

Тема: Обучение нейронных сетей.
Парадигмы обучения.

Процесс обучения - это процесс, в котором свободные параметры нейронной сети настраиваются посредством моделирования среды, в которую сеть встроена.

Последовательность обучения:

1. В нейронную сеть поступают стимулы из внешней среды.
2. В результате этого изменяются свободные параметры нейронной сети.
3. После изменения внутренней структуры нейронная сеть отвечает на возбуждения уже иным образом.

В теории нейронных сетей выделяют две фундаментальные парадигмы обучения нейронных сетей: обучение **с учителем и без учителя**.

Модели обучения ИНС:

1. на основе коррекции ошибок;
2. с использованием памяти;
3. Хеббовское обучение;
4. конкурентное обучение;
5. обучение Больцмана.

Парадигма обучения нейронной сети с учителем

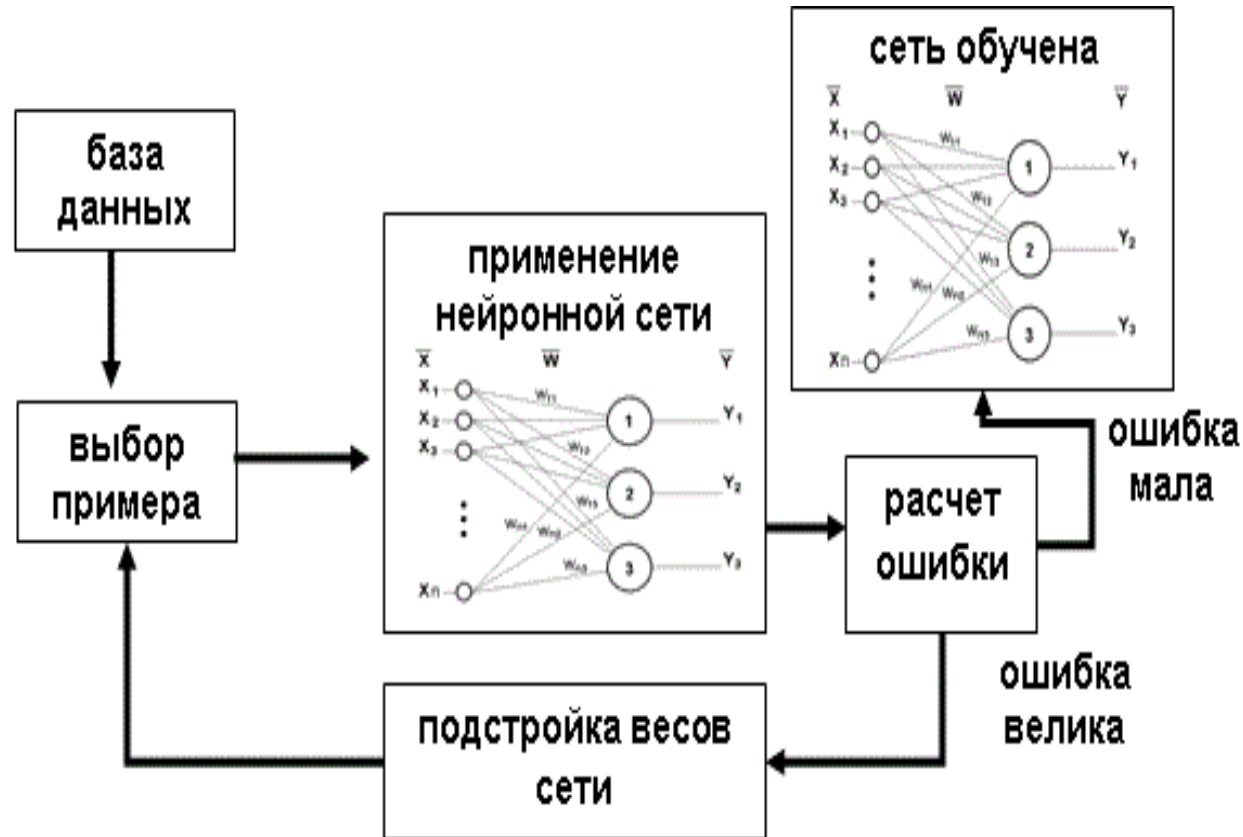


Рис.1 Процесс обучения нейронной сети

Парадигма	Правило обучения	Архитектура	Алгоритм обучения	Решаемая задача
С учителем	Коррекция ошибки	Однослойный и многослойный персептрон	Алгоритм обучения персептрона Обратное распространение ошибки	Классификация Распознавание образов Аппроксимация функций Прогнозирование Управление
С учителем	Больцмана	Рекуррентная	Алгоритм обучения Больцмана	Классификация Распознавание образов
С учителем	Хебба	Многослойная прямого распространения	Линейный дискриминантный анализ	Анализ данных Классификация Распознавание образов
С учителем	Соревнование	Сеть ART	ARTmap	Классификация Распознавание образов

