

# 第八章 机器学习

- 机器学习的定义和发展历史
- 机器学习的主要策略与基本结构
- 机器学习方法的分类
- 目前流行的机器学习方法

# 前言

- 机器能否象人类一样能具有学习能力呢？1959年美国的塞繆尔(Samuel)设计了一个下棋程序，这个程序具有学习能力，它可以在不断的对奕中改善自己的棋艺。4年后，这个程序战胜了设计者本人。又过了3年，这个程序战胜了美国一个保持8年之久的常胜不败的冠军。这个程序向人们展示了机器学习的能力，提出了许多令人深思的社会问题与哲学问题。
- 机器学习是继专家系统之后人工智能应用的又一重要研究领域，也是人工智能和神经计算的核心研究课题之一。现有的计算机系统和人工智能系统没有什么学习能力，至多也只有非常有限的学习能力，因而不能满足科技和生产提出的新要求

# 机器学习的定义和发展历史

- 1. 学习的基本概念

- 按照人工智能大师西蒙的观点，学习就是系统在不断重复的工作中对本身能力的增强或者改进，使得系统在下一次执行**同样任务或类似**任务时，会比现在做得更好或效率更高。

- 2. 机器学习的定义

- 机器学习是**研究如何使用机器来模拟人类学习活动的一门学科**。稍为严格的提法是：机器学习是一门研究机器获取新知识和新技能，并识别现有知识的学问。

机器学习就是计算机自动获取知识，它是知识工程的三个分支（**使用知识、知识表示、获取知识**）之一。

# 机器学习的发展史

- 机器学习是人工智能应用研究较为重要的分支，它的发展（国内）过程大体上可分为4个时期：
- 1. 第一阶段是在50年代中叶到60年代中叶，属于热烈时期。在这个时期，所研究的是“没有知识”的学习，即“**无知**”**学习**；其研究目标是各类自组织系统和自适应系统；指导本阶段研究的理论基础是早在40年代就开始研究的神经网络模型。在这个时期，我国研制了数字识别学习机。

- 2. 第二阶段在60年代中叶至70年代中叶，被称为机器学习的冷静时期。本阶段的研究目标是模拟人类的概念学习过程，并采用逻辑结构或图结构作为机器内部描述。这个时期正是我国“史无前例”的十年，对机器学习的研究不可能取得实质进展。
- 3. 第三阶段从70年代中叶至80年代中叶，称为复兴时期。在这个时期，人们从学习单个概念扩展到学习多个概念，探索不同的学习策略和各种学习方法。本阶段已开始把学习系统与各种应用结合起来，中国科学院自动化研究所进行质谱分析和模式文法推断研究，表明我国的机器学习研究得到恢复。1980年西蒙来华传播机器学习的火种后，我国的机器学习研究出现了新局面。
- 4. 机器学习的最新阶段始于1986年。一方面，由于神经网络研究的重新兴起，另一方面，对实验研究和应用研究得到前所未有的重视。我国的机器学习研究开始进入稳步发展和逐渐繁荣的新时期。

# 机器学习的目前状况

机器学习进入新阶段的表现：

1. 机器学习已成为新的**边缘学科**并在高校形成课程。
2. 综合各种学习方法的**集成学习**系统正在兴起。
3. 机器学习与人工智能问题的统一性观点正在形成。
4. 各种学习方法的应用范围不断扩大。
5. 数据挖掘和知识发现的研究已形成热潮。
6. 与机器学习有关的学术活动空前活跃。

# 机器学习的主要策略

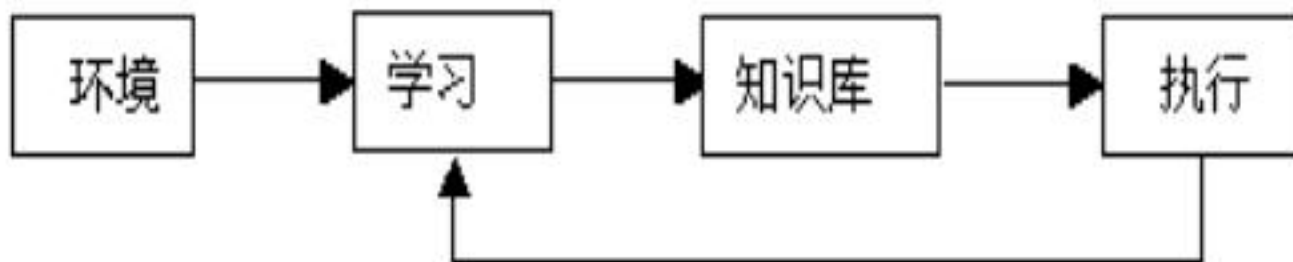
- 学习过程与推理过程是紧密相连的，按照学习中使用推理的多少，机器学习所采用的策略大体上可分为4种：**机械学习、示教学习、类比学习和示例学习**。学习中所用的**推理**越多，系统的能力越强。

- 1. **机械学习**就是记忆，是最简单的学习策略。这种学习策略不需要任何推理过程。
- 2. 比机械学习更复杂一点的学习是**示教学习策略**。系统在接受外部知识时需要一点推理，翻译和转化工作。
- 3. **类比学习系统**只能得到完成类似任务的有关因此，他比上述两种学习策略需要更多的推理。
- 4. 采用**示例学习策略**的计算机系统，事先完全没有完成任务的任何规律性的信息，因此需要推理是最多的。

