机器学习的范围

**机器学习跟模式识别，统计学习，数据挖掘，计算机视觉，语音识别，自然语言处理等领域有着很深的联系。**



**1模式识别**

**模式识别 = 机器学习；**模式识别源自 *工业界* ，而机器学习来自于 *计算机学科* 。不过，它们中的活动可以被视为同一个领域的两个方面，同时在过去的10年间，它们都有了长足的发展。

**2数据挖掘**

**数据挖掘 = 机器学习 + 数据库；大部分数据挖掘中的算法是机器学习的算法在数据库中的优化**；数据挖掘仅仅是一种 *思考方式* ，告诉我们应该尝试从数据中挖掘出知识，但不是每个数据都能挖掘出金子的；一个系统绝对不会因为上了一个数据挖掘模块就变得无所不能；恰恰相反，一个拥有 *数据挖掘思维* 的人员才是关键，而且还必须 *对数据有深刻的认识* ，这样才可能从数据中导出模式指引业务的改善。

**3统计学习**

**统计学习 近似等于 机器学习**；统计学习者重点关注的是统计模型的开发与优化，偏数学；而机器学习者更关注的是能够解决问题，偏实践，其重点研究学习算法在计算机上执行的效率与准确性的提升。

**4计算机视觉**

**计算机视觉 = 图像处理 + 机器学习。** *图像处理技术* 用于将图像处理成适合进入机器学习模型中的 *输入* ，*机器学习* 则负责从图像中识别出相关的 *模式* 。随着机器学习的新领域 *深度学习* 的发展，大大促进了计算机图像识别的效果，因此未来计算机视觉界的发展前景不可估量。

**5语音识别**

**语音识别 = 语音处理 + 机器学习。**语音识别就是 *音频处理技术* 与机器学习的结合。语音识别技术一般不会单独使用，一般会 *结合自然语言处理* 的相关技术。

**6自然语言处理**

**自然语言处理 = 文本处理 + 机器学习**。在自然语言 ***处理*** 技术中，大量使用了*编译原理*相关的技术，例如*词法分析*，*语法分析* 等等。除此之外，在 ***理解*** 这个层面，则使用了*语义理解，机器学习* 等技术。如何利用机器学习技术进行自然语言的的***深度理解***，一直是工业和学术界关注的焦点。

机器学习-大数据

大数据的核心是利用数据的价值，机器学习是利用数据价值的关键技术，对于大数据而言，机器学习是不可或缺的。

对于机器学习而言，越多的数据会越可能提升模型的精确性，同时，复杂的机器学习算法的计算时间也迫切需要分布式计算与内存计算这样的关键技术。

单从分析方法来看，大数据也包含以下四种分析方法：

1. **大数据，小分析**：即数据仓库领域的OLAP分析思路，即多维分析思想。  
   **2. 大数据，大分析**：这个代表的就是数据挖掘与机器学习分析法。  
   **3. 流式分析**：这个主要指的是事件驱动架构。  
   **4. 查询分析**：经典代表是NoSQL数据库。

也就是说，机器学习并不是是大数据下的唯一的分析方法。

机器学习模型的数据越多，机器学习的预测的效率就越好。

成功的机器学习应用不是拥有最好的算法，而是拥有最多的数据！

在大数据的时代，有好多优势促使机器学习能够应用更广泛。例如随着物联网和移动设备的发展，我们拥有的数据越来越多，种类也包括图片、文本、视频等非结构化数据，这使得机器学习模型可以获得越来越多的数据。同时大数据技术中的分布式计算Map-Reduce使得机器学习的速度越来越快，可以更方便的使用。种种优势使得在大数据时代，机器学习的优势可以得到最佳的发挥。

机器学习子类-深度学习

虽然深度学习这四字听起来颇为高大上，但其理念却非常简单，就是传统的神经网络发展到了多隐藏层的情况。

具有多个隐藏层的神经网络被称为深度神经网络，基于深度神经网络的学习研究称之为深度学习。

**1.多隐层的神经网络具有优异的特征学习能力，学习得到的特征对数据有更本质的刻画，从而有利于可视化或分类；**

**2.深度神经网络在训练上的难度，可以通过“逐层初始化” 来有效克服。**

机器学习父类-人工智能

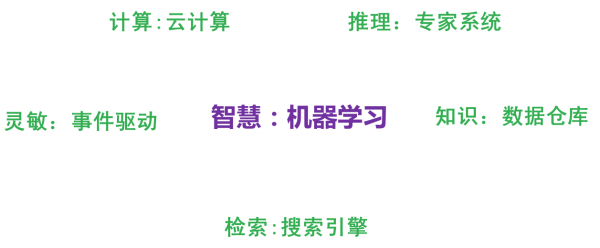
总结起来，人工智能的发展经历了如下若干阶段:

从早期的逻辑推理，到中期的专家系统；直到现在的机器学习（通过经验获取规律，指导人生与未来；即归纳与感悟能力）。

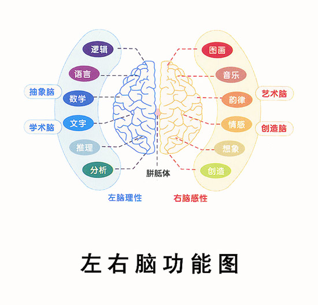
人工智能的发展可能不仅取决于机器学习，更取决于前面所介绍的深度学习，深度学习技术由于深度模拟了人类大脑的构成，在视觉识别与语音识别上显著性的突破了原有机器学习技术的界限，因此极有可能是真正实现人工智能梦想的关键技术。同时也要对人工智能系统制定严谨的规则（法律）；规律可以调整，但规则是不能改变的；有效的结合规律与规则的特点，说不定可以引导出一个合理的，可控的学习型人工智能。

机器学习-潜意识与直觉

其实直觉也是你在潜意识状态下思考经验后得出的规律。这与直接的训练模型，明意识的思考与归纳不同。如果计算机也拥有潜意识，那么会怎么样？譬如让计算机在工作的过程中，逐渐产生了自身的潜意识，于是甚至可以在你不需要告诉它做什么时它就会完成那件事。

******

机器学习-LR跃迁思考



*如果说：*

*左脑可以匹配为机器学习（或是更宽泛的说成是人工智能），右脑可以匹配为数据库；*

*左右脑发生的跃迁思考是不是可以看成机器学习与数据库之间的“关联规则”。而这种规则关键在于如何把抽象的概念用具象的事物表现出来。说不定，高级的记忆法会告诉我们这些。*