

Глубокое обучение и вообще

Соловей Влад

9 марта 2021 г.

Посиделка 15: Воспоминания

Agenda

- Нейронные ячейки - наши кубики
- Хаки специфичные под нейронки
- Технические хаки

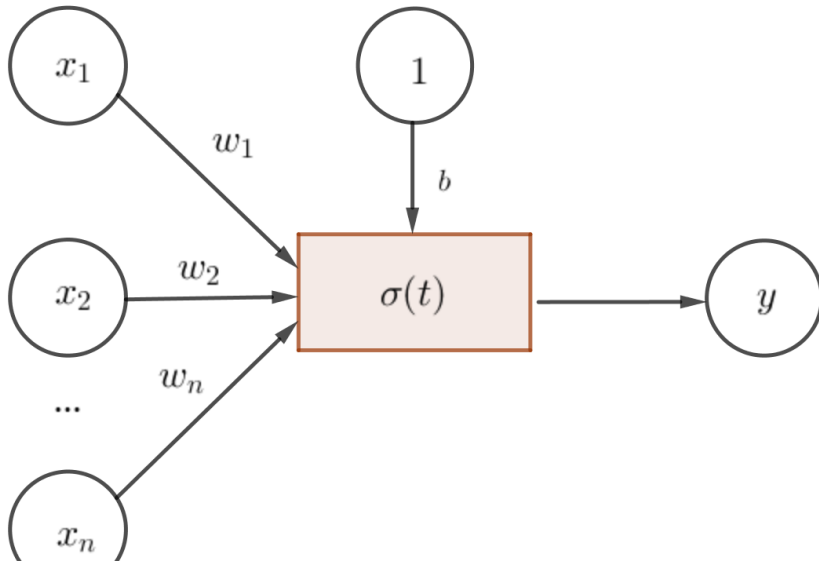
**А что такое нейросеть? Что мы можем назвать нейросетью в
текущем понимании индустрии?**

Давайте наберем нейронные ячейки? Накидываем!

Наша база - полносвязная ячейки. Почему она полносвязная? Чем она характеризуется? сколько нам надо слоев на практике?

Dense

Наша база - полносвязная ячейки.



Функции активации

А какие бывают функции активации? Как их выбирать? На что влияет их выбор?

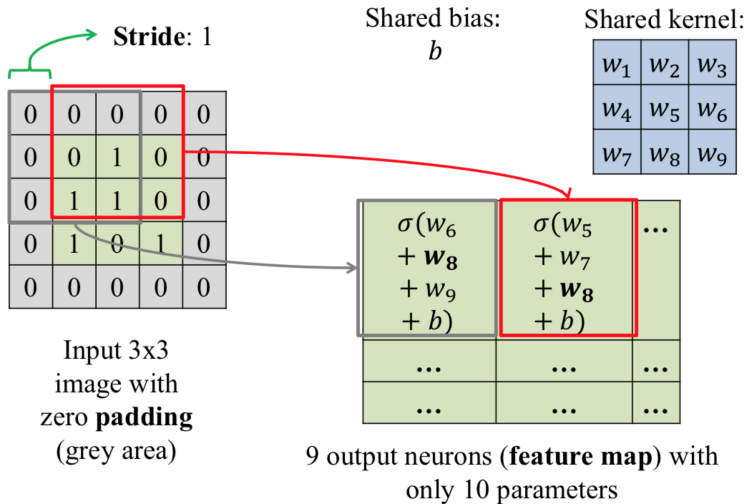
Функции активации

Наша база - полносвязная ячейки.

Название функции	Формула $f(x)$	Производная $f'(x)$
Логистический сигмоид σ	$\frac{1}{1+e^{-x}}$	$f(x)(1-f(x))$
Гиперболический тангенс \tanh	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	$1 - f^2(x)$
SoftSign	$\frac{x}{1+ x }$	$\frac{1}{(1+ x)^2}$
Ступенька (функция Хевисайда)	$\begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$	0
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	$\begin{cases} \alpha(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}$	$\begin{cases} f(x) + \alpha, & x < 0 \\ 1, & x \geq 0 \end{cases}$

