



大数据与量子计算

王书浩^{①②}, 龙桂鲁^{①②③*}

① 清华大学物理系, 低维量子物理国家重点实验室, 北京 100084;

② 量子物质科学协同创新中心, 北京 100084;

③ 清华大学信息科学技术国家实验室(筹), 北京 100084

* 联系人, E-mail: gllong@tsinghua.edu.cn

2014-07-31 收稿, 2014-09-16 接受, 2015-01-21 网络版发表

国家自然科学基金(91221205, 11175094)和国家重点基础研究发展计划(2011CB9216002)资助

摘要 大数据技术的迅猛发展对计算效率提出了更高的要求. 由于量子系统的独特性质, 量子计算具有经典计算不具有的量子超并行计算能力, 能够对某些重要的经典算法进行加速. 人们发现, 除了大数分解算法, 量子计算的更多用途是对量子体系的仿真计算和在数据分析领域的应用. 近年来, 大数据和量子计算开始融合. 虽然实际使用的量子计算机尚未建成, 量子计算在大数据的应用在理论上已经取得了一些重要的进展. 实验上也有了一些发展. 本文首先介绍量子计算的基本原理和Grover量子算法. 随后以量子机器学习作为切入点, 介绍了量子计算在数据挖掘领域的应用.

关键词

大数据
量子计算
量子机器学习
量子信息处理

大数据给当今的社会发展带来了无数的机遇, 也带来了巨大的挑战^[1]. 随着数据库越来越庞大^[2,3], 所需要的数据分析算法越来越复杂^[4], 加之科学研究和某些商业领域对高精度的要求, 计算资源成为大数据技术发展一个瓶颈. 由于量子系统的独特性质, 量子计算具有经典计算不具有的量子超并行计算能力, 能够对某些重要的经典算法进行加速. 人们发现, 除了大数分解算法, 量子计算的更多用途是对量子体系的仿真计算和在数据分析领域的应用. 近年来, 大数据和量子计算开始融合. 虽然实际使用的量子计算机尚未建成, 量子计算在大数据的应用在理论上已经取得了一些重要的进展. 实验上, 量子搜索算法在一些小规模量子体系中已经得到演示, 而目前还有争议的D-Wave公司出品的超导计算装置, 其主要用途也是数据搜索^[5]. 本文首先介绍量子计算的基本原理和Grover量子算法. 随后以量子机器学习作为切入点, 介绍了量子计算在数据挖掘领域的应用.

1 量子计算

1.1 量子计算的特性

量子计算是量子力学与计算机科学结合的产物. 根据Moore定律, 芯片上集成的晶体管数目随时间呈指数增长. 当计算机的存储单元达到原子尺度时, 显著的量子效应将会严重影响其性能, 传统计算机发展遇到根本性的困难. 计算机科学的进一步发展须借助于新的原理和方法, 量子计算为解决这一问题提供了一个可能途径.

态叠加原理是量子力学的理论基础, 也是量子并行性的来源, 并贯穿于量子算法之中, 是量子算法相比于经典算法效率高的来源. 量子的并行性可以用一个简单的例子(图1)来理解, 我们给定一个酉操作 $U_f: |x, y\rangle \rightarrow |x, y \oplus f(x)\rangle$, 其中 $f(x): \{0, 1\} \rightarrow \{0, 1\}$. 对于如图1所示的初态, 经过有 U_f 操作之后的结果为

引用格式: 王书浩, 龙桂鲁. 大数据与量子计算. 科学通报, 2015, 60: 499–508

Wang S H, Long G L. Big data and quantum computation (in Chinese). Chin Sci Bull, 2015, 60: 499–508, doi: 10.1360/N972014-00803

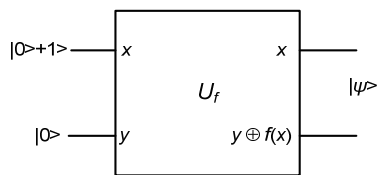


图1 量子并行计算线路示意图, 其中输入态的归一化常数已被忽略
Figure 1 Scheme for quantum parallel computation, where the normalization constants have been ignored

$$|\psi\rangle = \frac{|0, f(0)\rangle + |1, f(1)\rangle}{\sqrt{2}}, \quad (1)$$

可以看到, 经过一个 U_f 操作, 我们便获得了同时包含 $f(0)$ 和 $f(1)$ 的量子态 $|\psi\rangle$. 当有多个量子比特的时候, 这种量子并行计算的能力就更加明显.

除了叠加特性, 量子系统还具有一个更加奇妙的量子纠缠特性. 纠缠量子态指的是其波函数不能写成子系统波函数直积的量子态. 例如, Bell态可以写作

$$|\text{BELL}\rangle = \frac{1}{\sqrt{2}}|0_A 0_B\rangle + |1_A 1_B\rangle, \quad (2)$$

其中 A 和 B 表示两个粒子, $|0\rangle$ 和 $|1\rangle$ 是 σ_z (具体表达式见式(3))的本征态, 分别表示自旋向上和自旋向下. 若我们使用 σ_z 算符对 A 粒子进行测量, 当我们测得自旋向上时, 那么 B 粒子的自旋也一定是向上的; 反之亦然. 因此在这个过程中, A , B 粒子的自旋状态是纠缠在一起的, 这种现象叫做量子纠缠. 研究表明, 拥有量子纠缠的系统可以携带更多的信息.

量子计算的核心是量子算法. 与经典算法类似, 量子算法是通过量子逻辑门来实现的. 几个常见的单量子比特门为

$$\begin{aligned} U_x &= \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ U_y &= -i\sigma_y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \\ U_z &= \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ H &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \end{aligned} \quad (3)$$

其中 H 叫做Hadamard门.

量子控制非门(CNOT门), 即与或门, 是最常用的两比特量子门:

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (4)$$

量子算法是一系列量子逻辑门的顺序组合, 是一个实际问题从数据输入(量子态的初始化)到数据输出(即测量)完整的解决方案. 量子体系模拟算法主要是根据具体的量子体系, 模拟其动力学演化的性质, 偏重于物理. 到目前, 除了对量子体系模拟, 根据文献[6], 主要的量子算法可归类于量子搜索算法和Fourier变换两大类算法, 其他方法都与这两个算法有关系, 如Shor^[7]的大数分解算法、Deutsch-Jozsa算法^[8]等. Shor算法展示了量子计算机的优势, 它可以在多项式步骤里分解一个大数, 而相应的经典算法则需要指数多的步骤. 但是随着大数据技术的发展, 以Grover搜索算法为代表的量子搜索算法很可能是量子计算最主要和最重要的量子算法.

