

Санкт-Петербургский государственный университет
телекоммуникаций им. проф. М.А. Бонч-Бруевича

Искусственный интеллект в сетях связи

Машинное обучение

Выборнова А.И., доц. каф. ССиПД

СПб ГУТ)))

Что такое машинное обучение?

Машинное обучение – это обеспечение для какой-либо системы возможности решать задачи, обучаться и изменяться без непосредственного программирования системы, а на основе опыта.

Машинное обучение служит для обучения нейронных сетей, но может применяться и к другим системам.

Что такое машинное обучение?

Машинное обучение по сути позволяет анализировать некоторый набор данных и находить в нем паттерны, без непосредственного программирования логики паттернов и без знания закономерностей, лежащих в основе данных.

Классификация

Классификация методов машинного обучения:

- Классификация/кластеризация:
 - Два варианта ответа (да или нет).
 - Больше двух вариантов ответа (человек, дерево, здание).
- Регрессия (ответ – любое число).

Классификация

Классификация методов машинного обучения:

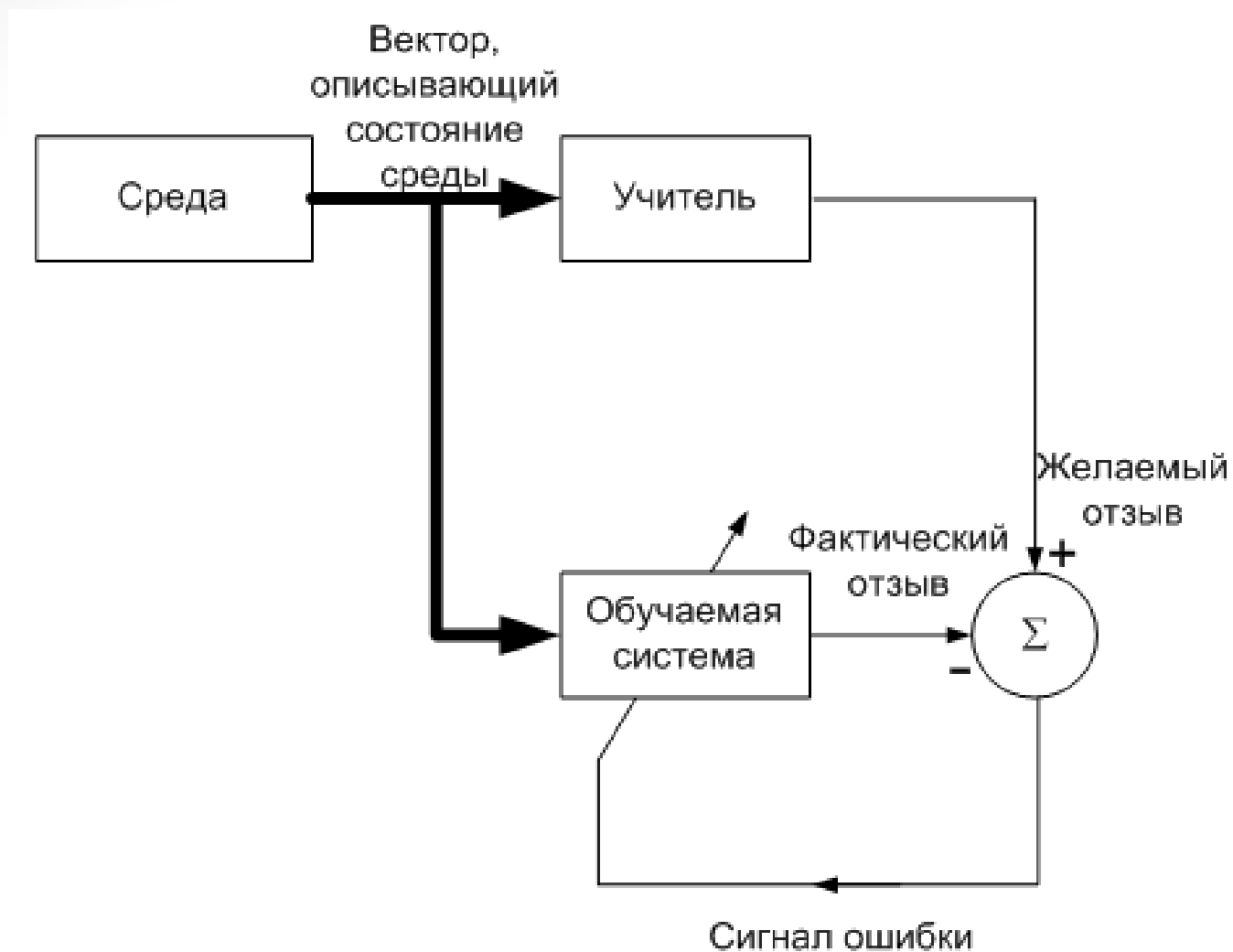
- Обучение с учителем:
 - Классификация.
 - Регрессия.
- Обучение без учителя:
 - Кластеризация.
- Обучение с частично известными ответами/метками («частично с учителем»).
- Обучение с подкреплением.

