Multitask Learning

1. 单任务学习 VS 多任务学习

单任务学习：一次只学习一个任务（task），大部分的机器学习任务都属于单任务学习。

多任务学习：把多个相关的(related)的任务放在一起学习，同时学习多个任务。

多任务学习产生的原因：现在大多数的机器学习任务都是单任务学习。对于复杂的问题，可以分解为简单且相互独立的子问题来单独解决，然后再合并结果。这样做看似合理，其实是不正确的，因为现实世界中很多问题不能分解为一个一个独立的子问题，即使可以分解，各个子问题之间也是相互关联的，通过一些共享因素或共享表示(shared representation)联系在一起。把现实问题当做一个个独立的单任务处理，忽略了问题之间所富含的丰富的关联信息。

多任务学习就是为了解决这个问题而诞生的，把多个相关(related)任务(task)放在一起学习，多个任务之间共享一些因素，他们可以在学习过程中，共享它们所学到的信息，这是单任务学习所不能具备的。相关联的多任务学习比单任务学习能有更好的泛化（generalization）效果。

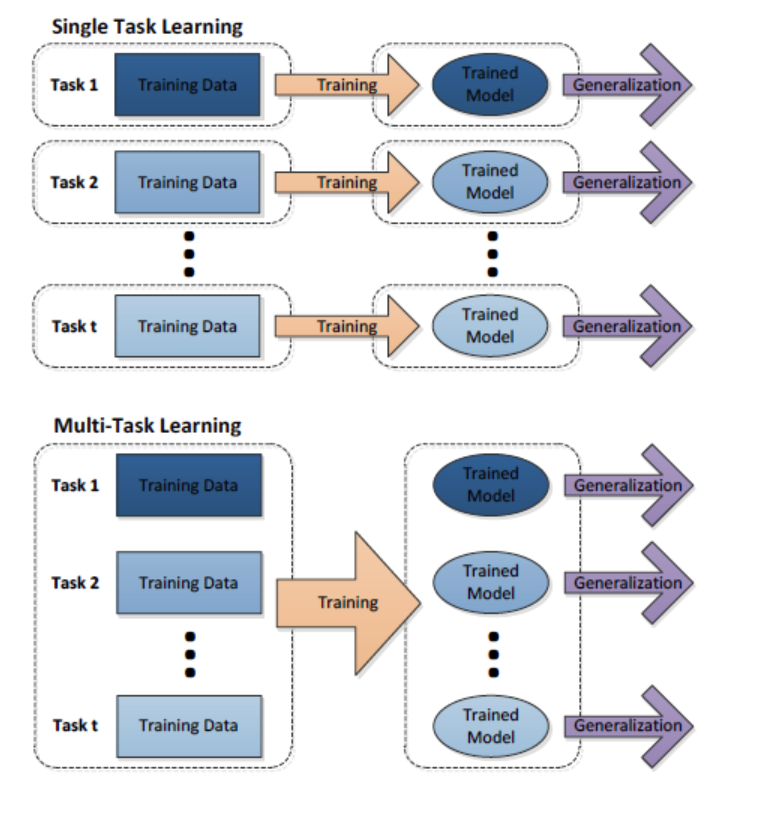


图1

单任务和多任务之间的对比如上图所示。从图中可以发现，单任务学习时，各个任务之间的模型空间（Trained Model）是相互独立的。多任务学习时，多个任务之间的模型空间（Trained Model）是共享的。

