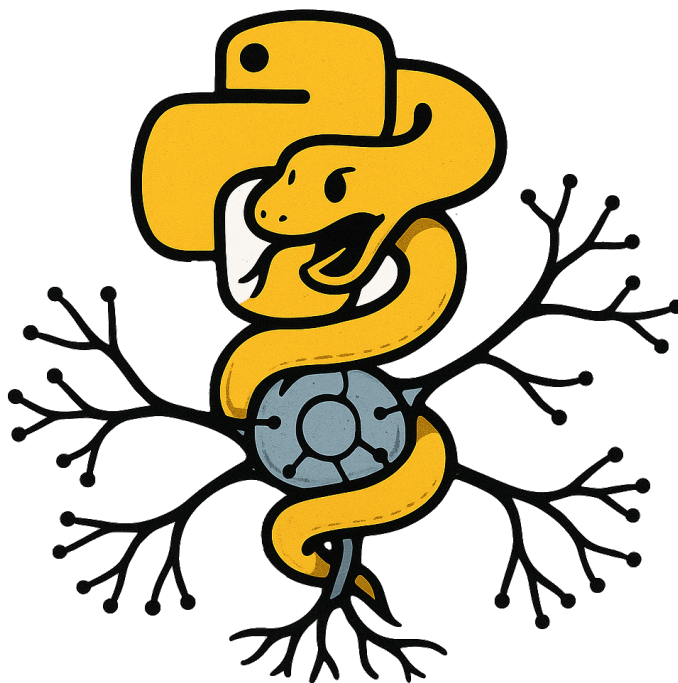


# Основы Машинного Обучения



**Автор:** Марьичев Алексей

Нижегородский государственный университет им. Лобачевского

# Содержание:

### Основные задачи ML:

- **Классификация** — определение объектов к определённым классам по общим признакам
  - **Регрессия** — прогнозирование величин, функций или событий
  - **Ранжирование** — упорядочивание входного набора данных
- 

## 1.1 Обучающая выборка

Представление объектов в виде различных векторов данных:

$$X = [x_1, x_2, \dots, x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

X - вектор входных данных

Допустим, у нас дана матрица:

$$A = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

Здесь A - матрица входных данных,  $n$  — количество признаков объекта, а  $m$  — количество самих объектов.

Таким же видом представлены и выходные данные:

$$Y = [y_1, y_2, \dots, y_n]^T = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{bmatrix}$$

Y - вектор выходных данных

Теперь мы рассмотрим важный вопрос: как же такие объекты как изображения, звук и т. д. могут представляться в виде векторов?

Допустим, на вход задаче подается Изображение:



Теперь важное замечание-**размерность вектора  $n$  будет зависеть от количества пикселей в изображении**

Например, если изображение 1024 на 256, то размерность вектора будет  $1024 * 256 = 262144$

$$X = [x_1, x_2, \dots, x_{262144}]$$

Теперь **объединим** эти понятия:

$$X' = \{(x_i, y_i) | 0 < i < m\} - \text{размеченные данные (обучающая выборка)}$$

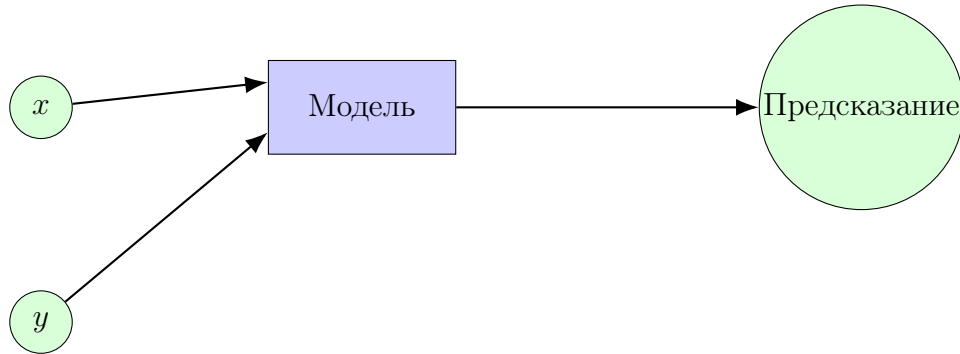
Это и является одним из важнейших понятий в области машинного обучения, с которым вы будете встречаться всюду.

---

## §2 Постановка задачи для модели

А теперь разберемся с тем, как же модель будет "обучаться":

Допустим, у нас есть размеченные данные  $(x_i, y_i)$ ,



В результате из исходных данных мы получили некое предсказание, которое на первых этапах обучения может не иметь ничего общего с правильным ответом.

Теперь представим нашу модель как линейную функцию:

$$y(x) = \phi(x, \Delta)$$

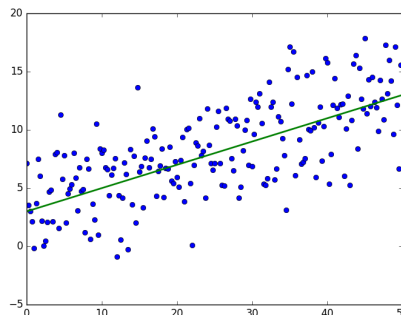
Здесь  $\Delta$  – **постоянно меняющийся параметр**

Его мы будем подстраивать для наиболее точного ответа нашей модели

Для лучшего понимания перейдем к задаче линейной регрессии: Задана функция:

$$y(x, k, b) = kx + b + \psi$$

Здесь  $k$  и  $b$  - параметры от которых зависит угол поворота прямой а так же ее сдвиг Т.е. получается, что эта прямая может проходить как угодно, но за счет размеченных данных мы задаем модели желаемый результат:



И получается, что во время обучения модель дает прогнозы все точнее и точнее к желаемому результату.

Но как же наш алгоритм понимает, что ответ надо корректировать?  
Сейчас мы подошли к еще одному очень важному определению в области ML:

**Функция потерь - функция, которая характеризует потери при неправильном предсказании модели** Примеры таких функций:

$$L(x, a) = |a(x) - y(x)|$$

- абсолютная ошибка

$$L(x, a) = |a(x) - y(x)|^2$$

- квадратичная ошибка