



# 基于混合量子-经典神经网络模型的股价预测

张晓旭<sup>1\*</sup>, 高振涛<sup>1</sup>, 吴磊<sup>1</sup>, 李鑫<sup>1</sup>, 卢明静<sup>2</sup>

(1. 建信金融科技有限责任公司 上海 浦东新区 200120; 2. 四川元匠科技有限公司 成都 611730)

**【摘要】**股票预测本质上是数据挖掘的问题, 大盘走势是一个很好的股票买卖时机抉择信号。在量化分析中, 常用深度学习技术对大盘历史数据进行拟合与特征提取, 为股票投资提供决策参考。该文首先训练了一个经典深度神经网络对沪深 300 的日 K 量价数据进行监督学习, 实现了一个输出“涨跌”概率的二分类预测器, 并以此制定策略进行模拟交易, 利用测试集数据计算累积收益率, 从而评估投资策略的优劣。此外, 还构造了一种混合量子-经典神经网络模型, 充分利用量子计算的线路模型特点, 构造参数化变分量子线路, 实现了量子前馈神经网络。在量子线路学习框架中, 将股票的特征因子编码到量子态的振幅上, 通过训练量子神经网络  $U$  的参数  $\theta$ , 迭代得到一个最优的分类器。量子算法的运行时间比经典算法少了 7.7%, 预测准确率更高, 回报率高出 3%, 因此证明了量子算法的表达力强、鲁棒性高的特点。

**关键词** 深度学习; 混合量子-经典神经网络; 量子金融; 股价预测; 技术面因子  
中图分类号 TP389.1 文献标志码 A doi:10.12178/1001-0548.2021394

## Stock Price Prediction Based on a Hybrid Quantum-Classical Neural Network Model

ZHANG Xiaoxu<sup>1\*</sup>, GAO Zhentao<sup>1</sup>, WU Lei<sup>1</sup>, LI Xin<sup>1</sup>, and LU Mingjing<sup>2</sup>

(1. CCB Fintech Co. Ltd. Pudong New Area, Shanghai 200120; 2. Sichuan Yuanjiang Technology Co. Ltd. Chengdu 611730)

**Abstract** Stock price prediction is essentially a problem of data mining. The market trend is a good signal for stock trading timing. In quantitative analysis, deep learning technology is often used to fit and extract the characteristics of the market historical data, so as to provide decision-making references for stock investment. In this paper, a classical deep neural network is presented and trained to implement supervised learning based on the daily candlestick chart's volume and price data of CSI 300; a binary predictor is realized to output the "rise" or "fall" labels and simulate the transaction with the strategy of the predictor; and then the test data set is utilized to calculate the cumulative rate of return, so as to estimate the advantages and disadvantages of the investment strategy. In addition, another hybrid quantum-classical neural network model is constructed, it makes full use of the characteristics of the circuit model of quantum computing to form the parameterized variational quantum circuit and realize the quantum implementation of feedforward neural network. In the frame of quantum circuit learning, the stock indicators are encoded into the amplitude of quantum states and the parameters  $\theta$  of the quantum neural network  $U$  are trained, and finally an optimal classifier is obtained after iterations. The results show the characteristics of strong expression and high robustness: the run-time of the quantum algorithm is 7.7% shorter than that of the classical algorithm and the prediction accuracy is higher, resulting in a return rate advantage of 3%.

**Key words** deep learning; hybrid quantum-classical neural network; quantum finance; stock market prediction; technical indicators

近年来, 随着计算机计算能力的提升及各行业应用大数据的积累, 深度学习得到迅速发展。凭借人工神经网络自动提取特征的特点, 深度学习在众多应用场景中展现了巨大应用价值, 尤其在图像识别、自然语言处理、棋牌游戏等方面甚至超过了人