Парадигма обучения нейронной сети без учителя

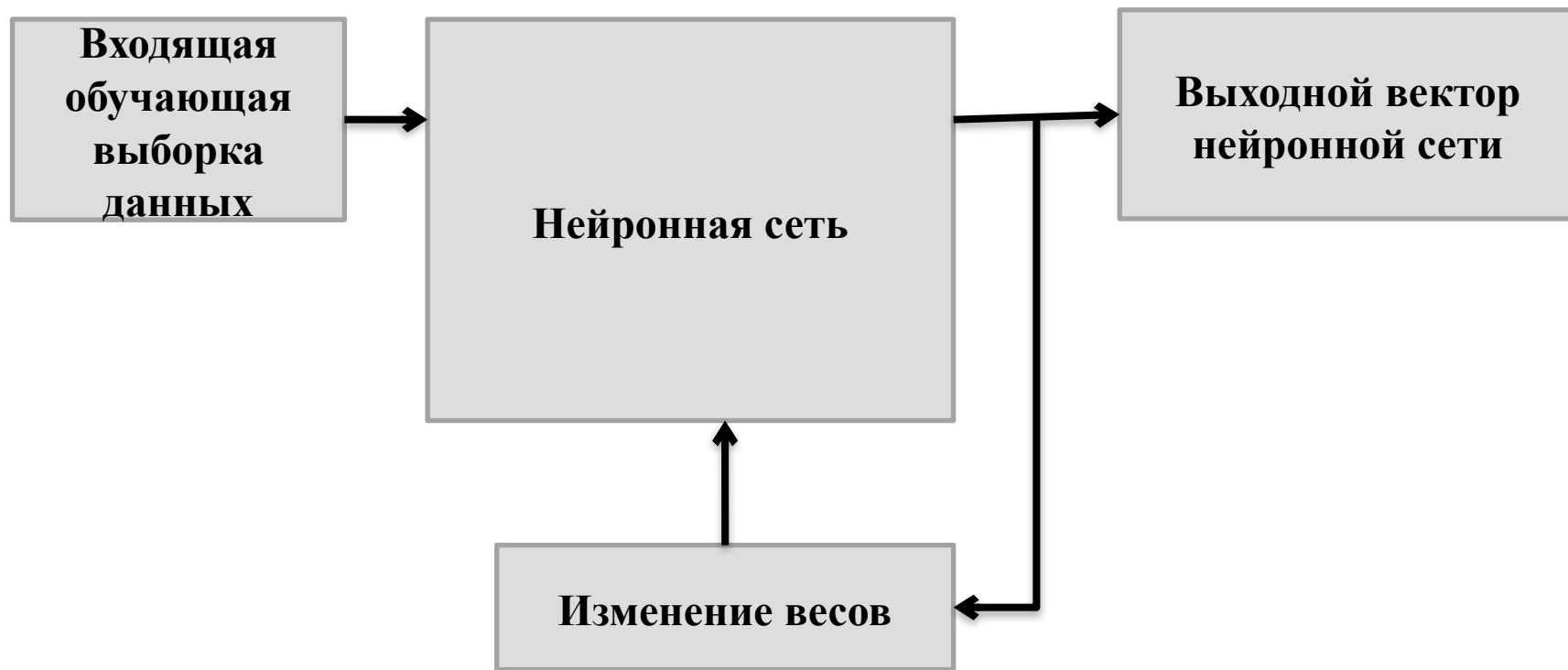


Рис.2 Процесс обучения нейронной сети

Парадигма	Правило обучения	Архитектура	Алгоритм обучения	Решаемая задача
Без учителем	Коррекция ошибки	Многослойная прямого распространения	Проекция Саммона	Категоризация в средние класса Анализ данных
Без учителем	Хебба	прямого распространения или соревнование Сеть Хопфилда	Анализ главных компонент Обучение ассоциативной памяти	Анализ данных Сжатие данных Ассоциативная память
Без учителем	Соревнование	Соревнование SOM, Кохоннена Сеть ART	Векторное квантование SOM, Кохоннена ART1, ART2	Категоризация Сжатие данных Категоризация Анализ данных Категоризация
Смешанная	Коррекция ошибки Соревнование	Сети RBF	Алгоритмы обучения RBF	Классификация Распознавание образов Аппроксимация функций Прогнозирование Управление

