

《神经网络与深度学习》



模型独立的学习方式

<https://nndl.github.io/>

模型独立的学习方式

- ▶ 这些学习方式不限于具体的模型
 - ▶ 前馈神经网络、循环神经网络还是其他模型
- ▶ 然而一种学习方式往往会对符合某种特性的模型更加青睐
 - ▶ 集成学习往往和方差大的模型组合时效果显著。

内容

- ▶ 集成学习
- ▶ 协同学习
- ▶ 多任务学习
- ▶ 迁移学习
- ▶ 终身学习
- ▶ 元学习

集成学习

►三个臭皮匠赛过诸葛亮

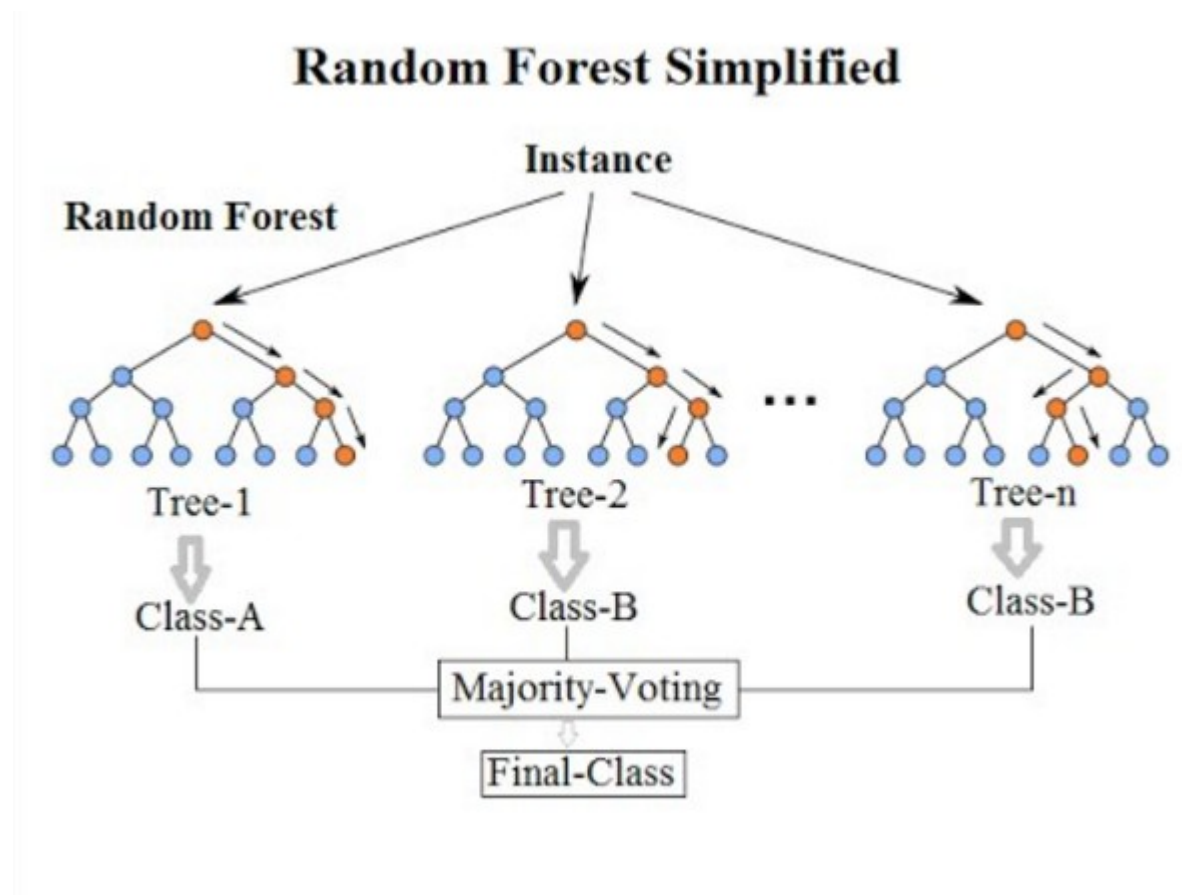
定理 10.1: 对于 M 个不同的模型 $f_1(\mathbf{x}), \dots, f_M(\mathbf{x})$, 平均期望错误为 $\bar{\mathcal{R}}(f)$ 。基于简单投票机制的集成模型 $f^{(c)}(\mathbf{x}) = \frac{1}{M} \sum_{m=1}^M f_m(\mathbf{x})$, 其期望错误在 $\frac{1}{M} \bar{\mathcal{R}}(f)$ 和 $\bar{\mathcal{R}}(f)$ 之间。