Обучение с учителем

Обучение с учителем – это вид МО, при котором для всех обучающих данных имеются «правильные» ответы (метки), собранные заранее при помощи других инструментов.

В этом случае у нас есть возможность сравнить реакцию обучаемой системы на входной сигнал с известным правильным ответом, и перенастроить систему исходя из результата сравнения.

Обучение с учителем



Обучение с учителем

Обучение с учителем может использоваться для построения как регрессионных (один численный выход), так и классифицирующих (два или несколько бинарных/вероятностных выходов) моделей.

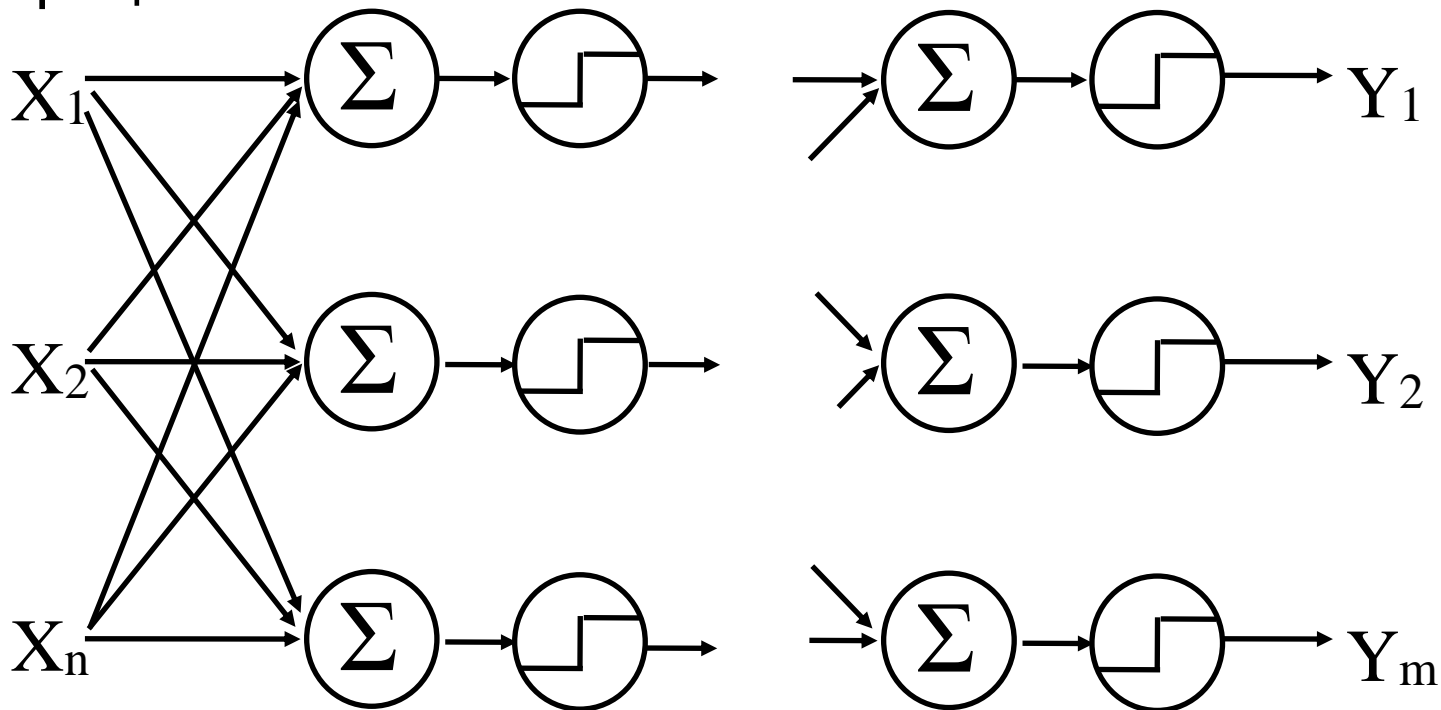
Обучение с учителем

Для обучения с учителем нейронных сетей (в отличие от других систем) может использоваться несколько разных методов, наиболее известным и популярным из которых является **метод обратного распространения ошибки (backpropagation)**.

Суть обучения нейронной сети заключается в поиске таких значений весов всех связей внутри сети, которое дает наиболее близкую к правильной реакцию сети.

Обучение с учителем

При использовании метода обратного распространения ошибки разница между ответом нейронной сети и правильным ответом передается в обратном направлении по сети и управляет изменением коэффициентов.



Обучение с учителем

- Как численно определяется ошибка сети?