Эпоха обучения - это один просмотр всех примеров обучающей выборки с одновременной коррекцией весов сети (на этих примерах, в зависимости от правильности их решения сетью).

Чтобы сеть обучилась, может потребоваться несколько (и даже несколько десятков, и даже несколько сотен,...) эпох.

Характеристика алгоритмов обучения

Адаптация и самоорганизация искусственных нейронных сетей достигается в процессе их обучения. Обучение нейронной сети может вестись с учителем или без него.

В первом случае сети предъявляются значения как входных, так и желательных выходных сигналов, и она по некоторому внутреннему алгоритму подстраивает веса своих синаптических связей.

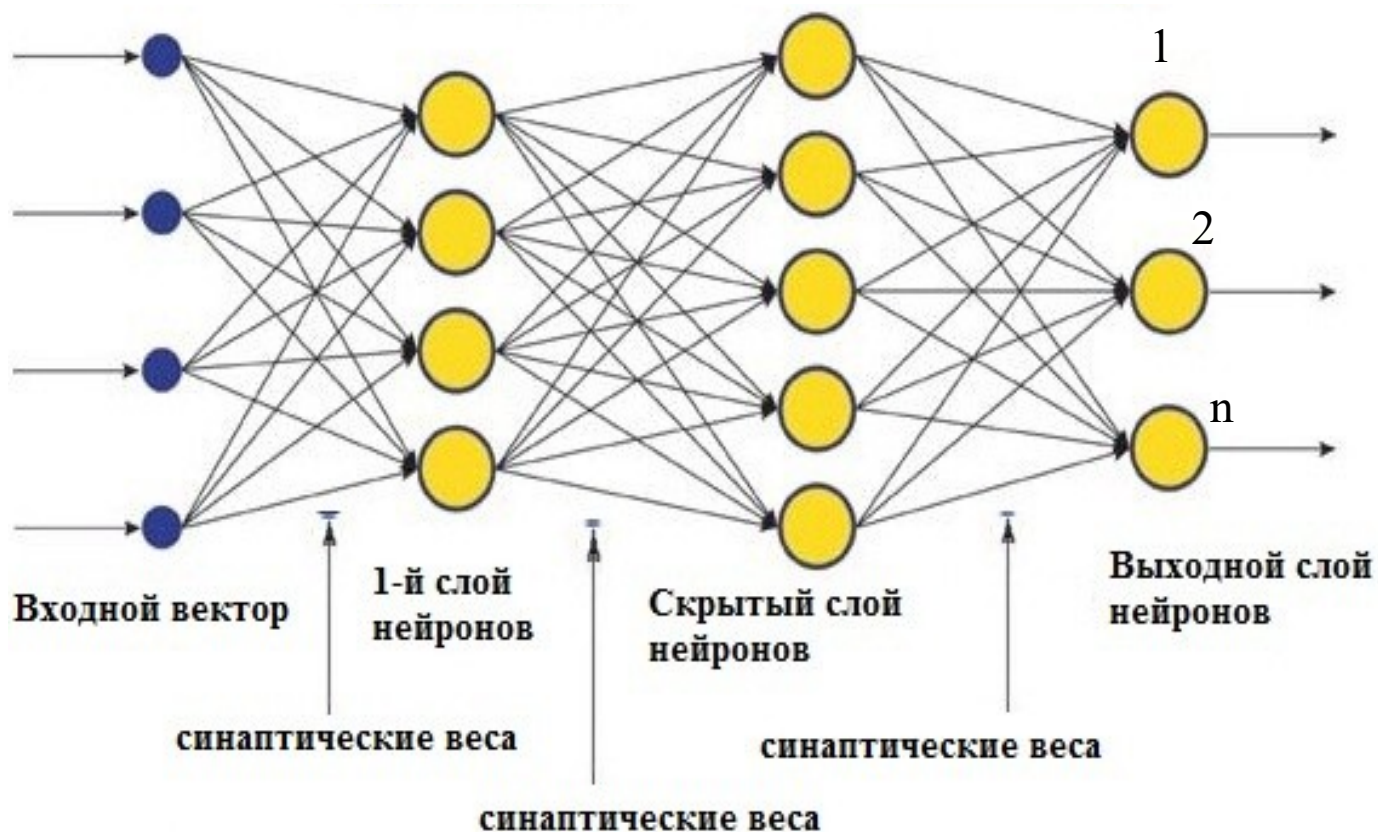
Во втором случае выходы нейронной сети формируются самостоятельно, а веса изменяются по алгоритму, учитывающему только входные и производные от них сигналы