类。量化投资作为金融领域中最依赖大数据的部分, 天然契合人工智能技术的应用场景<sup>[1-2]</sup>。

股票市场预测一直是一个经典而又富有挑战性的问题, 同时受到经济学家和计算机科学家的关注。但股票价格数据仍然很难预测, 市场受到投资

收稿日期: 2021-11-01; 修回日期: 2021-12-30

作者简介: 张晓旭, 男, 博士, 工程师, 主要从事量子金融算法方面的研究。

\*通信作者: 张晓旭, E-mail: zhangxiaoxu.zb@ccbft.com

者基于情感和认知因素的影响, 会产生如短期投机行为、恐慌性抛售和非理性繁荣等现象, 直接催生了价格泡沫这类高噪声数据。所以从总体上看, 大盘的股价数据是非线性、非平稳、非结构和高噪声的。得益于中国金融市场的逐步完善, 有效的金融数据得到大量积累, 深度学习在量化投资中的运用在国内也得以快速发展。最近用于预测股价数据的热门方法之一就是深度学习<sup>[3]</sup>。本文成功地训练了一个 24 维特征的深度神经网络 (deep neural network, DNN), 实现了对每日股价涨跌的二分类预测。

尽管深度学习在解决各方面问题时有很好的表现, 但深度学习必须使用大量数据进行训练, 否则这种方法将经历过拟合。因此, 为了提取海量大数据的特征, 包含上亿权重参数的神经网络被广泛地采用, 其对经典计算机的运算能力提出越来越高的要求。然而随着摩尔定律的失效, 集成电路上晶体管密度接近工艺制造的物理极限, 导致经典算力的瓶颈到来, 人工智能需要寻找新型算力平台<sup>[4]</sup>。量子计算作为全新的计算模式, 比起经典计算, 在一些特定算法上有着指数加速的能力, 有望给深度学习提供足够的计算能力。开发适用于量子计算机的量子神经网络, 使得更复杂的神经网络模型得以训练, 并加速模型的训练速度<sup>[5]</sup>, 这些潜在的量子优势使得量子人工智能成为了一个活跃的研究领域<sup>[6-8]</sup>。

本文针对量子计算的特性, 充分利用量子比特的叠加性和纠缠性, 搭建了混合量子-经典神经网络模型 (hybrid quantum-classical neural network), 然后在量子线路学习 (quantum circuit learning, QCL) 框架下, 对股价特征数据进行监督学习 (supervised learning)。训练后的量子神经网络可以对输入的股票特征进行准确分类, 给出涨跌标签。其中, 量子神经网络 (quantum neural network, QNN) 仅指在量子计算机上执行的参数化量子线路 (parameterized quantum circuit, PQC)。

此外, 本文证明了量子神经网络对股票数据具有良好的表达性, 能够实现与经典深度神经网络类似的分类能力。量子神经网络成功地对股票数据进行拟合, 证明了量子金融领域的巨大潜力。

## 1 金融量化分析

股票分析常用的手段有基本面分析 (fundamental analysis) 和技术面分析 (technical analysis)<sup>[9]</sup>。基本

面分析是从宏观上评估企业价值或预测股票价格的走势, 考虑因素一般为政府政治决策、企业负面舆情、国家宏观经济指数发布、产业分析、自然灾害、恐怖袭击等不可抗力。技术面分析是利用金融量化模型和拟合工具, 依靠股票市场历史上的交易价格 (price) 和成交量 (volume) 及其衍生出的一系列指标 (indicators) 来定量分析预测股票价格。

量化模型是研究股票价格波动的数学原理, 建立股票价格  $S$  与时间  $t$  之间的数量关系。金融学中用于描述股价变化的模型是几何布朗运动 (geometric Brownian motion, GBM), 即式 (1) 中的  $W_t$  项。根据随机波动假设下  $S(t)$  连续但不可导的性质, 应用伊藤引理后可以得到描述股价的随机偏微分方程。随机波动模型需要市场波动率  $\sigma$  和收益率  $\alpha$  两个重要参数, 但真实市场中  $\sigma(t, S_t)$  和  $\alpha(t, S_t)$  也是时间与股价的函数, 所以股票方程只是理论模型, 实际预测能力有限。

$$\frac{dS(t)}{S(t)} = \alpha dt + \sigma dW_t \quad (1)$$

一般来说, 金融工程是一门专业性很强的学科, 传统量化模型中要求专业人员提出合理解释市场的特征因子, 这需要很强的证券投资知识的储备。但量价时间序列数据具有非线性、非平稳、非结构和高噪声的特点, 大部分公共事件, 如舆论、政策, 甚至天气、群体心理等都会通过大家的逐利行为反映到一个单一的变量, 即股价上面。因此需要量化分析专家深刻理解金融数学模型的原理和数据分析应用场合, 他们的先验经验是非常重要的。