Для регрессионного анализа – метод средних квадратов (для одной тренировочной итерации):

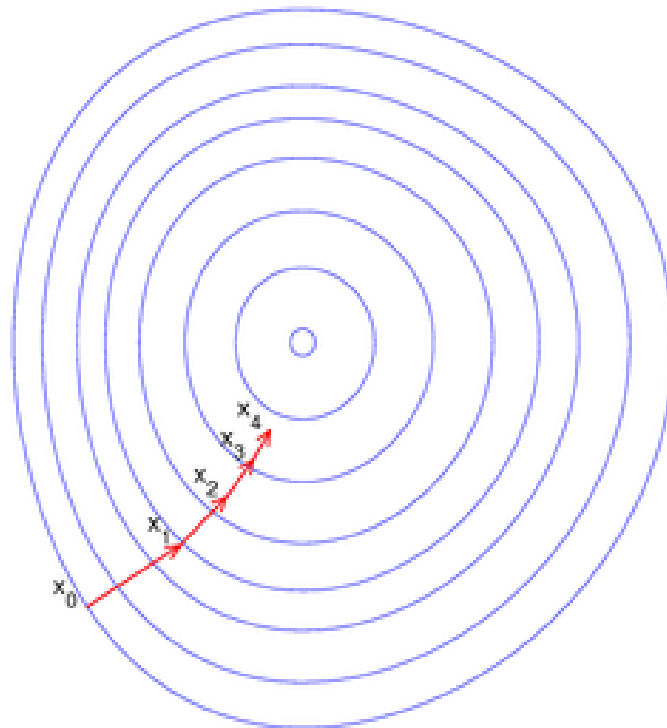
$$E(y, y') = \frac{1}{2} (y - y')^2$$

Для классификации – коэффициент корреляции между ответами нейронной сети и правильными ответами, или другие методы анализа схожести рядов.

Обучение с учителем

Как определяется насколько менять вес каждой связи?

В основе – метод градиентного спуска.



Обучение с учителем

■

Всю нейронную сеть можно представить как последовательное перемножение матриц:

X – матрица входных параметров.

W_i – матрица весов i -ого слоя, $i=1..n$,

$w_{k,l}^i$ – вес связи между k нейроном слоя $i-1$ и l нейроном слоя i .

f_i – функция активации слоя i .

Тогда для нейронной сети с одним скрытым слоем:

$$y' = f_2(W_2 * f_1(W_1 * X))$$

Обучение с учителем

- Изменение каждого отдельного коэффициента можно вычислить как:

$$\Delta b_{i,j} = -\eta \frac{\partial E(y - y')}{\partial b_{i,j}}$$

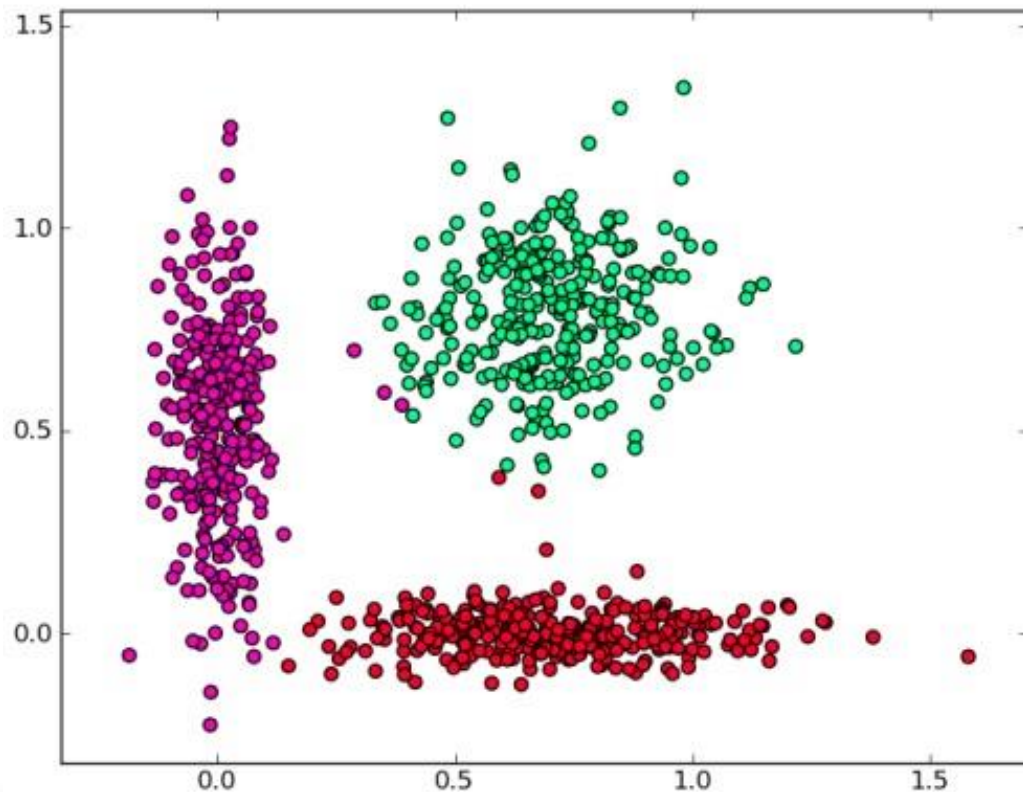
Где:

$b_{i,j}$ – текущий вес связи,

$0 < \eta < 1$ – множитель, задающий скорость «движения».

