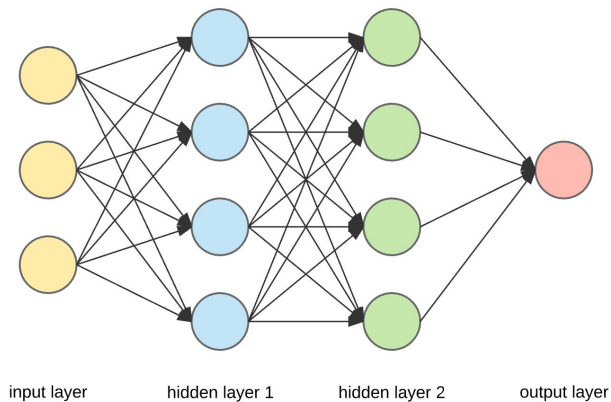


Модель нейрона

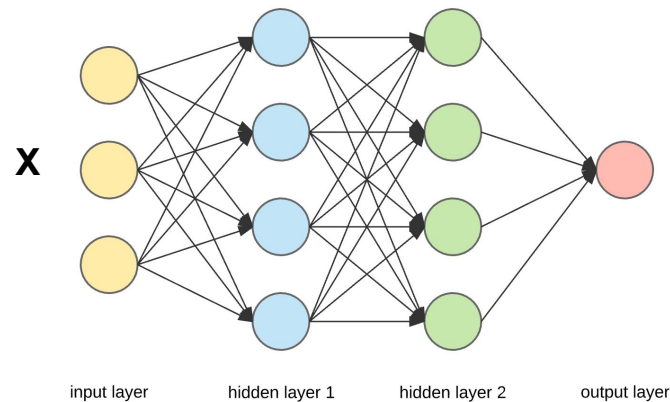
Основы Deep Learning

Полносвязная нейросеть



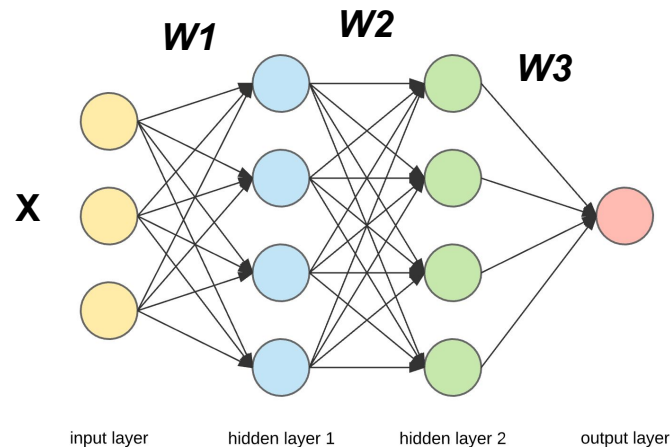
Полносвязная нейросеть

- Вход: числовая матрица **X**



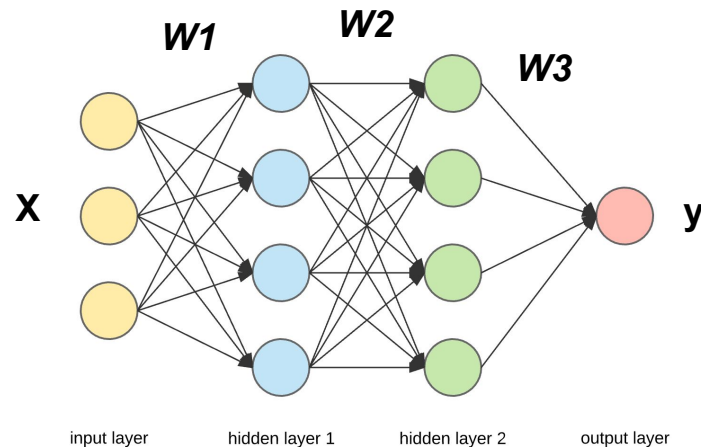
Полносвязная нейросеть

- Вход: числовая матрица X
- Внутри: матрицы параметров (веса нейросети)



