### Лекция 1

### НЕЙРОНЫ И ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

Изучение и использование искусственных нейронных сетей, в принципе, началось уже достаточно давно – в начале 20 века, но по настоящему широкую известность они получили несколько позже. Связано это, в первую очередь, с тем, что стали появляться продвинутые (для того времени) вычислительные устройства, мощности которых были достаточно велики для работы с искусственными нейронными сетями. По сути, на данный момент можно легко смоделировать нейронную сеть средней сложности на любом персональном

компьютере.

Нейронная сеть представляет из себя совокупность нейронов, соединенных друг с другом определенным образом. Рассмотрим один нейрон:

**Нейрон** представляет из себя элемент, который вычисляет выходной сигнал (по определенному правилу) из совокупности входных сигналов. То есть основная последовательность действий одного нейрона такая:

* + - * Прием сигналов от предыдущих элементов сети
			* Комбинирование входных сигналов
			* Вычисление выходного сигнала
			* Передача выходного сигнала следующим элементам нейронной

сети

Между собой нейроны могут быть соединены абсолютно по-разному, это определяется структурой конкретной сети. Но суть работы нейронной сети остается всегда одной и той же. По совокупности поступающих на вход сети сигналов на выходе формируется выходной сигнал (или несколько выходных сигналов). То есть нейронную сеть упрощенно можно представить в виде черного ящика, у которого есть входы и выходы. А внутри этого ящика сидит

огромное количество нейронов

Мы перечислили основные этапы работы сети, теперь давайте остановимся на каждом из них в отдельности.

###

### Комбинирование входных сигналов

Поскольку к каждому нейрону могут приходить несколько входных сигналов, то при моделировании нейронной сети необходимо задать определенное правило комбинирования всех этих сигналов. И довольно-таки часто используется правило суммирования взвешенных значений связей. Что значит взвешенных? Сейчас разберемся…

Каждую связь в сети нейронов можно полностью охарактеризовать при помощи трех факторов:

* + - * первый – элемент, от которого исходит связь
			* второй – элемент, к которую связь направлена
			* третий – вес связи.

Сейчас нас в большей степени интересует именно третий фактор. Вес связи определяет, будет ли усилен или ослаблен сигнал, передаваемый по данной связи. Если объяснять просто, “на пальцах”, то давайте рассмотрим такой пример:



Выходной сигнал нейрона 1 равен 5. Вес связи между нейронами равен 2. Таким образом, чтобы определить входной сигнал нейрона 2, приходящий от нейрона 1, необходимо умножить значение этого сигнала на вес связи (5\*2).

А если сигналов много, то они все суммируются. В итоге на входе нейрона мы получаем следующее:

𝑛𝑒𝑡𝑗 = ∑𝑁 𝑥𝑖 ∗ 𝑤𝑖𝑗

𝑖=1

В этой формуле 𝑛𝑒𝑡𝑗 – это результат комбинирования всех входных сигналов для нейрона j (комбинированный ввод нейрона). N – количество элементов, передающих свои выходные сигналы на вход сигнала j. А 𝑤𝑖𝑗– вес связи, соединяющей нейрон i с нейроном j. Суммируя все взвешенные входные сигналы, мы получаем комбинированный ввод элемента сети.

Чаще всего структура связей между нейронами представляется в виде

матрицы W, которую называют весовой матрицей. Элемент матрицы 𝑤𝑖𝑗, как и в формуле, определяет вес связи, идущей от элемента i к элементу j. Для того, чтобы понять, как составляются весовые матрицы, давайте рассмотрим простую нейронную сеть:



**Весовая матрица** такой нейронной сети будет иметь следующий вид:

𝑊 =

0

1.2

𝖥

0

−0.4

0

0

−5.6

3

0

4.7

0

−3.3

0

2.41

7.5

I−2.4 0 0 0 2.5I

[ 0 0 0 0 0 ]

Например, от второго элемента к третьему идет связь, вес которой равен

1. Смотрим на матрицу, вторая строка, третий столбец – число 3, все верно.

### Функция активации элемента.

Рассмотрим выходные сигналы. Для каждого элемента сети имеется определенное правило, в соответствии с которым из значения комбинированного ввода элемента вычисляется его выходное значение. Это правило называется функцией активации. А само выходное значение называется активностью нейрона. В роли функций активации могут выступать абсолютно любые математические функции, приведем в качестве примера несколько из наиболее часто использующихся:

* + пороговая функция – если значение комбинированного ввода ниже определенного значения (порога), то активность равна нулю, если выше – единице.
	+ логистическая функция.

Рассмотрим еще один небольшой пример, который очень часто используется в литературе для объяснения сути работы нейронных сетей.

Задача примера заключается в том, чтобы при помощи нейронной сети вычислить отношение XOR. То есть на вход мы будем подавать разные варианты сигналов, а на выходе должны получить результат операции XOR для поданных на вход значений:



Элементы 1 и 2 являются входными, а элемент 7 – выходным. Нейроны 5 и 6 называются скрытыми, поскольку они не связаны с внешней средой. Таким образом, мы получили три слоя – входной, скрытый и выходной. Элементы 3 и 4 называют элементами смещения. Их выходной сигнал (активность) всегда равен 1. Для вычисления комбинированного ввода в этой сети мы будем использовать правило суммирования взвешенных связей, а в качестве функции активности будет выступать пороговая функция. Если комбинированный ввод элемента меньше 0, то активность равна 0, если ввод больше 0, то активность – 1.

Подадим на вход нейрона 1 – единицу, а на вход нейрона 2 – ноль. В этом случае на выходе мы должны получить 1 (0 XOR 1 = 1). Рассчитаем выходное значение вручную для демонстрации работы сети.

Комбинированный ввод элемента 5:net5 = 1 \* (-1) + 0 \* (-1) + 1 \* 1.5 = 0.5.

Активность элемента 5: 1 (0.5 > 0).

Комбинированный ввод элемента 6: net6 = 1 \* (-1) + 0 \* (-1) + 1 \* 0.5 = -

0.5.

Активность элемента 6: 0.

Комбинированный ввод элемента 7: net7 = 1 \* (1) + 0 \* (-1) + 1 \* (-0.5) =

0.5.

Активность элемента 7, а в то же время и выходное значение сети равно

1. Что и требовалось доказать.

Можно попробовать использовать в качестве входных сигналов все возможные значения (0 и 0, 1 и 0, 0 и 1, 1 и 1), на выходе мы всегда будем видеть значение, соответствующее таблице истинности операции XOR.

В данном случае все значения весовых коэффициентов нам были известны заранее, но главной особенностью нейронных сетей является то, что они могут сами корректировать значения веса всех связей в процессе обучения сети.

### ИСТОРИЯ НЕЙРОННЫХ СЕТЕЙ

Основные этапы в истории исследования и применения искусственных нейронных сетей:

* 1943 — У. Маккалок и У. Питтс формализуют понятие нейронной сети в фундаментальной статье о логическом исчислении идей и нервной активности.
* 1948 — Н. Винер вместе с соратниками публикует работу о кибернетике. Основной идеей является представление сложных биологических процессов математическими моделями.
* 1949 — Д. Хебб предлагает первый алгоритм обучения.
* В 1958 Ф. Розенблатт изобретает однослойный перцептрон и демонстрирует его способность решать задачи классификации. Перцептрон обрёл популярность — его используют для распознавания образов, прогнозирования погоды и т. д.
* В 1960 году Уидроу совместно со своим студентом Хоффом на основе дельта-правила (формулы Уидроу) разработали Адалин, который сразу начал использоваться для задач предсказания и адаптивного управления.

Сейчас Адалин (адаптивный сумматор) является стандартным элементом многих систем обработки сигналов.

* В 1963 году в Институте проблем передачи информации АН СССР. А. П. Петровым проводится подробное исследование задач «трудных» для перцептрона.
* В 1969 году М. Минский публикует формальное доказательство ограниченности перцептрона и показывает, что он неспособен решать некоторые задачи (проблема «чётности» и «один в блоке»), связанные с инвариантностью представлений. Интерес к нейронным сетям резко спадает.
* В 1972 году Т. Кохонен и Дж. Андерсон независимо предлагают новый тип нейронных сетей, способных функционировать в качестве памяти.
* В 1973 году Б. В. Хакимов предлагает нелинейную модель с синапсами на основе сплайнов и внедряет её для решения задач в медицине, геологии, экологии.
* 1974 — Пол Дж. Вербос и А. И. Галушкин одновременно изобретают алгоритм обратного распространения ошибки для обучения многослойных перцептронов
* 1975 — Фукусима представляет когнитрон — самоорганизующуюся сеть, предназначенную для инвариантного распознавания образов, но это достигается только при помощи запоминания практически всех состояний образа.
* 1982 — после периода забвения, интерес к нейросетям вновь возрастает. Дж. Хопфилд показал, что нейронная сеть с обратными связями может представлять собой систему, минимизирующую энергию (так называемая сеть Хопфилда). Кохоненом представлены модели сети, обучающейся без учителя (нейронная сеть Кохонена), решающей задачи кластеризации, визуализации данных (самоорганизующаяся карта Кохонена) и другие задачи предварительного анализа данных.
* 1986 — Дэвидом И. Румельхартом, Дж. Е. Хинтоном и Рональдом Дж. Вильямсом и одновременно c С. И. Барцевым и В. А. Охониным (Красноярская группа) переоткрыт и существенно развит метод обратного распространения ошибки. Начался взрыв интереса к обучаемым нейронным сетям.
* 2007 Джеффри Хинтоном в университете Торонто созданы алгоритмы глубокого обучения многослойных нейронных сетей. Успех обусловлен тем, что Хинтон при обучении нижних слоев сети использовал ограниченную машину Больцмана (RBM — Restricted Boltzmann Machine).

### КЛАССИФИКАЦИЯ НЕЙРОННЫХ СЕТЕЙ

### Классификация нейронных сетей по характеру обучения делит их на:

* нейронные сети, использующие обучение с учителем;
* нейронные сети, использующие обучение без учителя.

**Нейронные сети, использующие обучение с учителем.** Обучение с учителем предполагает, что для каждого входного вектора существует целевой вектор, представляющий собой требуемый выход. Вместе они называются обучающей парой. Обычно сеть обучается на некотором числе таких обучающих пар. Предъявляется выходной вектор, вычисляется выход сети и сравнивается с соответствующим целевым вектором. Далее веса изменяются в соответствии с алгоритмом, стремящимся минимизировать ошибку. Векторы обучающего множества предъявляются последовательно, вычисляются ошибки и веса подстраиваются для каждого вектора до тех пор, пока ошибка по всему обучающему массиву не достигнет приемлемого уровня.

**Нейронные сети, использующие обучение без учителя.** Обучение без учителя является намного более правдоподобной моделью обучения с точки зрения биологических корней искусственных нейронных сетей. Развитая Кохоненом и многими другими, она не нуждается в целевом векторе для

выходов и, следовательно, не требует сравнения с предопределенными идеальными ответами. Обучающее множество состоит лишь из входных векторов. Обучающий алгоритм подстраивает веса сети так, чтобы получались согласованные выходные векторы, т. е. чтобы предъявление достаточно близких входных векторов давало одинаковые выходы. Процесс обучения, следовательно, выделяет статистические свойства обучающего множества и группирует сходные векторы в классы.

### Классификация нейронных сетей по типу настройки весов делит их на:

* сети с фиксированными связями – весовые коэффициенты нейронной сети выбираются сразу, исходя из условий задачи;
* сети с динамическими связями – для них в процессе обучения происходит настройка синаптических весов.

### Классификация нейронных сетей по типу входной информации делит их на:

* аналоговые – входная информация представлена в форме действительных чисел;
* двоичные – вся входная информация в таких сетях представляется в виде нулей и единиц.