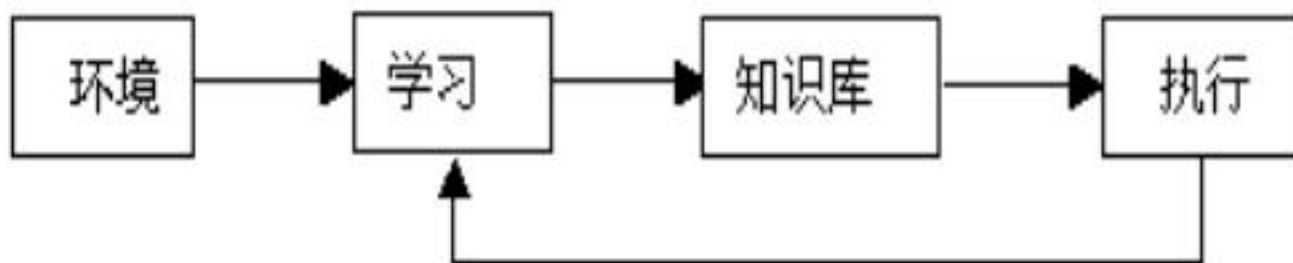


# 机器学习系统的基本结构

- 学习的过程是建立理论、形成假设和进行归纳推理。以Simon的关于学习的定义作为出发点，建立起如下机器学习系统的基本模型：



# 机器学习系统的基本结构



- 环境和知识库是以**某种知识表示形式表达的信息的集合**，分别代表外界信息来源和系统具有的知识。“学习环节”和“执行环节”代表两个过程。“**环境**”向系统的“**学习环节**”**提供某些信息**，而“学习环节”则利用这些信息**对系统的“知识库”进行改进**，以增进系统“执行环节”完成任务的效能。“执行环节”根据知识库中的知识来完成某种任务，同时把**获得的信息反馈给“学习环节”**。

# 影响学习系统设计的重要因素

- 1. 影响学习系统设计的最重要的因素是环境向系统提供的信息。整个过程要遵循“取之精华，弃之糟粕”的原则，同时谨记“实践是检验真理的唯一标准”。
- 2. 知识库是影响学习系统设计的第二个因素。知识的表示有多种形式，在选择表示方式时要兼顾以下4个方面：

- (1) **表达能力强**
- 所选择的表示方式能很容易地表达有关的知识。
- (2) **易于推理**
- 为了使学习系统的计算代价比较低，希望知识表示方式能使推理较为容易。
- (3) **容易修改知识库**
- 学习系统的本质要求它不断地修改自己的知识库，当推广得出一般执行规则后，要加到知识库中。
- (4) **知识表示易于扩展**
- 学习系统不能在全然没有任何知识的情况下凭空获取知识，每一个学习系统都要求具有某些知识理解环境提供的信息，分析比较，做出假设，检验并修改这些假设。因此，更确切地说，学习系统是对现有知识的扩展和改进

# 机器学习方法的分类

## • 基于学习方法的分类

➤ 机械学习

➤ 归纳学习

符号归纳（实例学习、决策树……）

函数归纳（神经网络、示例学习、统计学习）

➤ 类比学习

——》通过对相似事物进行比较所进行的学习。

➤ 解释学习

——》使用先验知识来对实例进行解释，从而得到实例成立的一般假设。

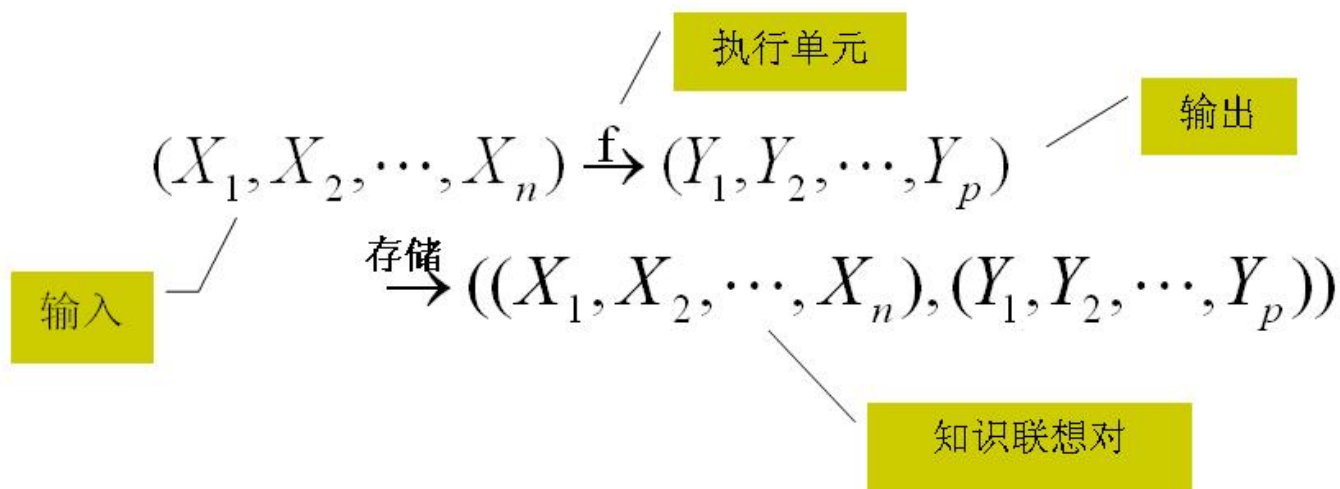
# 机械学习 (Rote Learning)

- 1、机械学习
- 又称为记忆学习或死记硬背式的学习。这种学习方法**直接记忆或存储环境提供的新知识**，并在以后通过对知识库的检索来直接使用这些知识，而**不再需要进行任何的计算和推导**。
- 机械学习是一种基本的学习过程，虽然它没有足够的能力独立完成智能学习，但存储对于任何智能型的程序来说，都是**必要的和基本的**。

记忆学习是任何学习系统的一部分，任何学习系统都要将它所获取的知识存储在知识库中，以便使用这些知识。

# 机械学习的过程

- 执行机构每解决一个问题，系统就记住这个问题和它的解。简单的机械学习模型：



# 例子：汽车修理成本估算系统

- **输入：**有关待修理汽车的描述，包括制造厂家、出厂日期、车型、汽车损坏的部位以及它的损坏程度
- **输出：**该汽车的修理成本
- **为了进行估算，系统必须在其知识库中查找同一厂家，同一出厂日期、同一车型、同样损坏情况的汽车，然后把知识库中对应的数据作为修理成本的估算数据输出给用户。如果在系统的知识库中没有找到这样的汽车，则系统将请求用户给出大致的费用并进行确认，系统则会将该车的描述和经过确认的估算费用存储到知识库中，以便将来查找使用。**



# 机械学习的主要问题

- 对于机械学习，需要注意3个重要的问题：存储组织，稳定性和存储与计算之间的权衡。
- (1) 存储组织信息：采用适当的存储方式，使检索速度尽可能地快，是机械学习中的重要问题。
- (2) 环境的稳定性与存储信息的适用性问题：机械学习系统必须保证所保存的信息适应于外界环境变化的需要，这也就是所谓的信息适用性问题。
- (3) 存储与计算之间的权衡：对于机械学习来说很重要的一点是它不能降低系统的效率。