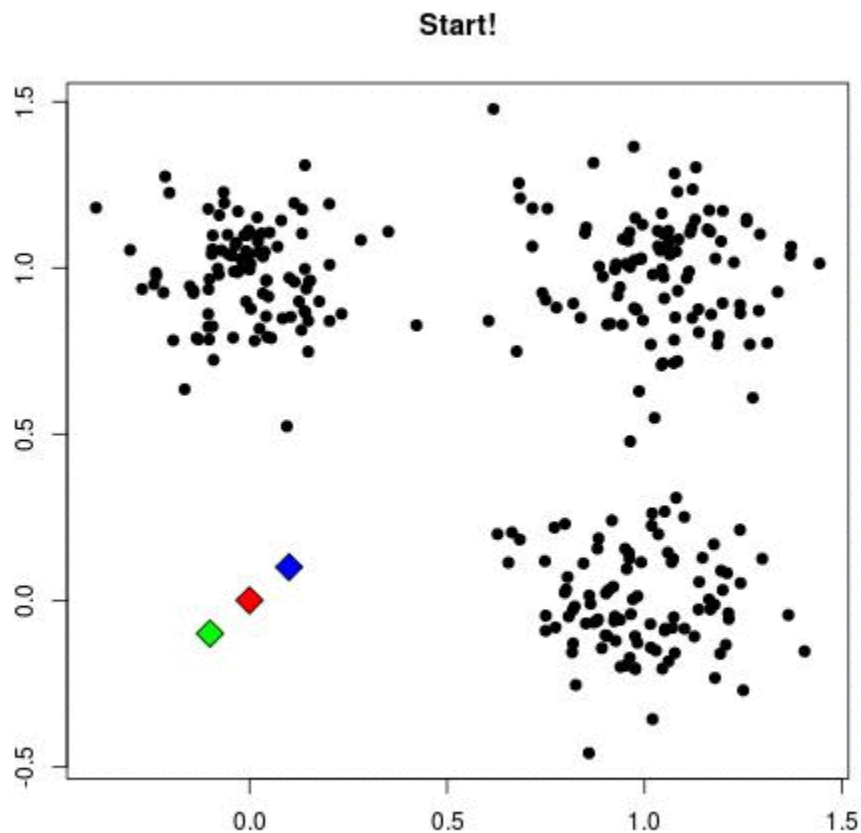
Обучение без учителя

Обучение без учителя – обучение на основе только входных данных, при котором обучаемая система сама кластеризует входные данные тем или иным способом.



Обучение без учителя

Обучение без учителя может использоваться как для обучения нейронных сетей, так и для других систем. Наиболее популярный способ – **метод k-средних**.



Обучение без учителя

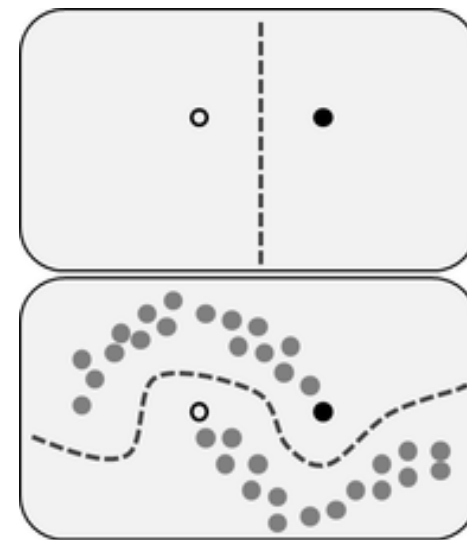
Для нейронных сетей может использоваться один из «нечетких» методов k-средних, в которых на выходах нейронной сети будут вероятности того, что данный входной набор параметров относится к той или иной категории.

При этом изменение коэффициентов определяется не сигналом ошибки, который невозможно вычислить ввиду отсутствия правильных ответов, а другими методами, например, величиной сигнала, проходящего через каждую связь на каждой итерации.

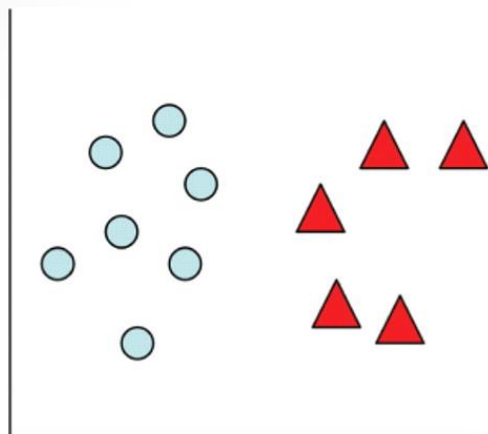
Обучение с частично известными метками

Подбор данных с известными ответами для обучения с учителем – достаточно трудоемкая задача и часто для сбора достаточных по объему обучающих данных требуется много времени и других ресурсов.

В таких ситуациях на помощь может прийти **обучение с частично известными метками/ответами** – когда часть строк в обучающей выборке имеет правильные ответы (размеченные данные), а часть (обычно БОльшая) – не имеет (неразмеченные данные).