但是深度学习几乎不需要处理原始数据, 深度神经网络模型有良好的非线性拟合能力, 能自动学习最有预测性的特征, 这使得数据分析师在基本了解证券市场知识的情况下, 也能像金融工程专家一样寻找证券市场波动背后的规律。同时, 深度神经网络模型还具有强大的表达能力和泛化能力。应用深度学习方法寻找股市规律正逐渐受到业界重视。

具体来说, 如果目标是预测股票的具体价值, 这属于回归问题 (regression); 如果目标是预测价格的运动方向, 如上升或下降, 这属于分类问题 (classification)。设第  $k$  组数据中的特征 (feature) 为一个向量  $\mathbf{X}^{(k)}$ , 对应的标签 (label) 为标量  $y^{(k)} \in \{0, 1\}$ 。输入  $M$  个训练样本  $[\mathbf{X}^{(k)}, y^{(k)}], k = 0, 1, \dots, M-1$ , 得到训练后的神经网络, 再输入训练集外任意一天的股市特征  $\mathbf{X}^{(h)}$ , 都能输出涨跌的预测标签  $y^{(h)}$ 。

## 2 经典深度神经网络预测器

由于股民的短期投机行为会引起层出不穷的价格波动, 所以金融数据样本越小则噪音越大, 必须基于较长时间段的多维度特征向量才能训练出表达性强、准确率高的神经网络。

当前市场中的证券数目是相当大的, A 股市场的股票数目现在有 3000 多只, 如果再加上期货、基金可能会有上万吨投资标的。基于大盘数据就可以充分利用神经网络对高维大数据进行拟合或分类。本章以沪深 300 股指期货为例, 选取一段时间内的日 K 线数据进行训练, 根据预测结果, 设定每日交易策略: 1) 看涨, 今日继续持有, 明日买入; 2) 看跌, 今日卖出, 明日买回; 3) 趋势不明, 保存现金, 股票不操作。

### 2.1 经典深度神经网络模型

神经网络模型是深度学习算法的核心, 神经网络最简单的形式是一种两层神经网络, 也叫感知机 (perceptron), 它有若干个输入值和一个输出值, 最早在 1957 年由 R. Rosenblatt 提出, 其结构如图 1 所示。输入值向量先经过仿射变换  $Z$  变为一个标量, 然后由激活函数  $a$  做非线性映射, 完成特征提取。

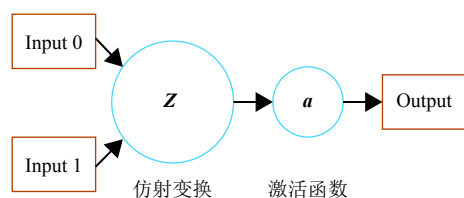


图 1 感知机模型

多层感知机构成了神经网络, 神经网络层可分为 3 类, 即输入层、隐藏层和输出层。而 DNN 就是有多层隐藏层的神经网络, 它可以提取更复杂的数据特征。整个 DNN 中的信息是朝着一个方向传播的, 第  $l$  层感知机把处理的结果传递给第  $l+1$  层, 并且每层感知机都包含自己的仿射变换  $Z$  和激活函数  $a$ , 形如:  $a^{(0)} \rightarrow Z^{(1)} \rightarrow a^{(1)} \rightarrow \dots \rightarrow Z^{(L)} \rightarrow a^{(L)}$ 。一般化的 DNN 可以用一个有向图来表示, 如图 2 所示。

通过简单非线性函数的多次复合, 实现输入空间到输出空间的复杂映射, 映射过程可由式 (2) 计算。容易看出, 各层神经网络的权重矩阵  $W^{(l)}$  和偏置向量  $b^{(l)}$  就是待训练的神经网络参数<sup>[10]</sup>。

$$\begin{aligned} z^{(l)} &= W^{(l)} a^{(l-1)} + b^{(l)} \\ a^{(l)} &= \text{activate}(z^{(l)}) \end{aligned} \quad (2)$$

