

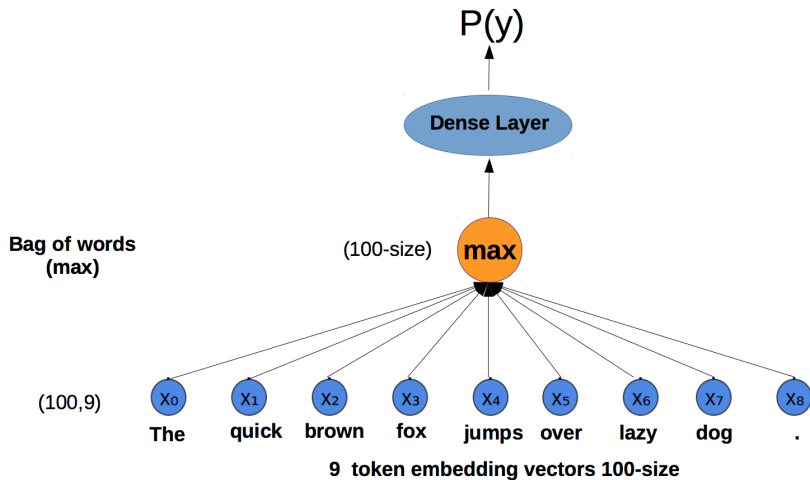
Глубокое обучение и вообще

Никита Бекезин

24 ноября 2021 г.

Лекция 13: Сверточные сети для последовательностей

Нейросеть для текстов



Проблемы

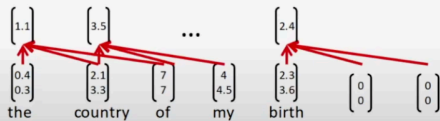
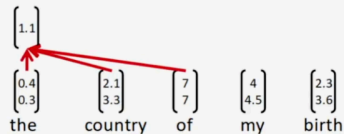
- Теряем информацию о порядке слов
- Решить эту проблему можно обучив эмбединги для биграмм, трграм и тд, но это расширит пространство признаков
- Можно решить эту проблему с помощью рекурентных нейросеток
- Можно решить эту проблему с помощью свёрточного слоя

One layer CNN

- Simple convolution + pooling
- Window size may be different (2 or more)
- The feature map based on bigrams:

$$\mathbf{c} = [c_1, c_2, \dots, c_{n-h+1}] \in \mathbb{R}^{n-h+1}$$

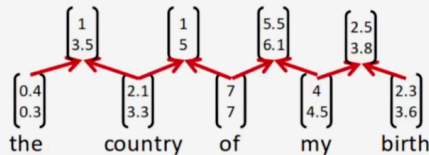
$$c_i = f(\mathbf{w}^T \mathbf{x}_{i:i+h-1} + b)$$



Источник: Machine Learning course at MIPT

One layer CNN

- Imagine using only bigrams
- Same operation as in RNN, but for every pair



$$p = \tanh \left(W \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + b \right)$$

Источник: Machine Learning course at MIPT

Max overtime pooling

- Feature representation is based on some applied filter:

$$\mathbf{c} = [c_1, c_2, \dots, c_{n-h+1}] \in \mathbb{R}^{n-h+1}$$

- Let's use pooling:

$$\hat{c} = \max\{\mathbf{c}\}$$

- Now the length of \mathbf{c} is irrelevant!
- So we can use filters based on unigrams, bigrams, tri-grams

Источник: Machine Learning course at MIPT

Свёрточная сетка

- Feature representation is based on some applied filter:

$$\mathbf{c} = [c_1, c_2, \dots, c_{n-h+1}] \in \mathbb{R}^{n-h+1}$$

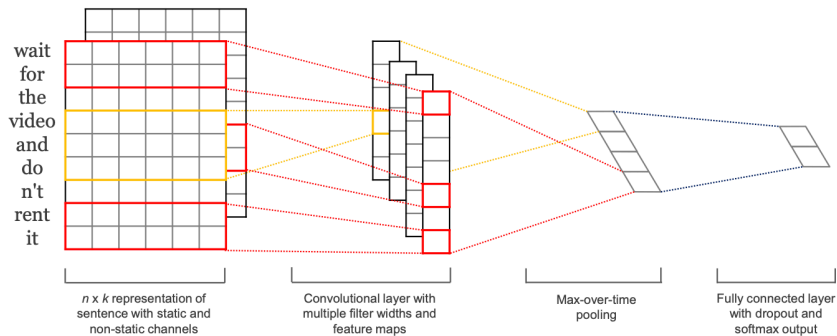
- Let's use pooling:

$$\hat{c} = \max\{\mathbf{c}\}$$

- Now the length of \mathbf{c} is irrelevant!
- So we can use filters based on unigrams, bigrams, tri-grams