集成方式

► Bagging 类

- Bagging (Bootstrap Aggregating) 是一个通过不同模型的训练数据集的独立性来提高不同模型之间的独立性。我们在原始训练集上进行有放回的随机采样，得到 M 比较小的训练集并训练 M 个模型，然后通过投票的方法进行模型集成。
- 随机森林 (Random Forest) 是在 Bagging 的基础上再引入了随机特征，进一步提高每个基模型之间的独立性。在随机森林中，每个基模型都是一棵决策树。

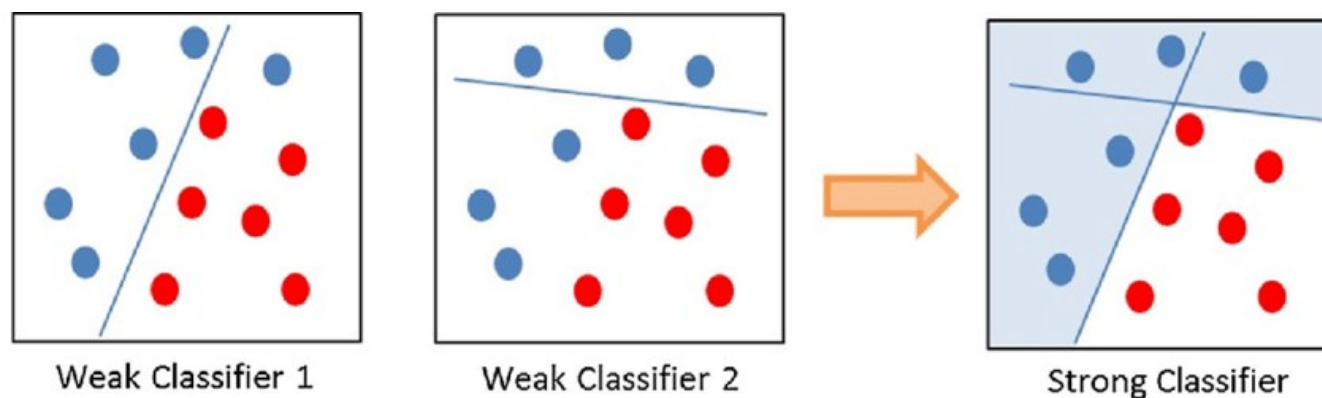
随机森林



集成方式

► Boosting 类

- 按照一定的顺序来先后训练不同的基模型，每个模型都针对前序模型的错误进行专门训练。根据前序模型的结果，来调整训练训练样本的权重，从而增加不同基模型之间的差异性。
- AdaBoost （ Adaptive Boosting ） 算法



AdaBoost

Input :

- A training set $S = ((\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m))$.

Initialization :

- Maximum number of iterations T ;
- initialize the weight distribution $\forall i \in \{1, \dots, m\}, D^{(1)}(i) = \frac{1}{m}$.

for $t = 1, \dots, T$ **do**

- Learn a classifier $f_t : \mathbb{R}^d \rightarrow \{-1, +1\}$ using distribution $D^{(t)}$
- Set $\epsilon_t = \sum_{i: f_t(\mathbf{x}_i) \neq y_i} D^{(t)}(i)$
- Choose $a_t = \frac{1}{2} \ln \frac{1-\epsilon_t}{\epsilon_t}$
- Update the weight distribution over examples

