величины S так, что при любых значениях S значения OUT принадлежат некоторому конечному интервалу, то F называется *«сжимающей» функцией.* В качестве «сжимающей» функции часто используется логистическая или «сигмоидальная» (S-образная) функция, показанная на рисунке 2.3. Эта функция математически выражается следующим образом:

$$OUT(S) = \frac{1}{1 + \exp(-S)}. \qquad (2.2)$$

По аналогии с электронными системами активационную функцию можно считать нелинейной усилительной характеристикой искусственного нейрона. Коэффициент усиления вычисляется как отношение приращения величины OUT к вызвавшему его небольшому приращению величины S. Он выражается наклоном кривой при определенном уровне возбуждения и изменяется от малых значений при больших отрицательных возбуждениях (кривая почти горизонтальна) до максимального значения при нулевом возбуждении и снова уменьшается, когда возбуждение становится большим положительным. Гроссберг (1973) обнаружил, что подобная нелинейная характеристика решает поставленную им дилемму шумового насыщения. Каким образом одна и та же сеть может обрабатывать как слабые, так и сильные сигналы? Слабые сигналы нуждаются в большом сетевом усилении, чтобы дать пригодный к использованию выходной сигнал. Однако усилительные каскады с большими коэффициентами усиления могут привести к насыщению выхода шумами усилителей (случайными флуктуациями), которые присутствуют в любой физически реализованной сети. Сильные входные сигналы в свою очередь также будут приводить к насыщению усилительных каскадов, исключая возможность полезного использования выхода. Центральная область логистической функции, имеющая большой коэффициент усиления, решает проблему обработки слабых сигналов, в то время как области с падающим усилением на положительном и отрицательном концах подходят для больших возбуждений. Таким образом, нейрон функционирует с большим усилением в широком диапазоне уровня входного сигнала:

$$F(S) = \frac{1}{1+\exp(-S)} . \tag{2.3}$$

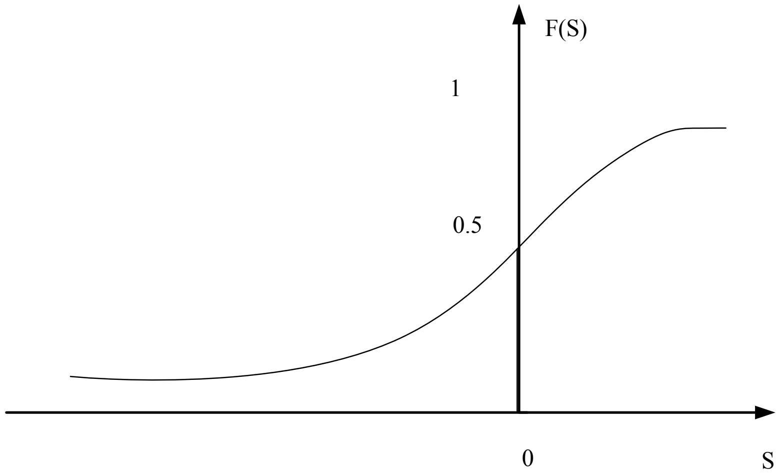

Рис. 2.3
Сигмоидальная логистическая функция

Другой широко используемой активационной функцией является гиперболический тангенс. По форме она сходна с логистической функцией и часто используется биологами в качестве математической модели активации нервной клетки. В качестве активационной функции искусственной нейронной сети она записывается следующим образом:

$$F(S) = th(S). \tag{2.4}$$

Подобно логистической функции гиперболический тангенс является S-образной функцией, но он симметричен относительно начала координат, и в точке $S = 0$ значение выходного сигнала OUT равно нулю (см. рис. 2.4).

В отличие от логистической функции гиперболический тангенс принимает значения различных знаков, что оказывается выгодным для ряда нейронных сетей.

Рассмотренная простая модель искусственного нейрона игнорирует многие свойства своего биологического двойника. Например, она не принимает во внимание задержки во времени, которые воздействуют на динамику системы. Входные сигналы сразу же порождают выходной сигнал. Она не учитывает воздействий функции частотной модуляции или синхронизирующей функции биологического нейрона, которые ряд исследователей считают решающими.

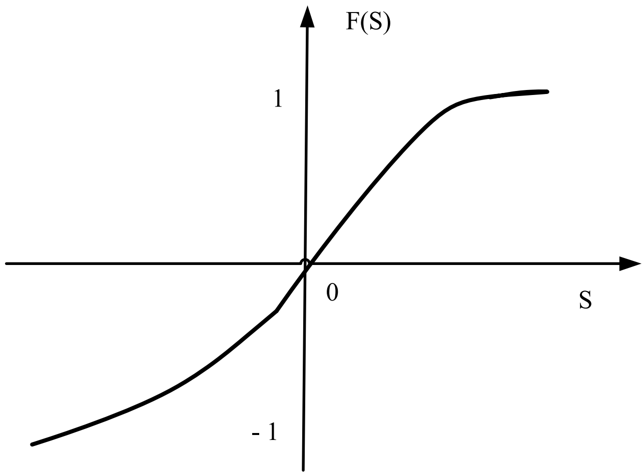

Рис. 2.4
Функция гиперболического тангенса

Несмотря на эти ограничения, сети, построенные из этих нейронов, обнаруживают свойства, сильно напоминающие биологическую систему.

Только время и исследования смогут ответить на вопрос, являются ли подобные совпадения случайными или следствием того, что в модели верно отражены важнейшие черты биологического нейрона.

2.3. Однослойные искусственные нейронные сети

Хотя один нейрон и способен выполнять простейшие процедуры распознавания, сила нейронных вычислений проистекает от соединений нейронов в сетях. Простейшая сеть состоит из группы нейронов, называемых однослойным персептроном, представлена на рисунке 2.5.

Отметим, что вершины-круги слева служат лишь для распределения входных сигналов. Они не выполняют каких-либо вычислений и поэтому не будут считаться входным слоем. По этой причине они обозначены кругами, чтобы отличать их от вычисляющих нейронов, обозначенных квадратами. Каждый элемент из множества входов **X** отдельным весом соединен с каждым искусственным нейроном.

Каждый нейрон выдает взвешенную сумму входов в сеть. В искусственных и биологических сетях многие соединения могут отсутствовать, все соединения показаны в целях общности. Могут иметь место также соединения между выходами и входами элементов в слое.

Удобно считать веса элементами матрицы **W**. Матрица имеет m строк и n столбцов, где m — число входов, а n — число нейронов. Например, $w_{2,3}$ — это вес, связывающий третий вход со вторым нейроном.

Таким образом, вычисление выходного вектора **N**, компонентами которого являются выходы *OUT* нейронов, сводится к матричному умножению **N** = **XW**, где **N** и **X** — векторы-строки.