Ошибка обучения для конкретной конфигурации сети определяется после прогона через сеть всех имеющихся наблюдений и сравнения выходных значений с целевыми значениями в случае обучения с учителем.

Соответствующие разности позволяют сформировать так называемую **функцию ошибок**.

Функции ошибок

Ошибка сети, выходной слой которой имеет n нейронов, есть разность между реальным (y) и желаемым (d) сигналами уа выходе i -го нейрона.



В качестве функций ошибок могут быть использованы следующие функции:

➤ **сумма квадратов ошибок** ²

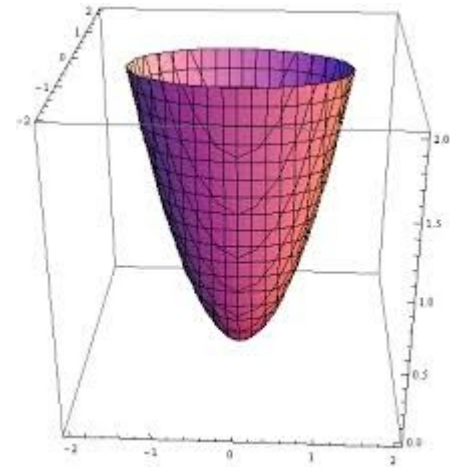
➤ **средняя квадратичная ошибка** ²

➤ **регулируемая или комбинированная ошибка**
^{2 2},

где - параметр регуляции,

➤ **средняя абсолютная ошибка**

В случае линейной модели сети и функции ошибок в виде суммы квадратов такая поверхность будет представлять собой параболоид, который имеет единственный минимум, и это позволяет отыскать такой минимум достаточно просто.



В случае нелинейной модели поверхность ошибок имеет гораздо более сложное строение и обладает рядом неблагоприятных свойств, в частности может иметь локальные минимумы, плоские участки, седловые точки и длинные узкие овраги.



Алгоритмы локальной оптимизации с вычислением частных производных первого порядка:

- градиентный алгоритм;
- метод Ньютона,
- методы оптимизации с разреженными матрицами Гессе,
- метод Левенберга—Марквардта;
- поиск в случайном направлении;
- алгоритм обратного распространения ошибки.

Свойства нейросетей

Способность сети, обученной на некотором множестве данных выдавать правильные результаты для достаточно широкого класса новых данных, в том числе и не представленных при обучении, называется *свойством обобщения* нейронной сети.

Для настройки параметров нейронных сетей широко используется также процедура *адаптации*, когда подбираются веса и смещения с использованием произвольных функций их настройки, обеспечивающие соответствие между входами и желаемыми значениями на выходе.