# 归纳学习

- 归纳学习的定义
- (1) 归纳 (induction) 是人类拓展认识能力的重要方法，是一种**从个别到一般的，从部分到整体的推理行为**。
- (2) 归纳推理是应用归纳方法，从足够多的**具体事例**中归纳出**一般性知识**，提取事物的一般规律；它是一种从个别到一般的推理。
- (3) **归纳学习 (induction learning)** 是应用归纳推理**进行学习的一种方法**。根据归纳学习可分为示例学习和观察与发现学习。前者属于有师学习，后者属于无师学习。

# 归纳学习的模式和规则

- 归纳学习的一般模式为：
- 给定：
- (1) 观察陈述（事实） $F$ ，用以表示有关某些对象、状态、过程等的特定知识；
- (2) 假定的初始归纳断言（可能为空）；
- (3) 背景知识，用于定义有关观察陈述、候选归纳断言以及任何相关问题领域知识、假设和约束，其中包括能够刻画所求归纳断言的性质的优先准则。

- 求：归纳断言（假设） $H$ ，能重言蕴涵或弱蕴涵观察陈述，并满足背景知识。
- 假设 $H$ 永真蕴涵事实 $F$ ，说明 $F$ 是 $H$ 的逻辑推理，则有：
  - $H \mid > F$ （读作 $H$ 特殊化为 $F$ ）
  - 或  $F \mid < H$ （读作 $F$ 一般化或消解为 $H$ ）
  - 这里，从 $H$ 推导 $F$ 是演绎推理，因此是保真的；而从事实 $F$ 推导出假设 $H$ 是归纳推理。

- **归纳学习系统的模型如图所示：**



- **实例规划过程通过对实例空间的搜索完成实例选择，并将这些选中的活跃实例提交解释过程。解释过程对实例加以适当转换，把活跃实例变换为规则空间中的特定概念，以引导规则空间的搜索。**

# 归纳学习方法

## • 1、示例学习

- 示例学习 (learning from examples) 又称为实例学习，它是通过环境中**若干与某概念有关的例子**，经归纳得出一般性概念的一种学习方法。
- 在这种学习方法中，外部环境提供的是一组例子（正例和反例），示例学习就是要从这些特殊知识中归纳出适用于更大范围的一般性知识，以覆盖所有的正例并排除所有反例。
- 例：给出肺炎与肺结核两种病的部分病例。每个病历都有五种症状：发烧（无，低高），X光所见阴影（点状，索条状，片状，空洞）血沉（正常，快）听诊（正常，干鸣音，水泡音）
- 通过示例学习，可以从病例中归纳如下诊断规则：
  - (1)  $\text{血沉} = \text{正常} \wedge (\text{听诊} = \text{干鸣音} \vee \text{水泡音}) \rightarrow \text{诊断} = \text{肺炎}$
  - (2)  $\text{血沉} = \text{快} \rightarrow \text{诊断} = \text{肺结核}$

# 归纳学习方法

- **2、观察发现学习**
- **观察发现学习又称为描述性概括，其目标是确定一个定律或理论的一般性描述，刻画观察集，指定某类对象的性质。观察发现学习可分为观察学习与机器发现两种。前者用于对事例进行聚类，形成概念描述；后者用于发现规律，产生定律或规则。**

# 类比学习

- 1. 类比推理和类比学习形式
- 类比推理是由新情况与已知情况在某些方面的相似来推出它们在其它相关方面的相似。显然，类比推理是在两个相似域之间进行的：类比推理的目的是从源域中选出与当前问题最近似的问题及其求解方法以求解决当前的问题，或者建立起目标域中已有命题间的联系，形成新知识。



# 类比学习

- 类比学习是利用二个不同领域（源域、目标域）中的知识相似性，可以通过类比，从源域的知识（包括相似的特征和其它性质）推导出目标域的相应知识，从而实现学习。
- 例如：
  1. 一个从未开过货车的司机，只要他有开小车的知识就可完成开货车的任务。
  2. 若把某个人比喻为很像消防车，则可通过观察消防车的行为，推断出这个人的性格。
- 所以，类比学习系统可以使一个已有的计算机应用系统转变为适应于新的领域，来完成原先没有设计的相类似的功能。

# 类比推理

- 类比推理是在两个相似域之间进行的：
  - 源域S（已经认识的域）
  - 目标域T（当前尚未完全完全认识的域）

- 推理目的：

从S中选出与当前问题最近似的问题及其求解方法以解决当前的问题，或者建立起目标域中已有命题间的联系，形成新知识。

# 类比推理过程

- 1 回忆与联想
- 遇到新情况或新问题 $T$ 时，首先通过回忆与联想在 $S$ 中找出与当前情况相似的情况，这些情况是过去已经处理过的，有现成的解决方法及相关的知识。
- 2 选择
- 从找出的相似情况中选出与当前情况最相似的情况及其有关知识。
- 3 建立对应映射
- 在 $S$ 与 $T$ 的相似情况之间建立相似元素的对应关系，并建立起相应的映射。
- 4 转换
- 在上一步建立的映射下，把 $S$ 中的有关知识引到 $T$ 中来，从而建立起求解当前问题的方法或者学习到关于 $T$ 的新知识。

## 2. 类比学习过程与研究类型

- 类比学习主要包括如下四个过程：
- (1) 输入一组已知条件（已解决问题）和一组未完全确定的条件（新问题）。
- (2) 对输入的两组条件，根据其描述，按某种相似性的定义寻找两者可类比的对应关系。
- (3) 按相似变换的方法，将已有问题的概念、特性、方法、关系等映射到新问题上，以获得待求解新问题所需的新知识。
- (4) 对类推得到的新问题的知识进行校验。验证正确的知识存入知识库中，而暂时还无法验证的知识只能作为参考性知识，置于数据库中。

# 类比学习研究类型

- 问题求解型

求解一个新问题时，先回忆以前是否求解过类似问题，若是，则以此为依据求解新问题。

- 预测推理型

- 传统的类比法

用来推断一个不完全确定的事物可能还有的其他属性

- 因果关系型

已知因果关系  $S1: A \rightarrow B$ ,  
现有  $A' \cong A$ , 则可能有  $B'$   
满足  $A' \rightarrow B'$

