基于深度学习的自然语言处理-大纲与阅读资料

> 下面部分DL表示《Deep Learning》这本书，NNM4NLP表示《Neural Network Method for Natural Language Processing》这本书。

1. 前馈神经网、计算图与自动求导

- 茶会内容
  - 前馈神经网的定义与组成部分：损失函数、激活函数、结构
  - 基于梯度下降的迭代优化
  - 实验：
    - 计算图与PyTorch基础：张量运算
    - 基于PyTorch的前馈神经网用于回归与多分类
    - TensorFlow基础
- 推荐阅读
  - 前馈神经网：NNM4NLP 第3、4章；
  - 基于梯度的优化：DL第6.2节；
  - 计算图：NNM4NLP第5.1节；DL第6.5节；
  - Neubig 01 intro 幻灯片, P13-45.
- 阅读材料
  - NNM4MLP 1-5章（NN相关的基础，写作风格十分实用）

2. 词级别表示：基于前馈神经网的语言模型、Word2vec与词向量

- 茶会内容
  - 从One-hot表示到词的稠密向量表示：NNLM的副产品
  - Word2vec的两种模型：CBOW与Skip-gram模型，词向量的可视化
  - Word2vec的目标函数与针对Softmax的训练加速
  - 基于有监督目标得到的词向量（词的表示学习）
  - 实验：PyTorch/TensorFlow实现：Word2vec模型
- 推荐阅读
  - NNM4NLP 第8章，第9.1-9.4节；

    - A Neural Probabilistic Language Model, Yoshua Bengio的03年论文，开创神经网对语言数据的处理，阅读第1、2节，即第3节“Parallel Implementation”之前部分.

- 阅读材料
    - Word2vec Parameter Learning Explained. 一篇解释Word2vec参数学习的文章：涉及层次Softmax与负采样目标函数的讲解与具体实现.
    - "On Word Embeddings" Part 1, Part 2, Part 3, 是三篇Sebastian Ruder所撰写的综述性博文，内容十分丰富，涵盖了许多文献的解读；他还有一篇新的博文"Word embeddings in 2017: Trends and future directions", 可以仔细研读一下.

3. 词序列表示：卷积神经网与循环神经网

- 茶会内容
    - 卷积操作的意义：视觉识别中的卷积核
    - 池化操作的意义：不变性
    - 通过卷积、非线性变换、池化等操作组合起来的卷积神经网
    - 循环神经网的动机：朴素循环神经网，训练时的梯度问题
    - 循环神经网的改进：GRU、LSTM、双向循环神经网
    - 句子级别表示：用CNN或RNN将变长词序列嵌入为定长向量，并用于分类
    - 实验：PyTorch/TensorFlow实现：文本蕴含任务
- 推荐阅读
    - NNM4NLP 第13、14、15章
- 阅读材料
    - CNNs for Sentence Classification. 2014年EMNLP会议论文. CNN在文本建模领域的一次引人注目的尝试.
    - NNM4NLP 第16章.
    - 如何使用与评估句子表示部分Neubig的幻灯片，该幻灯片从：句子分类（问题分类、情感分类等）、复述识别、语义相似度计算、文本蕴含、文本检索等任务展示了用神经网络对句子进行嵌入（向量表示）的优势.

4. 结构的生成与预测：序列到序列模型与注意力机制、递归神经网

- 茶会内容
    - Seq2Seq模型（或Encoder-Decoder模型）：两个RNN的故事
    - 神经机器翻译与注意力机制
    - 用于建模句法结构的递归神经网
    - “万种任务皆Seq2Seq模型”
    - 实验：PyTorch/TensorFlow基于Seq2Seq模型的回复生成
- 推荐阅读
    - NNM4NLP 第17、18章
- 阅读材料
    - ICML 2017 Seq2Seq, Google研究院Oriol Vinyals在2017机器学习领域顶会ICML上的教程PDF，内容十分丰富.
    - 递归神经网的动机与句子表示应用, 斯坦福大学DL4NLP课程幻灯片. 详细介绍了递归神经网是如何用于带有句法树句子的向量表示的.

- Semi-Supervised Recursive Autoencoders for Predicting Sentiment Distributions, 11年EMNLP会议论文，Socher使用递归神经网对带有情感倾向的句子进行建模，并能预测出每一次句法树的非终结节点所覆盖的短语的情感倾向性. 十分漂亮的工作.
- Improved Semantic Representation from Tree-Structured LSTM Networks, 15年ACL会议论文，使用LSTM替换之前Socher提出的朴素RNN计算单元，使得性能进一步提升.

5*. 神经网络杂谈：各类注意力机制、门限机制与记忆网

- 茶会内容
  - 当下各类注意力机制汇总
  - 作为信息流控制的门（Gating mechanism）
  - 一种有别于符号存储的向量存储模式：作为句子/篇章表示的记忆矩阵
- 推荐阅读
  - Neubig关于注意力机制的幻灯片.
  - Neubig课程主页关于注意力部分的参考资料（所有标有Reference的链接）.
- 阅读材料
  - Language Modeling with Gated Convolutional Networks, 用CNN与门限机制做语言模型.
  - Memory Networks 与 End-to-End Memory Networks. 两篇论文为基于连续向量存储机制的内容检索提供了基础（PyTorch实现，TensorFlow实现.

6. （神经）句法分析技术选讲

- 茶会内容
  - 什么是句法分析：组分语法、依存语法与其对应的句法分析任务定义
  - 基于转移的依存句法分析
  - 基于图的依存句法分析
  - 基于Seq2Seq模型的组分句法分析
- 推荐阅读
  - [TO-DO]
- 阅读材料
  - [TO-DO]

附录.

- 上面的安排均参考Neubig的自然语言处理课程删减而定，该课程首页上，在每一个主题下均有许多推荐阅读与论文的链接，请大家有余力定去看看。
- 上面大纲中的6次主题可能会分为6茶会进行，内容比较多，请非配到的同学务必认真准备。
- 上面大纲中的“茶会内容”部分，是分配到的同学需要准备的部分，能够大致按照顺序去讲即可，具体同我商量；“推荐阅读”部分是强烈建议每一位参加茶会的同学在茶会开始前一周阅读的内容；“阅读材料”部分是算是一些额外的拓展，也推荐有余力的同学去阅读。