Явление переобучения

Переобучение в нейронных сетях — явление, когда построенная модель хорошо объясняет примеры из обучающей выборки, но достаточно плохо работает на примерах, не участвовавших в обучении (на примерах из тестовой выборки).

Методы предотвращения переобучения

- Перекрёстная проверка
- Ранняя остановка,
- Байесовское сравнение моделей
- д.р. методы

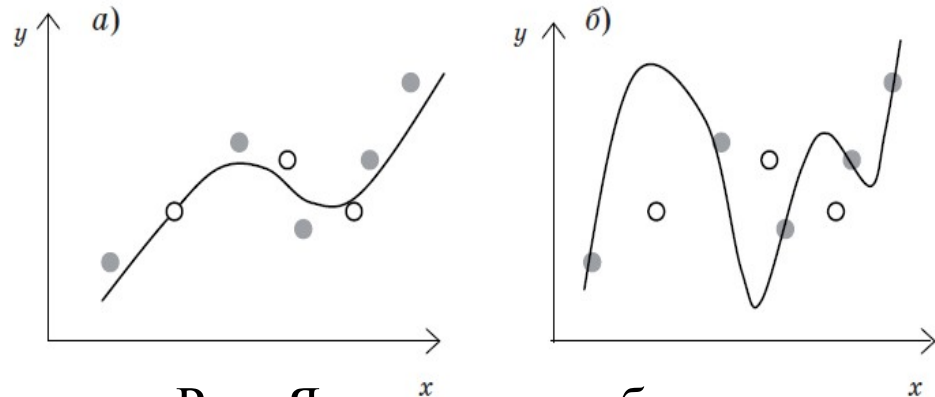
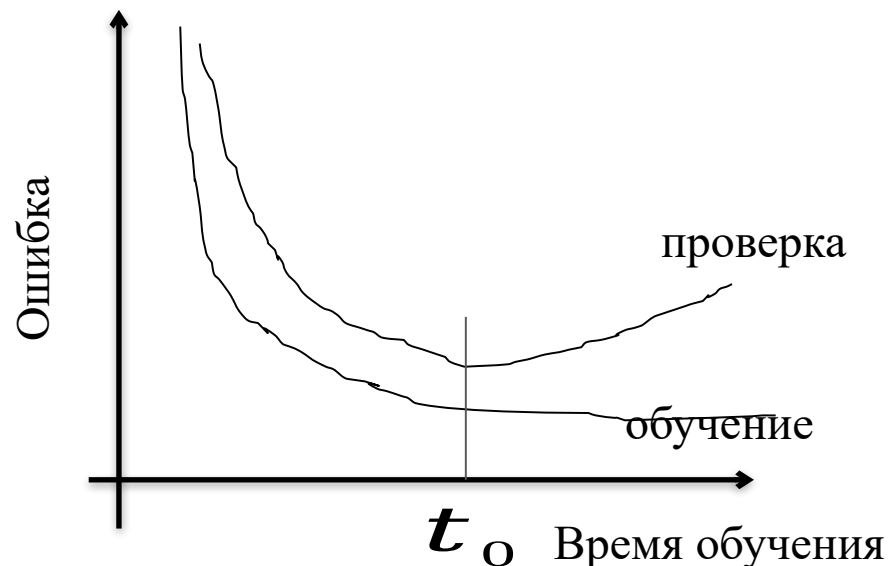


Рис. Явление переобучения
светлые кружки соответствуют тестовым данным,
а темные — обучающим.

Ранняя остановка обучения

Вся обучающая выборка разбивается на три набора:

- обучающий,
- проверочный,
- тестовый.



При обучении нейронной сети для настройки весовых коэффициентов используется обучающий набор

Ошибка обучения сети проверяется по проверочному набору

В момент обучения, когда сеть начинает настраиваться на шум, присутствующий только в наборе - ошибка, вычисляемая по совершенно другим данным начинает расти. (рис). Именно в этот момент, когда сеть работает наилучшим образом и имеет смысл прекратить обучение. Не обязательный набор служит для окончательной проверки эффективности работы.