# 解释学习

(Explanation-based learning, EBL)

- 解释学习兴起于20世纪80年代中期，根据任务所在领域知识和正在学习的概念知识，对当前实例进行分析和求解，得出一个表征求解过程的因果解释树，以获取新的知识。

# 解释学习

- 例如，学生根据教师提供的目标概念、该概念的一个例子、领域理论及可操作准则，首先构造一个解释来说明为什么该例子满足目标概念，然后将解释推广为目标概念的一个满足可操作准则的充分条件。

# 解释学习的一般性描述

- 1986年米切尔 (Mitchell) 等人为基于解释的学习提出了一个统一的算法EBG，该算法建立了基于解释的概括过程，并运用知识的逻辑表示和演绎推理进行问题求解。图表示EBG问题。

- EBG求解问题的形式可描述于下：

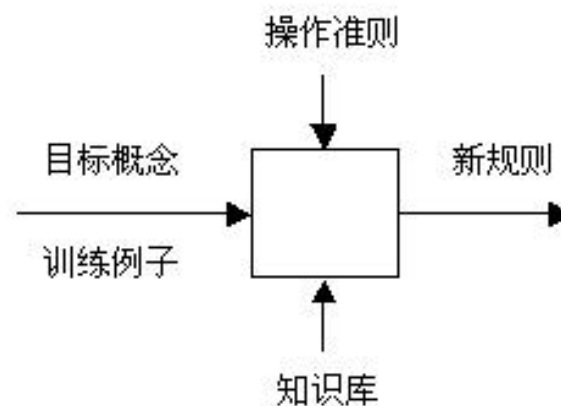
- 给定：领域知识DT

目标概念TC

训练实例TE

操作性准则OC

找出：满足OC的关于TC的充分条





# 解释学习的一般性描述

- 系统进行学习时，首先运用领域知识DT找出训练实例TE为什么是目标概念TC的实例的解释，然后根据操作性准则OC对解释进行推广，从而得到关于目标概念TC的一个一般性描述，即一个可供以后使用的形式化表示的一般性知识。

# 解释学习的学习过程与算法

EBG算法可概括为两步：

**1. 构造解释 2. 获取一般性的知识**

## 1. 构造解释

运用领域知识进行演绎，证明提供给系统的训练实例为什么是满足目标概念的一个实例。

例如：设要学习的目标概念是“一个物体 (Obj1) 可以安全地放置在另一个物体 (Obj2) 上”，即

Safe-To-Stack (Obj1, obj2)

# 解释学习的学习过程与算法

- 训练实例为描述物体Obj1与Obj2的下述事实：

$On(Obj1, Obj2)$

$Isa(Obj1, \text{book of AI})$

$Isa(Obj2, \text{table})$

$Volume(Obj1, 1)$

$Density(Obj1, 0.1)$

● ● ● ● ● ●

- 领域知识是把一个物体放置在另一个物体上面的安全性准则：

$Lighter(x, y) \rightarrow Safe\text{-}To\text{-}Stack(x, y)$

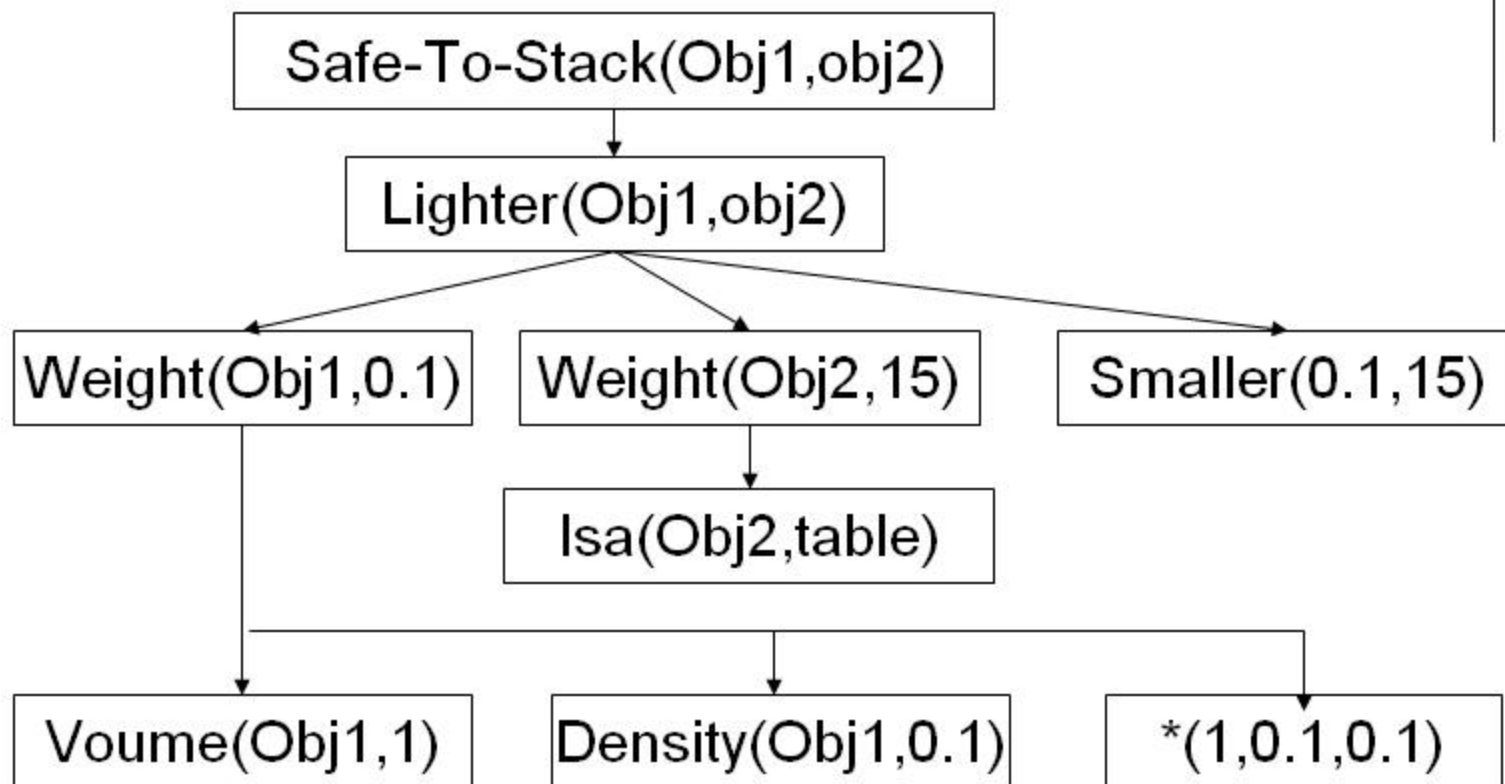
$Weight(p1, w1) \wedge Weight(p2, w2) \wedge Smaller(w1, w2) \rightarrow Lighter(p1, p2)$