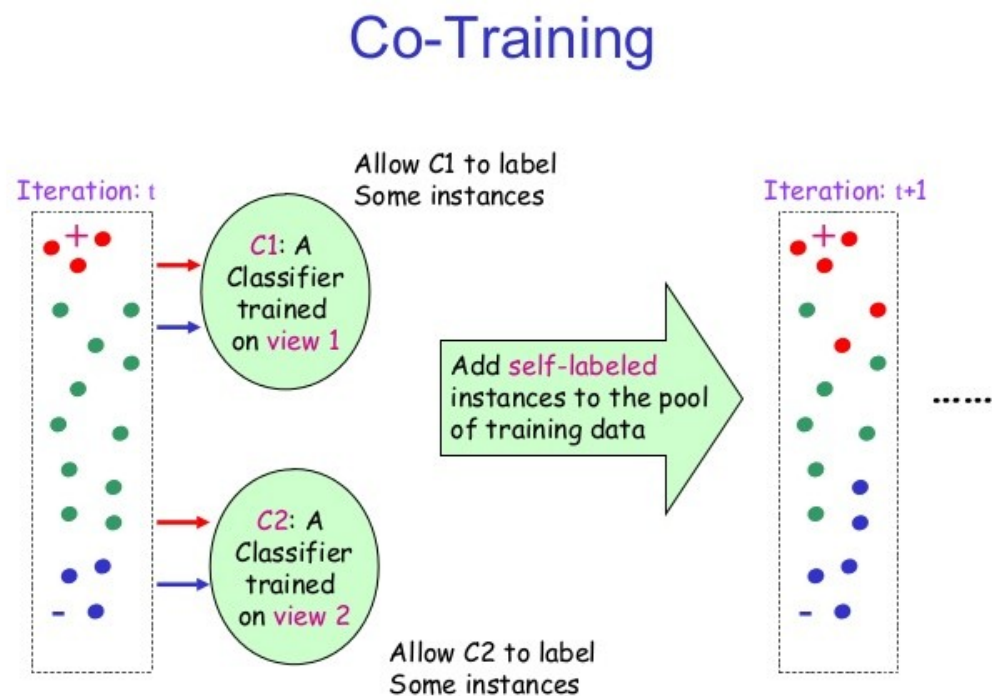
$$\forall i \in \{1, \dots, m\}, D^{(t+1)}(i) = \frac{D^{(t)}(i) e^{-a_t y_i f_t(\mathbf{x}_i)}}{Z^{(t)}}$$

where $Z^{(t)} = \sum_{i=1}^m D^{(t)}(i) e^{-a_t y_i f_t(\mathbf{x}_i)}$ is a normalization factor such that $D^{(t+1)}$ remains a distribution.

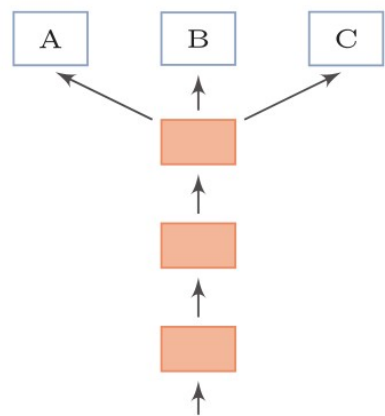
Output : The voted classifier $\forall \mathbf{x}, F(\mathbf{x}) = \text{sign} \left(\sum_{t=1}^T a_t f_t(\mathbf{x}) \right)$

协同训练 (Co-Training)

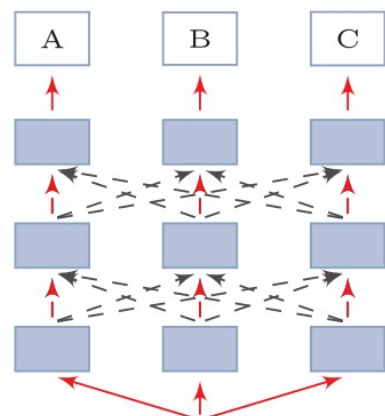
► Multi-View Learning



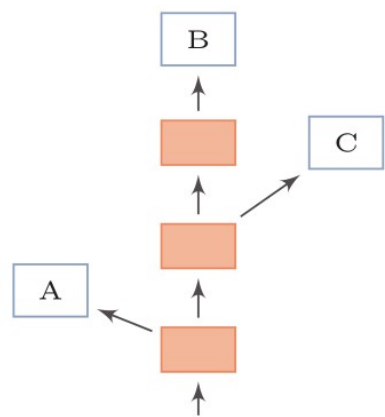
多任务学习 (Multitask Learning)



