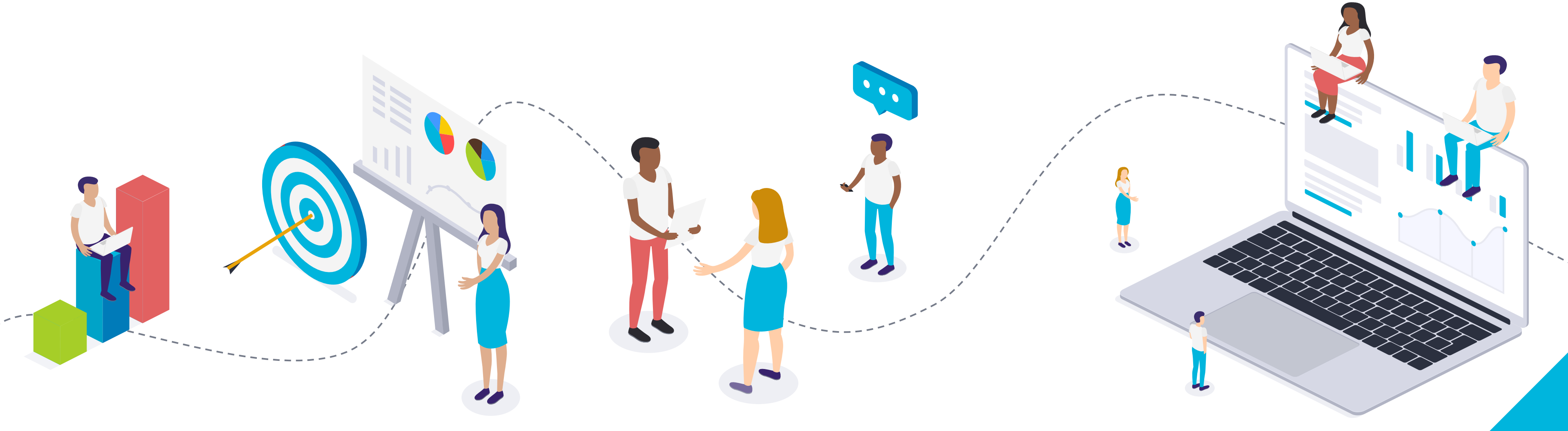
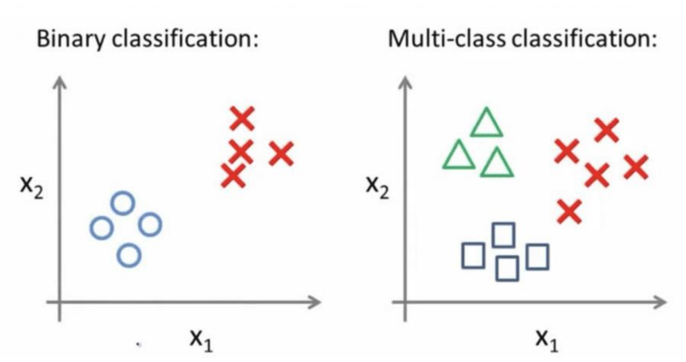
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
федеральное государственное бюджетное образовательное учреждение высшего образования
«АЛТАЙСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

# Программа профессиональной переподготовки «Технологии искусственного интеллекта, визуализации и анализа данных»



# Задача классификации

- **Классификация** заключается в присваивании объектам меток классов. Метка класса представляет собой дискретное, неупорядоченное значение, которое может пониматься как *принадлежность к группе объектов*.
  - Бинарная классификация: $Y = \{-1, 1\}$
  - Многоклассовая классификация: $Y = \{0, 1, 2, \ldots, n\}$



# Задача бинарной классификации

- Модель линейной регрессии:

$$\alpha(x) = \omega_0 + \sum_{j=1}^{d} x_j \omega_j$$

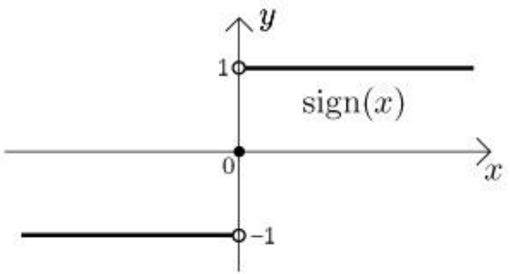
- Линейный классификатор:

$$\alpha(x) = \operatorname{sign}(\omega_0 + \sum_{j=1}^{d} x_j \omega_j)$$

$$\alpha(x) = \operatorname{sign}(\langle \omega, x \rangle)$$

$\alpha(x) = \operatorname{sign}(\langle \omega, x \rangle - t), t$ – порог

$$\operatorname{sign}(x) = \begin{cases} 1, & x > 0; \\ 0, & x = 0; \\ -1, & x < 0. \end{cases}$$