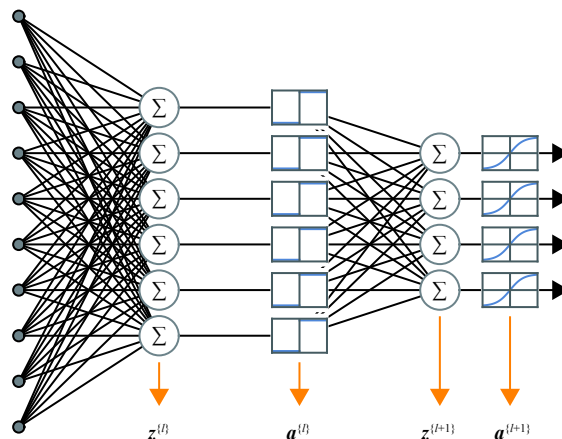


图 2 神经网络隐藏层间的连接映射过程

此外, 还需要用损失函数来衡量神经网络参数是否合理。损失函数可以定量描述由神经网络前向传播得到的结果与真实值之间的差距, 拟合常用均方误差 (mean squared error, MSE), 分类常用交叉熵 (cross entropy)。

训练过程就是不断修正神经网络参数使得损失函数越来越小的过程, 这需要高效的误差反向传播算法来进行参数调整, 一般这是一个迭代收敛的过程。训练前馈神经网络最著名的方法是梯度下降法 (gradient descent), 通常基于 TensorFlow、PyTorch、Jax 等高效可靠的微编程工具<sup>[11]</sup>。

### 2.2 股市数据的训练参数设计

首先是数据集的准备, 包括训练集 (training)、测试集 (testing)、特征 (features) 和标签 (labels)。数据来源为沪深 300 股指从 2019 年 1 月 2 日-2021 年 7 月 6 日的所有 K 线量价指标, 取前 400 天数据为训练数据, 剩下为测试数据。

如图 3 所示, 日 K 线原始数据包括 5 个量价指标, 即开盘价  $O$ 、收盘价  $C$ 、当日最高价  $H$ 、当日最低价  $L$  和总量  $V$ 。标签  $y$  的取值只有两个  $\{0, 1\}$ , 连续两个交易日的收盘价  $C[k+1] > C[k]$ , 则  $y[k]=1$ , 表示上涨; 若  $C[k+1] < C[k]$ , 则  $y[k]=0$ , 表示下跌。

特征  $X$  如果设为这 5 个参数  $\{x_0^{(k)}=V[k], x_1^{(k)}=O[k], x_2^{(k)}=C[k], x_3^{(k)}=H[k], x_4^{(k)}=L[k]\}$ , 根据实验, 这种 5 维的原始特征是有极大噪声的, 需要发掘更高维度的量化因子来防止过拟合。更有效的金融量价因子需要从原始数据中挖掘, 在量化领域被称为特征提取 (feature extraction)<sup>[12-13]</sup>。如一个较好的因子叫做价格变动百分比 (percentage price change, PPC), 是今日收盘价相当于昨日收盘价的变化率, 即  $PPC[k] = (C[k] - C[k-1]) / C[k-1] \times 100$ 。经过多次实验, 本文用 5 个原始 K 线参数生成了 24 维度的特



征因子  $X^{(k)} = [x_0^{(k)}, x_1^{(k)}, \dots, x_{23}^{(k)}]$ 。当然, 挖掘的因子其有效性也不是恒定的, 会随时间和环境发生变化。