假设用含一个隐含层的神经网络来表示学习一个任务，单任务学习和多任务学习可以表示成如图2所示：

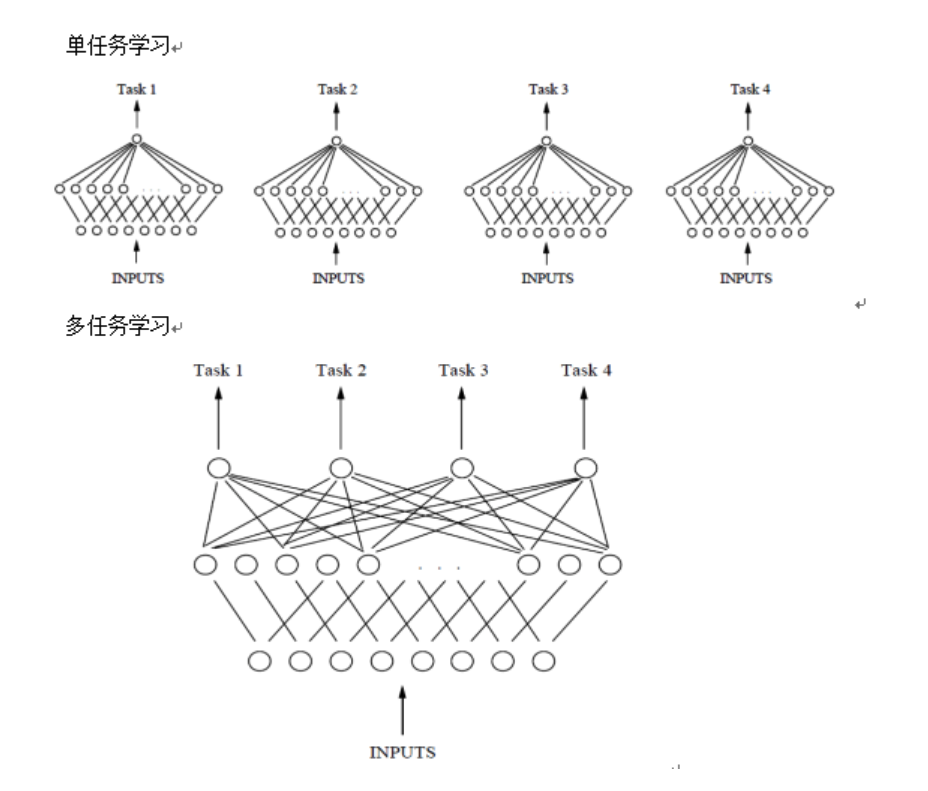


图2：基于单层神经网络的单任务和多任务学习对比

从图2可以看出，单任务学习时，各个task任务的学习是相互独立的，多任务学习时，多个任务之间的浅层表示共享（shared representation）。

1. 多任务学习的定义

多任务学习（Multitask Learning）定义：基于共享表示（shared representation），把多个相关的任务放在一起学习的一种机器学习方法。

多任务学习是一种推导迁移学习方法，主任务（main task）使用相关任务(related task)的训练信号(training signal)所拥有的领域相关信息(domain-specific information)，作为推导偏差来提升主任务(main tasks)的泛化效果(generalization performance)的一种机器学习方法。

多任务学习涉及多个相关的任务同时并行学习，梯度同时反向传播，多个任务通过底层的共享表示（shared representation）来互相帮助学习，提升泛化效果。

简单来说：多任务学习把多个相关的任务放在一起学习（注意，一定是相关的任务，后面会给出相关任务【related task】的定义，以及他们共享了哪些信息），学习过程(training)中通过一个浅层的共享(shared representation)表示来互相分享，互相补充学习到的领域相关的信息（domain information），互相促进学习，提升泛化的效果。

