

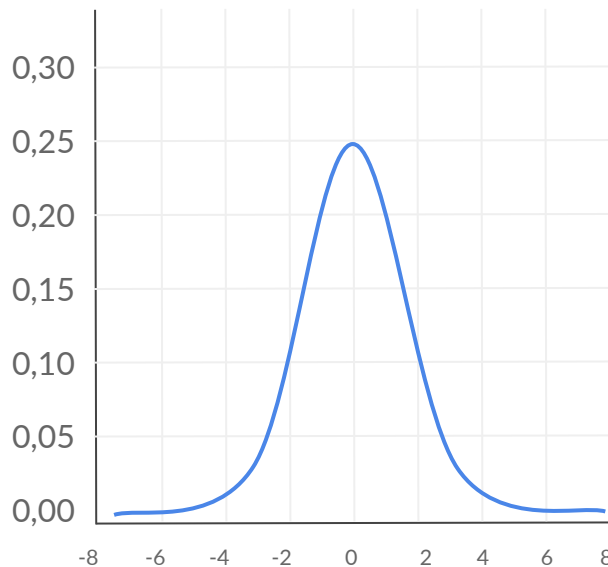
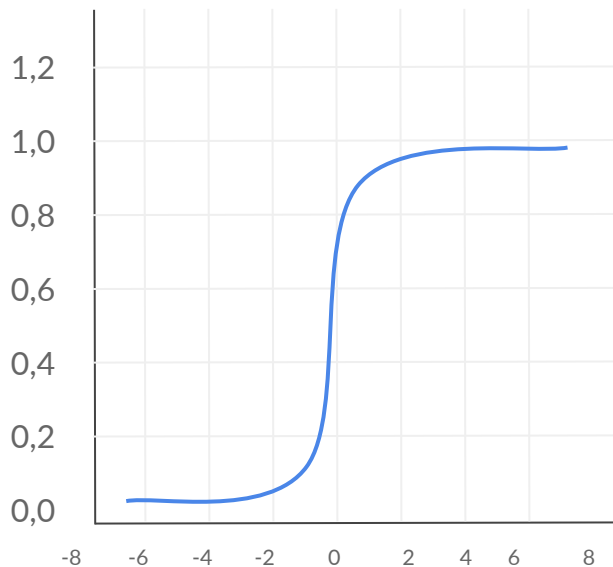


Функции активации

Тетерин Михаил Александрович

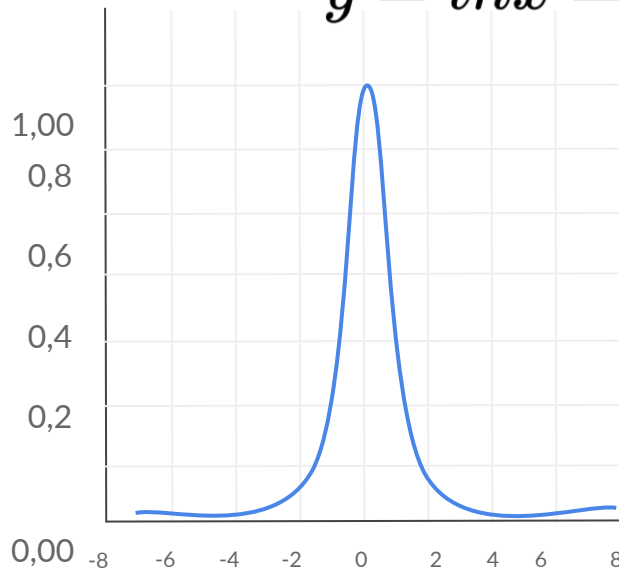
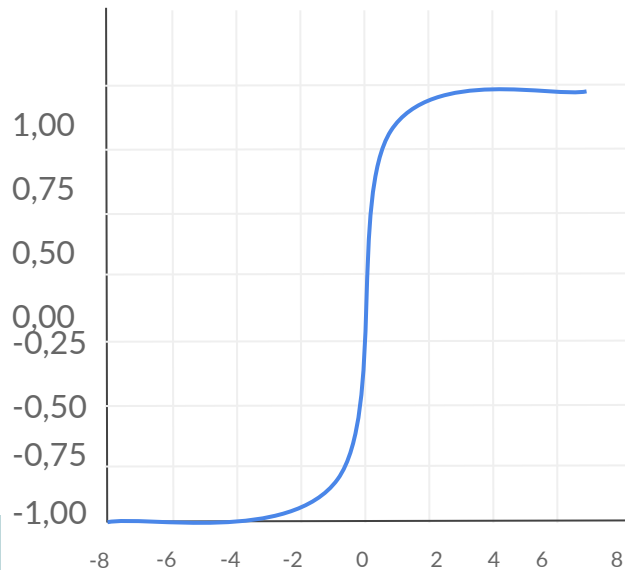
- сжимает числа до диапазона $[0, 1]$.
- исторически популярны
- максимум производной равен 0.25
- сигмоидальные выходы не центрированы по нулю
- \exp дорогая операция для вычисления