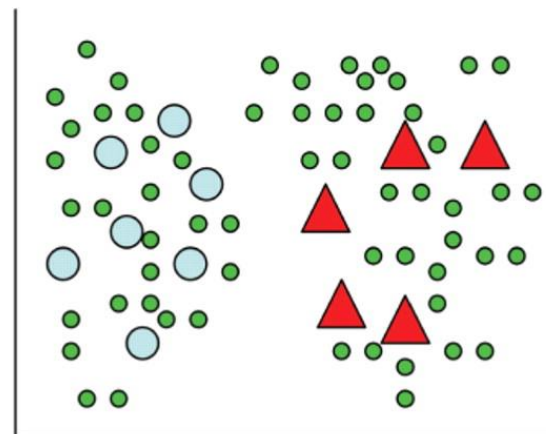


Обучение с частично известными метками



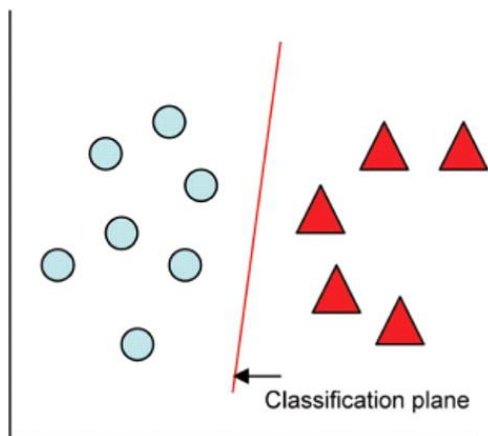
Labeled Data

(a)



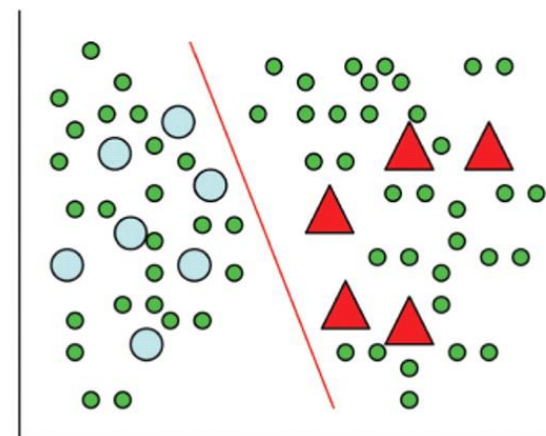
Labeled and Unlabeled Data

(b)



Supervised Learning

(c)



Semi-Supervised Learning

(d)

Обучение с частично известными метками

Обучение на частично размеченной выборке теоретически позволяет более точно классифицировать данные, как по сравнению с обучением без учителя, так и по сравнению с обучением с учителем на небольшом объеме данных.

Целью обучения при таком подходе является не только создание модели, которая могла бы успешно классифицировать новые данные, но также и добавление ответов к неразмеченным обучающим данным.

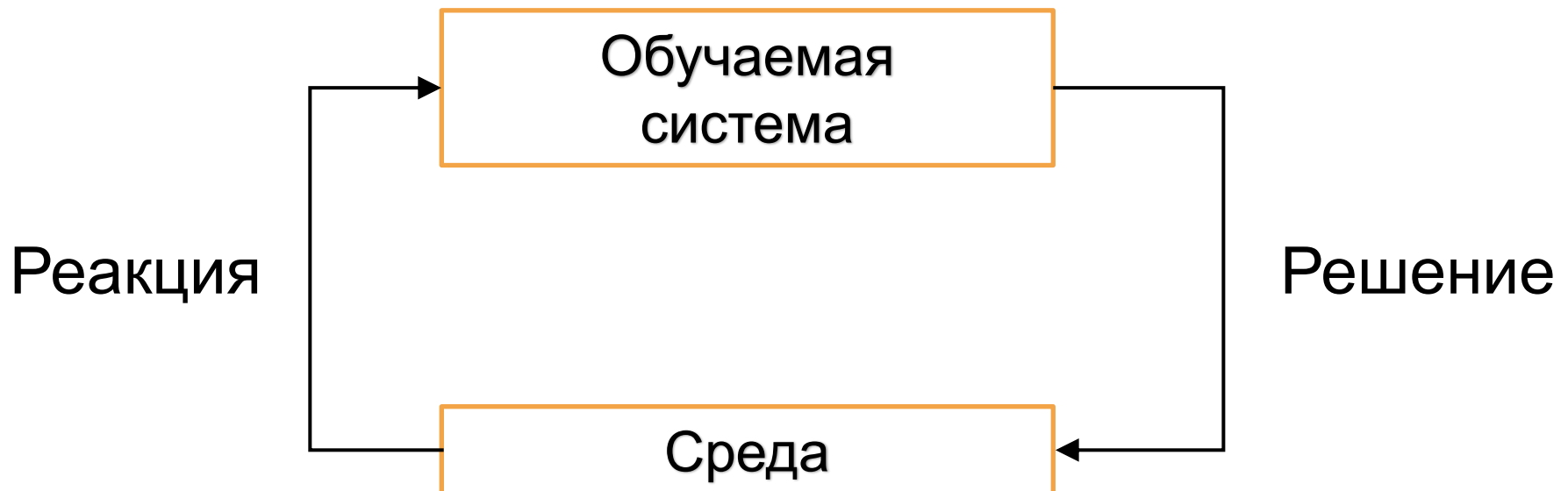
Обучение с частично известными метками

Для успешного обучения с частично известными ответами данные должны удовлетворять хотя бы одному из указанных допущений:

1. Непрерывность – данные, расположенные близко друг к другу, с большей вероятностью имеют одинаковую метку, чем данные, расположенные дальше друг от друга.
2. Кластеризация – данные, входящие в один кластер, с большей вероятностью имеют одинаковую метку, чем данные, входящие в разные кластеры.
3. Избыточность (размерностей) – закономерности в основе данных имеют значительно меньшую размерность, чем сами данные.