共享表示的目的就是为了提高泛化，MTL(多任务学习)中共享由两种方式：

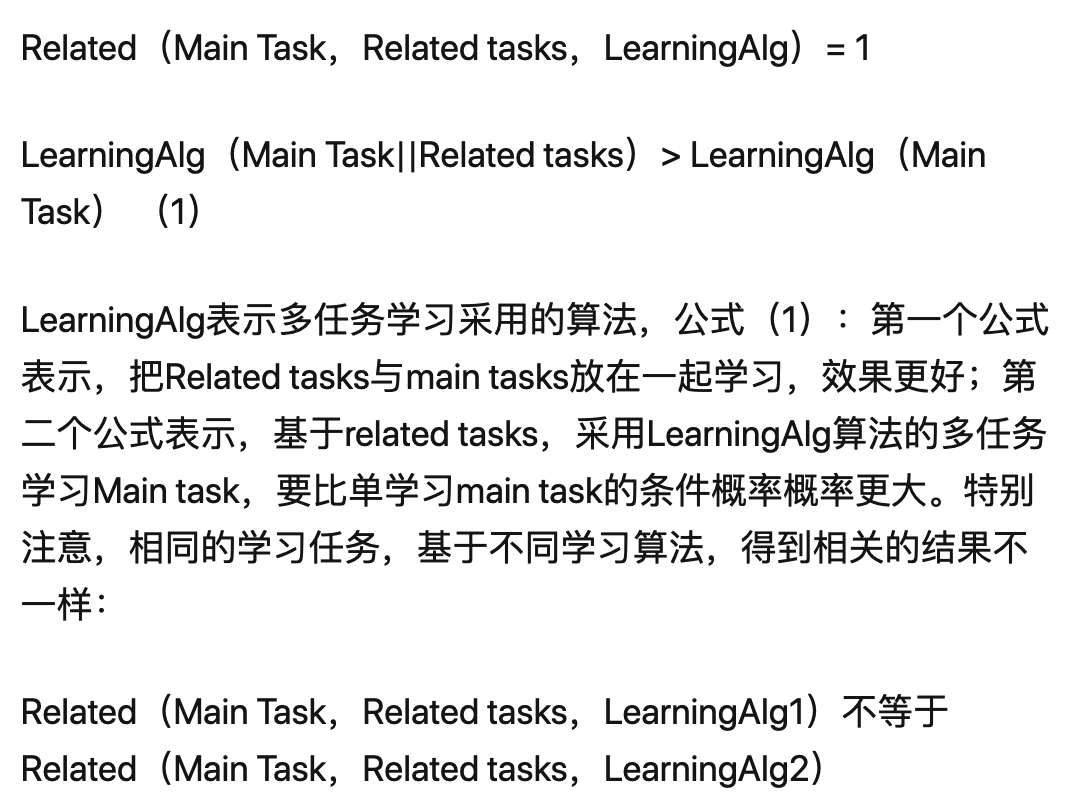
* 基于参数的共享(Parameter based)：比如基于神经网络的MTL，高斯处理过程。
* 基于约束的共享（regularization based）：比如均值，联合特征(Joint feature)学习（创建一个常见的特征集合）。

1. 多任务学习有效的原因

* 1）多任务学习时，有相关的部分，也有不相关的部分。当学习一个任务（Main task）时，与该任务不相关的部分，在学习过程中相当于噪声。引入噪声是可以提高学习的泛化效果的。
* 2）单任务学习时，梯度的反向传播倾向于陷入局部最小值，多任务学习中不同任务的局部最小值处于不同的位置，通过相互作用，可以帮助隐含层逃离局部最小值。
* 3）多个任务在浅层共享表示，可能削弱了网络的能力，降低网络的过拟合，提升了泛化效果。

1. 相关定义

相关的具体定义很难，但我们可以知道，在多任务学习中，related task可以提升main task的学习效果，基于这点得到相关的定义：



1. Multitask Learning方法

最简单的MTL是浅层隐含节点共享的神经网络，如上图2所示，还有基于特征（feature table）共享MTL。

常见的基于特征共享的MTL(联合特征学习，joint feature learning)，通过构建一个常见的特征集合来实现多个任务之间基于特征的共享表示，其共享表示如图3所示：