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

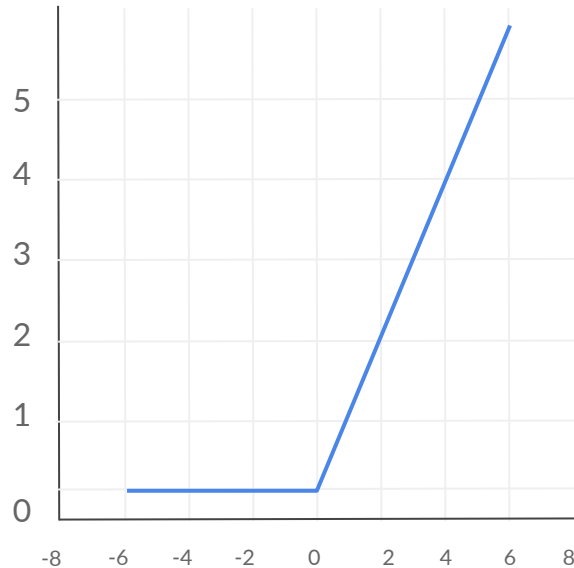


- сжимает числа до диапазона $[-1, 1]$
- значения центрированы относительно оси абсцисс
- проблема мертвых градиентов все равно существует
- \exp дорогая операция для вычисления

$$y = \tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{\sinh x}{\cosh x}$$

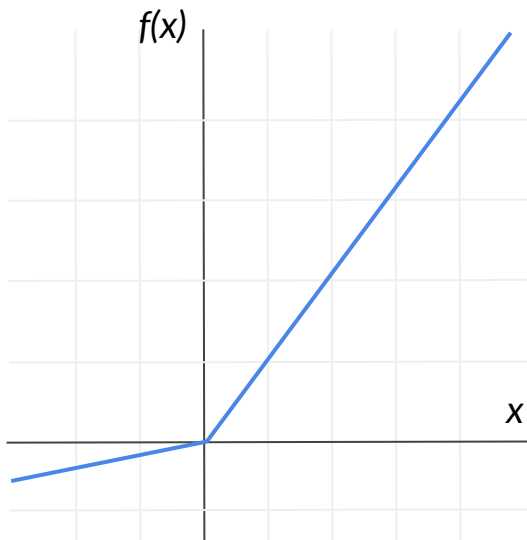


- не возникает проблем с обучением там где $x > 0$
- сходится в 6 раз быстрее чем гиперболический тангенс и сигмоида
- **сигмоидальные выходы не центрированы по нулю**
- **проблема мертвых градиентов**



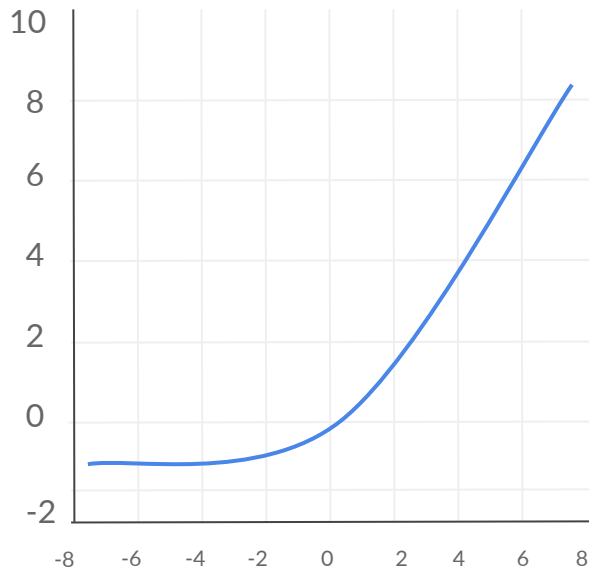
$$f(x) = \begin{cases} x, & \text{if } x > 0 \\ 0 & \end{cases}$$

- не возникает проблем с насыщением
- сходится быстрее чем гиперболический тангенс и сигмоида
- Вычислительно эффективна
- нет проблем мертвых градиентов



$$f(x) = \max(0.01x, x)$$

- Все преимущества RELU
- Нет проблемы мертвых градиентов
- Выходы центрированы относительно 0
- **exp дорогая операция для вычисления**



$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$

- Обобщение логистической функции для многомерного случая
- Используется в многоклассовых задачах

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}}$$

Название функции	Формула $f(x)$	Производная $f'(x)$
Логистический сигмоид σ	$\frac{1}{1 + e^{-x}}$	$f(x)(1 - f(x))$
Гиперболический тангенс \tanh	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	$1 - f^2(x)$
Soft Sign	$\frac{x}{1 + x }$	$\frac{1}{(1 + x)^2}$
Ступенька (функция Хевисайда)	$\begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$	$\begin{cases} 0 \\ 1 \end{cases}$
SoftPlus	$\log(1 + e^x)$	$\frac{1}{1 + e^{-x}}$
ReLU	$\begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$	$\begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$
Leaky ReLU, Parameterized ReLU	$\begin{cases} ax, & x < 0 \\ x, & x \geq 0 \end{cases}$	$\begin{cases} a, & x < 0 \\ 1, & x \geq 0 \end{cases}$
ELU	$\begin{cases} a(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}$	$\begin{cases} f(x) + a, & x < 0 \\ 1, & x \geq 0 \end{cases}$

Классификация	Формат метки	Функция активации	Функция потерь
Бинарная	Вероятность первого класса	Сигмоидальная	Бинарная кросс-энтропия
Бинарная	Унитарный код (One-hot encoding)	Softmax	Категориальная кросс-энтропия
Многоклассовая	Унитарный код (One-hot encoding)	Softmax	Категориальная кросс-энтропия
Многоклассовая	Номер класса	Softmax	Разряженная категориальная кросс-энтропия
Многометочная	Унитарный код (One-hot encoding)	Сигмоидальная	Категориальная кросс-энтропия

Спасибо за внимание

Тетерин Михаил Александрович

michail.teterin92@gmail.com

2022

ОБРАЗОВАТЕЛЬНЫЕ ПРОГРАММЫ В ОБЛАСТИ
ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА