



SCHOOL
OF ENERGY
& POWER ENGINEERING

Основы нейросетевых вычислений

Сергей Владимирович Аксёнов,

Доцент отделения информационных технологий ИШИТР,

Томский политехнический университет

Томск-2023

Искусственные нейронные сети

Нейронная сеть искусственная нейронная сеть, ИНС, или просто нейросеть) — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.

Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы.

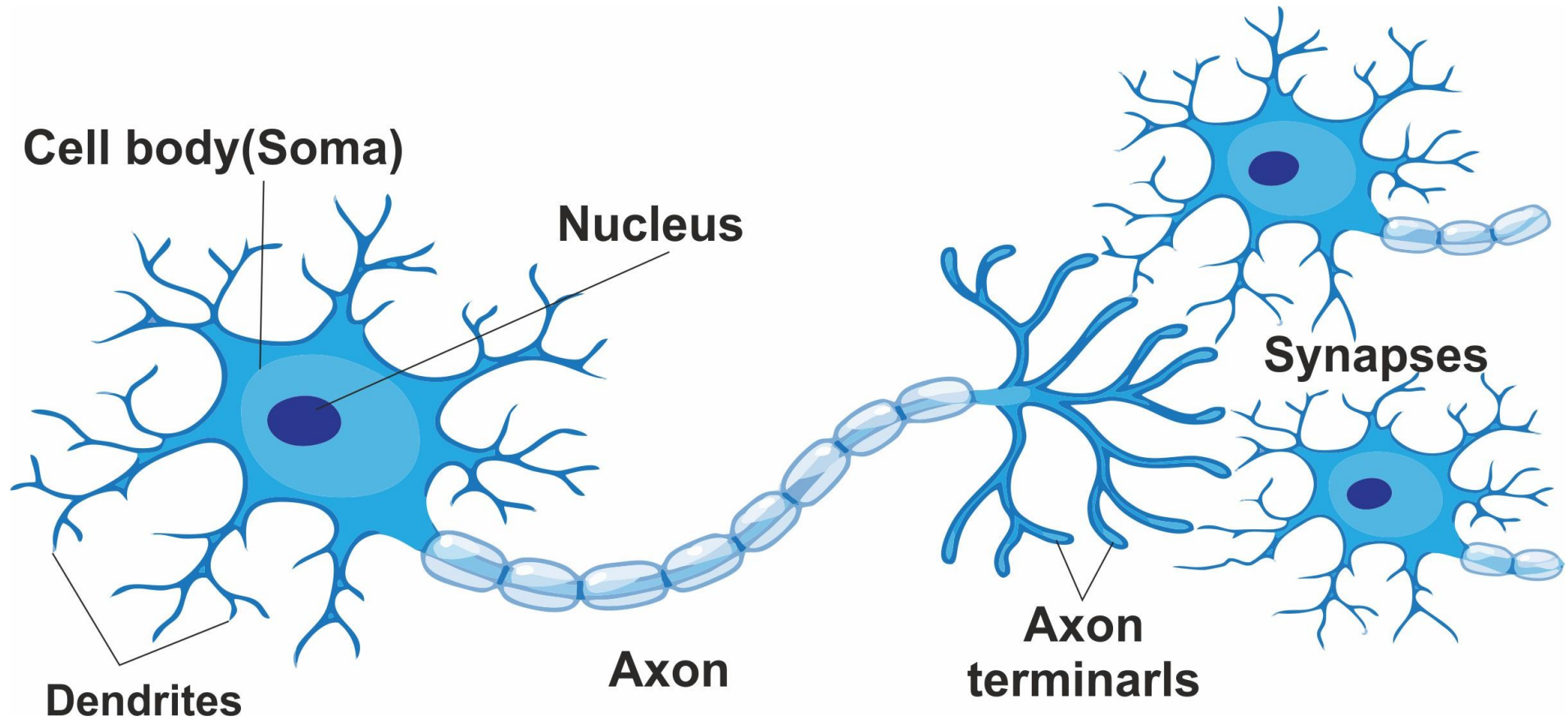
ИНС представляет собой систему соединённых и взаимодействующих между собой простых процессоров (нейронов).

Будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие по отдельности простые процессоры вместе способны выполнять довольно сложные задачи.