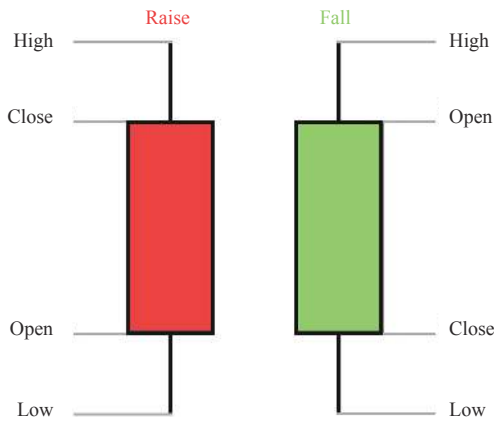


图 3 涨跌 K 线图

具体因子包括  $X = [O\_log, C\_log, H\_log, L\_log, V\_log\_1, V\_log\_2, V\_log\_3, V\_log\_4, V\_log\_5, PPC\_1, PPC\_2, PPC\_3, PPC\_4, PPC\_5, CO\_gap\_1, CO\_gap\_2, CO\_gap\_3, CO\_gap\_4, CO\_gap\_5, HL\_gap\_1, HL\_gap\_2, HL\_gap\_3, HL\_gap\_4, HL\_gap\_5]$ 。这 24 个因子定义如下:  $O\_log$  开盘价取对数;  $C\_log$  收盘价取对数;  $H\_log$  当日最高价取对数, 高点也叫阻力位;  $L\_log$  当日最低价取对数, 低点也叫支撑位;  $V\_log\_5$  是日交易量取对数的 5 日滑动平均值, 其余  $V\_log\_1-4$  含义类似;  $PPC\_5$  是 5 日价格变动百分比的滑动平均值, 其余  $PPC\_1-4$  含义类似;  $CO\_gap\_5$  是收盘价减去开盘价的 5 日滑动平均值, 其余  $CO\_gap\_1-4$  含义类似, 这也是著名的动量特征;  $HL\_gap\_5$  是单日震荡幅度的 5 日滑动平均值, 其余  $HL\_gap\_1-4$  含义类似。

取对数是为了保证特征因子的大小在同一数量级, 避免梯度下降过程中损失函数进入振荡。针对特别大和特别小的数据要进行恰当的缩放, 如成交量是远大于股价的。综上所述, 用于神经网络训练与测试的数据集如表 1 所示。

表 1 大盘数据预处理生成数据集

数据集 编号	Features					Labels	
	PPC_1	...	V_log_1	...	HL_garp_5	y	
训练集	0	15.82	...	18.08	...	47.19	0
	1	239.57	...	18.45	...	100.98	1
	...	...	...	...	...	...	...
	499	-21.60	...	18.27	...	16.98	0
测试集	500	60.70	...	18.43	...	25.84	1
	...	...	...	...	...	...	...
	698	-5.29	...	18.76	...	63.70	1

本文用 DNN 算法模型进行股价预测器的训练。如图 4 所示, 训练的 DNN 总共有 4 层: 输入层有 24 个神经元, 激活函数 ReLU; 隐藏层 1 有 12 个神经元, 激活函数 ReLU; 隐藏层 2 有 12 个神经元, 激活函数 ReLU; 隐藏层 3 有 6 个神经元, 激活函数 ReLU; 输出层有两个神经元, 经过激活函数 Sigmoid 以后给出涨的概率  $P(\text{raise} \equiv 1)$  和跌  $P(\text{fall} \equiv 0)$  的概率。

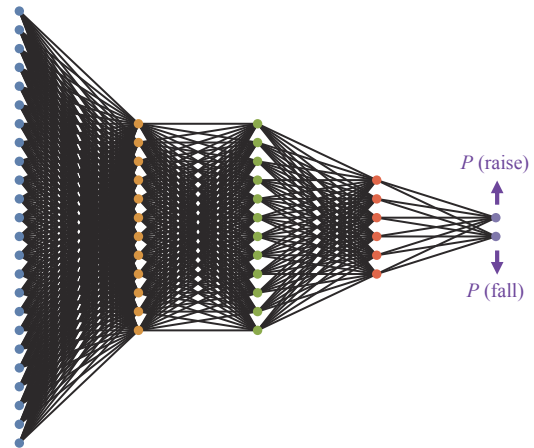


图 4 深度神经网络股价分类结构

### 2.3 预测器的盈利能力检测

神经网络的输出经过 sigmoid 函数的转换, 可以将数值压缩到  $(0, 1)$  之间, 如神经网络以概率  $P[k]$  把第  $k$  日的标签  $y^{(k)}$  归类于 1, 以概率  $P[k+1]$  把第  $k+1$  日的标签  $y^{(k)}$  归类于 0。

但  $P$  是连续的, 不能简单认为  $P > 0.5$  就看涨, 而是要在可能性很大的情况下才看涨。根据多次实验结果, 本文的策略是对神经网络输出结果进行修剪, 设定  $P > 0.95$  看涨,  $P < 0.05$  看跌, 其余区间属于趋势不明, 这样最终得到了 3 个标签, 恰好可以指导 3 种操作策略。

预测股价上涨则做多, 即明日开盘买入, 记作  $do=1$ 。预测股价下跌则做空, 记作  $do=-1$ , 即明日开盘卖出, 套现开盘的高价股价。无法预测则选择不操作, 即保持现金, 既不把现金换成股票, 也不把股票换成现金, 记作  $do=0$ 。