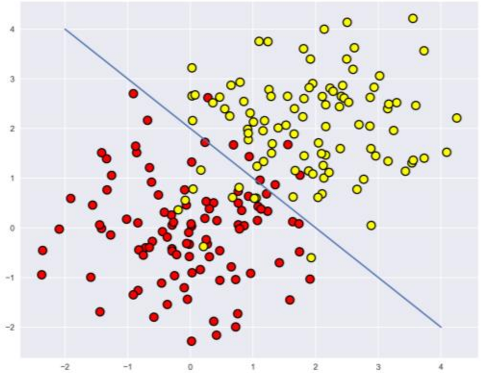

Что делать если 0?
Будем считать, что это отказ от классификации.

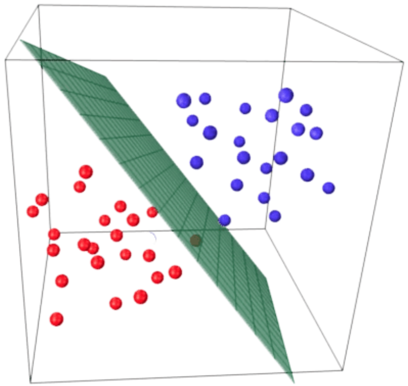
## Геометрический смысл

- Уравнение гиперплоскости $\langle \omega, x \rangle = 0$

Знак $\langle \omega, x \rangle$ показывает с какой стороны объект $x$ находится относительно гиперплоскости:

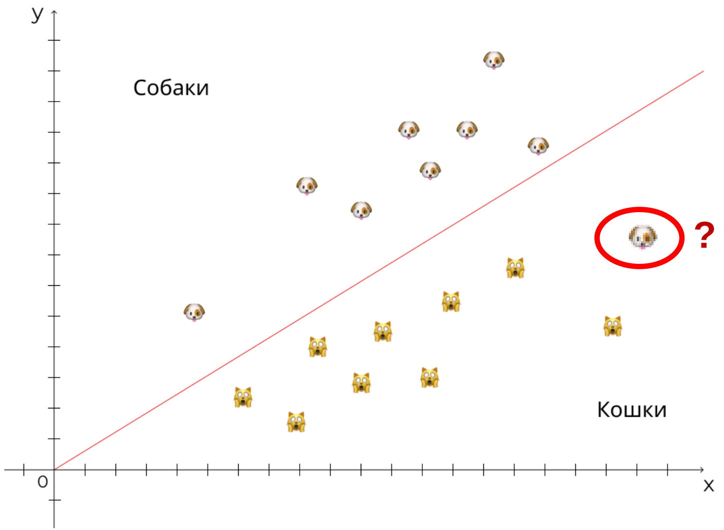
- $\langle \omega, x \rangle > 0$ – объект находится «справа»
- $\langle \omega, x \rangle < 0$ – объект находится «слева»
- $\langle \omega, x \rangle = 0$ – объект лежит на гиперплоскости





# Отступ

- $\frac{|\langle \omega, x \rangle|}{\|\omega\|}$ – расстояние от точки до гиперплоскости.
- Чем больше $\langle \omega, x \rangle$ тем дальше объект от гиперплоскости.
- Отступ (margin) $M_i = y_i \langle \omega, x_i \rangle$ определяет корректность классификатора: положительный отступ соответствует правильному ответу, а отрицательный – неправильному.
- Чем больше $|M_i|$ тем больше модель уверенна в ответе.



# Обучение классификатора

- Хотим увеличить долю правильных ответов алгоритма:

$$Q(\alpha, X) = \frac{1}{n}\sum_{i=1}^{n}[\alpha(x_i) = y_i] \ \to \max_{\omega}$$

- Или уменьшить долю неправильных ответов:

$$Q(\alpha, X) = \frac{1}{n}\sum_{i=1}^{n}[\alpha(x_i) \neq y_i] \ \to \min_{\omega}$$

Но возникает много проблем: нельзя использовать градиентные методы, может быть много глобальных минимумов.

- *Модифицируем*:

$$Q(\alpha, X) = \frac{1}{n}\sum_{i=1}^{n}[M_i < 0] \ \to \min_{\omega}$$

$$L(M) = [M < 0]$$

Отрицательный отступ на $i$-м объекте означает неправильный ответ алгоритма.

