**Лекция 2**

**Однослойные и многослойные нейроны. Принципы создания искусственных нейронов**

В **полносвязных нейронных сетях** каждый нейрон передает свой выходной сигнал остальным нейронам, в том числе и самому себе. Все входные сигналы подаются всем нейронам. Выходными сигналами сети могут быть все или некоторые выходные сигналы нейронов после нескольких тактов функционирования сети.

В **многослойных (слоистых) нейронных сетях** нейроны объединяются в слои. Слой содержит совокупность нейронов с едиными входными сигналами. Число нейронов в слое может быть любым и не зависит от количества нейронов в других слоях. В общем случае сеть состоит из слоев, пронумерованных слева направо. Внешние входные сигналы подаются на входы нейронов входного

слоя (его часто нумеруют как нулевой), а выходами сети являются выходные сигналы последнего слоя. Кроме входного и выходного слоев в многослойной нейронной сети есть один или несколько скрытых слоев. Связи от выходов нейронов некоторого слоя q к входам нейронов следующего слоя (q+1) называются последовательными.

### Типы многослойных нейронных сетей

**Монотонные**. Это частный случай слоистых сетей с дополнительными условиями на связи и нейроны. Каждый слой, кроме последнего (выходного), разбит на два блока: возбуждающий и тормозящий. Связи между блоками тоже разделяются на тормозящие и возбуждающие. Если от нейронов блока к нейронам блока ведут только возбуждающие связи, то это означает, что любой выходной сигнал блока является монотонной неубывающей функцией любого выходного сигнала блока. Если же эти связи только тормозящие, то любой выходной сигнал блока является невозрастающей функцией любого выходного сигнала блока. Для нейронов монотонных сетей необходима монотонная зависимость выходного сигнала нейрона от параметров входных сигналов.

**Сети без обратных связей**. В таких сетях нейроны входного слоя получают входные сигналы, преобразуют их и передают нейронам первого скрытого слоя, и так далее вплоть до выходного, который выдает сигналы для интерпретатора и пользователя. Если не оговорено противное, то каждый выходной сигнал q-го слоя подается на вход всех нейронов (q+1)-го слоя; однако возможен вариант соединения q-го слоя с произвольным -м слоем.

Среди многослойных сетей без обратных связей различают **полносвязные** (выход каждого нейрона q–го слоя связан с входом каждого нейрона (q+1)–го слоя) и частично **полносвязные**.

### Сети с обратными связями

В **сетях с обратными связями** информация с последующих слоев передается на предыдущие. Различают следующие типы нейронных сетей с обратными связями:

* + - **слоисто-циклические**, отличающиеся тем, что слои замкнуты в кольцо: последний слой передает свои выходные сигналы первому; все слои равноправны и могут как получать входные сигналы, так и выдавать выходные;
		- **слоисто-полносвязные** состоят из слоев, каждый из которых представляет собой полносвязную сеть, а сигналы передаются как от слоя к слою, так и внутри слоя; в каждом слое цикл работы распадается на три части: прием сигналов с предыдущего слоя, обмен сигналами внутри слоя, выработка выходного сигнала и передача к следующему слою;
		- **полносвязно-слоистые**, по своей структуре аналогичные слоисто-полно-связным, но функционирующим по-другому: в них не разделяются фазы обмена внутри слоя и передачи следующему, на каждом такте нейроны всех слоев принимают сигналы от нейронов как своего слоя, так и последующих.

На картике изображены сеть Элмана и сеть Жордана



Нейронные сети можно разделить по типам структур нейронов на гомогенные (однородные) и гетерогенные. **Гомогенные** сети состоят из нейронов одного типа с единой функцией активации, а в **гетерогенную** сеть входят нейроны с различными функциями активации.

Еще одна классификация делит нейронные сети на **синхронные** и **асинхронные***.* В первом случае в каждый момент времени лишь один нейрон меняет свое состояние, во втором – состояние меняется сразу у целой группы нейронов, как правило, у всего слоя. Алгоритмически ход времени в нейронных сетях задается итерационным выполнением однотипных действий над нейронами.

### ФОРМАЛЬНЫЙ НЕЙРОН

У нейрона есть несколько входных каналов и только один выходной канал. По входным каналам на нейрон поступают данные задачи, а на выходе формируется результат работы. Нейрон вычисляет взвешенную сумму входных сигналов, а затем преобразует полученную сумму с помощью заданной нелинейной функции. Множество, состоящее из порогового уровня и всех весов, называют параметрами нейрона.



Здесь введены следующие обозначения: X1,X2, …, Xn - входной сигнал

(паттерн), w1, w2,…,wn – весовые коэффициенты, b- порог нейрона

Сначала нейрон вычисляет взвешенную сумму

*S*  *wi Xi i*

* *b* **,** далее

применяя **функцию активации** F(S) вычисляет выходной сигнал Y.

**Функция активации нейрона** - это функция, которая вычисляет выходной сигнал нейрона. На вход этой функции подается сумма всех произведений сигналов и весов этих сигналов.

Рассмотрим наиболее часто используемые функции активации.

**а) Пороговая функция.** Это простая кусочно-линейная функция. Если входное значение меньше порогового, то значение функции активации равно минимальному допустимому, иначе – максимально допустимому.

**б) Линейный порог.** Это несложная кусочно-линейная функция. Имеет два линейных участка, где функция активации тождественно равна минимально допустимому и максимально допустимому значению и есть участок, на котором функция строго монотонно возрастает.

**в) Сигмоидальная функция** или **сигмоида (sigmoid)**. Это монотонно возрастающая дифференцируемая S-образная нелинейная функция. Сигмоида позволяет усиливать слабые сигналы и не насыщаться от сильных сигналов.