把 3 类预测结果  $y$  转化为对应的投资策略  $do$  以后就能进行评估分析了。这里采用了回测法, 在训练集的真实市场数据上检验本文投资策略盈利情况。不同于衡量拟合器性能的回测曲线, 针对分类器的性能指标是累积收益 (cumulative return)。设定首日累积收益率  $CR[0]=1$ , 第  $k$  日的累计收益率为  $CR[k]$ , 股价为  $S[k]$ , 那么  $k+1$  日的累计收益率计

算为:

$$CR[k+1] = CR[k] \left( 1 + do \times \frac{S[k+1] - S[k]}{S[k]} \right) \quad (3)$$

当策略  $do$  与真实市场涨跌相同时, 累积收益率就持续增长。如股价今日 100 元, 神经网络模型预测明日会跌, 即  $do=-1$ , 但是实际第二天股价涨到 110 元了, 错误的决策就导致第二天的累积收益率降低 10%。

训练集上检测的投资结果如图 5 所示。可以看出在 2020 年 9 月 1 日-2021 年 7 月 6 日测试时间段内, 大部分时间都是持续盈利的, 最高收益在 2021 年 2 月底出现, 达到了 25%。这证明了该神经网络分类器的准确性较高, 表现稳定。

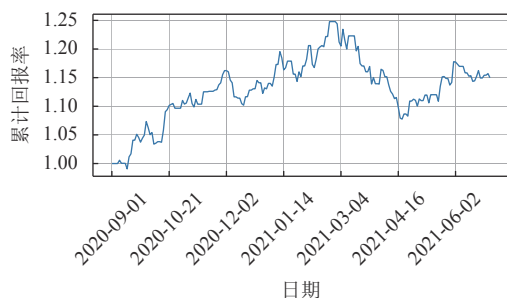


图 5 投资策略的累积收益率结果测试

### 3 混合量子-经典神经网络预测器

当前的量子算法研究者已经不是为完美的大规模量子计算设备设计算法, 而是充分利用短期内的含噪声小规模量子计算设备, 为计算过程提供局部加速的混合量子-经典计算模型<sup>[14]</sup>。混合量子-经典计算概念的发展引出一系列新算法, 如量子近似优化算法 (quantum approximate optimization algorithm, QAOA)<sup>[15]</sup>、变分量子本征求解器 (variational quantum eigensolver, VQE)<sup>[16]</sup>、量子分类器<sup>[17]</sup>、量子生成对抗网络 (quantum generative adversarial networks, QGANs)<sup>[18-19]</sup> 和量子神经网络 (QNN)<sup>[20]</sup>。量子神经网络既可以作为一种启发式算法, 利用量子信息原理把输入特征映射到输出分类上<sup>[21-22]</sup>; 也可以在真实量子硬件平台上运行, 如光子量子平台<sup>[23]</sup> 和钻石 NV 色心平台<sup>[24]</sup>。

如图 6 所示, 混合量子-经典算法模型包括量子信息处理部分和经典信息处理部分, 用一个经典优化器来训练一个含参量子线路 (parameterized quantum circuit, PQC)。

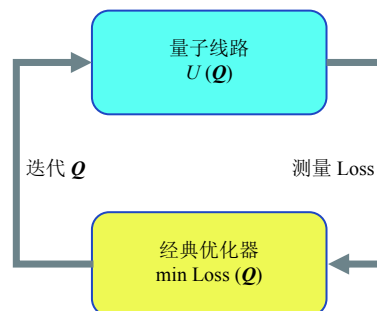


图 6 混合量子-经典方案

#### 3.1 量子感知器

与经典神经网络类似, 量子神经网络是由多层量子感知机 (quantum perceptron, QP) 组合而成。每个 QP 都包括了前处理和后处理参数化量子门, 分别对照仿射变换中的特征向量  $X$  与权重向量  $W$ 。