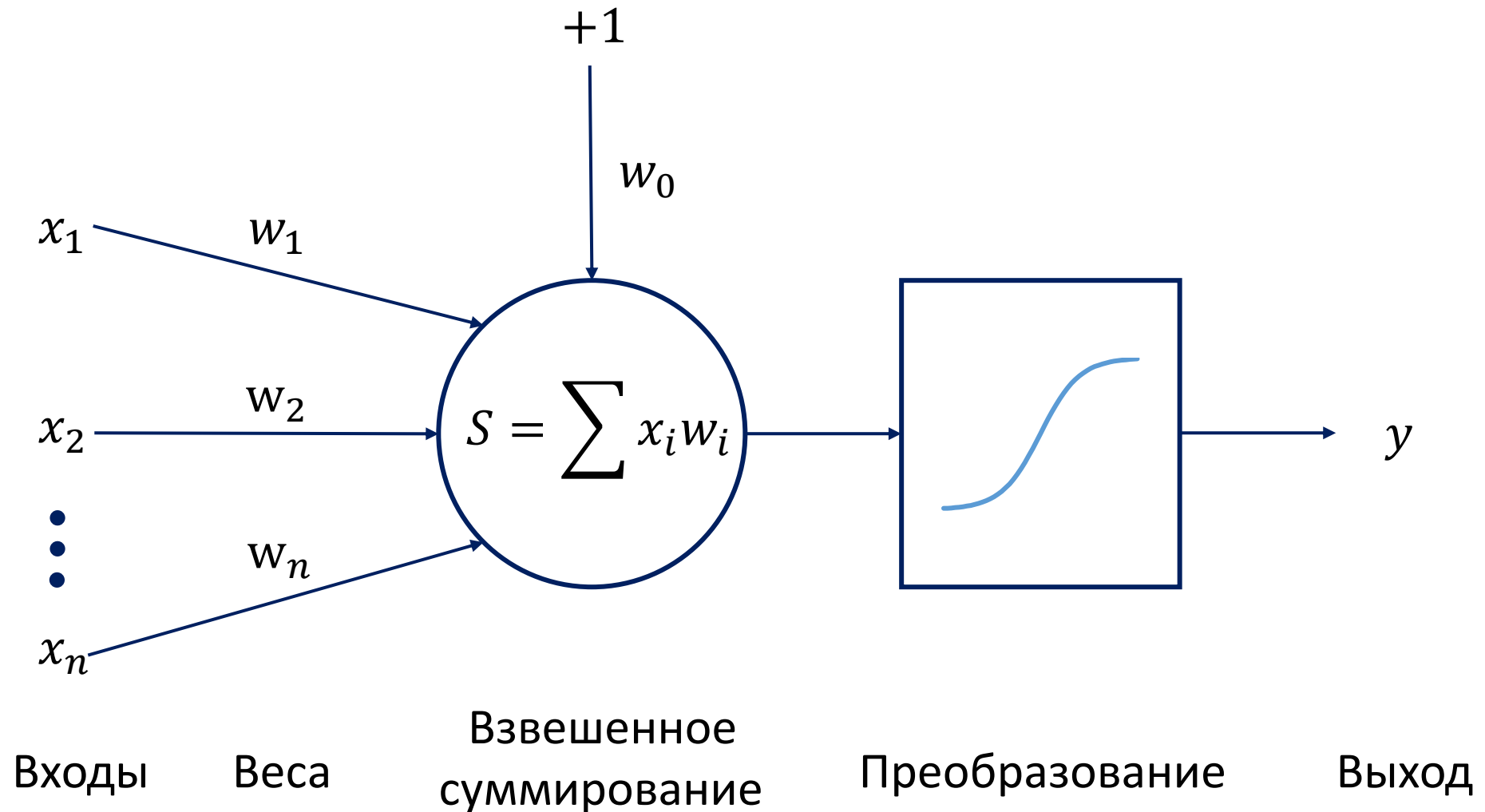
Преимущества нейросетевых вычислений

1. Устойчивость к шумам
2. Возможность дообучения
3. Отказоустойчивость
4. Параллельная обработка данных
5. Наличие продвинутых библиотек (Keras/Tensorflow, PyTorch)