1.2 Grover搜索算法及其改进

1996年, Grover^[9]提出了针对非结构化数据时间复杂度为 $O(\sqrt{N})$ 的量子搜索算法(经典搜索算法的时间复杂度为 $O(N)$). 假设非结构化数据库包含 n 个量子比特, 总共承载有 $N=2^n$ 个量子态 $|i\rangle$ ($i=1, 2, \dots, N$), Grover搜索算法主要解决的问题是以尽可能大的概率搜索目标态 $|\tau\rangle$, 使得查询函数 $f(\tau)=1$ (对于其他量子态, 有 $f(i)=0, i \neq \tau$).

Grover搜索算法中的数据库为所有计算基矢量子态的线性叠加, 如果初始时刻 n 个量子比特的寄存器处于 $|0\rangle=|0\rangle^{\otimes n}$ 态, 实施Hadamard操作 $W=H^{\otimes n}$ 后, 就可以得到均匀叠加的数据库初始量子态

$$|\psi_0\rangle = W|0\rangle = \frac{1}{\sqrt{N}} \sum_i |i\rangle = \sin \beta |\tau\rangle + \cos \beta |c\rangle, \quad (5)$$

其中

$$|c\rangle = \sqrt{\frac{1}{N-1}} \sum_{i \neq \tau} |i\rangle, \beta = \arcsin\left(\sqrt{\frac{1}{N}}\right). \quad (6)$$

Grover量子算法是进行 $O(\sqrt{N})$ 次Grover迭代后, 然后测量量子态的状态以获取目标量子态 $|\tau\rangle$. 其中Grover迭代包括4个步骤,

(1) 保持其他态不变, 反转 $|\tau\rangle$ 的相位, 相应的操作可以写成 $I_\tau = I - 2|\tau\rangle\langle\tau|$;

(2) 对系统作用 $W = H^{\otimes n}$;

(3) 反转 $|0\rangle$ 态, 其他态的相位不变, 即 $I_0 = I - 2|0\rangle\langle 0|$;

(4) 再次对系统作用 $W = H^{\otimes n}$.

Grover迭代能够整体表述为

$$G = WI_0WI_\tau = (2|\psi_0\rangle\langle\psi_0| - I)(I - 2|\tau\rangle\langle\tau|). \quad (7)$$

在以 $|\tau\rangle$ 和 $|c\rangle$ 为基矢的二维 Hilbert 空间中, Grover迭代能够表示为

$$G = \begin{bmatrix} \cos 2\beta & \sin 2\beta \\ -\sin 2\beta & \cos 2\beta \end{bmatrix}. \quad (8)$$

可以将Grover搜索算法用几何图像描绘出来^[6]. 如图2所示, 每次Grover迭代可以看作是在二维 Hilbert空间沿逆时针方向旋转 2β .

虽然Grover算法的搜索成功率很大, 但是其成功率一般不是100%, 只有在数据库中有4个数据时, 其成功率才是100%. 在大数据应用时, 由于数据数目很大, Grover搜索的成功率是很大的. 但是在多目标态的时候, 随着目标态的数量增多, 成功率下降. 而且在一些量子搜索策略中要不断地缩减数据库, Grover算法的这一缺陷就非常突出和经常出现.

龙桂鲁和合作者^[10,11]证明, 在一般的相位转动下, 2个转动的角度(即2个相位)须满足相位匹配条件. 一般相位转动下的量子搜索操作为

$$\begin{aligned} G &= -UR_0U^{-1}R_\tau, \\ R_0 &= I + (e^{i\theta} - 1)|0\rangle\langle 0|, \\ R_\tau &= I + (e^{i\phi} - 1)\sum_k |\tau_k\rangle\langle\tau_k|, \end{aligned} \quad (9)$$

其中 θ 和 ϕ 是2个相位转角. 相位匹配条件与数据库

的形式有关, 对于初始态均匀分布数据库, 相位匹配条件为 $\theta = \phi$. 而对于一般的情形形式如下的数据库:

$$|\psi_0\rangle = \sin\theta_0|\tau\rangle + \cos\theta_0 e^{i\delta}|c\rangle, \quad (10)$$

相位匹配条件是

$$\begin{aligned} &\tan\frac{\theta}{2}[\cos 2\beta + \tan\theta_0 \cos\delta \sin 2\beta] \\ &= \tan\frac{\phi}{2}\left[1 - \tan\theta_0 \sin\delta \sin 2\beta \tan\frac{\theta}{2}\right], \end{aligned} \quad (11)$$

其中 $\sin\beta = \langle\tau|U|0\rangle$, 对于Hadamard算符有 $\sin\beta = 1/\sqrt{N}$. 相位匹配条件是量子搜索算法成功的必要条件. 利用相位匹配条件, 龙桂鲁等人^[12,13]给出了成功率为100%的量子搜索算法, 改进了Grover搜索算法的成功率. 改进的Grover搜索算法的关键是将两个180°相位转变改为

$$\theta = \phi = 2\arcsin\left(\sqrt{\frac{N}{M}} \sin\left(\frac{\pi}{4J+6}\right)\right), \quad (12)$$

其中

$$J \geq J_{op} = \frac{\left(\frac{\pi}{2}\right) - \beta}{2\beta}, \beta = \arcsin\left(\sqrt{\frac{M}{N}}\right), \quad (13)$$

M 是标记态的数目, N 是数据库的数据个数.

2 大数据与量子计算的结合——量子机器学习

在大数据的研究中, MapReduce为数据挖掘的研究提供了优秀的并行计算模型. 除了技术层面的实现, 还需要设计优秀的数据挖掘算法. 数据挖掘过程一般是指从大量的数据中通过算法搜索其中的信息, 可以通过机器学习等诸多方法来实现. 大多数数据挖掘技术都来自机器学习领域, 本文将专注于使用机器学习的手段来进行数据挖掘的方法.