$Volume(p, v) \wedge Density(p, d) \wedge *(v, d, w) \rightarrow Weight(p, w)$

$Isa(p, \text{table}) \rightarrow Weight(p, 15)$



图1 Safe-To-Stack(Obj1,obj2)解释结构



这是一个由目标概念引导的逆向推理，最终得到了一个解释结构。

# 解释学习的学习过程与算法

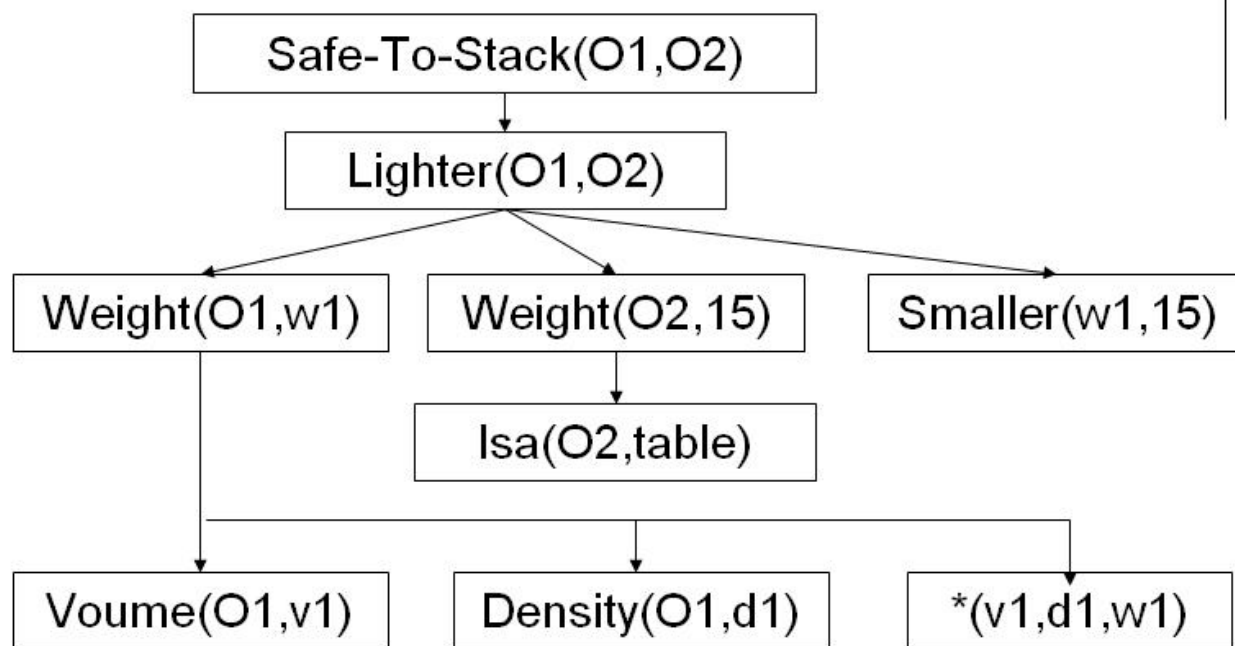
## 2. 获取一般性的知识

**任务：**对上一步得到的解释结构进行一般化的处理，从而得到关于目标概念的一般性知识。

**方法：**将常量换成变量，并把某些不重要的信息去掉，只保留求解问题必须的关键信息。图2为图1的一般化解释结构，可以得到如下一般性知识：

$$\text{Volume}(01, v1) \wedge \text{Density}(01, d1) \wedge *(v1, d1, w1) \wedge \text{Isa}(02, \text{table}) \wedge \text{Smaller}(w1, 15) \rightarrow \text{Safe-To-Stack}(01, 02)$$

图2 Safe-To-Stack(O1,O2)一般化解释结构



以后求解类似问题时，就可以直接利用这个知识进行求解，提到了系统求解问题的效率。

# 领域知识的完善性

- 领域知识对证明的形成起着重要的作用，只有完善的领域知识才能产生正确的学习描述。但是，不完善是难以避免的，此时有可能出现如下两种极端情况：
  1. 构造不出解释
  2. 构造出了多种解释
- 解决办法：
  1. **最根本的办法是提供完善的领域知识**
  2. **学习系统也应具有测试和修正不完善知识的能力，使问题能尽早地被发现，尽快地被修正。**

# 机器学习的主要内容

- 决策树
- 神经网络
- 支持向量机
- 贝叶斯分类器
- 半监督学习
- 强化学习
- 聚类
- .....



# 决策树

- 决策树 (Decision Tree)

别名：判定树、多级分类器。

- ◆ 模式识别中进行分类的一种有效方法。首先对数据进行处理，利用归纳算法生成可读的规则和决策树，然后使用决策对新数据进行分析。**本质上决策树是通过一系列规则对数据进行分类的过程。**
- ◆ 利用一系列查询回答来判断某一模式的类别是非常直观和自然的做法。后一个问题的提法依赖于前一个问题的回答。
- ◆ 这种问答方式对非度量数据特别有效，回答问题中不涉及任何距离度量。

名称	体温	表皮覆盖	胎生	水生动物	飞行动物	有腿	冬眠	类标号
人类	恒温	毛发	是	否	否	是	否	哺乳动物
海龟	冷血	鳞片	否	半	否	是	否	爬行类
鸽子	恒温	羽毛	否	否	是	是	否	鸟类
鲸	恒温	毛发	是	是	否	否	否	哺乳类

