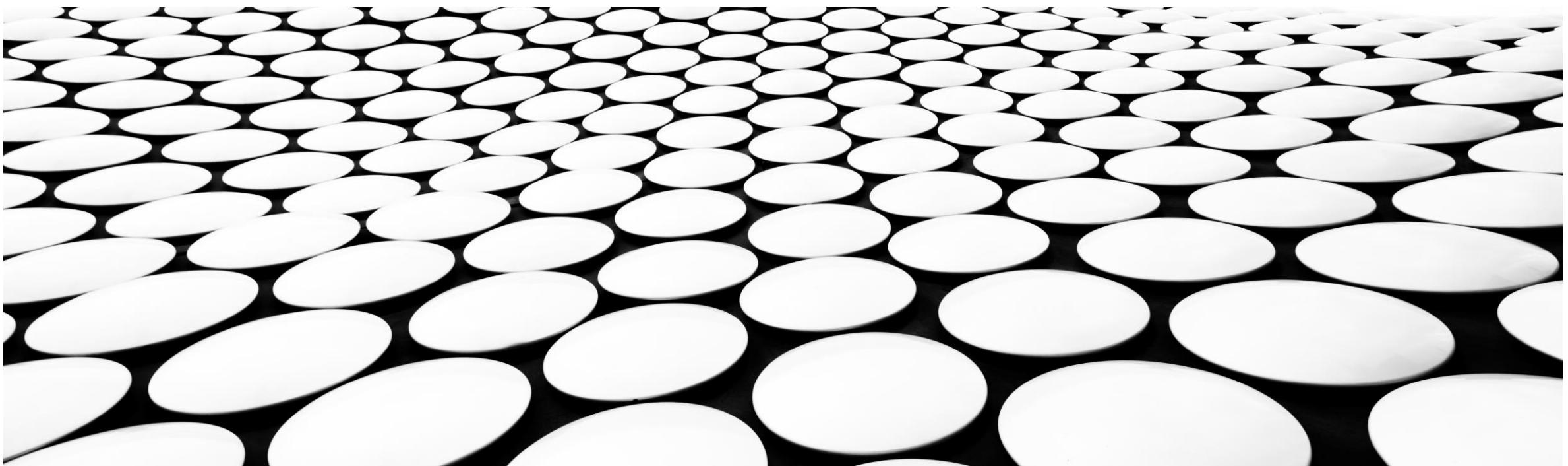

深度学习

邱怡轩



今天的主题

- 卷积神经网络实践
- 一般化建模方法



实践

练习

- 参见 [lec7-convnet.ipynb](#)



一般化建模方法

神经网络建模

- 从建模的角度看，CNN 与前馈神经网络并没有本质上的不同
- 神经元 + 参数 => 新的神经元

模块化

- 可以将各种类型的层抽象成具有统一接口的模块
- 输入 => 输出
- 可学习的参数
- 导数计算机制

模块化

- 全连接层
 - 线性变换
 - 参数为权重矩阵和偏置向量
- 激活函数层
 - 固定的非线性变换
 - 无参数 (某些会带少量参数)
- 卷积层
 - 卷积运算 (**也是一种线性变换**)
 - 参数为卷积核和偏置向量

一般流程

- 搭建模型
 - 模型 = 结构 + 参数
- 计算损失函数 (标量, 通常按输入取平均)
- 对参数求导
- 最优化、参数更新
- 重复、迭代

问题

- 模型怎么选?
- 损失函数怎么选?
- 导数怎么求?
- 最优化方法怎么选?

模型

- 根据数据和问题特点选择层的类型
- 例如卷积层的引入正是为了处理图片数据
- 神经元的数目、层的深度等超参数通常需要微调和试验，也可借鉴已有文献成果
- 一些最新的研究在关注自动化的神经网络架构搜索 (neural architecture search, NAS)

损失函数

- 本质上是一个统计问题
- 大部分情况下根据极大似然准则导出
- 也有许多其他的损失函数，如 SVM

导数

- 手动计算反向传播
- 现代软件框架的自动微分

最优化

- 基础的随机梯度下降
- 更多改良的优化算法