

Lecture 6

2018-May-13

Outline

- 1. Boolean Information Retrieval;
- 2. Tensorflow & Keras quick introduction;
- 3. Word Embedding;

用Tensorflow实现神经网络

用Tensorflow & Keras 实现神经网络



- 1. 神经网络的基本构成；
- 2. Tensorflow框架；

神经网络的构成

- 1. 训练数据
- 2. 参数
- 3. Operator 计算符
- 4. Loss 函数
- 5. 优化方法 (SGD, Adam等)

神经网络框架

- 1. 常用的方法框架已经定义好；
- 2. 可以自定义网络结构；
- 3. Tensorflow, Caffe, PyTorch
 - 自动计算导数\梯度
 - 自动根据梯度进行参数的迭代优化

TensorFlow—Tensor

tensor

't'
'e'
'n'
's'
'o'
'r'

tensor of dimensions [6]
(vector of dimension 6)

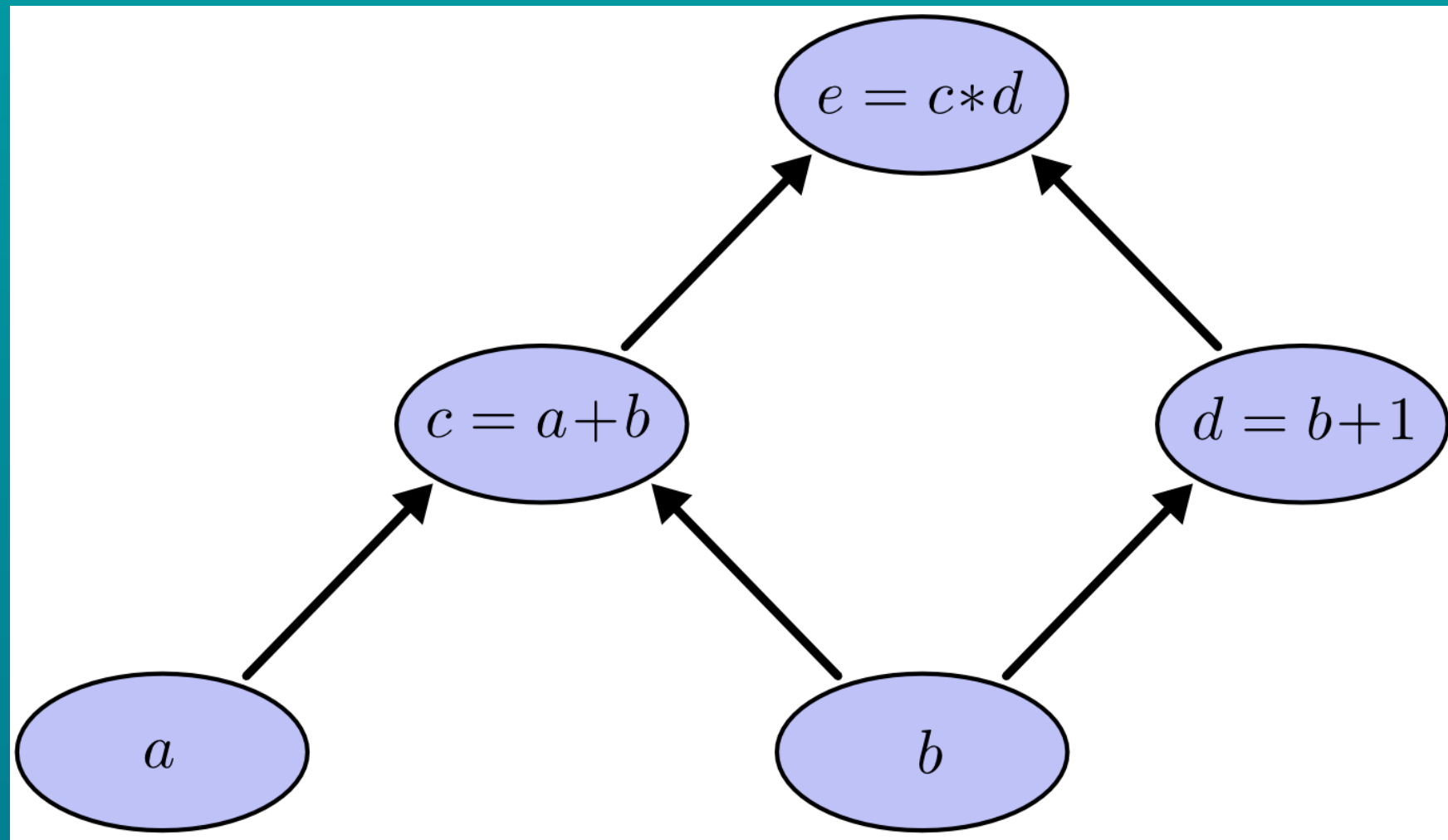
3	1	4	1
5	9	2	6
5	3	5	8
9	7	9	3
2	3	8	4
6	2	6	4

tensor of dimensions [6,4]
(matrix 6 by 4)

2	7	8	8	1	8
2	8	4	5	9	0
2	3	5	3	6	0
7	4	7	1	3	5

tensor of dimensions [4,4,2]

Tensorflow-GraphFlow



Session 和 Graph

- 所以Tensorflow中tensor的值， 需要在session中进行计算！
 - 实例： session, ipython-notebook
- Graph 定义了一个计算图， 一整个模型可以有好几个sub-graph构成， 不同的graph用来计算不同的模型图

TensorFlow的元素

- 1. 数据： 两种一种是利用feed_dict进行， 另一种是利用Data.pipeline 进行.
- 2. Variable： 变量， 能够在训练中改变；
- 3. placeholder： 位置的量， 例如训练数据， 需要每次导入；
- 4. Operator: 各种数学操作
- 5. Optimizer： 进行自动优化的函数

Word Embedding

outline

- 1. What is embedding;
- 2. Why we need word embedding;
- 3. How do we implement word embedding;
 - One-hot; PCA, SVD
- 4. What is the word embedding;