Обучение с подкреплением

Обучение с подкреплением – случай МО, когда ответы не известны, но на принятое обучаемой системой решение приходит ответ от среды.



Проблемы МО

В некоторых случаях МО не может дать приемлемого результата. Это может быть следствием одной или нескольких проблем:

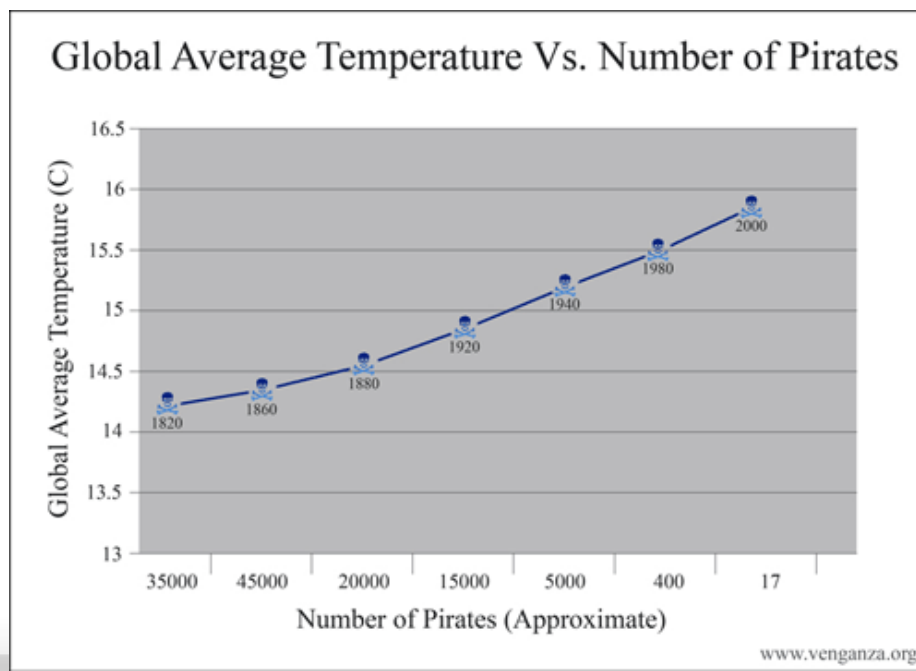
- Проблемы выбора данных:
 - Корреляция и причинно-следственная связь.
 - Скрытые переменные.
 - Сильная корреляция входных значений.
 - Человеческие систематические ошибки.
- Проблемы алгоритмов:
 - Некорректный выбор алгоритма и его ошибки.
 - Переобучение.

Bias – систематическая ошибка, искажение.

Корреляция и причинно-следственная связь

При выборе данных для анализа следует выбирать те данные, которые теоретически имеют объяснимую взаимосвязь с тем параметром, который мы хотим предсказать.

Данные, которые хорошо коррелируют между собой не всегда на самом деле имеют связывающую их непосредственную закономерность:



Скрытые переменные

Скрытые переменные в МО – это явление, когда обучающаяся система берет в расчет данные, которые попали в обучающие данные случайным образом и не предназначались для участия в обучении.

«Проблема танка»



Мультиколлинеарность

Если во входных данных есть две или более сильно коррелированных между собой величины, то это может привести к нестабильности в процессе обучения и, как следствие, к плохим результатам.

Решение:

Составление матрицы корреляций входных величин и исключение всех сильных корреляций.

