

1、引言

机器学习，尤其是基于人工神经网络的深度学习近年来得到了迅猛发展。一个自然思路是利用量子状态的叠加和量子算法的加速，来解决当前数据科学中数据量巨大，训练过程缓慢的困难。

从经典的二元概念分类可以将机器学习分为4类：

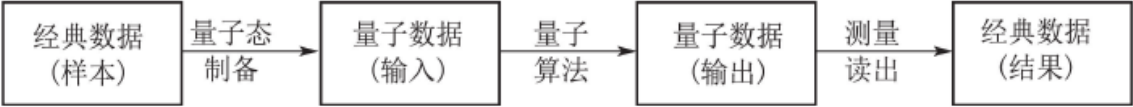
简称	算法类型	数据类型	应用举例
C-C	经典	经典	传统机器学习
C-Q	经典	量子	用机器学习解决量子多体物理问题; 量子优化控制
Q-C	量子	经典	量子支持向量机等
Q-Q	量子	量子	量子反馈控制

利用量子理论改进机器学习(Q - C)的方法大致可以分为两种：

- 1) 通过量子算法使某些在经典计算机上不可计算的问题变为可计算的，从而大幅降低机器学习算法的计算复杂度，如量子退火算法、Gibbs 采样等。
- 2) 量子理论的并行性等加速特点直接与某些机器学习算法深度结合，催生出一批全新的量子机器学习模型，如张量网络、概率图模等。

2、量子机器学习一般过程

量子机器学习的训练数据必须以某种可以为量子计算机识别的格式载入，经过量子机器学习算法处理以后形成输出，而此时的输出结果是量子格式的，流程如下图所示：



下面介绍数值型 Q - C 问题计算思路，计算机中所有数据都是有限字长的，假设某训练数据量 $x = [x_1 \cdot \cdot \cdot x_n]$ ，若每个元素都用长度为 m 的二进制表示，则整个向量空间需要 $n \cdot m$ 个比特位，原始的方式是用 $N \cdot M$ 个比特位存储这些数据，但这样耗费的资源无疑十分巨大。因此，利用量子系统的叠加特性，将这些数据以“叠加”的方式存储在 N 个比特上，用状态 $\sum_j c_j |x_j\rangle$ 表示。这样一来，数据存储所占用空间变小了，同时由

于支持并行计算，计算速度也大大提高，该过程称为量子加速。

量子计算有两种复杂度，分别为查询复杂度和门复杂度，前者指查询次数，后者指量子操作或门的数量。同时，量子计算机还可以识别出经典计算机无法在数据中识别的模式。

基于线性代数的量子机器学习是当下最流行也最成熟的量子机器学习算法。通过对高维向量执行矩阵运算，可以对各种数据分析和机器学习协议进行操作，而量子力学最擅长这一点。

该算法执行过程的关键步骤是， n 个量子比特的量子态是 2^n 维复向量空间中的一个向量，对量子位进行量子逻辑运算，是将相应的状态向量乘以 $2^n \times 2^n$ 个矩阵，这已经被证明可以执行常见的线性代数运算，算法时间复杂度仅为 n 的多项式时间。

3、量子数据集的转换

从经典数据到量子数据的转换，需要通过存储器来实现。

针对数值型数据，其实对应于一个量子化的数字逻辑电路，或者一个量子子程序，通过执行该子程序完成将寄存器的状态制备到训练数据对应的状态，这类存储器有一个专有名称—量子关联存储器 QAM。

量子态制备完毕并且也完成了计算后，需要对数据进行状态层析，这一步相对耗时。