

# Искусственные нейронные сети

Копылов Д.Е., Михайлов А.А.

<sup>1</sup>Институт динамики систем и теории управления им. В.М. Матросова  
Сибирского отделения Российской академии наук

<sup>2</sup>Институт системного программирования им. В.П. Иванникова  
Российской академии наук

<sup>3</sup>Институт математики и информационных технологий  
Иркутский государственный университет

13 августа 2024 г.

# Лекция 2

## формальная постановка задачи

Пусть есть два множества  $X$  - множества объектов,  $Y$  - множество ответов и предполагается, что существует функциональная зависимость

$$f : X \rightarrow Y. \quad (1)$$

Зависимость не известна, известна только совокупность  $S$  пар вида (объект, ответ)

$$S = \{(x_i, y_{x_i} = f(x_i)) \in X \times Y | i = 1, \dots, l\} \quad (2)$$

Требуется найти приближенный вид этой функции  $f$  путем построения аппроксимирующей функции  $f_S(x) : X \rightarrow Y$ , такую, что

$$\forall x \in X \quad f_S(x) \approx f(x). \quad (3)$$

# Переход от объектов к векторам

## Множества признаков

$I$  - множество признаков (features),

Каждому  $i \in I$  сопоставляется некоторое множество  $D_i$   
(множество значений признака  $i$ )

## Признаки объектов

У объекта зафиксируем конечное число признаков  $n$  и пронумеруем их  $1, 2, \dots, n$ .

$$x^i \in D_i$$

$$x = (x^1, x^2, \dots, x^n)^T \in D_1 \times D_2 \times \dots \times D_n$$

# Виды задач по возвращаемым ответам

## Множество ответов

1.  $Y = \{0, 1\}$  - бинарная классификация (фильтр Байеса)
2.  $Y = \{c_1, c_2, \dots, c_m\}$  - классификация
3.  $Y = \{ "a" "b" \dots, "z" \}$  - задача распознавания символов
4.  $Y = \{1, 2, \dots, N\}$  и  $i \neq j, y_i \neq y_j$  - задача ранжирования
5.  $Y = \{1, 2, 3, \dots N, \dots\}$  - кластеризация
6.  $Y = R^m$  - задача регрессии

# Функция потерь

## Функция потерь (loss function)

Функция потерь - функция, которая сопоставляет паре  $f_S, x$ , где  $x \in X$  число  $Loss(f_S, x)$  и характеризует величину ошибки аппроксимации  $f_S$  на объекте  $x \in X$ .

## Виды функций потерь

1.  $Loss(f_S, x) = \begin{cases} 0, & \text{if } f_S(x) = f(x); \\ 1, & \text{else} \end{cases}$
2.  $Loss(f_S, x) = |f_S(x) - f(x)|$   $Loss(f_S, x) = (f_S(x) - f(x))^2$

# Функция потерь для выборки

## Функция потерь для обучающей выборки

Если  $S' = \{(x'_i, y'_i), i = 1, 2, \dots, l'\}$  соответствует функции  $f : X \rightarrow Y$ , также как и  $S$ , то

$$Loss(f_S, S') = \frac{1}{l'} \sum_{i=1}^{l'} Loss(f_S, x'_i) \quad (4)$$

Для всего  $S$  это функция эмпирического риска

$$Q(f_S) = Loss(f_S, S) \quad (5)$$

# Гиперплоскости

Гиперплоскость в пространстве  $R^n$  называется множество точек  $x \in R^n$  удовлетворяющих уравнению

$$\langle x, w \rangle = w_0. \quad (6)$$

Гиперплоскостью можно разделить множество  $X$  на  $X^+$  и  $X^-$ , так что  $\forall x \in X^+ \langle x, w \rangle \geq w_0$  и  $\forall x \in X^- \langle x, w \rangle \leq w_0$ .

## Подбор параметров $w$

Если функция  $Y = \{-1, 1\}$ , то в качестве функции  $f_S(x)$  можно выбрать следующую функцию

$$f_S(x) = \text{sign}(\langle x, w \rangle - w_0) \quad (7)$$

$$(w, w_0) = \arg \min_{w, w_0} Q(f_S) \quad (8)$$