# Обучение классификатора

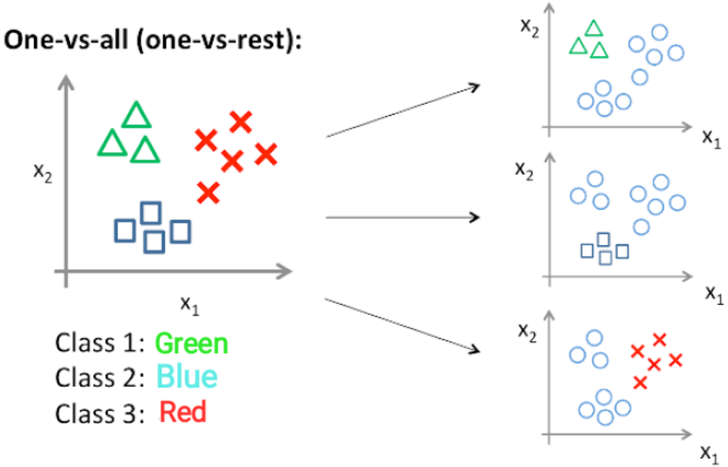
## Оценка сверху $L(M) \leq \tilde{L}(M)$

- $\tilde{L}(M) = \log(1 + e^{-M})$ (логистическая функция потерь используется в логистической регрессии)
- $\tilde{L}(M) = (1 - M)_+ = \max(0, 1 - M)$ (кусочно-линейная функция потерь используется методом опорных векторов)
- $\tilde{L}(M) = (-M)_+ = \max(0, -M)$ (тоже кусочно-линейная и соответствует персептрону Розенблатта)
- $\tilde{L}(M) = e^{-M}$ (экспоненциальная функция потерь используется в алгоритмах бустинга)
- $\tilde{L}(M) = \frac{2}{1+e^M}$ (сигмоидальная функция потерь)

# Многоклассовая классификация

- Один против всех (one-vs.-rest). Строим $k$ классификаторов, каждый из которых должен отличать $k$-й класс от всех остальных. Итоговый классификатор будет выдавать класс, соответствующий самому уверенному из бинарных.

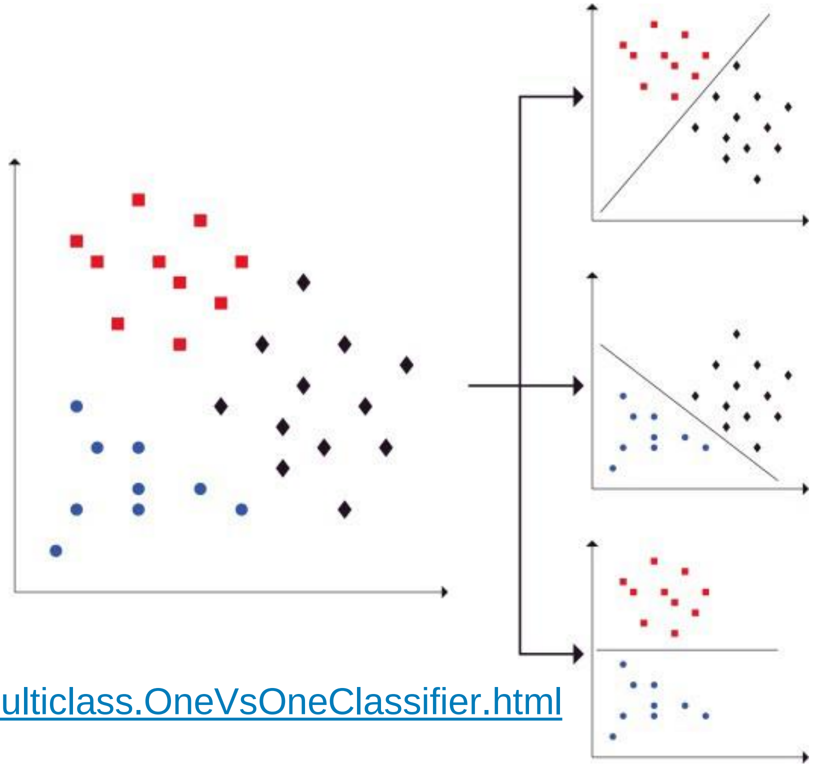
Это яблоко? Нет.
Это медведь? Нет.
Это конфета? Да.




https://scikit-learn.org/stable/modules/generated/sklearn.multiclass.OneVsRestClassifier.html

# Многоклассовая классификация

- Один против одного (one-vs.-one). Строим бинарные классификаторы, способные отличать пару классов. Итоговый классификатор выбирает ответ по наибольшему количеству голосов.



https://scikit-learn.org/stable/modules/generated/sklearn.multiclass.OneVsOneClassifier.html