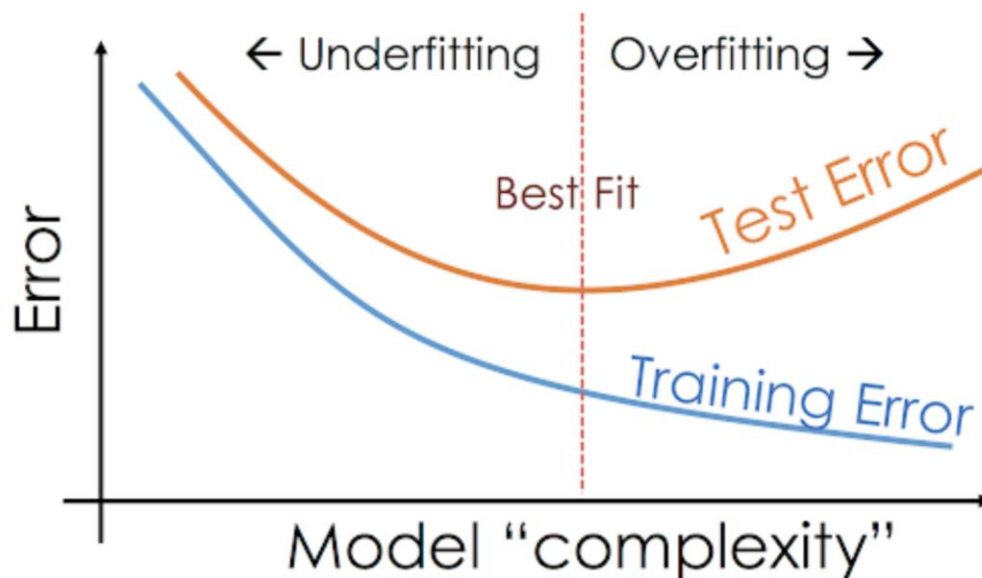
Человеческие систематические ошибки

В выборе данных для МО чаще всего участвует человек, а значит может привносить собственные систематические ошибки мышления в данные:

- Систематическая ошибка выжившего (Абрахам Вальд и американские бомбардировщики).
- Предвзятость подтверждения (поиска информации/интерпретации результатов).
- Генерализация частных случаев.
- И др.

Переобучение

Переобучение (overfitting) – это явление, когда из-за слишком высокой «сложности» обучаемой системы и методов обучения система показывает отличный результат на обучающей выборке, но значительно более плохой на проверочных (тестовых) данных, не входивших в данные для обучения.

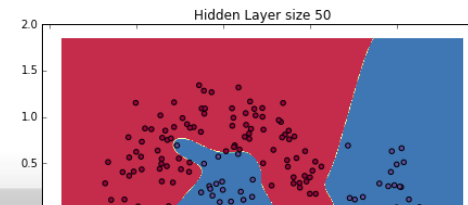
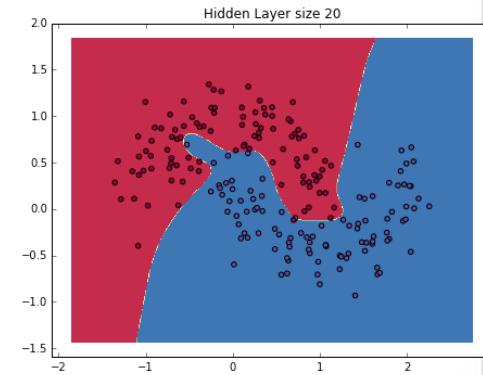
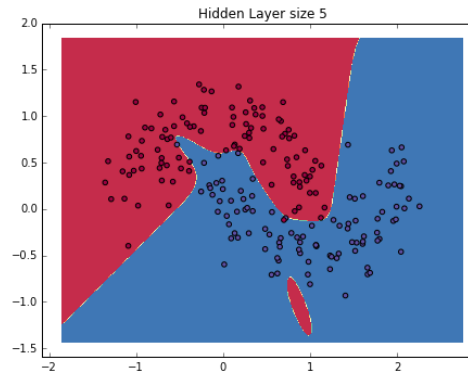
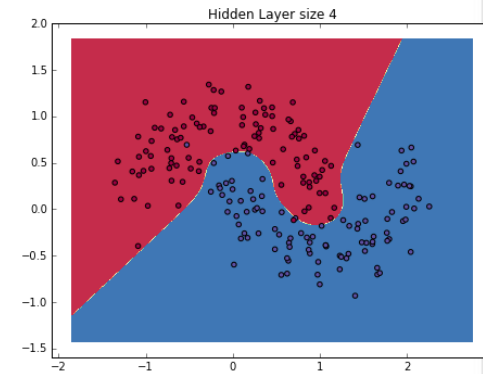
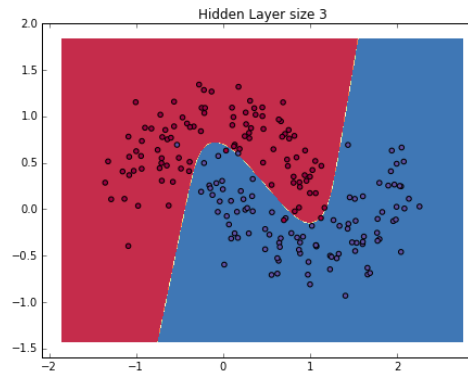
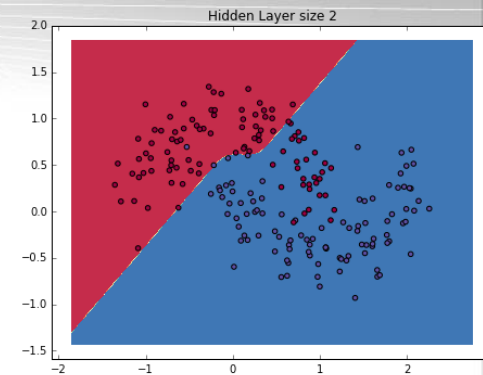
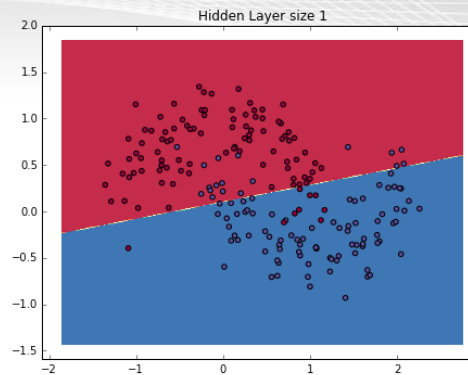


Переобучение

На переобучение влияет:

- Число узлов скрытого слоя.
- Число шагов (эпох) обучения.
- Размер допустимой ошибки.