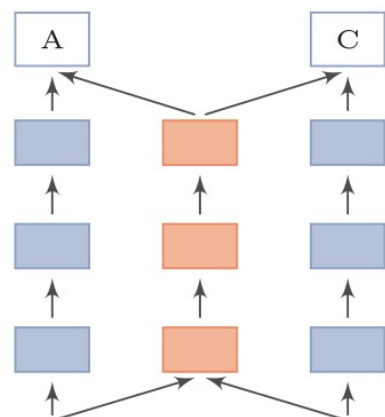
(a) 硬共享模式



(b) 软共享模式



(c) 层次共享模式



(d) 共享-私有模式

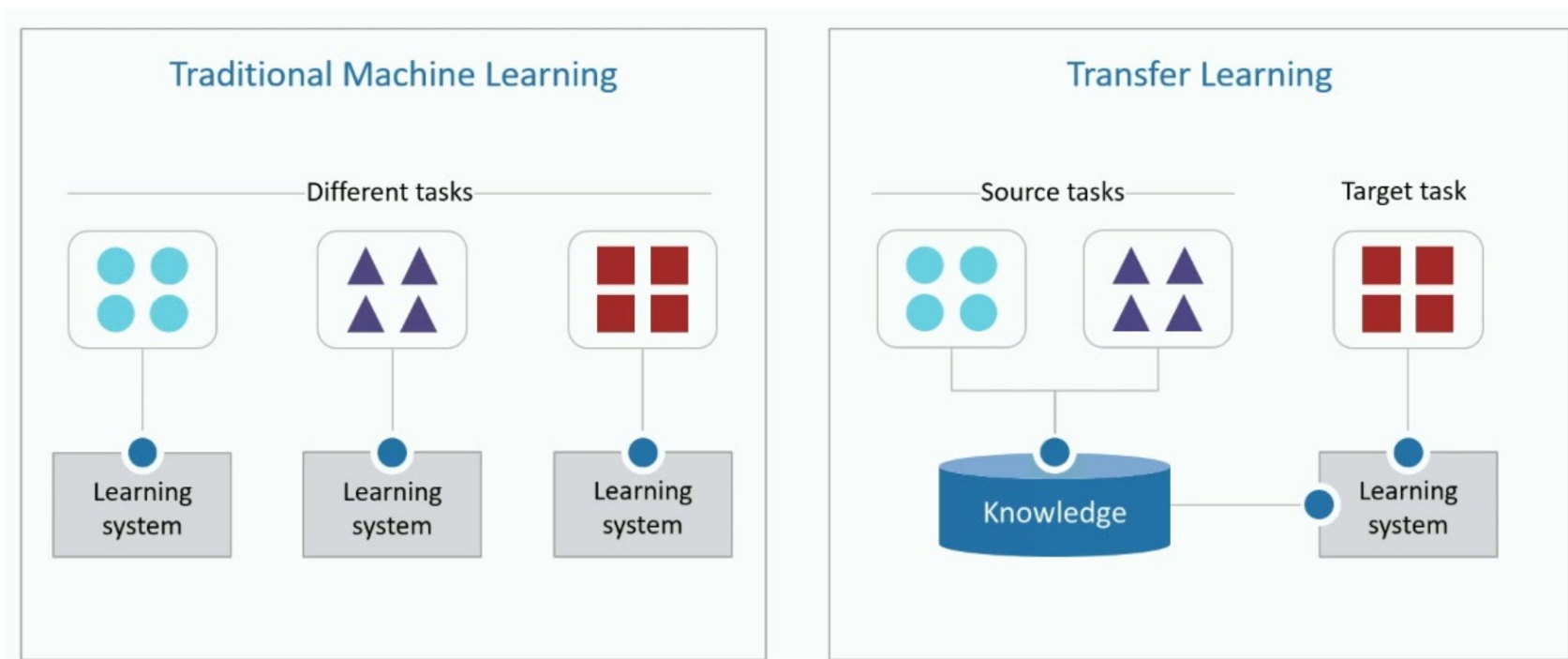
RICH CARUANA

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213

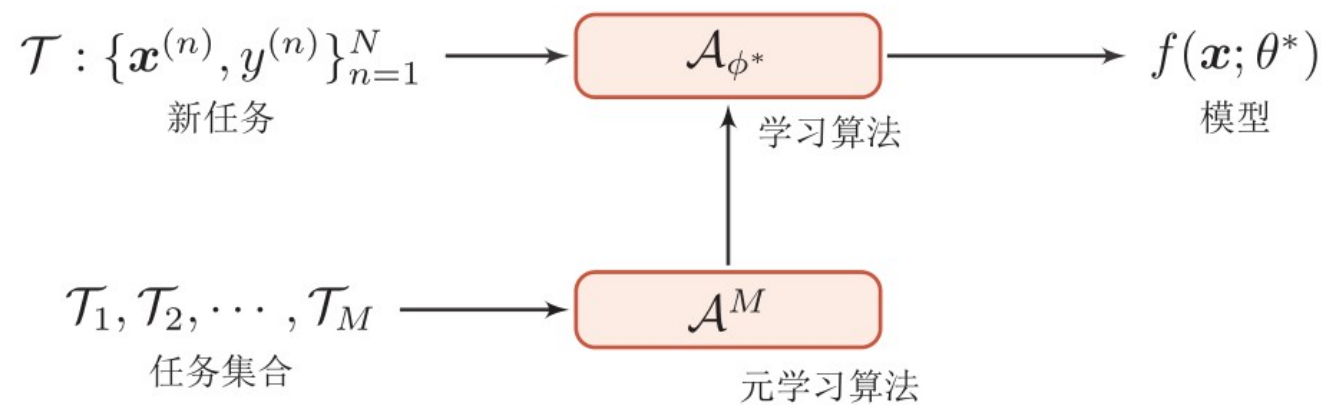
caruana@cs.cmu.edu

Multitask Learning is an approach to inductive transfer that improves generalization by using the domain information contained in the training signals of related tasks as an inductive bias. It does this by learning tasks in parallel while using a shared representation; what is learned for each task can help other tasks be learned better.