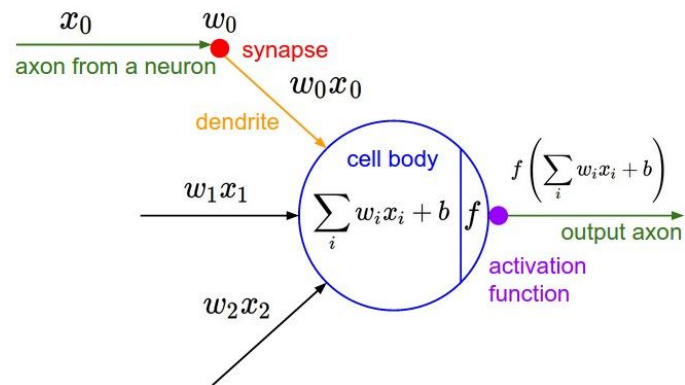
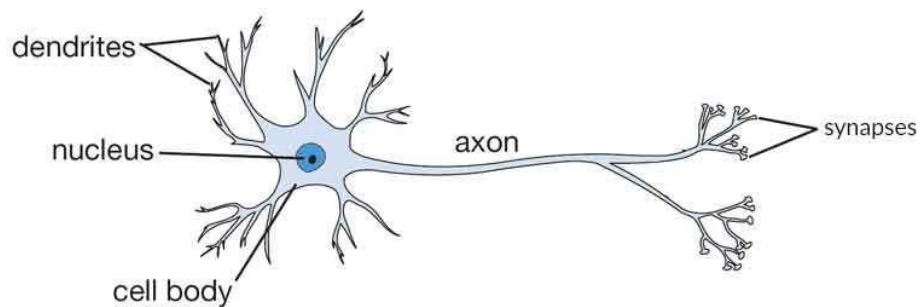
Полносвязная нейросеть

- Вход: числовая матрица X
- Внутри: матрицы параметров (веса нейросети)
- Выход: вектор ответов y
 - метки классов (классификация)
 - вещественные числа (регрессия)



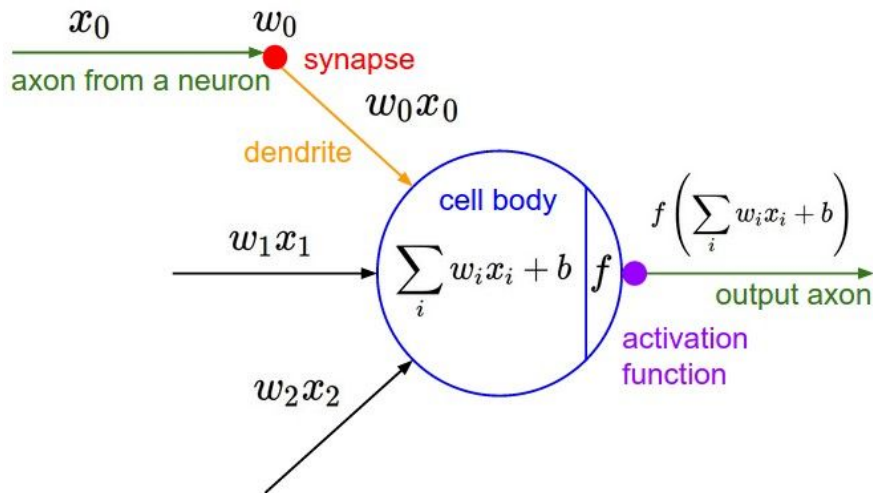
Один нейрон

Biological Neuron

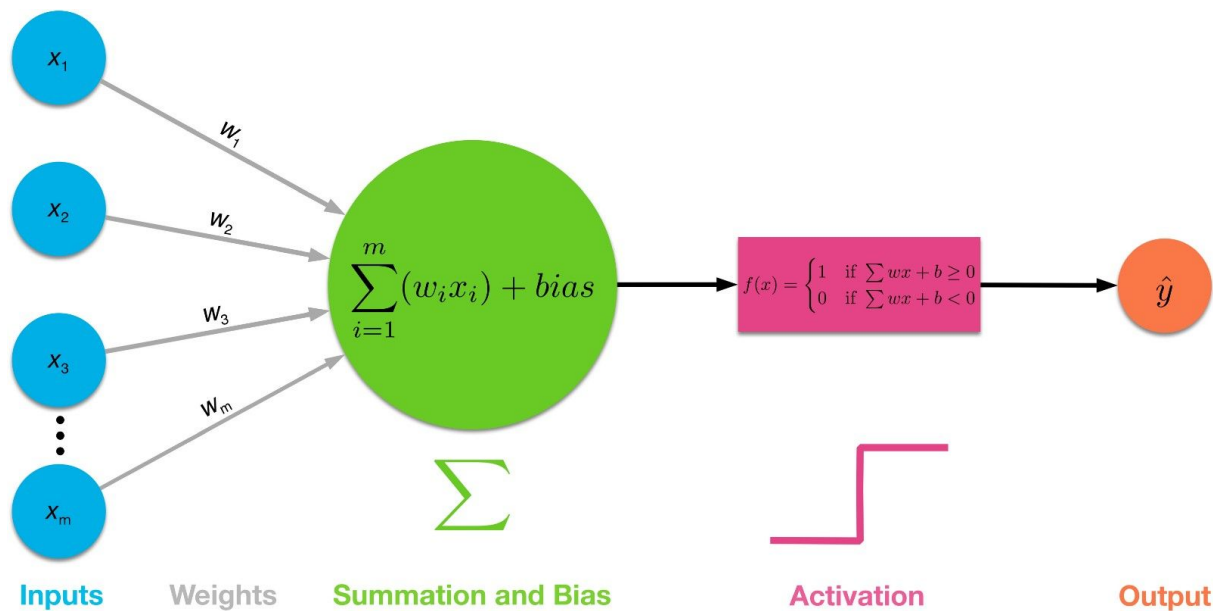


Искусственный нейрон

- Вход: вектор чисел \mathbf{x} размера $(k, 1)$
- Веса нейрона:
 - вектор чисел \mathbf{w} размера $(k, 1)$
 - скаляр \mathbf{b} (свободный член)
- Функция активации: \mathbf{f}
- Выход: метка класса \mathbf{y} (0 или 1)

