

预训练模型学习情况周报 1

姚凯

一、本周学习：

1. 《统计学习方法》

第一章：统计学习及监督学习概论

统计学习分为监督学习，无监督学习和强化学习

统计学习的问题及应用

第二章：感知机

感知机是二分类的线性分类模型，属于监督学习算法。输入为实例的特征向量，输出为实例的类别。旨在求出将训练数据进行线性划分的分离超平面

损失函数：所有误分类的点到超平面的总距离，用梯度下降法对损失函数进行极小化

附录 A: 梯度下降法

$$x^{(k+1)} = x^{(k)} - \alpha \nabla f(x^{(k)})$$

常见的有批量梯度下降法，随机梯度下降法，小批量梯度下降法

附录 B: 牛顿法和拟牛顿法

牛顿法：

$$x_{k+1} = x_k - H_k^{-1} g_k$$

拟牛顿法：

用正定对称阵近似黑塞矩阵的逆

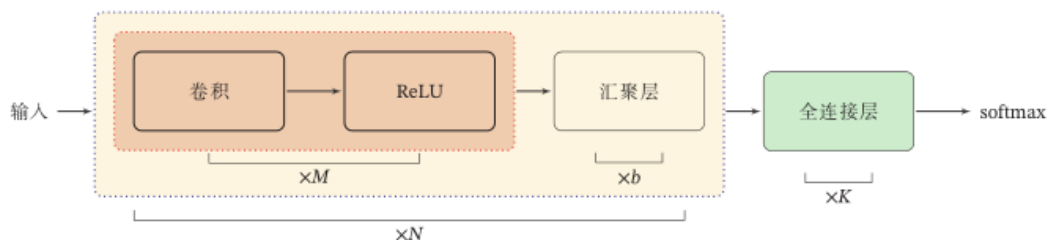
2. 《神经网络与深度学习》

第四章：前馈神经网络

$$\begin{aligned}z &= \sum_{d=1}^D w_d x_d + b \\&= \mathbf{w}^\top \mathbf{x} + b, \\a &= f(z),\end{aligned}$$

常见激活函数 f : Sigmoid, Relu, Switch, GELU

第五章：卷积神经网络



卷积层的作用是提取一个局部区域的特征

汇聚层作用是进行特征选择，降低特征数量，从而减少参数数量

卷积层虽然可以显著减少网络中连接的数量，但特征映射组中的神经元个数并没有显著减少。如果后面接一个分类器，分类器的输入维数依然很高，很容易出现过拟合。为了解决这个问题，可以在卷积层之后加上一个汇聚层，从而降低特征维数，避免过拟合。

二、下周学习：

《神经网络与深度学习》第六、七、八、十五章

目前进度：红色对勾为当前完成的任务，圆圈代表正在进行中的任务