随着机器学习技术的发展, 人们意识到很多传统的算法是可以通过量子计算机进行加速的. 通过恰当地采用量子算法, 同时利用量子系统的叠加性和量子纠缠, 可以对一些传统的机器学习算法来进行优化, 从而大幅提高计算的性能^[14-17]. 同时, 对于某些特定的问题, 可以由量子力学得到启发, 以优化算法的性能. 致力于数据挖掘的机器学习算法, 大致包括分类(classification)、聚类(clustering)以及神经网络(neural network)等. 接下来我们将对这些领域作简单的介绍, 并展示量子机器学习各个领域的发展.

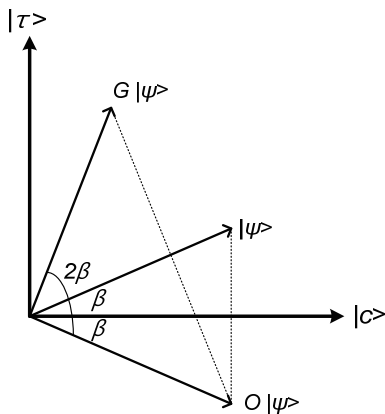


图2 Grover搜索算法的几何示意图^[6]

Figure 2 Geometrical interpretation of Grover algorithm^[6]

2.1 数据分类与量子计算

分类就是确定目标对象属于哪个预定的目标类. 详细一些说, 分类就是通过学习得到一个目标函数 f , 把每一个属性集 x 映射到一个预先定义的目标类标号 y 上去^[18]. 在本节中, 我们将介绍决策树(decision tree)分类算法、支持向量机(support vector machine)方法的量子版本, 并简要介绍一些其他的分类方法. 关于量子神经网络, 我们将在后面的小节中单独加以讨论.

在介绍量子分类问题之前, 我们先来介绍一下经典的决策树分类算法和支持向量机方法.

如图3所示, 基于信息熵的决策树分类方法可以概述如下^[19]:

- (1) 树从训练样本单个节点开始构建;
- (2) 若样本归属同一类, 则该节点成为树叶(标记为该类); 反之, 则使用信息熵作为启发信息选择更优属性;
- (3) 针对择优出的属性, 创建与其值对应的分支, 将样本归入对应的分支;
- (4) 递归以上过程, 出现在节点的属性将被该节点的所有后代所忽略.

若以下条件之一被满足, 则上述过程中断: (1) 结点仅包含同一类的样本; (2) 属性已被穷尽(此时采用多数表决); (3) 没有样本对应该属性(此时创建的树叶服从样本的多数类).

基于统计学的支持向量机方法是一种新型的数据分类技术, 经过几十年的发展, 支持向量机具有广泛的应用^[18]. 接下来我们介绍样本可分时(不可分的情况可以通过加入松弛变量来解决)线性支持向量机的工作原理. 考虑 N 个训练样本的二元分类问题, 每个样本可以表示为一个二元组 $(\vec{x}_i, y_i), i=1, \dots, N$, 令 $y_i \in \{-1, 1\}$ 作为类标号. 线性支持向量机的目标, 就是寻找拥有最大边缘(即距离两类数据点最远)的

超平面, 它满足:

$$y = \begin{cases} 1, \vec{w} \cdot \vec{z} + b > 0, \\ -1, \vec{w} \cdot \vec{z} + b < 0, \end{cases} \quad (14)$$

其中 $\vec{w} = (w_1, \dots, w_N)^T$, b 为模型的参数, $\vec{z}$ 为测试样本^[18].

根据Karush-Kuhn-Tucker(KKT)最优化条件, 应满足在约束 $\sum_{i=1}^M \alpha_i = 0$ 和 $y_i \alpha_i > 0$ 的条件下, 使得

$$L(\vec{w}) = \sum_{i=1}^N y_i \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j K_{ij}, \quad (15)$$

取最大值, 其中 $\vec{\alpha} = (\alpha_1, \dots, \alpha_M)^T$ 为Lagrange乘子, K 为核矩阵(kernel matrix), 有 $K_{ij} = k(\vec{x}_i, \vec{x}_j) = \vec{x}_i \cdot \vec{x}_j$.

由此我们可以得到下列分类器:

$$y(\vec{x}) = \text{sign} \left(\sum_{i=1}^M \alpha_i k(\vec{x}_i, \vec{x}) + b \right), \quad (16)$$

同时有

$$\vec{w} = \sum_{i=1}^M \alpha_i \vec{x}_i, b = y_i - \vec{w} \cdot \vec{x}_i. \quad (17)$$

作为经典的对应, 假定量子分类问题中的数据元素均为量子态(即 $|x\rangle, |y\rangle$)^[20]. 对于决策树的分类方法, 我们可以直接将经典信息熵替换为量子熵(von Neumann熵). 从而, 通过使用量子熵作为启发信息, 同时运用Grover搜索算法完成节点的搜索, 便可以得到量子决策树分类方法^[21].

Anguita等人^[22,23]指出, 支持向量机的训练过程也是可以通过Grover搜索算法来进行优化的. 最近Rebentrost等人^[24]提出了一种用另外一种量子计算的手段来进行支持向量机训练的方法, 能够将核矩阵计算的复杂度由多项式降低到对数. 在他们所提出的方法中, 首先将每条数据映射到量子态上:

$$|\vec{x}_i\rangle = \frac{1}{|\vec{x}_i|} \sum_{k=1}^N (\vec{x}_i)_k |k\rangle = \frac{1}{|\vec{x}_i|} ((\vec{x}_i)_1, \dots, (\vec{x}_i)_N)^T. \quad (18)$$

进而将数据的令牌(Oracle)定义为

$$|O\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^M |i\rangle. \quad (19)$$

我们便可以得到量子态

$$|\chi\rangle = \frac{1}{\sqrt{N_\chi}} \sum_{i=1}^M |\vec{x}_i\rangle |i\rangle, \quad (20)$$

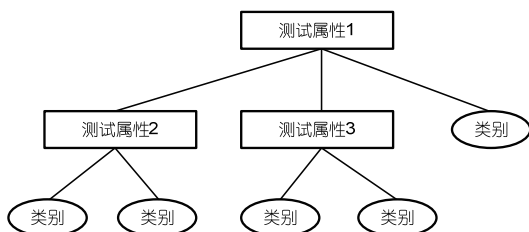


图3 决策树分类方法示意图^[19]

Figure 3 Scheme for decision tree classification technique^[19]

其中 $N_x = \sum_{i=1}^M |\bar{x}_i|^2$ 为归一化常数. 可以看到, 如果我们对 $|\chi\rangle$ 的密度矩阵求偏迹:

$$\text{Tr}_2(|\chi\rangle\langle\chi|) = \frac{1}{N_x} \sum_{i,j=1}^M \langle \bar{x}_i | \bar{x}_j \rangle |\bar{x}_i\rangle\langle\bar{x}_j| = \frac{K}{\text{Tr}K}, \quad (21)$$

就得到了归一化的核矩阵. 上述过程在量子计算机的实现方法详见文献[24].

