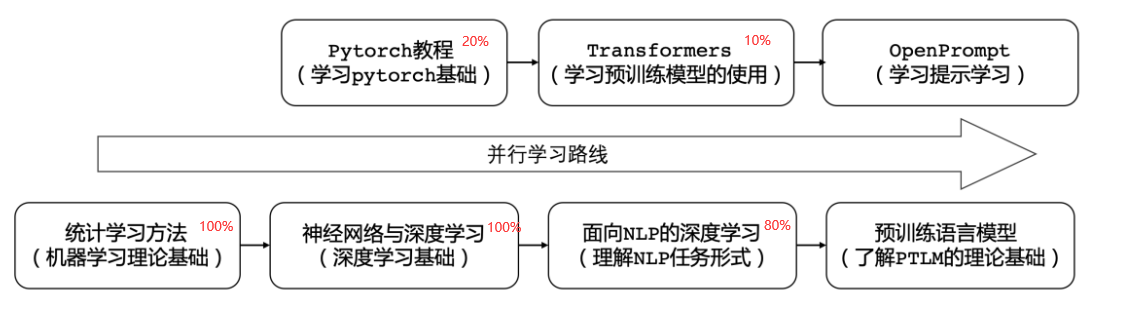
# 学习进度



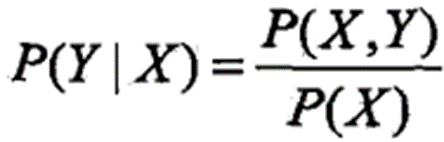
# 一、统计学习方法

研一上学期，我对李航老师的《统计学习方法》进行了完整的学习，并且每周和研一的同学们交流讨论。在这次预训练模型的学习任务中，我回顾了前两章和附录A、B的内容。

针对**第1章-统计学习及监督学习概论**，我主要对统计学习方法的分类，生成模型和判别模型这两个问题进行一下总结。

1. 统计学习方法可以大致分为：监督学习、无监督学习、半监督学习、强化学习这四类。

* 监督学习目的是学习一个由输入到输出的映射，称为模型。模式的集合就是假设空间。假设输入与输出的随机变量X和Y遵循联合概率分布P(X,Y)，对于学习系统来说，联合概率分布是未知的，训练数据和测试数据被看作是依联合概率分布P(X,Y)独立同分布产生的。
* 半监督学习的训练集是少量标注数据，大量未标注数据，利用未标注数据的信息，辅助标注数据，进行监督学习。
* 无监督学习的训练集是未标注数据，模型的目标是学习到一个条件概率分布函数。
* 强化学习问题可以描述为一个智能体从与环境的交互中不断学习以完成特定目标（比如取得最大奖励值)。强化学习就是智能体不断与环境进行交互，并根据经验调整其策略来最大化其长远的所有奖励的累积值。

1. 生成方法对应生成模型：如：朴素贝叶斯法和隐马尔科夫模型。判别方法由数据直接学习决策函数f(X)或条件概率分布P(Y|X)作为预测的模型，即判别模型。如：K近邻法、感知机、决策树、logistic回归模型、最大熵模型、支持向量机、提升方法和条件随机场。

* 生成方法可还原出联合概率分布P(X,Y)，而判别方法不能。生成方法的收敛速度更快，当样本容量增加的时候，学到的模型可以更快地收敛于真实模型；当存在隐变量时，仍可以使用生成方法，而判别方法则不能用。
* 判别方法直接学习到条件概率或决策函数，直接进行预测，往往学习的准确率更高；由于直接学习Y=f(X)或P(Y|X)，可对数据进行各种程度上的抽象、定义特征并使用特征，因此可以简化学习过程。

针对**第2章-感知机**，我对感知机模型的定义和方法概括如下：感知机模型是神经网络与支持向量机的基础。感知机对应于输入空间中将实例划分为正负两类的分离超平面，属于判别模型。输入为实例的特征向量，输出为实例的类别，取+1和-1。通常利用梯度下降法对损失函数进行极小化，感知机学习算法具有简单而易于实现的优点，分为原始形式和对偶形式。