属性

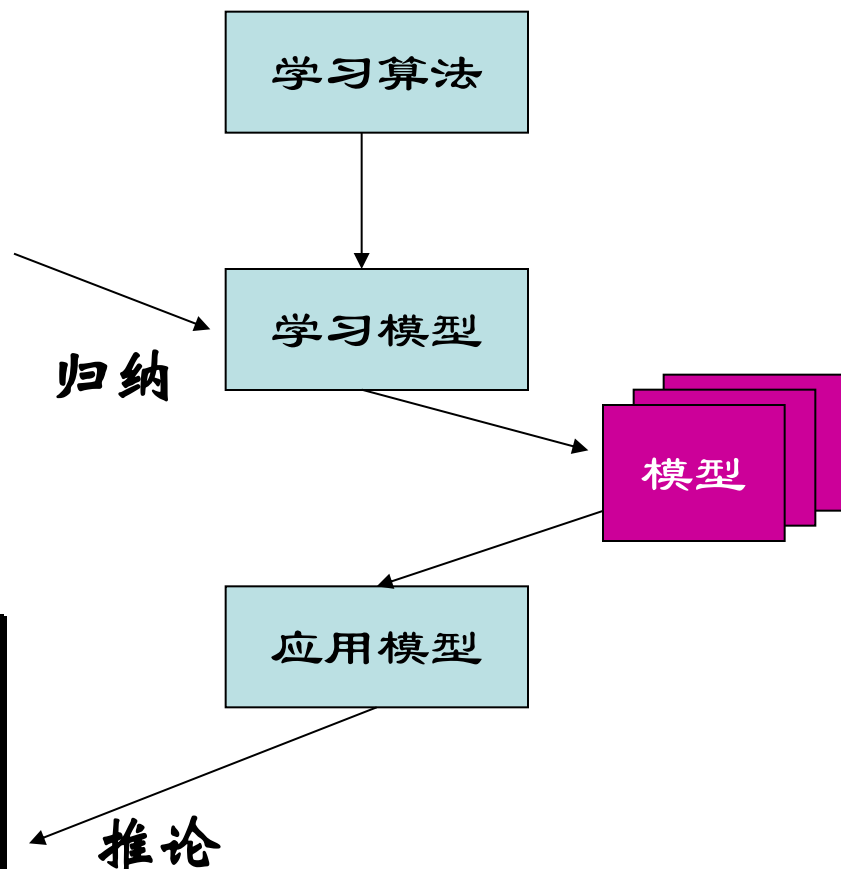
类别

训练集（类标号已知）

TID	A1	A2	A3	类
1	Y	100	L	N
2	N	125	S	N
3	Y	400	L	Y
4	N	415	M	N

检验集（类标号未知）

TID	A1	A2	A3	类
5	N	150	M	?
6	Y	225	L	?
7	N	320	S	?
8	N	100	L	?

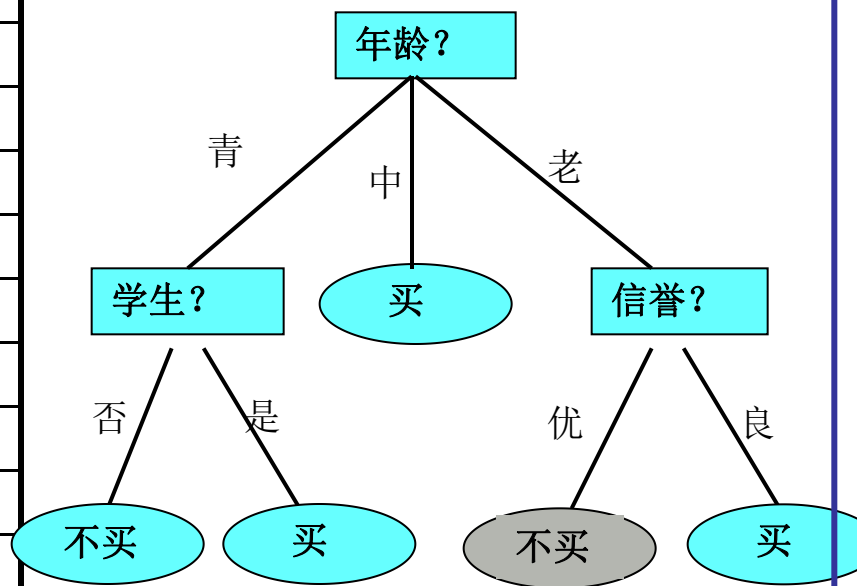


在训练样本集上归纳学习  
得到决策树，利用该决策  
树对测试样本进行分类。

## 决策树求解问题示例：

计数	年龄	收入	学生	信誉	归类：买计算机？
64	青	高	否	良	不买
64	青	高	否	优	不买
128	中	高	否	良	买
60	老	中	否	良	买
64	老	低	是	良	买
64	老	低	是	优	不买
64	中	低	是	优	买
128	青	中	否	良	不买
64	青	低	是	良	买
132	老	中	是	良	买
64	青	中	是	优	买
32	中	中	否	优	买
32	中	高	是	良	买
63	老	中	否	优	不买
1	老	中	否	优	买

## 谁在买计算机？



## 决策树的特点：

- 每个中间节点对应一个属性，节点下的分支为该属性的可能取值；
- 叶节点代表一个类别标记，每个叶节点到根节点的路径代表着一条判别规则；
- 决策树的产生规则是完备的，对于任何可分的问题，均可构造相应的判定树对其进行分类；

## 决策树的优点：

- 1、推理过程容易理解，决策推理过程可以表示成If Then形式。
- 2、推理过程完全依赖于属性变量的取值特点。
- 3、可自动忽略目标变量没有贡献的属性变量，也为判断属性变量的重要性，减少变量的数目提供参考。
- 4、分类速度快，当问题比较简单以及训练样本比较少时，利用专家知识十分有效。

# 贝叶斯分类器

- **先验概率：**  $P(\omega_i)$  **未获得观测数据之前类别的分布**
- **类条件概率：**  $P(\mathbf{x}|\omega_i)$  **观测数据在各类别种情况下的分布**
- **后验概率：**  $P(\omega_i|\mathbf{x})$  **X属于哪一类的概率**
- **贝叶斯公式**