除了上述方法, 人们提出还可以使用绝热量子计算和基于测量的量子学习算法来进行数据的分类(后者主要是针对两比特量子态的分类^[25,26]). 将绝热量子计算与数据分类联系在一起的初衷, 是希望用 D-Wave 量子计算机来进行分类问题的优化^[27-29]. 由于 D-Wave 量子计算机是通过物理退火过程, 对一个 Ising 函数进行优化, 以获取其最小值的, 因此要求系统的 Hamiltonian 最多只能包含到二次项^[27]. 模拟表明, 使用绝热量子计算能够使机器学习的效率得到较大的提升^[27,28,30].

2.2 量子聚类分析

将一组对象, 根据它们之间的相似程度, 将相似的对象构成一组, 进而分为若干组, 这一过程就称为聚类^[19]. 作为数据挖掘的一项功能, 聚类分析可以用来分析数据的分布、挖掘各数据类的特征、确定有价值的数据库等等. 由于在当今“大数据”的背景下数据库所包含的数据量越来越大, 聚类分析已成为数据挖掘研究中一个非常活跃的研究课题.

将聚类分析的问题用量子的方法进行优化, 主要采用两种办法. 第一种方法是直接在经典的聚类方法加入 Grover 搜索算法来改进^[31]; 第二种方法要利用统计力学^[32]或者量子力学的原理, 优化算法对数据的发掘性能. 对于前者, 举个最简单的例子, 假设我们要寻找数据集 $\{x_i, i=1, \dots, M\}$ 中距离最远的两点, 加入 Grover 搜索算法后的程序为

- (1) 随机选择两个下标 i 和 j ;
- (2) 将 $d_{\max}$ 定义为 $\text{Dist}(x_i, x_j)$, 其中 $\text{Dist}(x_i, x_j)$ 表示 x_i 与 x_j 之间的距离;
- (3) 利用 Grover 算法, 搜索新的下标 i 和 j , 使得 $\text{Dist}(x_i, x_j) > d_{\max}$, 并令 $d_{\max} = \text{Dist}(x_i, x_j)$;
- (4) 不断重复搜索, 直到没有更大的 $\text{Dist}(x_i, x_j)$ 出现, 此时的 i 和 j 即为所求.

上述方法可以直接应用到所有基于数据点距离

的经典聚类算法(特别是 K 均值算法)中^[31,33]. 由于 Grover 搜索算法较经典搜索算法具有更高的效率, 因此上述步骤可以提高聚类分析的速度.

借助于量子力学原理, Horn 等人^[34-36]提出了一种非常新颖的聚类分析算法, 并称之为(动力学)量子聚类分析. 在传统的尺度空间算法(scale-space algorithm)中, 人们提出可以使用 Parzen 窗估算量(Parzen-window estimator)的极大值来确定聚类中心. Parzen 窗估算量定义为 d 维 Euclidean 空间中的一个高斯函数:

$$\psi(\vec{x}) = \sum_i e^{-(\vec{x}-\vec{x}_i)^2/2\sigma^2}, \quad (22)$$

式中忽略了归一化常数, 其中 $\vec{x}_i$ 为数据点.

我们接下来将 ψ 视为 Schrödinger 方程的本征态, 即有

$$H\psi(\vec{x}) = \left(-\frac{\sigma^2}{2} \nabla^2 + V(\vec{x}) \right) \psi(\vec{x}) = E\psi(\vec{x}), \quad (23)$$

其中我们改写了 Schrödinger 方程中的 H 和 V , 使得方程中只有一个自由变量 σ . 容易得到

$$V(\vec{x}) = E + \frac{\frac{\sigma^2}{2} \nabla^2 \psi(\vec{x})}{\psi(\vec{x})}, \quad (24)$$

其中要求 $\min V = 0$, 即

$$V(\vec{x}) = E + \frac{\frac{\sigma^2}{2} \nabla^2 \psi(\vec{x})}{\psi(\vec{x})}. \quad (25)$$

通过对不同种类的数据集进行聚类分析可以发现, 相对于 $\psi(\vec{x})$, 通过找寻 $V(\vec{x})$ 的极小值点, 我们能够找到更多的聚类中心. 同时, 通过调节参数 σ , 我们能够进一步发掘数据中的聚类信息^[35]. 因此量子聚类分析方法在数据挖掘方面具有更大的优势.

2.3 量子神经网络

人工神经网络的灵感来源于生物学层面的神经系统, 人脑神经细胞内的神经元通过一个叫做突触的结构错综复杂地交织在一起, 这样神经脉冲就能够在神经元之间进行传递. 神经学家发现, 人脑通过改变突触之间连接的强度来进行学习. 如图4所示, 人工神经网络学习了这种感知技术, 其感知器由输入和输出节点构成, 这些节点被称为神经元或者感知器^[18].

