Нейронные сети

И. Е. Кураленок

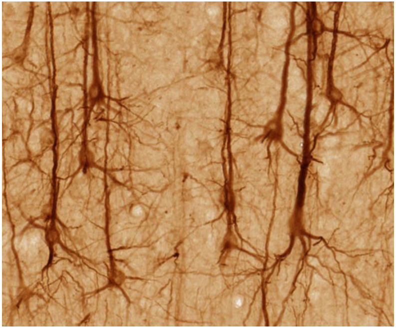
ikuralenok@gmail.com

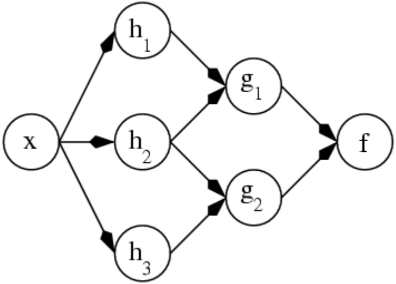
Нейронные сети: идея

Основная идея: “давайте научим машину ’думать’ по тем же принципам, что работает центральная нервная система”

Нейроны $\rightarrow$ Узлы
Аксоны $\rightarrow$ Связи (топология)
Синапс $\rightarrow$ Вес связи
Дендрит $\rightarrow$ Функция активации

Нейронные сети: как это выглядит





Обучение нейронных сетей

Как и в случае с генетикой, нейронные сети скорее язык описания решения, нежели местод обучения.

- Градиентные методы (когда производную легко взять $\rightarrow$ персептронные сети)
  - регрессии (оптимизация)
  - генеративные модели (графические модели)
- Минимизация функция энергии в заданых точках (сети Хопфилда)
- Смешанные (Bolzmann machine)
- EM
- etc.

Многоуровневые персептронные сети

Персептронная сеть – это искуственная нейронная сеть обладающая свойствами:

- ацикличность
- функция активации – сигмойд

$$\begin{array}{l} tanh(x), \\ \frac{1}{1+e^{-x}}, \\ \begin{cases} 1, & w_i x \geq \theta \\ 0 \; or -1, & otherwise \end{cases} \end{array}$$

- нейроны расположены слоями
  - нейроны в слое не связаны между собой
  - существуют связи только между нейронами, находящимися в соседних слоях
  - связи направлены от нижних слоев к верхним

МПС: мат модель

$$P = \left(\{W\}_1^L, K\right)$$
$$y = K\left(W_L K\left(W_{L-1} K\left(\ldots K\left(W_1(x)\right)\right)\right)\right)$$

По шагам:

- подаем сигнал x на входной уровень
- последовательно для каждого уровня применяем сумматор W_l и фунфцию активации K
- размерности W_l могут изменяться

МПС: обучение

Метод коррекции ошибок с обратной передачей сигнала
Зададим p_1, p_2, p_3 и ϵ

- Посчитаем модель с текущими весами
- Назначим ошибкой выходных точек отклонение от искомого значения $E_r = R - r$
- Для каждого элемента последнего внутреннего слоя
  - Поставим $E_i = 0$
  - Для каждой связи элемента с выходным слоем
    - Активен w_{ij} отличается от знака ошибки $\Rightarrow$ добавить -1 с вероятностью p_1;
    - Не активен w_{ij} совпадает со знаком ошибки $\Rightarrow$ добавить $+1$ с вероятностью p_2
    - Не активен w_{ij} отличается от знака ошибки или 0 $\Rightarrow$ добавить -1 с вероятностью p_3
  - Если $E_i \neq 0 \Rightarrow$ ко всем связям с элементом добавить $\nu = \epsilon a_i sign E_i$
- Повторить для всех слоев, считая только что обработанный выходным

МПС: обучение

Метод обратного распространения ошибки
Функция активации гладкая. Зададим ν. H – целевая функция.
Переход от уровня к уровню $y^l = s(W^l y^{l-1})$, $y^0 = x$.

- Посчитаем модель с текущими весами
- Честно посчитаем производную для каждого веса w^l_{ij}
  - Для выходного слоя так:

$$\frac{\partial H}{\partial y^L_i} \frac{\partial s}{\partial w^L_{ij}}$$

  - Для последнего скрытого слоя:

$$\left(\sum_t \frac{\partial H}{\partial y^L_t} \frac{\partial s}{\partial y^{L-1}_j} \right) \frac{\partial s}{\partial w^{L-1}_{ij}}$$

  - etc.
- Сходить против этой производной с шагом ν

МПС: обучение

Метод обратного распространения ошибки для $H = \frac{1}{2}\|y - b\|^2$ и $s = \frac{1}{1+e^{-x}}$

- Для выходного слоя так:

$$\frac{\partial H}{\partial y_i^L}\frac{\partial s}{\partial w_{ij}^L} = -y_i^{L-1}y_j^L(1 - y_j^L)(y_j^L - b)$$

- Для последнего скрытого слоя:

$$\left(\sum_t \frac{\partial H}{\partial y_t^L}\frac{\partial s}{\partial y_j^{L-1}}\right)\frac{\partial s}{\partial w_{ij}^{L-1}} = y_i^{L-2}y_j^{L-1}(1-y_j^{L-1})\sum_t w_{jt}^L y_t^L(1-y_t^L)(y_t^L-b)$$

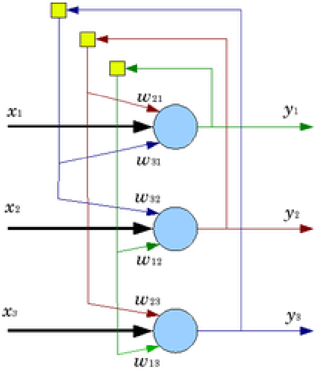
- etc.

Сеть Хопфилда (ассоциативная память)

Свойства:

- Рекурентная
- Входные узлы == выходные == "думающие"
- Полносвязная
- Работает до сходимости (не детерминированно)

Выглядит это так:



Сеть Хопфилда: Модель

$$s(x) = \begin{cases} 1, & w_i x \geq \theta_i \\ 0, & otherwise \end{cases}$$
$$x(t+1) = s\left(Wx(t)\right)$$

Хотим, чтобы при $t \to \infty$, при условии того, что начальная точка близка к одной из тренировочных x_i, сетка сходилась бы к этой x_i. Воспользуемся условием сходимости по Ляпунову.

$$\exists V:$$
$$V(x) \geq 0$$
$$\frac{dV}{dt} \leq 0$$
$$V(x) = 0 \Leftrightarrow x(t) = 0$$
$$\Rightarrow$$

Для нашей сети можно взять $V = E = \theta^T x - \frac{1}{2} x^T W x$

Сеть Хопфилда: Обучение

Чтобы E удовлетворяла условиям сходимости нужно:

$$w_{ij} = w_{ji}$$
$$w_{ii} = 0$$

А если мы хотим, чтобы сеть сходилась к набору точек x_i, то в окрестности этих точек E должна соответствовать условиям. В частноти $E(x_i) = 0$ для всех x_i из обучения.