

Skip-Thought Vectors

Reference: 《Skip-Thought Vectors》
2015,NIPS

论文目的

- 构建一个非监督的句子表示模型，将句子编码成向量
- 不针对某个特定任务而创建，具有通用性
- 可以应用在语义相似性、段落检测、文档分类等问题

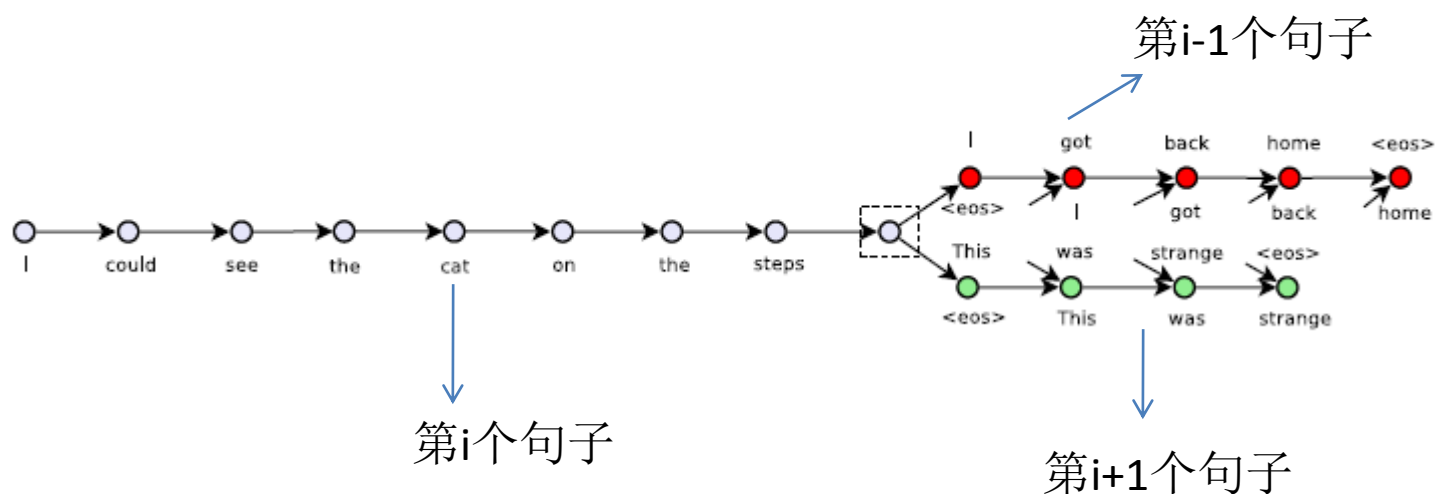
论文思想

- 提出一种通用的无监督句子表示模型，借鉴word2vec中skip-gram模型，通过一句话预测这句话的上一句和下一句。
- 采用端到端的模型seq2seq，通过训练数据集BookCorpus，将模型中的encoder部分作为feature extractor，可以为任何句子生成vector。

模型构建

- 三元组:

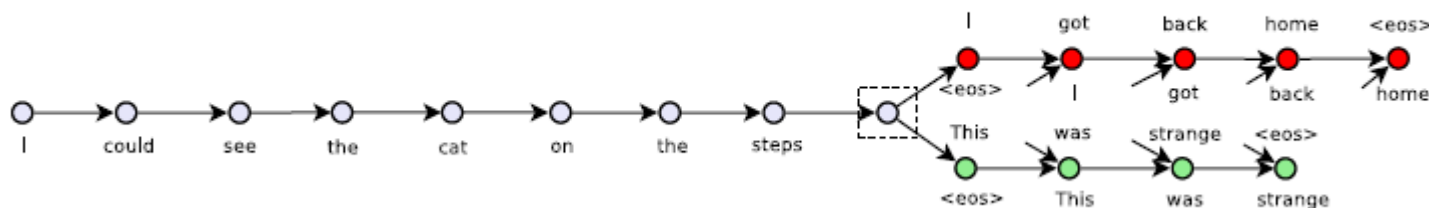
(I got back home. I could see the cat on the steps.
This was strange.)



模型构建

- Encoder: GRU-RNN
- Decoder: encoder部分的最后一个词的hidden state 作为decoder的输入来生成
- Objective: 模型中的目标函数分为两个部分，一个来自于预测下一句，一个来自于预测上一句

$$\sum_t \log P(w_{i+1}^t | w_{i+1}^{<t}, \mathbf{h}_i) + \sum_t \log P(w_{i-1}^t | w_{i-1}^{<t}, \mathbf{h}_i)$$



论文总结

基于句子的上下文，提出一种通用的句子表示模型架构。

- 实验结果：通用模型。不是最好的，但也不是最差的。
- 监督模型：针对特定任务，依据数据标签修正模型。
- 非监督模型：通用性，依据分布假说调节模型参数。