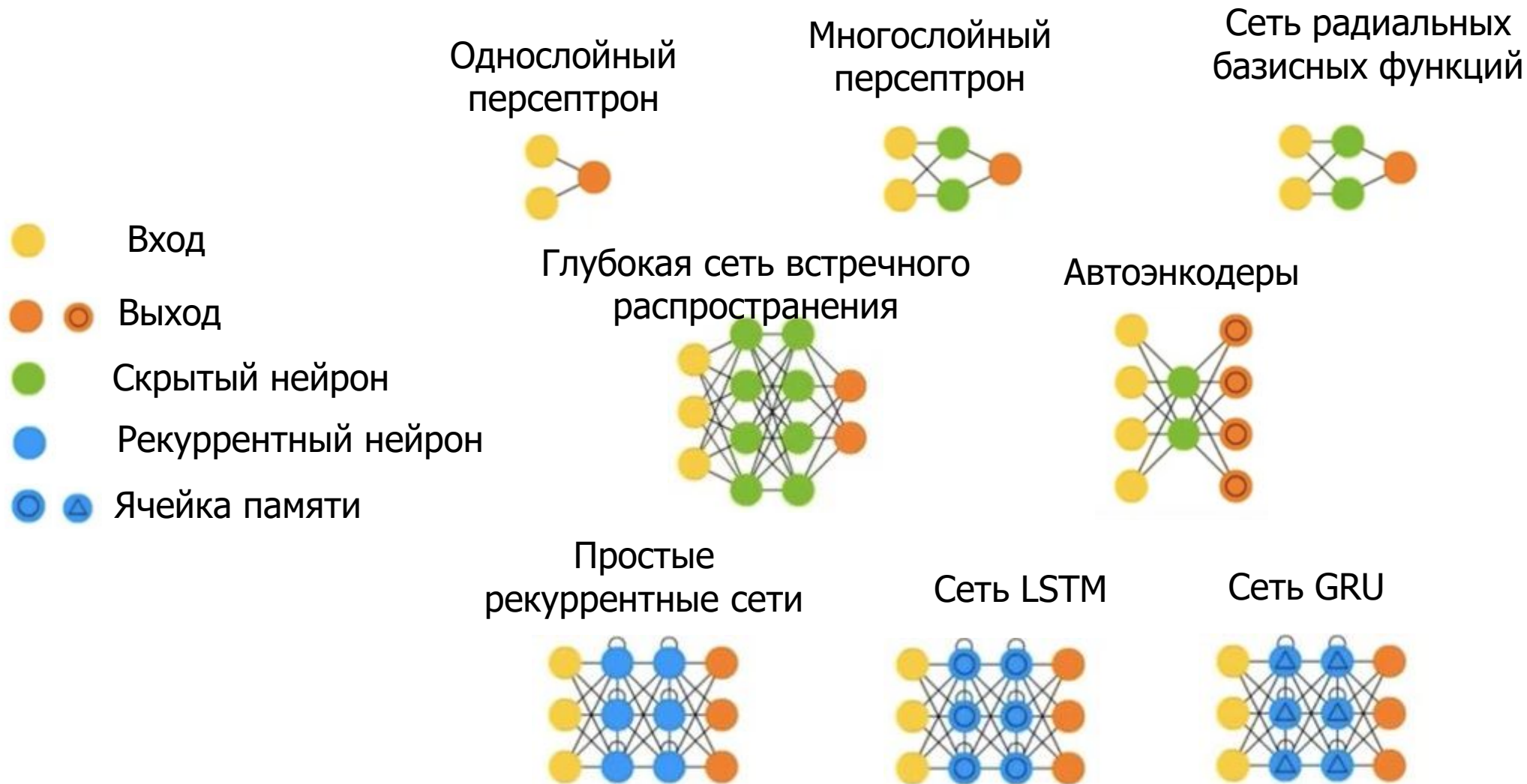
Биологический нейрон



Математическая модель нейрона

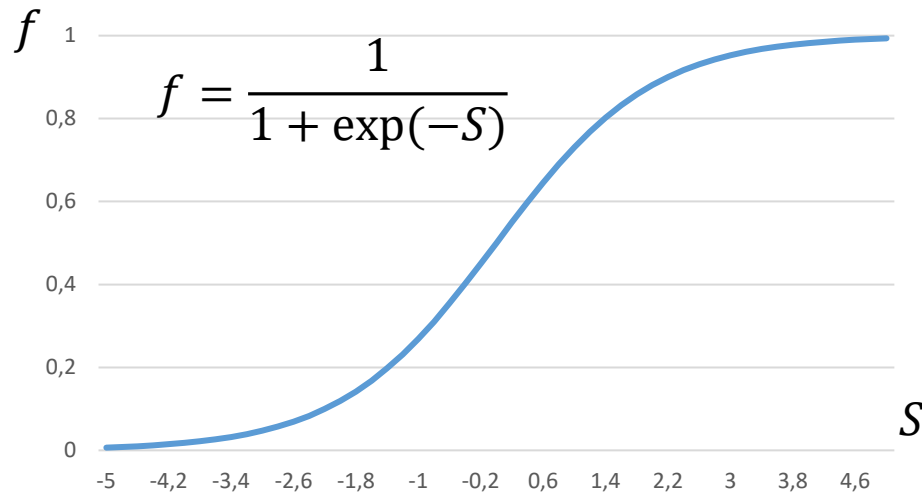


