机器学习的发展：

1. 概念来源

人在面对问题时的解决途径：

1. 采用自己已有的知识（如来自DNA—进化）；
2. 寻求帮助（他人、文献资料等）；
3. 总结过去面对类似问题的经验。

传统计算机通过接受人的指令来完成相应的动作，我们希望计算机能够不再仅仅根据人的指令来执行命令，而是能够像人一样进行自主决策。故机器学习不再接受指令，而是接受数据（像人一样接受知识），从数据中提取规律，再利用此规律解决未知问题，即机器学习是一种让计算机利用数据进行工作的方法（不显式编程地赋予计算机能力）。（图灵测试提到了机器学习的可能性，阿瑟.萨缪尔的西洋跳棋程序战胜了当时全美排名第四的棋手进一步证实了这种可能性。）

机器学习算法与传统算法相比，不仅在结构上更加简单，而且泛化能力更强（相同的算法可以完成不同的事）。因此我们希望通过对机器学习的研究能够寻找到一个尽量简单且适用范围尽量大的算法，最理想的就是终极算法——只要提供的数据足量且合适，就能完成所有功能的单个算法。神经科学、进化论、物理学、统计学和计算机科学等通过对现实的观察，都对实现终极算法的可能性进行了论证，也因此诞生了机器学习的五大学派。

参考：

《终极算符》. 佩德罗多明戈斯

<https://www.cnblogs.com/subconscious/p/4107357.html#nine>

1. 发展阶段

1、“推理期”：

人们通过软件编程使计算机具有逻辑推理能力，计算机根据人的指令完成相应动作，如 “Logistic Theorist”，但仅具有逻辑推理能力并不能使机器智能。

2、“知识期”：

人们从自身的思维中提取规则，并将此规则教给计算机，计算机根据这些规则完成决策行为，如专家系统、“知识工程”，但知识是无穷无尽的，人们总结出的规律有限，故计算机的决策能力也非常有限。

3、“学习期”：

机器能够通过数据学习规律，数据量、数据采集能力和计算能力的巨大增加使机器学习成为可能。

参考：

《机器学习综述》. 赵晨阳

《机器学习》. 周志华

1. 主要学派
2. 连接学派：

人的学习和决策行为是通过大脑完成的，故要实现机器学习就要模拟大脑的结构和决策过程。大脑的神经网络由神经元组成，大脑通过神经元之间不同的连接结构和改变连接强度进行学习，人们希望通过模拟这一过程让机器具备学习能力。大脑的神经网络中含有的神经元数量非常庞大，导致连接学派的学习过程涉及大量的参数，而参数的设置缺乏严谨的理论指导；同时，大脑进行信息处理的过程包含很多生物和化学过程，这是算法很难模拟的，人们只能对其中的物理过程进行逆向解析。

主算法：反向传播（根据反馈不断对神经元之间的连接的权值进行微调，以降低误差）

2、符号学派（？？？）

符号主义起源于数学逻辑，认为所以的信息都可以简化为符号，对信息的处理可以等效为对相应符号的操作。“天下没有免费的午餐”，要想计算机具备我们所需要的能力，必须使其事先具有一定的知识。因此，符号主义的知识分为两部分：手动输入的知识和归纳得出的知识，前者供学习算法使用，后者用来加入知识库中。符号主义既不需要我们理解进化论和大脑的工作原理，也有效避免了贝叶斯主义的数学复杂性；但由于逆向演绎涉及到非常密集的计算，因此这种方法很难扩展到海量数据集中。

主算法：逆向演绎（归纳，学习每个规则之后，会排除该规则包含的正面例子，则下一个规则会尽可能地包含剩下的正面例子，以此类推，直到所有的例子都被包含在内）

1. 进化学派

所有形式的学习都源于自然选择。自然选择造就了人类，那么它就可以造就一切，人们要做的就是在计算机上进行模仿。自然界中的有性生殖包括父本和母本染色体之间的材料交换（染色体交叉），从而产生两条新的染色体。进化学派即对这一过程进行模拟，它为每一代两个适应力最强的个体进行配对，经过突变过程产生新的个体，每一个新个体都会反馈对环境的适应程度；重复这个过程，则每一代的个体都会比前一代的适应度更高。进化主义不只是像反向传播那样调整参数，它还创造大脑（相当于自然的角色）对参数进行微调。

主算法：遗传算法（输入一个适应度函数，当给定一个特定程序和某个设定目标后，适应度函数会给该程序打分，反映它与目标的契合度）

1. 贝叶斯学派

当你接收到新的论据时，它会改变你对某个假设的信任度：如果论据和假设一致，假设成立的概率上升，反之则下降，即所有掌握的知识具有不确定性。只有运用概率推理的方法，计算出各个事件发生的可能性，才能让我们在面对嘈杂、不完整甚至自相矛盾的信息时能够做出正确的决策。

主算法：贝叶斯定理及其衍生定理（通过新证据修正你之前相信的东西，得到后来相信的东西）

1. 类推学派

学习的关键就是要能够认识到不同场景的相似性，从而将此相似性应用到对其他场景的预测，其中的关键就是如何判断两个场景的相似程度。由于类推主义的基础是找到物体的相似性，随着数据维数的增加，相似性的概念就会逐渐失效，故类推主义受“维数灾难”的影响最大。但当给定一定的约束条件，即避免数据维数过高时，类推主义会致力于寻找该条件下的最优解。

主算法：支持向量机

参考：

《机器学习》. 周志华

《终极算符》. 佩德罗多明戈斯

<https://blog.csdn.net/wspba/article/details/77621112>

<https://www.zhihu.com/question/59800121/answer/184888043>

<https://blog.csdn.net/qq_39384184/article/details/81232818>

<https://blog.csdn.net/shandianke/article/details/79599120>

<https://blog.csdn.net/shandianke/article/details/80489787>