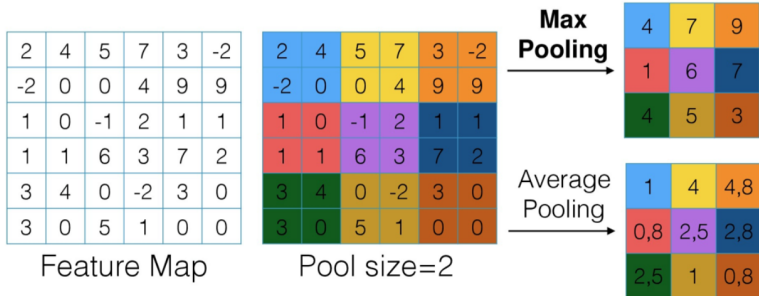
Сверточная часть

Давайте обсудим тут чем мы можем управлять:

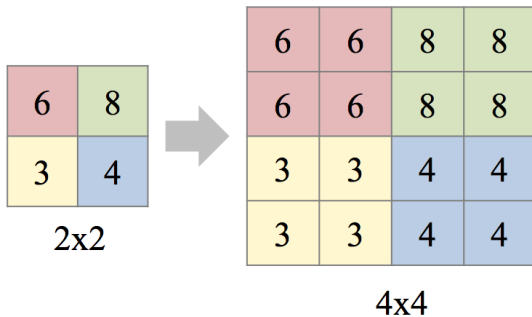


Сверточная часть

Когда и зачем мы можем использовать каждый из пулингов?

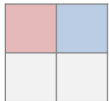


Анпулинг

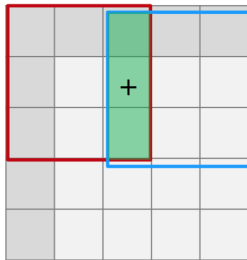


Анпулинг

Input: 2x2



Input gives
weight for
filter



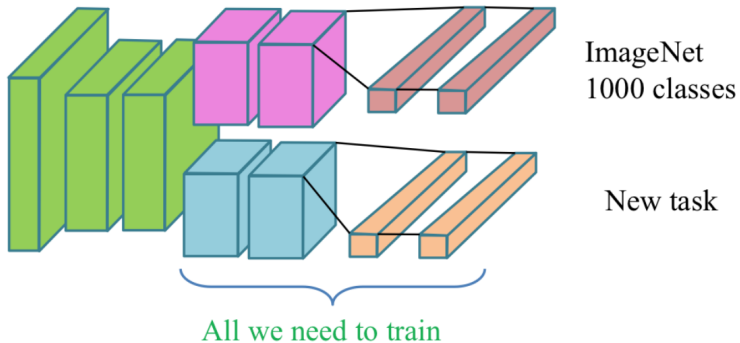
Stride: 2

Output: 4x4

Сверточная часть

А это что за зверь и почему?

ImageNet features extractor



Реккурентная часть

А в чем идея рекуррентных ячеек? Что нам дает рекуррентность?

Обучение и градиентный спуск

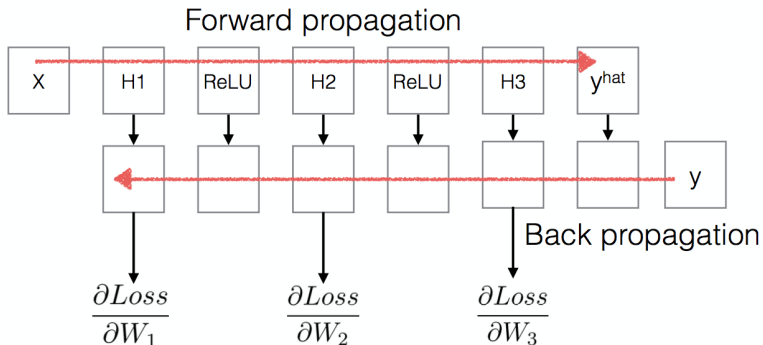
Градиентные спуски

Какой главный параметр градиентных спусков?

1. SGD
2. Momentum SGD
3. RMSprop
4. Adam

Обратное распространение ошибки

А это о чем ? Чем он лучше прямого?



Эвристики

Эвристики

Ага... А что это такое и какие они есть ?

Архитектуры

Архитектуры

Вспоминаем архитектуры, которые мы смотрели? Чем архитектура отличается от ячейки?

1. Inception
2. Resnet
3. Autoencoders
4. Unet
5. GAN
6. W2V (Embeddings)

Что-то еще?

Техническая часть

Умеем

1. Задавать модели по разному
2. Использовать разные callback (а какие?)
3. Брать куски моделей и их применять
4. Использовать кастомные кусочки
5. Дебажить код (ну мы точно пытались)