

Lecture 1. Intro to NLP and Deep Learning

Lecture 1主要讲解了下课程的大纲，NLP、Deep Learning的一些基础知识和背景，可能稍微觉得枯燥，了解这部分的背景可以跳过。

教学大纲

- Lecture 1. Intro to NLP and Deep Learning
- Lecture 2. Simple Word Vector representations: word2vec, Glove
- Lecture 3. Advanced word vector representations language models, softmax, single layer networks
- Lecture 4. Neural Networks and backpropagation — for named entity recognition
- Lecture 5. Project Advice, Neural Networks and Back-Prop
- Lecture 6. Practical tips: gradient checks, overfitting, regularization, activation functions, details
- Lecture 7. Introduction to Tensorflow
- Lecture 8. Recurrent neural networks — for language modeling and other tasks
- Lecture 9. GRUs and LSTMs — for machine translation
- Lecture 10. Recursive neural networks —for parsing
- Lecture 11. Recursive neural networks — for different tasks
- Lecture 12. Review Session for Midterm
- Lecture 13. Convolutional neural networks — for sentence classification
- Lecture 14. Guest Lecture with Andrew Maas: Speech recognition
- Lecture 15. Guest Lecture with Thang Luong: Machine Translation
- Lecture 16. Guest Lecture with Quoc Le: Seq2Seq and Large Scale DL
- Lecture 17. The future of Deep Learning for NLP: Dynamic Memory Networks
- three problem sets
- Midterm Exam
- Final Course Project

知识背景

在学习CS224D之前，需要有一些基础的知识背景，例如了解SVD等。

- python，所有的课堂作业代码都需要用python实现
- 微积分、线性代数
- 基础的概率统计知识
- CS229《Machine Learning》同等水平的知识点，例如损失函数、推导、梯度下降的性能优化

NLP是什么？

NLP是一个结合了计算机科学、人工智能和语言学的领域。

NLP的研究目标是让计算机来处理或者了解自然语言，从而能够解决类似智能问答的任务。

然而，让计算机完全理解或者表达语言的意义，是不可能的。

NLP Levels

NLP主要等级有语音方面的Phonological Analysis, Morphological Analysis形态学分析, Syntactic Analysis句法分析, Semantic Interpretation语义解释, 文本方面的OCR/Tokenization, Discourse Processing。NLP更关注的是前面四个方向, 它们的传统方法和Deep NLP表示分别为:

- Phonology:

DL从声音的特征中训练语音 (词), 并把它们表征为向量。

- Morphology:

传统的形态学分析将词分为“prefix”, “stem”, “suffix”三个部分来训练; DL将每个形态作为一个向量, 神经网络将两个向量结合为一个向量。

paper见Thang et al.2013

- Syntax:

传统句法分析将短语训练为NP, VP等离散类别; DL为每个词和每个短语训练向量, 同样, 神经网络将两个向量结合为一个向量。

paper见Socher et al.2011

- Semantics:

传统语义分析是使用lambda计算, lambda的公式需要很仔细的计算, 容易出错, 并且无法计算相似短语的相关性; DL将每个词、短语或者逻辑表达训练为一个向量, 神经网络同样能对向量进行联合。

paper见Bowman et al.2014

NLP应用

从简单到复杂, NLP有以下几种应用实例:

- 拼写检查、关键词提取、查找同义词
- 信息提取, 例如从网页上提取商品价格、日期、地点、人名、公司名等
- 分类, 正负情感分析
- 机器翻译
- 对话系统
- 复杂问答系统

NLP在工业界的应用有:

- 搜索 (读写)
- 在线广告
- 自动/辅助翻译
- 市场或者经济贸易的情感分析
- 语音识别

为什么NLP困难?

- 在表达、学习和使用语言学、场景、真实世界、视觉知识时有复杂性
- 例如“Jane hit June and then she [fell/ran]”这句话中, “she”指的是谁?
- 语义的模糊性

Deep Learning是什么？

Deep learning是机器学习的子领域。大多数的机器学习方法效果好是因为算法的输入是人们设计好的特征，通过优化权重来达到最好的越策效果。

DL算法则尝试自动学习输入数据的特征或者表达（多层），输入通常是原始数据，例如文本中的词。

在本课程中，我们将会关注的是不同的神经网络，以及深度学习中具有优势模型的算法家族。我们更多关注的是深度学习应用于NLP的算法，而不是深度学习模型的历史，如果了解历史，可以参考：
Deep Learning in Neural Networks: An Overview

探索DL的原因

手动选择特征存在着不完整、过度指定和花费时间来设计和验证的问题，而自动学习特征适应简单且学习快速；

DL提供一种几乎稳定、通用、可学习的框架来表达世界、视觉和语言信息；

DL既可以处理无监督学习，也可以处理监督学习；

另外，DL现在火速发展的三个主要的原因：大量的数据、快速的机器和多核CPU/GPU设备、新的模型和算法；

上述的所有，使得DL在语音、视觉领域得到了很好的效果。

- Speech

论文：Context-Dependent Pre-trained Deep Neural Networks for Large Vocabulary Speech Recognition Dahl et al.(2010)

- Computer Vision

论文：ImageNet Classification with Deep Convolutional Neural Networks by Krizhevsky et al.2012

NLP的典型应用

- 情感分析

传统方法：通常使用情感词典结合BOW词袋或者结合指定否定特征来实现。

DL演示demo：<http://nlp.stanford.edu/sentiment>

- 问答系统

传统方法：使用大量的特征工程来捕捉世界和知识之间的关系，例如利用正则表达式计算。

- 机器翻译

DL：源句子映射为向量，然后生成目标句子。

论文：Sequence to Sequence Learning with Neural Networks by Sutskever et al.2014

下节课程，将会讲解所有NLP等级都会用到的向量表示，我们将学习如何生成词的向量、词向量表示的是什么，以及如何将词向量用于NLP任务中。