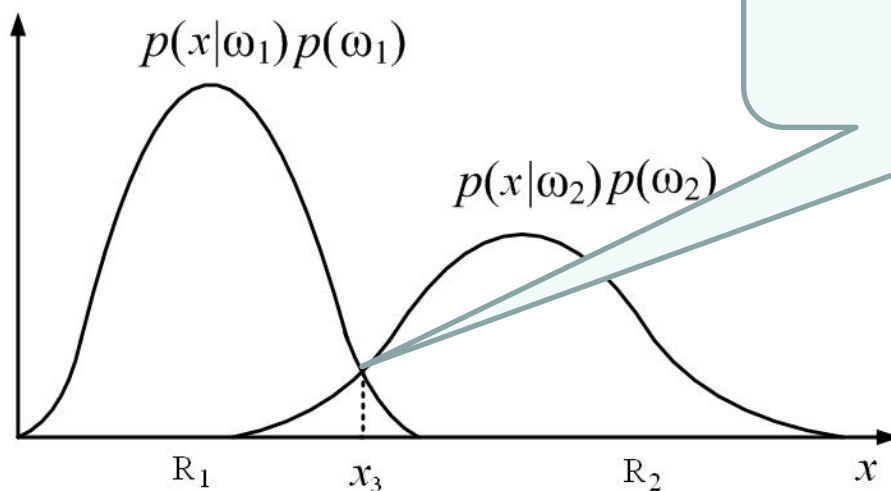
$$P(\omega_i|\mathbf{x}) = \frac{P(\mathbf{x}|\omega_i)P(\omega_i)}{P(\mathbf{x})}$$

**其中：**  $P(\mathbf{x}) = \sum_{i=1}^c P(\mathbf{x}|\omega_i)P(\omega_i)$

# 贝叶斯分类器

贝叶斯分类器是一类分类算法的总称，贝叶斯定理是这类算法的核心，因此统称为贝叶斯分类。贝叶斯决策论通过**相关概率已知**的情况下利用**误判损失**来选择最优的类别分类。

- 最小错误率准则
- 最小风险准则
- Neyman-Pearson准则
- 最小最大决策准则



基于最小错误率准则， $x_3$ 位置是分类界面

# SVM分类器

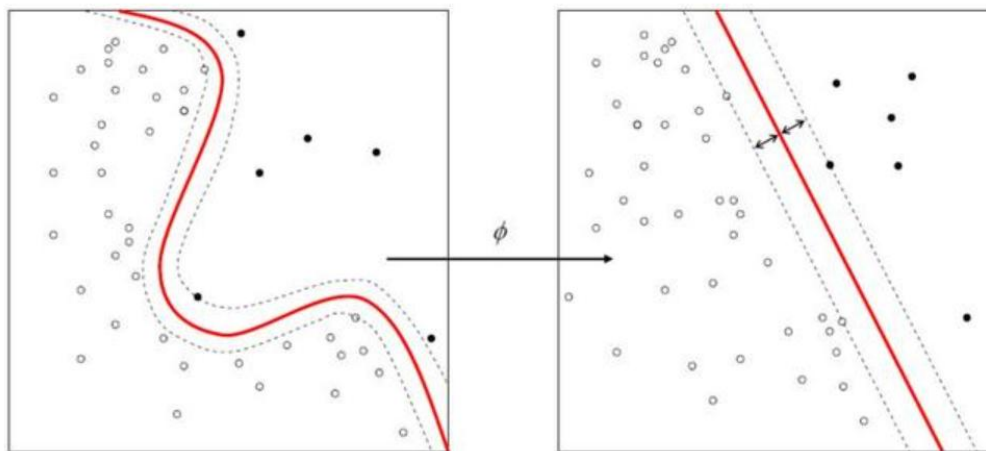
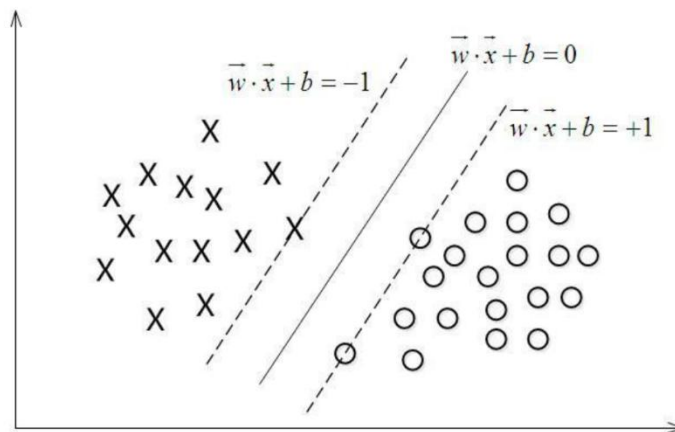
SVM本身是一个线性分类器，最开始解决的是线性可分的问题，然后拓展到非线性可分的问题，甚至扩展到非线性函数中去。

## 支撑向量

距离分类界面最近的一些样本点

## 核映射 —— 核函数

将低维空间线性不可分的数据投影到线性可分的高维空间中。





# 流形机器学习方法介绍

- 集成学习
- 增量学习
- 主动学习
- 迁移学习
- 强化学习
- 元学习

.....

我爱学习,  
学习使我快乐



# 集成学习 (ensemble learning)

集成学习就是组合多个弱监督模型以期得到一个更好更全面的强监督模型，集成学习潜在的思想是即便某一个弱分类器得到了错误的预测，其他的弱分类器也可以将错误纠正回来。

- **序列集成方法**，其中参与训练的基础学习器按照顺序生成（例如 AdaBoost）。序列方法的原理是利用基础学习器之间的依赖关系。通过对之前训练中错误标记的样本赋值较高的权重，可以提高整体的预测效果。
- **并行集成方法**，其中参与训练的基础学习器并行生成（例如 Random Forest）。并行方法的原理是利用基础学习器之间的独立性，通过平均可以显著降低错误。