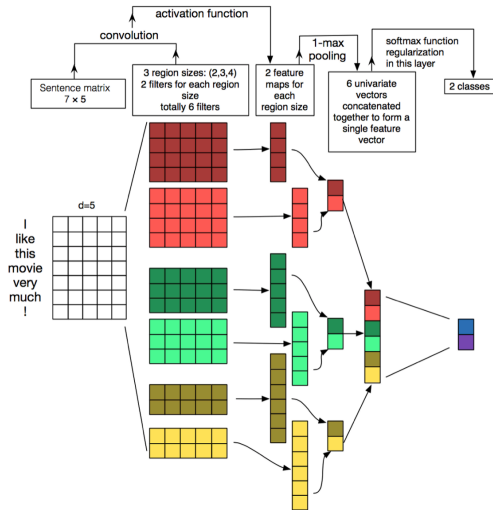
Источник: Machine Learning course at MIPT

Оригинальная статья Yoon Kim

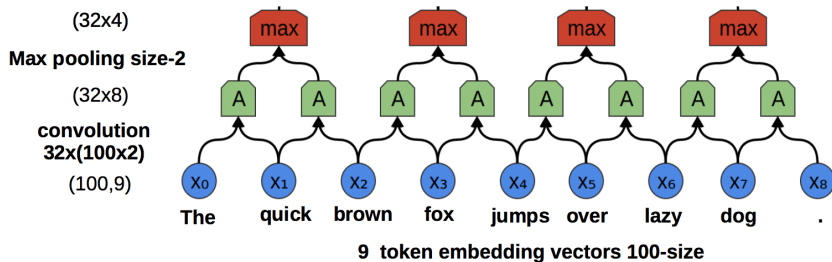


Статья

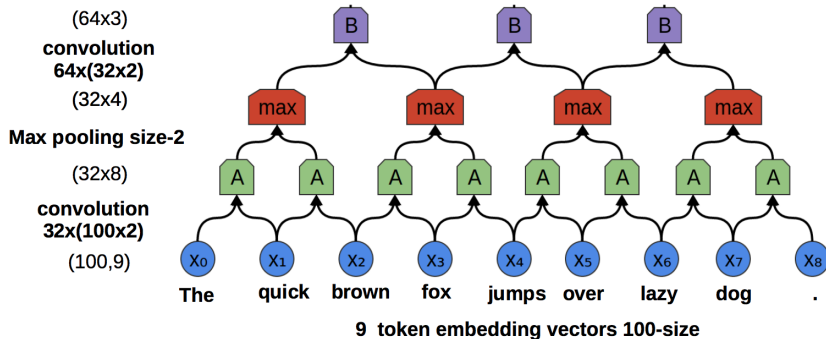
Еще одно пояснение



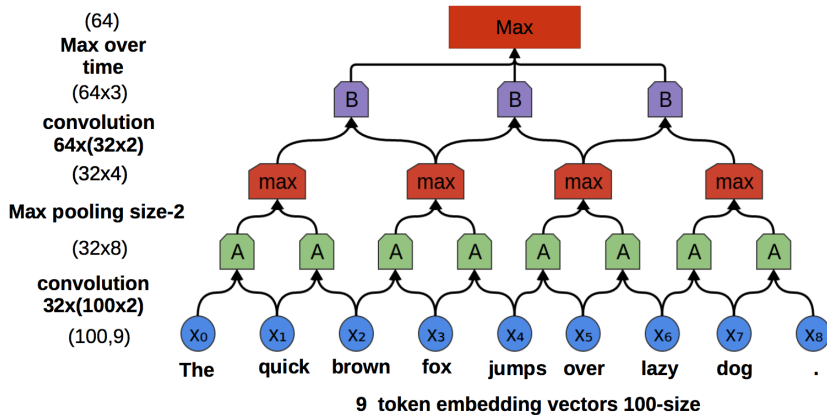
Свёрточная сетка



Свёрточная сетка



Свёрточная сетка



Свёрточная сетка

