机器学习理解： 浅谈对自己对机器学习理解和自己认为的局限性

（全文手打）

机器学习是人工智能中十分重要的环节，不过让机器学习确实十分困难。就我认为，机器相比较人来说最大的优势就是其计算的速度及记忆性，如果一台机器的数据容量足够大，那么他就可以在较短的时间内重复出所有的可能性，而且不会忘记之前计算过的任意种可能性，基于此得出最佳方案。而因为不存在无限大的机器容量，上述的仅仅存在于理想之中。如何减少机器的运算量就是机器学习的核心。

目前的机器学习以数学的方式做到，包括概率论和逼近论等，这些数学的方法使机器可以在浩如烟海的模拟运算中总结出最普遍的规律，从而在遇到问题时可以极大减少运算量得到最佳的方案。就像阿尔法狗一样，其在几千万次的对局数据中通过每一步对胜利的影响来总结出对于每一步最佳的方案，比如说在一个50步走完的对局中，若胜利，则将每一步的位置记录并加分，输则扣分，这样做就能使得其在自己走过的各种比赛中总结出大概的规律。

不过我们知道围棋的可能性实在太多，19\*19的19\*19次方，不是及千万次的对局就能模拟完的，所以这里又引入了一个学习的思想，就是记忆棋谱。很多人可能也有这点体会，当我们背了即使篇新概念后，我们看到某些句子就能想都不想就直接填空，这里新概念起了一个培养语感的作用。同样的道理，给机器背诵“谱”，那么当其碰到类似的局面时，他能够首先“想”到棋谱上类似的局面，基于此来减少运算量。

可以想象，能够学习的机器在更多“经验”的作用下，其对于某项工程的能力一定会越来越大。但是，机器有其本身固有的缺陷，那就是他将模拟信号转换为了数字信号。虽说数字信号不容易失真，但就我理解来说，将模拟信号转化为数字信号其本身就是一种失真。我们都知道自然界是连续不断地，我们在某个地点向另一个地点走时，我们每一时刻都是在变化的，但是将其数据化后便不是这样的。可以把这两者比喻函数图像和数列图像，前者接连不断，而后者仅仅是孤立的点集，对于数列Xn，即使我们将n无限取小，甚至可以取小数，其终究还是孤立的点。同样的道理，无论我们如何优化模拟信号转换为数字信号，其终究会趋向一个极限值，而人类的思想则可以无限细化。就像国际象棋中的“象”，无论如何走，他终究只能在黑色或白色的格子中走。当我们让机器人从一个地点走向另一个地点时，他总是只能行走“距离子”的N倍，换句话说，如果我们要精细的让机器走到某个位置，比如行走到10.5个“距离子”是做不到的。还有一点就是机器在学习的过程中，其因为存在学习人类的先进知识这一学习过程，使得其对于一些常人想不到的方案缺乏涉及，就像阿尔法狗，如果我们第一步直接走围棋盘上的角落，可能就出现了其涉及的空缺，不过这样的一步确实是没有什么意义的，这里提及只是为了说明：机器学习存在饱和，他虽然能够战胜世界著名的高手，却可能会被名不见经传的民间高手所打败。这两点是我在了解机器学习时突然想到的，若有不对，希望学长学姐能够指出。