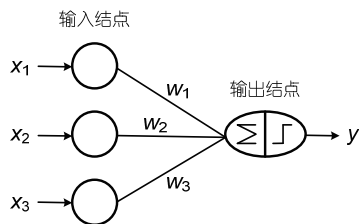


图4 神经元示意图, 其中 $x_i = \pm 1, y = \pm 1, i = 1, 2, 3$ ^[18]

Figure 4 Illustration of the artificial neuron, where $x_i = \pm 1, y = \pm 1, i = 1, 2, 3$ ^[18]

考虑线性神经元模型:

$$y = \text{sign}(\bar{w} \cdot \bar{x}), \quad (26)$$

其中 $\bar{w} = (w_0, \dots, w_N)^T$ 为权重向量 (N 为数据总条数, $w_0 = -t$, t 为阈值). 对于给定的数据集 $D = (\bar{x}_i, y_i)$, $i = 0, \dots, N$ ($x_0 = 1$), 神经元的训练步骤为: 首先将权重向量分量初始化为随机值, 记为 $\bar{w}^{(0)}$. 然后对于每条数据 $\bar{x}_i$, 根据权重向量计算预测的 $\hat{y}_i^{(k)}$, 根据结果更新权重向量分量:

$$w_j^{k+1} = w_j^k + \lambda(y_i - \hat{y}_i^{(k)})(\bar{x}_i)_j, \quad (27)$$

其中 $\hat{y}_i^{(k)}$ 表示迭代次数, λ 为学习速率. 当程序的误差小于某个规定值时, 停止迭代, 得到最终的权重向量.

相比于神经元, 人工神经网络具有非常复杂的结构. 如图5所示, 人工神经网络可以分成输入层、隐藏层和输出层. 除了符号函数外, 人工神经网络存在很多其他的激活函数, 其中最著名的即Sigmoid函数: $g(x) = 1/(1 + e^{-x})$. 由于隐藏层的存在, 神经元的优化方法不再适用于复杂的人工神经网络, 由此我们需要设计新的优化算法.

定义最小化误差平方和函数

$$E(\bar{w}) = \frac{1}{2} \sum_i (y_i - \hat{y}_i)^2, \quad (28)$$

此时梯度下降给出的更新公式为^[18]

$$w_i \leftarrow w_i - \lambda \frac{\partial E(\bar{w})}{\partial w_i}. \quad (29)$$

由于 $\frac{\partial E(\bar{w})}{\partial w_i}$ 在实际运用中难以估计, 人们提出了Back-propagation的方法, 其基本思想是利用第 $k+1$ 层的误差来估计第 k 层的误差.

从20世纪末开始, 量子神经网络受到了广泛的关注, 研究人员提出了一些非常新颖的实现方案,

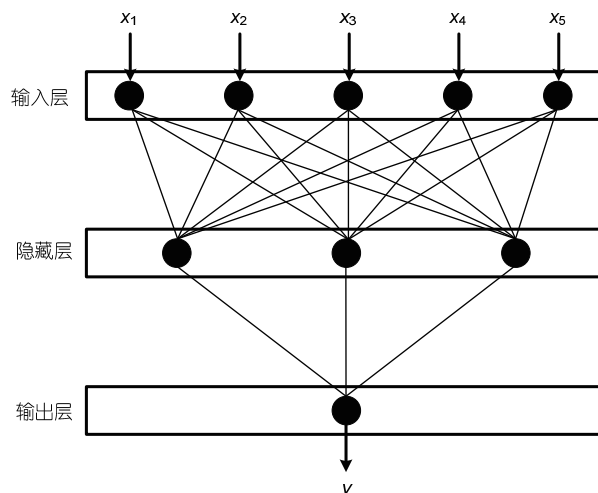


图5 人工神经网络示意图^[18]

Figure 5 Illustration of the artificial neural network^[18]

包括测量模型^[37-41]、量子点模型^[42-45]、量子神经线路模型^[46-50]、量子感知器模型^[51-54]、量子联想记忆模型^[55-61]、量子比特神经网络模型^[62]、量子行走模型^[63]以及一些其他模型^[64-69]. 利用量子神经网络能够完成语音识别、图像压缩等任务^[70-73].

作为一个例子, 接下来我们简要介绍一下量子比特神经网络模型.

Kouda等人指出, 神经元的激活与非激活状态, 分别可以对应到两个量子态 $|1\rangle$ 和 $|0\rangle$. 因此对于任意的量子比特神经元, 其量子态便为二者的叠加

$$|\psi\rangle = \alpha|1\rangle + \beta|0\rangle, \quad (30)$$

其中归一化条件要求 $|\alpha|^2 + |\beta|^2 = 1$. 量子比特的量子态还可以写作 $f(\phi) = e^{i\phi} = \cos\phi + i\sin\phi$, 其中 ϕ 表示 $|1\rangle$ 和 $|0\rangle$ 的相对相位. 可以看到, $f(\phi)$ 满足

$$\begin{aligned} f(\phi_1 + \phi_2) &= f(\phi_1) \cdot f(\phi_2), \\ f\left(\frac{\pi}{2}\gamma - \phi\right) &= \begin{cases} \sin\phi + i\cos\phi, & \gamma = 1, \\ \cos\phi - i\sin\phi, & \gamma = 0, \end{cases} \end{aligned} \quad (31)$$

所以当 $\gamma=1$ 时, 对应相位的翻转; 当 $\gamma=0$ 时, 相位不发生翻转. 由此可见, 通过控制 γ , 我们便可以实现可控的相位门.

将信息加载在相位上, 我们便可以得到如图6所示的量子比特神经元. 其中各参数的定义为

$$\begin{aligned} u &= \sum_i f(\theta_i) \cdot x_i - f(\eta) = \sum_i f(\theta_i) \cdot f(y_i) - f(\eta), \\ y &= \frac{\pi}{2} \cdot g(\delta) - \arg(u), \\ z &= f(y), \end{aligned} \quad (32)$$

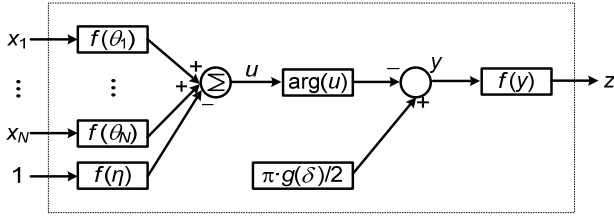
图6 量子比特神经元示意图^[62]

Figure 6 Illustration of the qubit neuron

上式中 g 为 Sigmoid 函数, 值域为 $[0, 1]$. y_i 为 x_i 量子态的相位, $\arg(u) = \arctan(\text{Im}(u) / \text{Re}(u))$ ^[62].