Слишком увеличивая первые два параметра и требования к точности системы можно получить ситуацию переобучения.



Переобучение



**THE BEST WAY TO
EXPLAIN OVERFITTING**

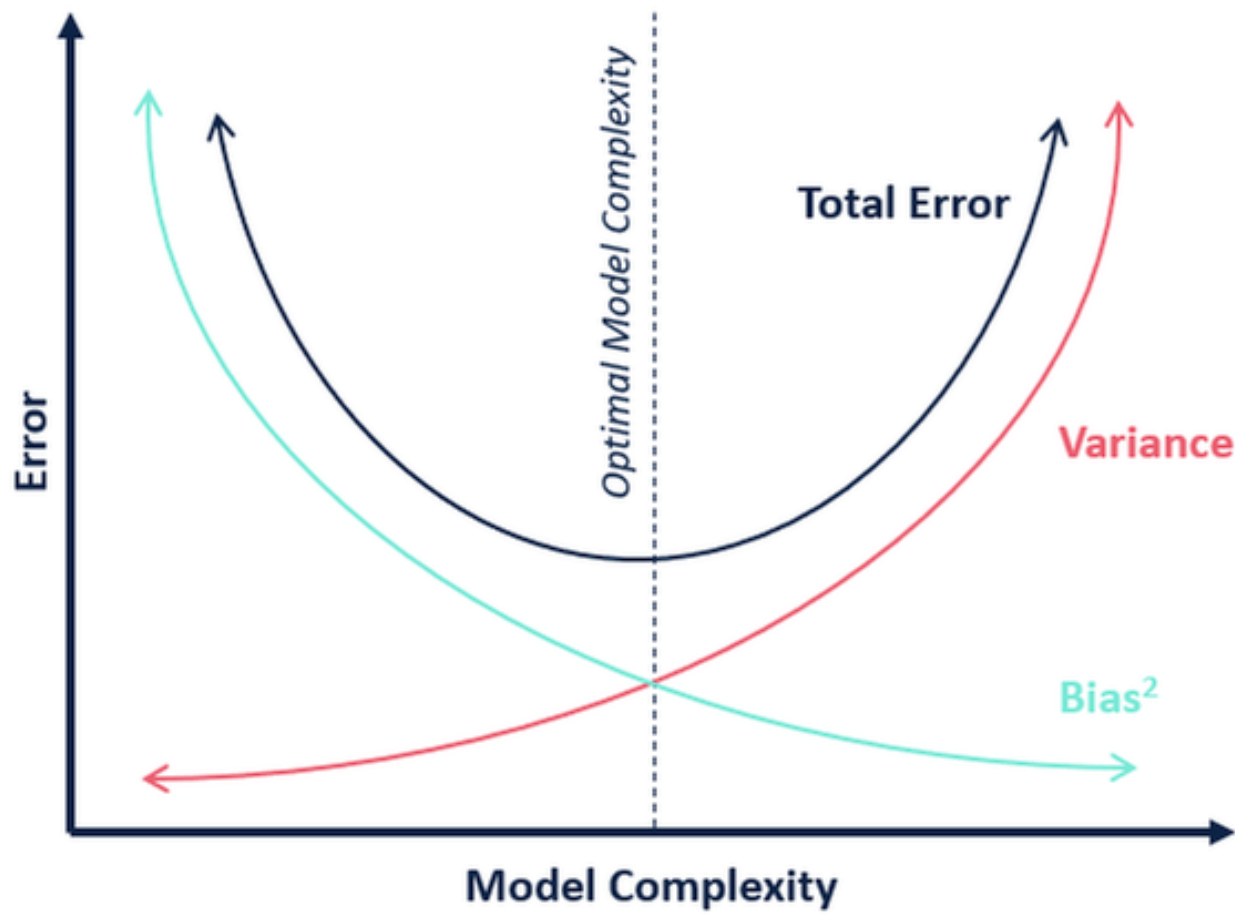
Дилемма «отклонение – дисперсия»

Переобучение также тесно связано с дилеммой «отклонение – дисперсия» (**Bias-Variance Tradeoff**).

Отклонение (смещение, bias error) – это ошибочное предположение алгоритма МО, т.е. ситуация, когда алгоритмом не были определены все закономерности в данных. Отклонение обратно коррелирует с условной «сложностью» системы.

Дисперсия (Variance) в контексте МО – это ситуация, когда алгоритм находит мнимые закономерности в «шумах», то есть случайных отклонениях данных. Дисперсия прямо коррелирует со «сложностью» системы

Дилемма «отклонение – дисперсия»



Дилемма «отклонение – дисперсия»

