

深度学习讲习班报告

2017-8-23



深度学习与神经网络

- 神经网络基础知识
- 基础神经网络模型：前馈、卷积、循环
- 改进神经网络模型：注意力模型、GAN等



词法句法语义分析

- 结构化预测任务（序列分割、序列标注、句法分析）
- 基于图的方法、基于转移的方法
- 结合神经网络的图方法、转移方法



知识表示与知识获取

- 基于深度学习的词表示方法
- 知识图谱表示方法
- 基于深度学习的知识抽取：实体关系抽取、开放域关系抽取、事件抽取



深度学习与神经网络

1.深度学习

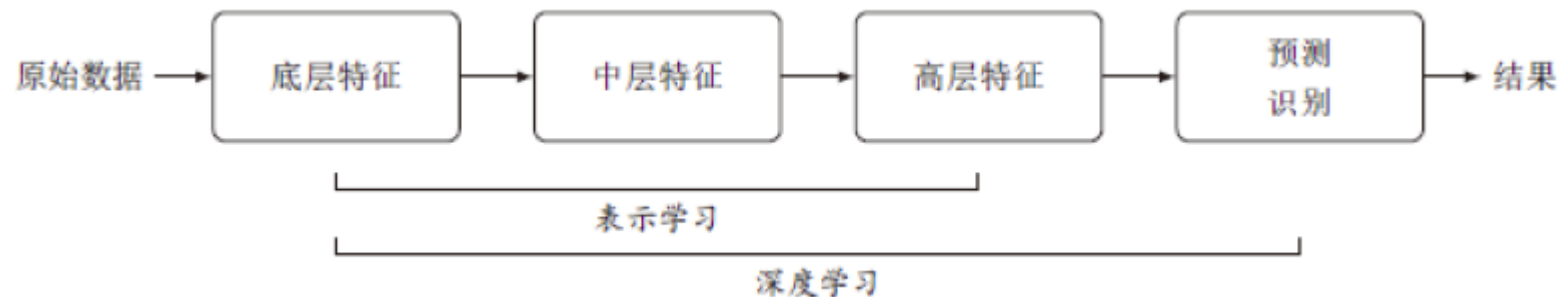
2. 基础神经网络模型

3.网络模型优化

4.进阶模型&未来研究

深度学习

- **深度学习 = 表示学习（特征表示）+ 浅层学习（线性分类器）**
- **难点：贡献度分配问题（每一步对结果的贡献）**
- **深度学习&神经网络：深度学习天然不是神经网络，而神经网络天然
是深度学习。**（神经网络可以解决贡献度分配问题，图模型也可以做深度学习的问题）





基础神经模型：

➤前馈神经网络：

- 信号正向传播，误差反向传播
- 全连接网络，权重矩阵参数很多

➤卷积神经网络：

- 局部连接、权重共享、空间或时间的次采样
- 节点之间无连接，输入输出维度固定，输出仅依赖当前输入

➤循环神经网络：

- 使用自反馈的神经元，能处理任意长度序列
- 因梯度爆炸/消失，只能学习短周期的依赖关系



深度神经网络：

➤难点：

- 参数多，训练时间长
- 非凸优化：存在局部最优而非全局最优，影响迭代
- 下层参数难调
- 参数解释性差

➤要求：

- 计算资源大
- 数据多
- 算法效率要好，能够快速收敛



模型优化（经验）

- 用ReLU作为激活函数（梯度=1，不会梯度爆炸/消失）
- 分类时使用交叉熵作为损失函数
- 每次迭代都重新随机排序
- 数据预处理（归一化，加速模型收敛）
- 动态学习率（越来越小）
- Dropout（训练时随机舍弃一些结点，防止过拟合）

进阶模型：

➤注意力机制与外部记忆：

解决简单RNN存在长期依赖关系

注意力机制：指人在工作时注意力往往会集中在整个场景中的某些焦点上。
鸡尾酒会效应

➤深度生成模型

对抗生成式网络（GAN，2014）

变分自动编码器（VAE，2015）

生成模型：用于随机生成可观测数据的模型。

生成数据 x ：

□ 根据隐变量的先验分布 $p(z;\theta)$ 进行采样，得到样本 z 。

□ 根据 z 生成数据 x 。
强化学习：智能体与环境交互，根据反馈调整策略以最大化其得到长远的所有奖励的积累值。

➤深度强化学习

结合深度学习，用强化学习来定义问题和优化目标，用深度学习来解决状态表示、策略表示等问题。（AlphaGo）

□



未来研究方向：

- 任意结构网络
- 非监督学习（如何评价？）
- 长期以来问题
- 优化
- 自然语言理解和推理

- 分布式训练
- 生物学

" Thanks "