Нейронные сети

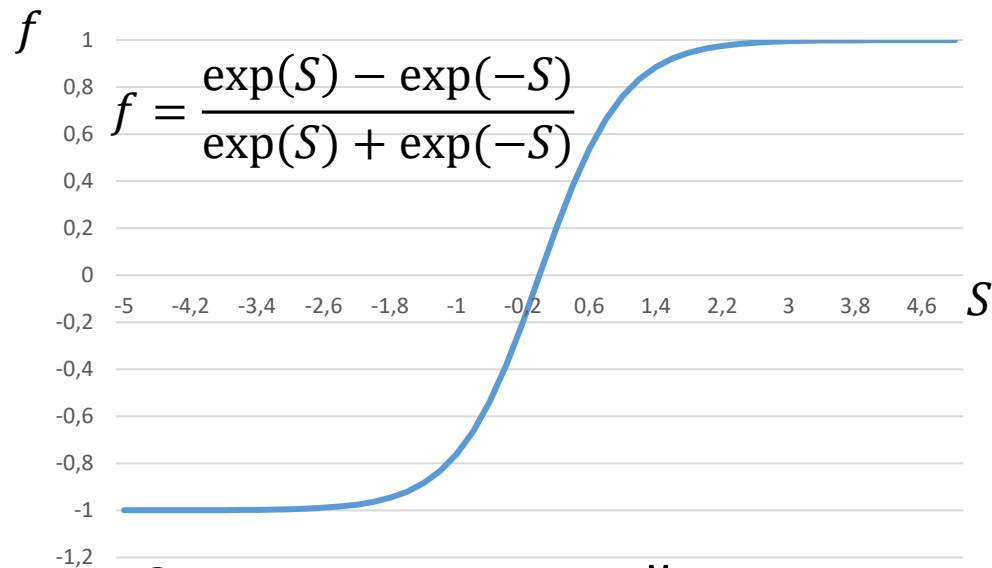
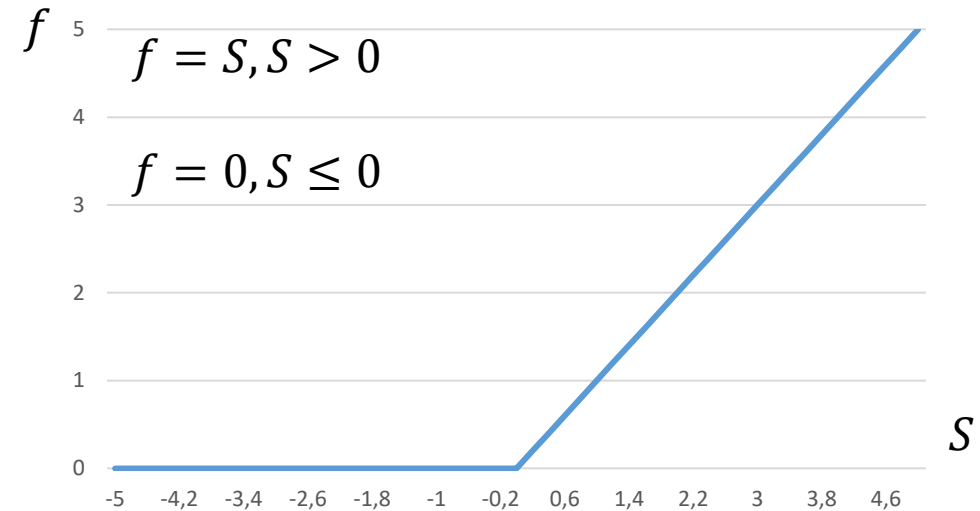