针对**附录A和B**，梯度下降法，牛顿法和拟牛顿法这三个算法都是求解无约束最优化问题的常用方法。

* 梯度下降法是一种迭代算法。选取适当的初值x(0)，不断迭代，更新x的值，进行目标函数的极小化，直到收敛。由于负梯度方向是使函数值下降最快的方向，在迭代的每一步，以负梯度方向更新α 的值，从而达到减少函数值的目的。当目标函数是凸函数时，梯度下降法的解是全局最优解。一般情况下，其解不保证是全局最优解。梯度下降法的收敛速度也未必是很快的。
* 牛顿法是迭代算法，每一步需要求解目标函数的黑塞矩阵的逆矩阵，计算比较复杂。牛顿法比梯度下降法的收敛速度快。
* 拟牛顿法通过正定矩阵近似黑塞矩阵的逆矩阵或黑塞矩阵，简化了这一计算过程。

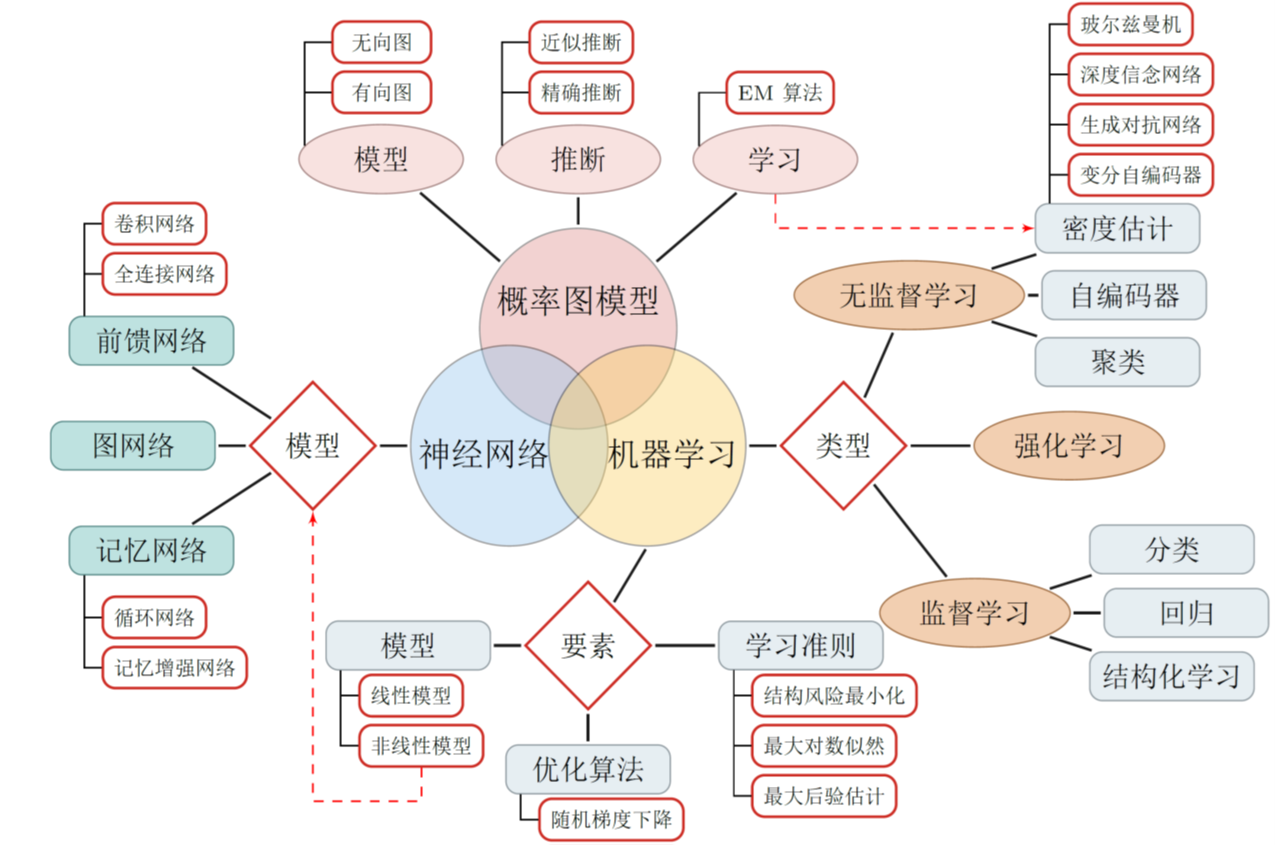
# 二、神经网络与深度学习

本周认真阅读了邱锡鹏老师的《神经网络与深度学习》3小时课程概要PPT，我复习了神经网络基础，前馈神经网络，卷积神经网络，循环神经网络，图网络等知识。

前馈网络存在一些不足，比如连接存在层与层之间，每层的节点之间是无连接的。（无循环）输入和输出的维数都是固定的，不能任意改变。无法处理变长的序列数据。假设每次输入都是独立的，也就是说每次网络的输出只依赖于当前的输入。循环神经网络通过一些策略解决了这些不足，循环神经网络通过使用带自反馈的神经元，能够处理任意长度的序列。循环神经网络比前馈神经网络更加符合生物神经网络的结构。循环神经网络已经被广泛应用在语音识别、语言模型以及自然语言生成等任务上。