# 增量学习

## Motivation :

随着时间的推移，数据分布和数据特征会发生变化。一段时间后，基于历史数据训练的分类模型可能不适用于一些新的数据。

重新训练模型

在原模型基础上  
学习更新模型

## Methods :

### 样本增量学习

需要确保在现有知识的情况下，通过新样本的增量学习来提取新知识，融合新旧知识以提高分类的准确性。

### 类别增量学习

识别新类，并将其加入现有类别的集合中，提升分类的准确性和智能。

### 特征增量学习

一些新的属性特征能够将分类提升到一个很大的程度，并提升分类准确率。

朴素贝叶斯，支持向量机，决策树，随机森林，人工神经网络，K-最近邻，模糊粗糙集理论等。

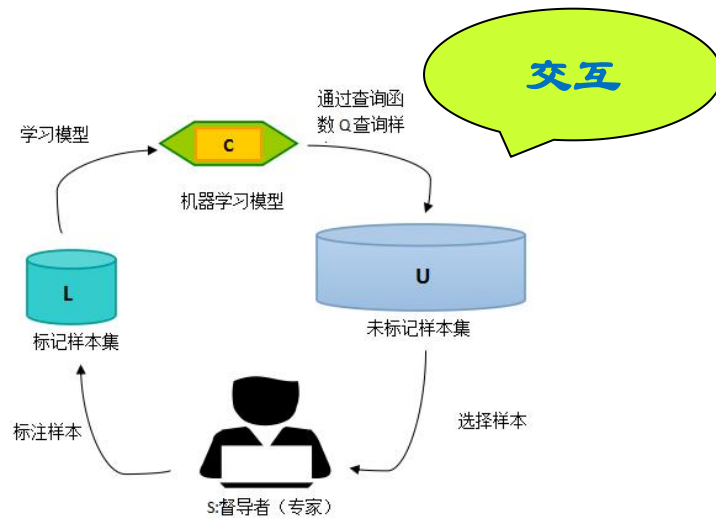
# 主动学习

## Motivation :

现实生活的很多场景中，标记样本的获取是比较困难的，这需要领域内的专家来进行人工标注，所花费的时间成本和经济成本都是很大的。使用较少的训练样本来获得性能较好的分类器呢？

## Methods :

主动学习通过一定的算法查询最有用的未标记样本，并交由专家进行标记，然后用查询到的样本训练分类模型来提高模型的精确度。



## 什么样的样本是有用的呢？

- 不确定性准则 (uncertainty)
- 差异性准则 (diversity)。

# 迁移学习

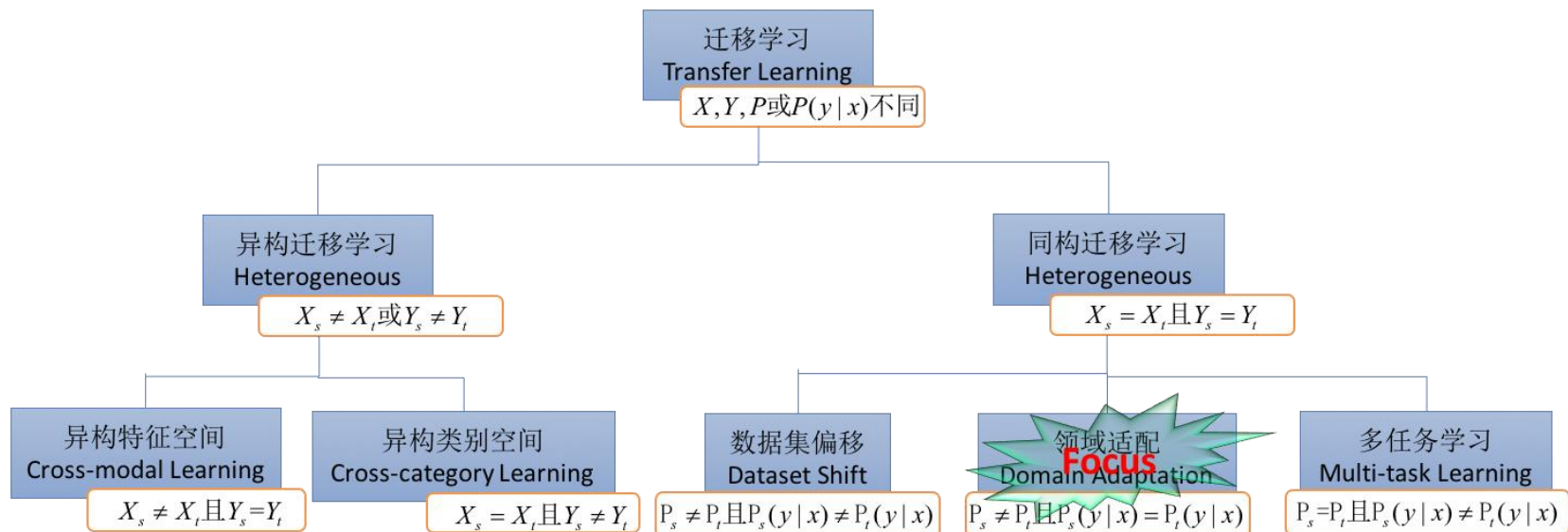
## Motivation :

受人类跨领域迁移知识能力的启发，迁移学习旨在利用来自相关领域（称为源域）的知识来提高学习性能或最小化目标域所需的标记示例的数量。

## Methods :

域(domain)  $D = \{\mathcal{X}, P(X)\}$   $\mathcal{X}$  表示特征空间,  $P(X)$  表示边际概率分布, 其中  $X = \{x_1, \dots, x_n\} \in \mathcal{X}$

任务(task)  $T = \{y, f(\cdot)\}$   $y$  表示标签空间,  $f(\cdot)$  表示目标预测函数



# 元学习

## Motivation :

人类能够以最少的信息同时获取完成多个任务的能力。元学习 (Meta Learning) 或者叫做“学会学习” (Learning to learn), 它是要“学会如何学习”, 即利用以往的知识经验来指导新任务的学习, 具有学会学习的能力。

## Methods :

**小样本元学习**: “小样本元学习”的理念是创建深层神经网络, 从最简单的数据集中学习, 比如模仿婴儿如何通过只看到一两张图片来学习识别物体。

**优化器元学习**: 优化器元学习模型的重点是学习如何优化神经网络从而更好地完成任务。这些模型通常包括一个神经网络, 该神经网络将不同的优化应用于另一个神经网络的超参数, 从而改进目标任务。

**度量元学习**: 度量元学习的目标是确定一个高效率学习的度量空间。

# 思考题

- 1. 什么是学习和机器学习？
- 2. 试述机器学习系统的基本结构并说明各部分的作用？
- 3. 试解释机械学习的模式？机械学习有那些重要问题需要加以研究？
- 4. 什么是类比学习？其推理和学习过程为何？