量子比特神经网络需要首先将输入态转换成相位, 即 $f(\pi/2) \cdot x_i$, 然后通过隐藏层后进入输入层进行输出, 最后测得量子态 $|1\rangle$ 的概率为 $|\text{Im}(y)|^2$ (见图6). 需要指出的是, 量子比特神经网络的优化过程采用的是量子版本的 Back-propagation 方法. 各个优化参数的更新公式为

$$\begin{aligned}\theta_i &\leftarrow \theta_i - \lambda \frac{\partial E}{\partial \theta_i}, \\ \eta &\leftarrow \eta - \lambda \frac{\partial E}{\partial \eta}, \\ \delta &\leftarrow \delta - \lambda \frac{\partial E}{\partial \delta},\end{aligned}\quad (33)$$

其中 E 为总的误差平方和函数.

2.4 其他量子机器学习算法

首先介绍量子模式匹配算法. 模式匹配是将一个给定的模式, 根据其特征划入指定的分类中去. 模式匹配技术在图像、语音、信号识别等领域有着诸多的应用. 虽然对于量子算法, 测量无法获取量子态的全部信息, 但是由于量子系统的并行性, 我们依然能够采取合理的方式有效获取模式的结构信息, 从而提高算法的效率. 经典模式匹配算法中的关联函数可以用 Fourier 变换来计算, 人们自然想到最直接的推广是将其改为量子快速 Fourier 变换. 这样的话相对于经典算法, 我们将会有较大的概率获得效率的提升^[74]. 另外一种方法是把不同的模式类定义为模式量子态(template state), 通过比较未知量子态与模式量子态的匹配程度(即量子态的交叠程度), 实现未

知模式的分类^[75,76].

下面介绍解决 NOT 的 k 次方根的量子算法. 假设输入态为 Boolean 值(即 1 和 0), 我们将 NOT 的 k 次方根(简记为 $\sqrt[k]{\text{NOT}}$)操作作用到输入态 $n \times k$ 次, 如果 n 为偶数, 则结果为输入态本身; 反之, 则结果与输入态相反. 不同于经典算法, 采用量子算法并利用 Gaussian 随机行走, 仅需要一个量子比特就能够进行任意 $\sqrt[k]{\text{NOT}}$ 的学习, 而且学习的速度是经典算法的 $\log 2k$ 倍^[77].

除了上面介绍的工作之外, 量子机器学习算法还包括通过量子手段来研究概率近似正确(probably approximately correct, PAC)模型等^[78,79].

特别应该指出的是, 除了量子计算对机器学习的贡献, 机器学习方面的理论也在促进着量子信息领域的发展. 在量子控制方面, 机器学习可以让系统获得更加精细的控制能力^[80-83], 也能够增加人们对量子系统的学习能力^[84-89].

3 总结与展望

本文从大数据与量子计算的基本原理出发, 介绍了量子机器学习领域的发展. 可以发现, 通过量子计算机来处理经典算法的加速问题通常有以下3种方式:

(1) 直接使用量子算法代替经典算法实现传统机器学习算法的加速, 比如使用 Grover 搜索算法进行决策树节点的搜寻以及寻找数据集中距离最远的两个点等;

(2) 使用量子系统的独特性质设计新的量子机器学习算法, 比如用量子手段进行支持向量机的训练、利用 Schrödinger 方程进行动力学量子聚类分析等;

(3) 基于已有的量子体系, 设计适用于该体系的量子优化算法, 比如专门为 D-Wave 量子计算机设计的量子退火算法.

随着大数据和量子信息领域的蓬勃发展, 我们期待量子机器学习领域能涌现出更多的工作. 当然, 除了理论研究, 我们也期待更多的实验工作, 将对未来的构想转变为现实.