Функции активации нейрона

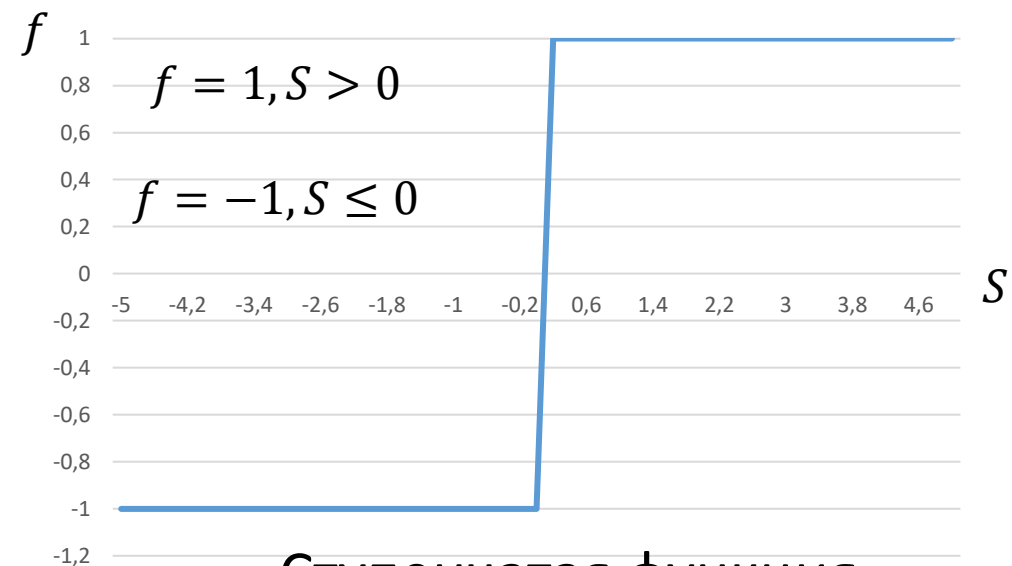
Логистическая сигмоида



ReLu

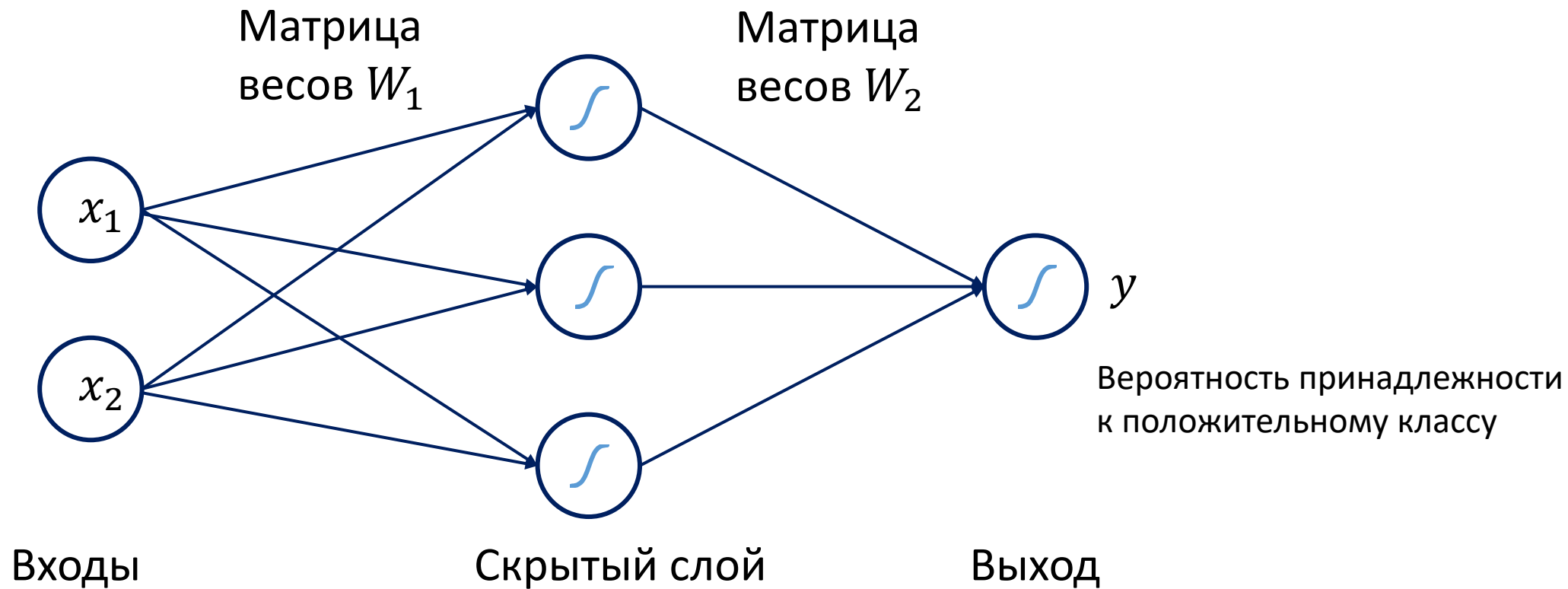


Экспоненциальный тангенс



Ступенчатая функция

Полносвязная сеть для бинарной классификации



Популярные функции потерь

Средняя квадратичная ошибка (MSE) : Регрессия

$$J_{MSE} = \frac{1}{2N} \sum_{i=1}^N |y^{(i)} - d^{(i)}|^2$$

N — число выходных нейронов

d — желаемый выход

Бинарная кросс-энтропия: Бинарная классификация

$$y^i = \sigma(W^T h^{(i)} + b)$$

h — выход нейронов скрытого слоя

$$J_{BinXENT} = \frac{1}{N} \sum_{i=1}^N [-d^{(i)} \log p(y^{(i)}) - (1 - d^{(i)}) \log(1 - y^{(i)})]$$

Категориальная кросс-энтропия: Многоклассовая классификация

$$y_j^{(i)} = \text{softmax}(W_j^T h^{(i)} + b_j) = \frac{e^{W_j^T h^{(i)} + b_j}}{\sum_k e^{W_k^T h^{(i)} + b_k}}$$

$$J_{CatXENT} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C d_j^{(i)} \log y_j^{(i)}$$

Обучающий, тестовый и валидационные наборы

Обучающий набор – множество примеров, использующихся для настройки параметров модели. На основании результата, полученного путем активации модели обучающим примером, производится корректировка весов сети.