**г) Гиперболический тангенс (hyperbolic tangent, tanh).** Эта функция принимает на входе произвольное вещественное число, а на выходе дает вещественное число в интервале от –1 до 1. Подобно сигмоиде, гиперболический тангенс может насыщаться. Однако, в отличие от сигмоиды, выход данной функции центрирован относительно нуля.

**Недостатки формального нейрона:**

* + Предполагается, что нейрон мгновенно вычисляет свой выход, поэтому с помощью таких нейронов нельзя моделировать непосредственно системы с внутренним состоянием.
	+ Формальные нейроны, в отличие от биологических, не могут обрабатывать информацию синхронно.
	+ Нет четких алгоритмов выбора функции активации.
	+ Невозможно регулировать работу всей сети.
	+ Излишняя формализация понятий «порог» и «весовые коэффициенты». У реальных нейронов порог меняется динамически, в зависимости от активности нейрона и общего состояния сети, а весовые коэффициенты изменяются в зависимости от проходящих сигналов.

###

### ОДНОСЛОЙНАЯ НЕЙРОННАЯ СЕТЬ

Один нейрон может выполнять простейшие вычисления, но основные функции нейросети обеспечиваются не отдельными нейронами, а соединениями между ними. Однослойный перцептрон представляет собой простейшую сеть, которая состоит из группы нейронов, образующих слой. Входные данные кодируются вектором значений, каждый элемент подается на соответствующий вход каждого нейрона в слое. В свою очередь, нейроны вычисляют выход независимо друг от друга. Размерность выхода (то есть количество элементов) равна количеству нейронов, а количество синапсов у всех нейронов должно быть одинаково и совпадать с размерностью входного сигнала.



Здесь X1,X2,X3- называется входной паттерн, Y1,Y2,Y3- выходной паттерн, а wi,j- это j-ый весовой коэффициент i-го нейрона

### ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

**Обучение нейронной сети**- это процесс, в котором параметры нейронной сети настраиваются посредством моделирования среды, в которую эта сеть встроена. Тип обучения определяется способом подстройки параметров. Различают алгоритмы обучения с учителем и без учителя. Процесс обучения с учителем представляет собой предъявление сети выборки обучающих примеров. Каждый образец подается на входы сети, затем проходит обработку внутри структуры НС, вычисляется выходной сигнал сети, который сравнивается с соответствующим значением целевого вектора, представляющего собой требуемый выход сети.

Для того, чтобы нейронная сети была способна выполнить поставленную задачу, ее необходимо обучить. Различают алгоритмы обучения с учителем и без учителя. Процесс обучения с учителем представляет собой предъявление сети выборки обучающих примеров. Каждый образец подается на входы сети, затем проходит обработку внутри структуры НС, вычисляется выходной сигнал сети, который сравнивается с соответствующим значением целевого вектора, представляющего собой требуемый выход сети. Затем по определенному правилу вычисляется ошибка, и происходит изменение весовых коэффициентов связей внутри сети в зависимости от выбранного алгоритма. Векторы обучающего множества предъявляются последовательно, вычисляются ошибки и веса подстраиваются для каждого вектора до тех пор, пока ошибка по всему обучающему массиву не достигнет приемлемо низкого уровня.



При обучении без учителя 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. Предъявление на вход вектора из данного класса даст определенный выходной вектор, но до обучения невозможно предсказать, какой выход будет производиться данным классом входных векторов. Следовательно, выходы подобной сети должны трансформироваться в некоторую понятную форму, обусловленную процессом обучения. Это не является серьезной проблемой. Обычно не сложно идентифицировать связь между входом и выходом, установленную сетью. Для обучения нейронных сетей без учителя применяются сигнальные метод обучения Хебба и Ойа.

Математически процесс обучения можно описать следующим образом. В процессе функционирования нейронная сеть формирует выходной сигнал Y, реализуя некоторую функцию Y = G(X). Если архитектура сети задана, то вид функции G определяется значениями синаптических весов и смещенной сети.

Пусть решением некоторой задачи является функция Y = F(X), заданная параметрами входных-выходных данных (X1, Y1), (X2, Y2), …, (XN, YN), для которых Yk = F(Xk) (k = 1, 2, …, N).

Обучение состоит в поиске (синтезе) функции G, близкой к F в смысле некоторой функции ошибки E.

Если выбрано множество обучающих примеров – пар (XN, YN) (где k = 1, 2, …, N) и способ вычисления функции ошибки E, то обучение нейронной сети превращается в задачу многомерной оптимизации, имеющую очень большую размерность, при этом, поскольку функция E может иметь произвольный вид обучение в общем случае – многоэкстремальная невыпуклая задача оптимизации.

Для решения этой задачи могут использоваться следующие (итерационные) алгоритмы:

1. алгоритмы локальной оптимизации с вычислением частных производных первого порядка:
* градиентный алгоритм (метод наискорейшего спуска),
* методы с одномерной и двумерной оптимизацией целевой функции в направлении антиградиента,
* метод сопряженных градиентов,
* методы, учитывающие направление антиградиента на нескольких шагах алгоритма;
1. алгоритмы локальной оптимизации с вычислением частных производных первого и второго порядка:
* метод Ньютона,
* методы оптимизации с разреженными матрицами Гессе,
* квазиньютоновские методы,
* метод Гаусса-Ньютона,
* метод Левенберга-Марквардта и др.;
1. стохастические алгоритмы оптимизации:
* поиск в случайном направлении,
* имитация отжига,
* метод Монте-Карло (численный метод статистических испытаний);
1. алгоритмы глобальной оптимизации (задачи глобальной оптимизации решаются с помощью перебора значений переменных, от которых зависит целевая функция).