参考文献

- 1 Mayer-Schönberger V, Cukier K. Big Data: A Revolution That Will Transform How We Live, Work and Think. New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013
- 2 Ghemawat S, Gobioff H, Leung S T. The Google file system. In: Proceedings of the 17th ACM symposium on Operating Systems Principles, New York, 2003, 37: 29–43
- 3 Chang F, Dean J, Ghemawat S, et al. Bigtable: A distributed storage system for structured data. ACM T Comput Sys, 2008, 26: 4
- 4 Dean J, Sanjay G. MapReduce: Simplified data processing on large clusters. Commun ACM, 2008, 51: 107–113
- 5 Grant A. Quantum computer fails challenge: D-Wave two shows no speed gain over traditional machine. Sci News, 2014, 186: 6
- 6 Nielsen M A, Chuang I L. Quantum Computation and Quantum Information. Cambridge: Cambridge University Press, 2010
- 7 Shor P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J Comput, 1997, 26: 1484–1509
- 8 Deutsch D, Jozsa R. Rapid solution of problems by quantum computation. P Roy Soc Lond A Mat, 1992, 439: 553–558
- 9 Grover L K. A fast quantum mechanical algorithm for database search. In: Proceedings of the 28th Annual ACM Symposium on Theory of Computing, Pennsylvania, 1996, 212–219
- 10 Long G L, Li Y S, Zhang W L, et al. Phase matching in quantum searching. Phys Lett A, 1999, 262: 27–34
- 11 Long G L, Li X, Sun Y. Phase matching condition for quantum search with a generalized initial state. Phys Lett A, 2002, 294: 143–152
- 12 Long G L. Grover algorithm with zero theoretical failure rate. Phys Rev A, 2001, 64: 022307
- 13 Long G L, Liu Y. Search an unsorted database with quantum mechanics. Front Comput Sci China, 2007, 1: 247–271
- 14 Bonner R, Freivalds R. A survey of quantum learning. In: Proceedings of the 3rd Workshop on Quantum Computation and Learning, Riga, Latvia, 2003, 106–119
- 15 Servedio R A, Gortler S J. Equivalences and separations between quantum and classical learnability. SIAM J Comput, 2004, 33: 1067–1092
- 16 Aïmeur E, Brassard G, Gambs S. Machine learning in a quantum world. Adv Artif Intell, 2006, 431–442
- 17 Lloyd S, Mohseni M, Rebentrost P. Quantum algorithms for supervised and unsupervised machine learning. arXiv:1307.0411, 2013. <http://arxiv.org/abs/1307.0411>
- 18 Tan P T, Steinbach M, Kumar V. Introduction to Data Mining. Boston: Addison-Wesley Publishing Company, 2014
- 19 Zhu M. Introduction to Data Mining (in Chinese). Beijing: China Science and Technology Press, 2012. [朱明. 数据挖掘导论. 北京: 中国科学技术出版社, 2012]
- 20 Gambs S. Quantum classification. arXiv:0809.0444, 2008. <http://arxiv.org/abs/0809.0444>
- 21 Lu S, Braunstein S L. Quantum decision tree classifier. Quantum Inf Process, 2014, 13: 757–770
- 22 Anguita D, Ridella S, Riviello F, et al. Quantum optimization for training support vector machines. Neural Netw, 2003, 16: 763–770
- 23 Yoo S, Bang J, Lee C, et al. A quantum speedup in machine learning: finding an N-bit Boolean function for a classification. New J Phys, 2014, 16: 103014
- 24 Rebentrost P, Mohseni M, Lloyd S. Quantum support vector machine for big feature and big data classification. Phys Rev Lett, 2014, 113: 130503
- 25 Guță M, Kotłowski W. Quantum learning: Asymptotically optimal classification of qubit states. New J Phys, 2010, 12: 123032
- 26 Sentís G, Calsamiglia J, Muñoz-Tapia R, et al. Quantum learning without quantum memory. Sci Rep, 2012, 2: 708
- 27 Neven H, Denchev V S, Rose G, et al. Training a binary classifier with the quantum adiabatic algorithm. arXiv:0811.0416, 2008. <http://arxiv.org/abs/0811.0416>
- 28 Neven H, Denchev V S, Rose G, et al. Training a large scale classifier with the quantum adiabatic algorithm. arXiv:0912.0779, 2009. <http://arxiv.org/abs/0912.0779>
- 29 Denchev V S, Ding N, Vishwanathan S V N, et al. Robust classification with adiabatic quantum optimization. In: Proceedings of the 29th International Conference on Machine Learning, Edinburgh, 2012
- 30 Pudenz K L, Lidar D A. Quantum adiabatic machine learning. Quantum Inf Process, 2013, 12: 2027–2070
- 31 Aïmeur E, Brassard G, Gambs S. Quantum clustering algorithms. In: Proceedings of the 24th International Conference on Machine Learning, Oregon, 2007. 1–8
- 32 Blatt M, Wiseman S, Domany E. Superparamagnetic clustering of data. Phys Rev Lett, 1996, 76: 3251
- 33 Aïmeur E, Brassard G, Gambs S. Quantum speed-up for unsupervised learning. Mach Learn, 2013, 90: 261–287
- 34 Horn D, Gottlieb A. The method of quantum clustering. In: Proceedings of the 25th Annual Conference on Neural Information Processing Systems, Granada, 2001. 769–776

- 35 Horn D, Gottlieb A. Algorithm for data clustering in pattern recognition problems based on quantum mechanics. *Phys Rev Lett*, 2001, 88: 018702
- 36 Weinstein M, Meirer F, Hume A, et al. Analyzing big data with dynamic quantum clustering. *arXiv:1310.2700*, 2013. <http://arxiv.org/abs/1310.2700>
- 37 Kak S. On quantum neural computing. *Inf Sci*, 1995, 83: 143–160
- 38 Zak M, Williams C P. Quantum neural nets. *Int J Theor Phys*, 1998, 37: 651–684
- 39 Menneer T S I. Quantum artificial neural networks. Doctor Dissertation. Exeter: University of Exeter, 1999
- 40 Peruš M. Neural networks as a basis for quantum associative networks. *Neural Netw World*, 2000, 10: 1001–1013
- 41 Narayanan A, Menneer T. Quantum artificial neural network architectures and components. *Inform Sci*, 2000, 128: 231–255
- 42 Behrman E C, Niemel J, Steck J E, et al. A quantum dot neural network. In: *Proceedings of the 4th Workshop on Physics and Computation*, Boston, 1996, 22–24
- 43 Tóth G, Lent C S, Tougaw P D, et al. Quantum cellular neural networks. *Superlattice Microst*, 1996, 20: 473–478
- 44 Behrman E C, Nash L R, Steck J E, et al. Simulations of quantum neural networks. *Inf Sci*, 2000, 128: 257–269
- 45 Behrman E C, Chandrashekar V, Wang Z, et al. A quantum neural network computes entanglement. *arXiv: quant-ph/0202131*, 2002. <http://arxiv.org/abs/quant-ph/0202131>
- 46 Gupta S, Zia R K P. Quantum neural networks. *J Comput Syst Sci*, 2001, 63: 355–383
- 47 Li P C, Li S Y. Learning algorithm and application of quantum BP neural networks based on universal quantum gates. *J Syst Eng Electron*, 2008, 19: 167–174
- 48 Silva A, de Oliveira W, Ludermit T. A weightless neural node based on a probabilistic quantum memory. In: *Proceedings of the Eleventh Brazilian Symposium on Neural Networks*, São Paulo, 2010, 259–264
- 49 Panella M, Martinelli G. Neural networks with quantum architecture and quantum learning. *Int J Circ Theor App*, 2011, 39: 61–77
- 50 Da Silva A J, de Oliveira W R, Ludermit T B. Classical and superposed learning for quantum weightless neural networks. *Neurocomputing*, 2012, 75: 52–60
- 51 Altaisky M V. Quantum neural network. *arXiv: quant-ph/0107012*, 2001. <http://arxiv.org/abs/quant-ph/0107012>, 2001
- 52 Li F, Zheng B. A study of quantum neural networks. In: *Proceedings of the International Conference on Neural Networks and Signal Processing*, Nanjing, 2003, 539–542
- 53 Siomau M. A quantum model for autonomous learning automata. *Quantum Inf Process*, 2014, 13: 1211–1221
- 54 Sagheer A, Zidan M. Autonomous quantum perceptron neural network. *arXiv:1312.4149*, 2013. <http://arxiv.org/abs/1312.4149>
- 55 Ventura D, Martinez T. Quantum associative memory. *Inf Sci*, 2000, 124: 273–296
- 56 Trugenberger C A. Probabilistic quantum memories. *Phys Rev Lett*, 2001, 87: 067901
- 57 Long G L, Sun Y. Efficient scheme for initializing a quantum register with an arbitrary superposed state. *Phys Rev A*, 2001, 64: 014303
- 58 Trugenberger C A. Quantum pattern recognition. *Quantum Inf Process*, 2002, 1: 471–493
- 59 Trugenberger C A. Phase transitions in quantum pattern recognition. *Phys Rev Lett*, 2002, 89: 277903
- 60 Andrecut M, Ali M K. Quantum associative memory. *Int J Mod Phys B*, 2003, 17: 2447–2472
- 61 Zhou R, Wang H, Wu Q, et al. Quantum associative neural network with nonlinear search algorithm. *Int J Theor Phys*, 2012, 51: 705–723
- 62 Kouda N, Matsui N, Nishimura H, et al. Qubit neural network and its learning efficiency. *Neural Comput Appl*, 2005, 14: 114–121
- 63 Schuld M, Sinayskiy I, Petruccione F. Quantum walks on graphs representing the firing patterns of a quantum neural network. *Phys Rev A*, 2014, 89: 032333
- 64 Lewenstein M. Quantum perceptrons. *J Mod Opt*, 1994, 41: 2491–2501
- 65 Peruš M. Neuro-quantum parallelism in brain-mind and computers. *Informatica*, 1996, 20: 173–184
- 66 Purushothaman G, Karayiannis N B. Quantum neural networks (QNNs): Inherently fuzzy feedforward neural networks. *IEEE Trans Neural Netw*, 1997, 8: 679–693
- 67 Bonnell G, Papini G. Quantum neural network. *Int J Theor Phys*, 1997, 36: 2855–2875
- 68 Kasabov N. *Future Directions for Intelligent Systems and Information Sciences*. Heidelberg: Physica Verlag, 2000
- 69 Graliewicz P. Quantum computing in neural networks. *arXiv: quant-ph/0401127*, 2004. <http://arxiv.org/abs/quant-ph/0401127>
- 70 Li F, Zhao S, Zheng B. Quantum neural network in speech recognition. In: *Proceedings of the 6th International Conference on Signal Processing Proceedings*, Beijing, 2002, 2: 1267–1270
- 71 Kouda N, Matsui N, Nishimura H. Image compression by layered quantum neural networks. *Neural Process Lett*, 2002, 16: 67–80
- 72 Kouda N, Matsui N, Nishimura H. Control for swing-up of an inverted pendulum using Qubit neural network. In: *Proceedings of the 41st SICE Annual Conference*, Osaka, 2002, 1: 805–810