循环神经网络可以很容易地扩展到两种更广义的记忆网络模型：递归神经网络和图网络。递归神经网络是是在一个有向图无循环图上共享一个组合函数。如果将循环神经网络按时间展开，每个时刻的隐状态 ht 看做一个节点，那么这些节点构成一个链式结构，而链式结构是一种特殊的图结构。

下面这幅图可以帮助我们梳理神经网络、机器学习、强化学习等学习任务及其常用算法。

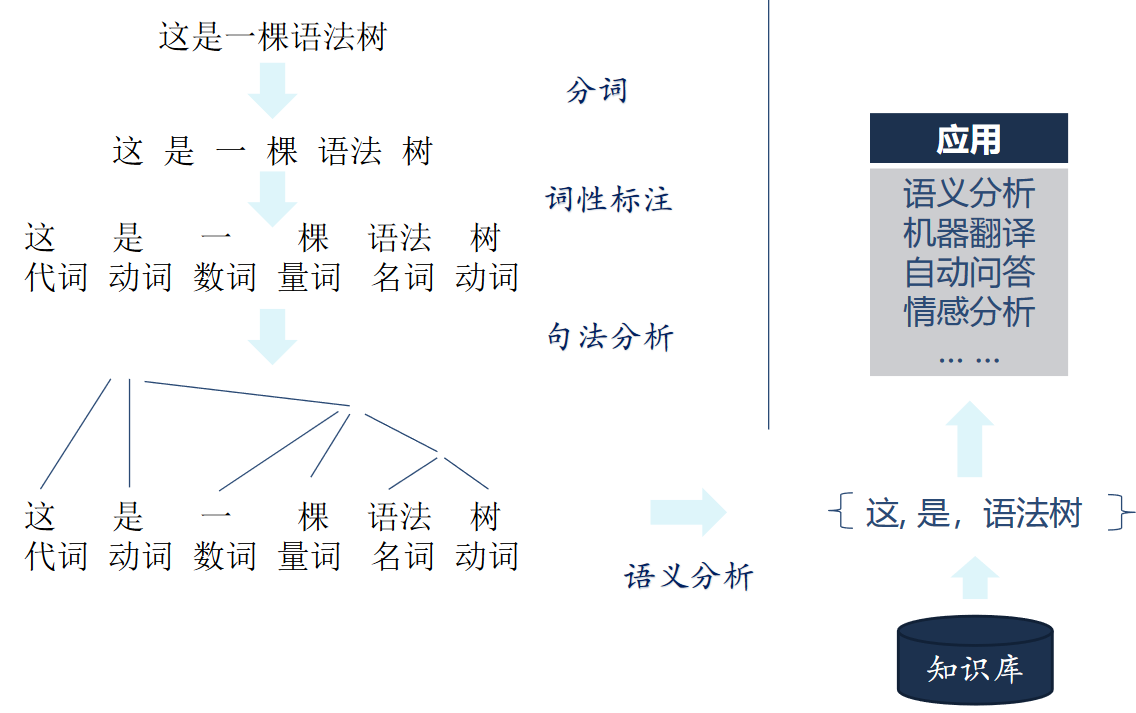


# 三、面向自然语言处理的深度学习基础

和第二个任务相比，该任务重点关注将神经网络与深度学习的知识应用在自然语言处理任务中。自然语言处理中的典型任务可以分为：分类，序列标注和生成。我的研究方向是多媒体事件抽取，属于序列标注任务。

其中令我印象深刻的是理想中的NLP技术路线和实际的NLP技术路线这两者的对比。结合华为项目的实际经历，我认识到在深度学习模型落地到实际应用的过程中，数据预处理阶段发挥着重要的作用。

1. 理想中的NLP技术路线



1. 实际的NLP技术路线