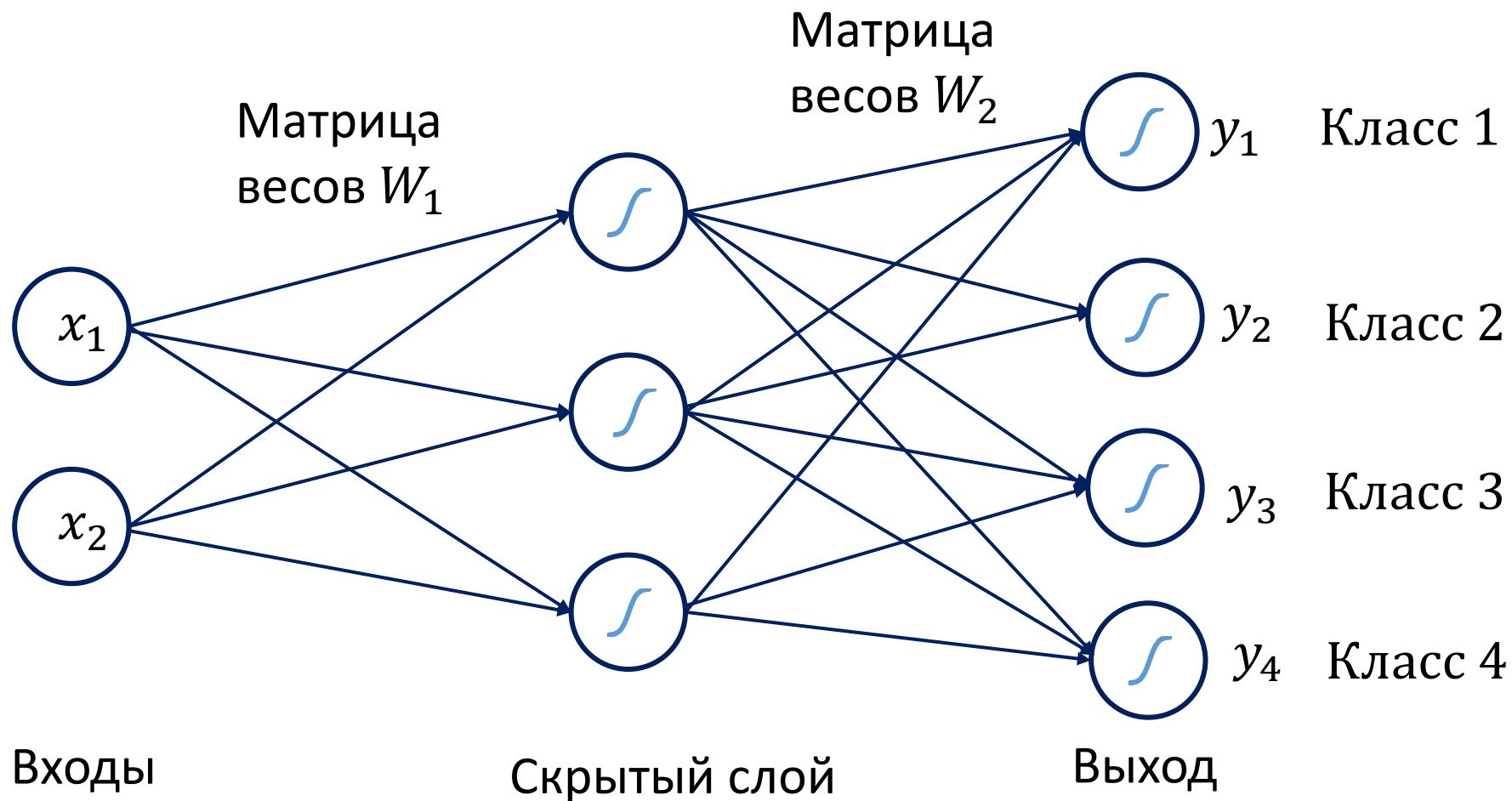
Валидационный набор – выборка, используемая для беспристрастной оценки соответствия модели обучающему набору данных при настройке гиперпараметров модели. Используется для минимизации переобучения.

Тестовый набор – данные, используемые для обеспечения объективной оценки окончательной модели, настроенной на обучающем наборе.

Популярные библиотеки (на 2023 г.)

- **TensorFlow**: Вычислительная библиотека, основанная на графах потоков данных. Эти графы имеют узлы, представляющие математические операции, и ребра, представляющие тензоры, которые протекают между ними. Написанный на Python, TensorFlow был разработан командой Google Brain.
- **Keras**: Простая высокоуровневая библиотека Python, разработанная для быстрого создания прототипов и проведения экспериментов. Он может работать поверх TensorFlow и CNTK и теперь является частью базовой библиотеки TensorFlow. Keras содержит реализации общих компонентов нейронной сети и многочисленные примеры архитектуры.
- **PyTorch**: Пакет Python для быстрого прототипирования нейронных сетей. Он основан на Torch, чрезвычайно быстрой вычислительной среде, и обеспечивает динамическое вычисление графа во время выполнения. PyTorch был разработан исследовательской группой Facebook AI.
- **Caffe**: Высокопроизводительная среда C++ для создания архитектур глубокого обучения, которые могут изначально поддерживать распределенное выполнение и выполнение с использованием нескольких графических процессоров. Текущая версия Caffe2 — это серверная часть, используемая Facebook в своих процессах.
- **CNTK**: Вычислительная среда, основанная на ориентированных графах. Она поддерживает языки Python, C# или C++ и была разработана Microsoft Research..
- **MXNet**: высокопроизводительная вычислительная платформа, написанная на C++ с собственной поддержкой графического процессора, в рамках проекта Apache.

Полносвязная сеть для многоклассовой классификации



Вероятность принадлежности
к классу

Полносвязная сеть - регрессор