- 73 Fei L, Baoyu Z. A study of quantum neural networks. In: Proceedings of the International Conference on Neural Networks and Signal Processing, Nanjing, 2003, 1: 539–542
- 74 Curtis D, Meyer D A. Towards quantum template matching. In: Proceedings of the SPIE's 48th Annual Meeting, San Diego, 2004, 134–141
- 75 Sasaki M, Carlini A, Jozsa R. Quantum template matching. Phys Rev A, 2001, 64: 022317
- 76 Sasaki M, Carlini A. Quantum learning and universal quantum matching machine. Phys Rev A, 2002, 66: 022303
- 77 Manzano D, Pawłowski M, Brukner Č. The speed of quantum and classical learning for performing the k -th root of NOT. New J Phys, 2009, 11: 113018
- 78 Servedio R A, Gortler S J. Quantum versus classical learnability. In: Proceedings of the 16th Annual IEEE Conference on Computational Complexity, Chicago, 2001, 138–148
- 79 Orejas F, Spirakis P G. Automata, Languages and Programming. Berlin: Springer, 2001
- 80 Pearson B J, White J L, Weinacht T C, et al. Coherent control using adaptive learning algorithms. Phys Rev A, 2001, 63: 063412
- 81 Sen P. Random measurement bases, quantum state distinction and applications to the hidden subgroup problem. In: Proceedings of the 21st Annual IEEE Conference on Computational Complexity, Prague, 2006. 287
- 82 Hentschel A, Sanders B C. Machine learning for precise quantum measurement. Phys Rev Lett, 2010, 104: 063603
- 83 Bang J, Ryu J, Yoo S, et al. Strategy for quantum algorithm design assisted by machine learning. New J Phys, 2014, 16: 073017
- 84 Ziman M, Plesch M, Bužek V, et al. Process reconstruction: From unphysical to physical maps via maximum likelihood. Phys Rev A, 2005, 72: 022106
- 85 Hayashi A, Hashimoto T, Horibe M. Reexamination of optimal quantum state estimation of pure states. Phys Rev A, 2005, 72: 032325
- 86 Gammelmark S, Mølmer K. Quantum learning by measurement and feedback. New J Phys, 2009, 11: 033017
- 87 Bisio A, Chiribella G, D'Ariano G M, et al. Optimal quantum learning of a unitary transformation. Phys Rev A, 2010, 81: 032324
- 88 Granade C E, Ferrie C, Wiebe N, et al. Robust online Hamiltonian learning. New J Phys, 2012, 14: 103013
- 89 Wiebe N, Granade C, Ferrie C, et al. Hamiltonian learning and certification using quantum resources. Phys Rev Lett, 2014, 112: 190501

Big data and quantum computation

WANG ShuHao^{1,2} & LONG GuiLu^{1,2,3}

¹ State Key Laboratory of Low-Dimensional Quantum Physics and Department of Physics, Tsinghua University, Beijing 100084, China;

² Innovative Center of Quantum Matter, Beijing 100084, China;

³ Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China

With the explosion of big data, higher requirements for computational efficiency have emerged. Compared with classical computing, quantum computing possesses quantum parallelism due to the unique nature of quantum systems. It has been found that many classical algorithms can be accelerated using quantum computing. In addition to factorizing a large integer, quantum computers can be used for data processing and analysis. In recent years, two frontiers, i.e., big data and quantum computing have begun to merge. Though practical quantum computers have not yet been built, theoretical studies have made some important progress. In this review, we introduce the basic principles of quantum computing. As a representative example, we describe the Grover search algorithm and its important generalizations. Quantum machine learning is the entry point for the integration of big data with quantum computation. We review in detail, the applications of quantum computation in data mining, the main application of machine learning. Other aspects of quantum computing in big data are also briefly summarized.

big data, quantum computation, quantum machine learning, quantum information processing

doi: 10.1360/N972014-00803