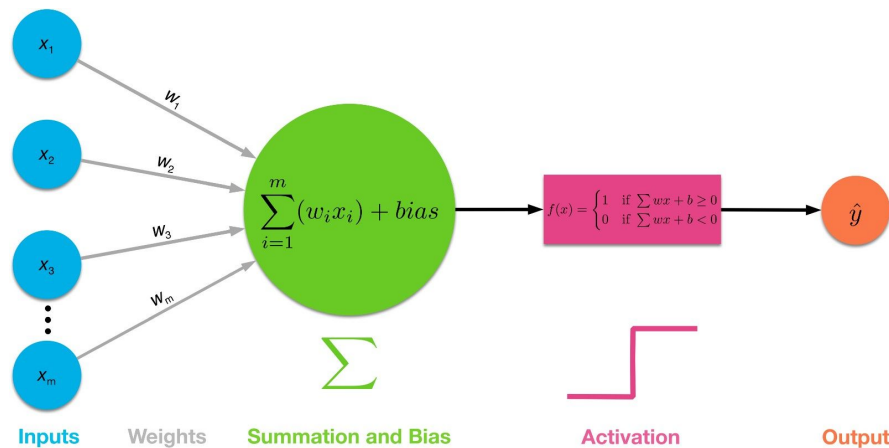


Перцептрон Розенблатта



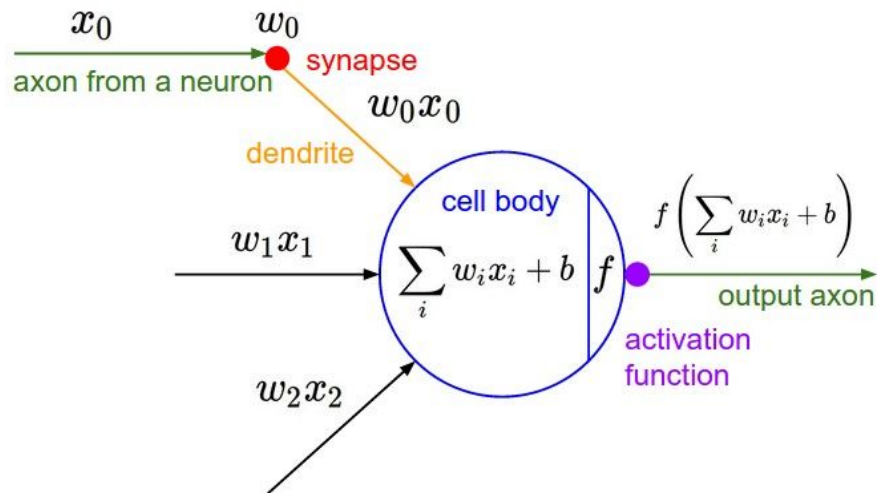
Перцептрон Розенблатта: обучение

- Метод коррекции ошибки :
метод обновления весов
- “Двигаем веса в соответствии с ошибками”
- Ранняя версия градиентного спуска для частного случая
- Теорема о сходимости перцептрона



Нейрон

- Модель та же, что и у перцептрона

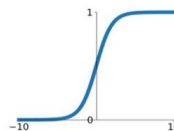


Нейрон

- Модель та же, что и у перцептрона
- Другие функции активации

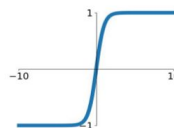
Sigmoid

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



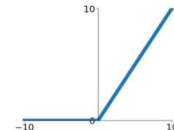
tanh

$$\tanh(x)$$



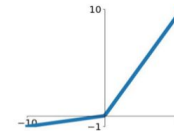
ReLU

$$\max(0, x)$$



Leaky ReLU

$$\max(0.1x, x)$$

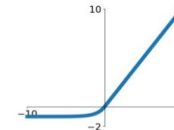


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

ELU

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



<https://arxiv.org/pdf/1710.05941.pdf>

Нейрон: обучение

- Задан Loss: функция потерь
- Задача обучения с учителем: есть ответы на каждом из объектов
- Оптимизируем градиентным спуском

$$w^{j+1} = w^j - \alpha \frac{\partial Loss}{\partial w}(w^j)$$