图3：基于特征的共享表示示意图

基于特征共享的MTL的输入输出关系如图4所示，其中采用L1正则来保证其稀疏性。

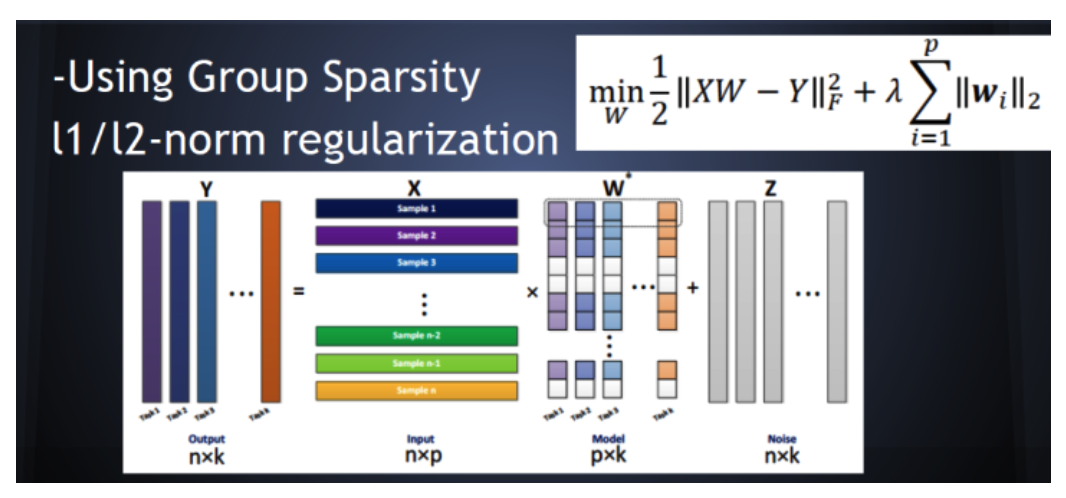


图4：基于联合特征的MTL示意图

图4的大体意思为：

* 总共有k个任务，每个任务用一个n维向量表示
* 等号左边n\*p的输入矩阵，表示有n个sample，每个sample用一个p维特征表示。然后再加上一个n\*k的噪声矩阵（不相关部分），就得到了左边的输出矩阵。

1. 多任务学习与其他学习算法之间的关系

多任务学习（Multitask Learning）是迁移学习算法的一种。定义一个源领域source domain和一个目标领域target domain，在source domain学习，并把学习到的知识迁移到target domain，提升target domain的学习效果。

多标签学习（Multilabel Learning）是多任务学习中的一种，建模多个label之间的相关性，同时对多个label进行建模，多个类别之间共享相同的数据/特征。

多类别学习(Multiclass Learning)是多标签学习任务中的一种，对多个相互独立的类别(classes)进行建模。这几个学习之间的关系如图5所示：

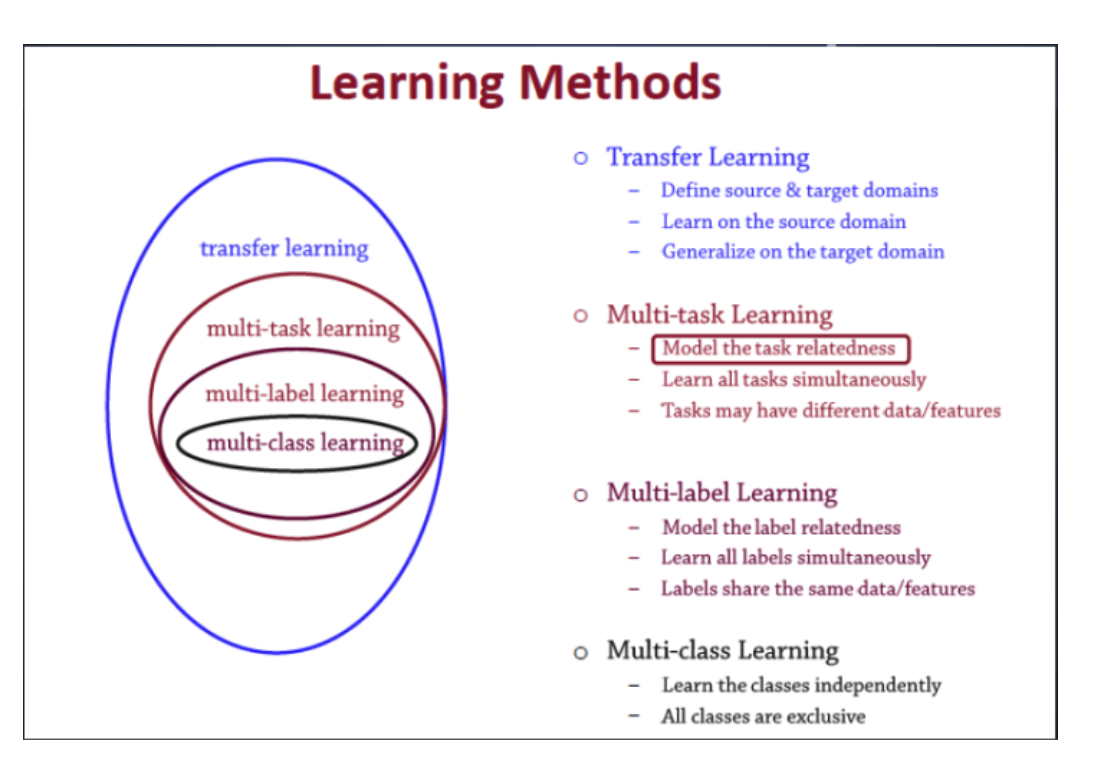


图5：多任务学习与其他机器学习方法之间的关系