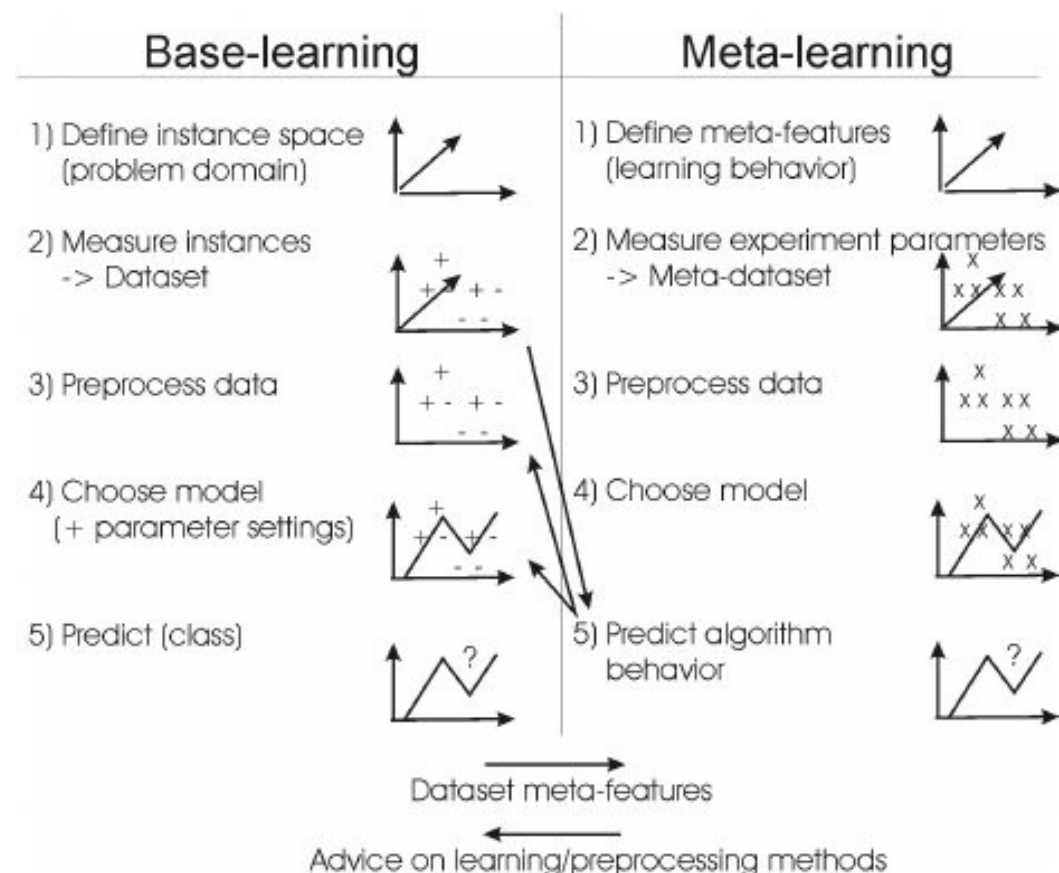
迁移学习（Transfer Learning）



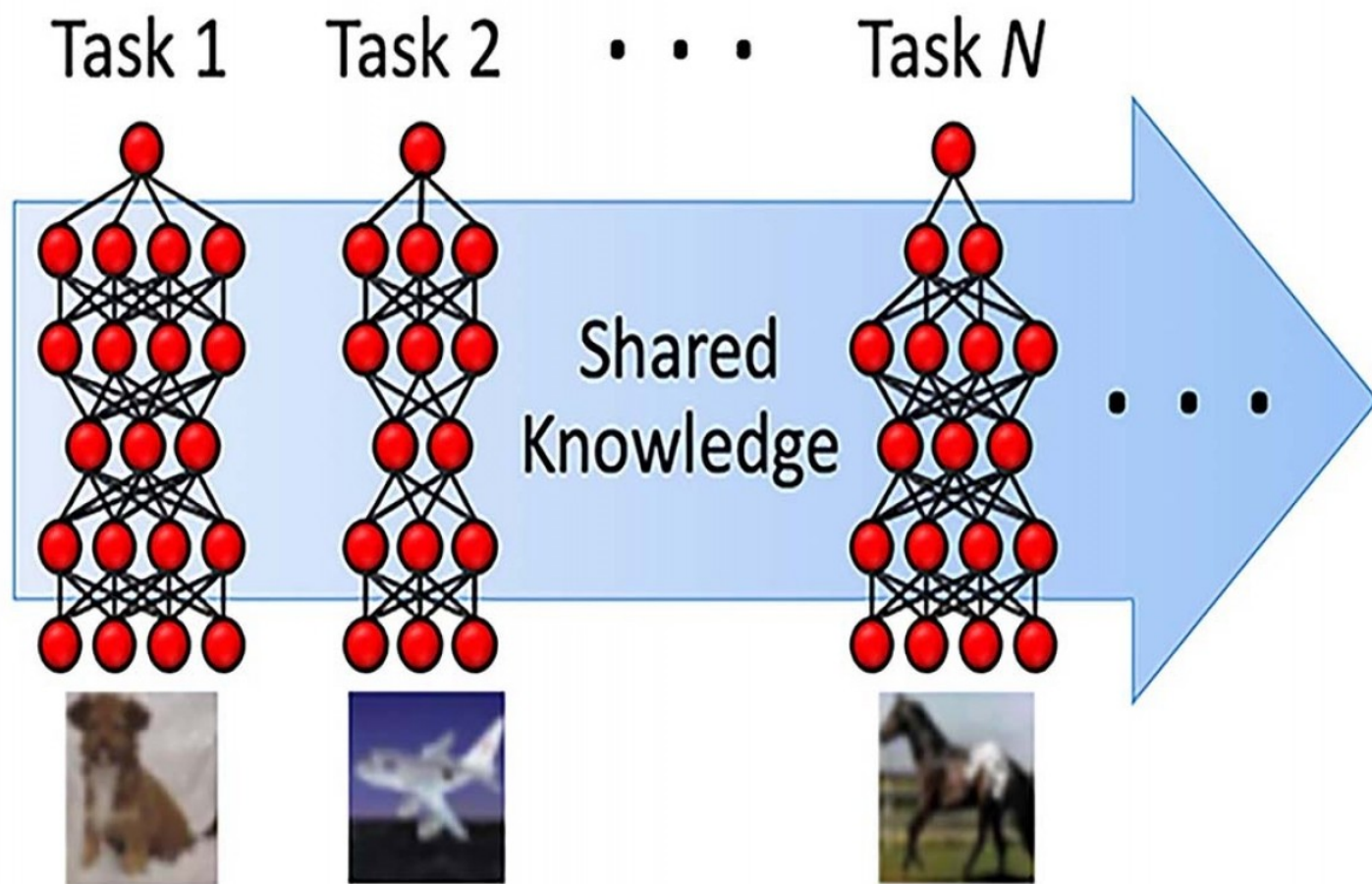
元学习 (Meta Learning)



元学习 (Meta Learning)



终身学习



谢 谢

<https://nndl.github.io/>