图 7 为一个量子感知机的实现方案, 对比图 1 的经典感知机结构, 它充分利用量子计算的两个特点来实现感知机的功能: 1) 计算是可逆的。经典感知机把  $m$  维特征向量  $X$  通过仿射变换得到 1 维标量  $z$ , 而量子计算的酉变换需要输入量子比特数等于输出量子比特数, 这里引入辅助量子比特 (ancilla qubits) 来寄存仿射变换结果, 这里的量子比特数为  $\text{ceil}(\log_2 m)+1$ ; 2) 量子逻辑门是酉矩阵的线性变换, 而测量算符是埃米尔特矩阵 (Hermitian matrix) 的非线性变换。经过测量就会引入非线性, 起到激活函数的作用。

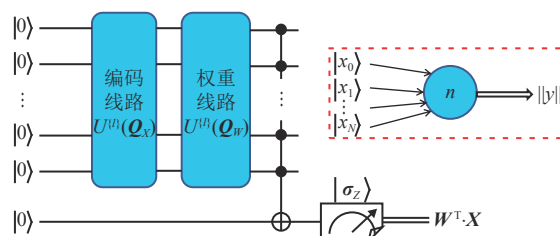


图 7 混合量子-经典感知机的部署方案, 量子线路的多个输入态经过酉变换后, 测量得到经典输出信息。虚线框内是等效的感知机原理图。

量子感知机完成仿射变换的原理如图 8 所示。在  $N$  个量子比特张开的  $m=2^N$  高维希尔伯特空间 (Hilbert space) 中, 携带特征向量  $X$  的量子态  $|\psi_X\rangle$  向携带权重向量  $W$  的量子态  $|\psi_W\rangle$  上进行投影, 就得到了仿射变换结果  $z = \langle \psi_X | \psi_W \rangle$ 。然后通过参数化量子线路  $U(\theta_W)$ , 把  $z$  编码到任意计算基底  $|j\rangle$  上, 这里  $0 < j < m-1$ , 如式 (4) 所示。只要该  $m \times m$  维的酉矩阵  $U(\theta_W)$  的第  $j$  行是经典权重向量  $W^T$  就可以了。

$$U(\theta_W) |\psi_X\rangle = \sum_{j=0}^{n-1} c_j |j\rangle \quad (4)$$

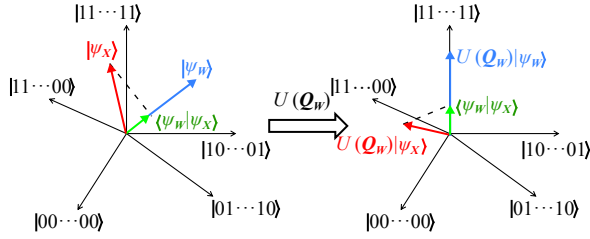


图 8 把仿射变换结果编码到 $|11\dots 11\rangle$ 基底上

式 (5) 表示把投影结果  $\langle\psi_x|\psi_w\rangle$  编码到基底  $|11\dots 11\rangle = |m-1\rangle$  上, 即态矢  $U(\theta_w)|\psi_w\rangle$  与基矢  $|11\dots 11\rangle$  共线, 那么参数化量子线路的酉矩阵只要保证最后一行为  $W^T$  就可以了。

$$U(Q_w) = \begin{pmatrix} & & & \vdots \\ W_1 & W_2 & \dots & W_m \end{pmatrix} \quad (5)$$

在经典的神经元模型中, 仿射变换后还需要经过一个激活函数为神经元的输出提供非线性, 保证了多层网络的表达能力不会退化成单层线性网络。一个适当的激活函数可以大大增加分类任务的准确性, 但由于量子线路内部是一系列酉矩阵相乘构成的, 矩阵相乘必然不存在非线性, 所以在量子感知机中, 要通过变基底测量 (computational basis changed measurement) 产生不可逆的非线性映射<sup>[25-26]</sup>。如图 8 中的酉矩阵  $U(\theta_w)$  起到了旋量子态  $|\psi_w\rangle$  到计算基底  $|j\rangle$  的作用, 也等价于坐标系基底的旋转。

通过多个纠缠量子门把前  $N$  个量子比特与辅助比特相连, 就能把  $c_{m-1}$  制备到辅助比特上。

根据式 (5) 发现, 参数化酉矩阵  $U(\theta_w)$  的设计不是唯一的, 一个表达能力强的  $U(\theta_w)$  搭配合适的泡利测量  $P$  就构造出了一个包含一些先验知识的假设线路 (ansatz circuit)。如何根据目标构造有效的 ansatz 是量子人工智能的重点研究方向, 通常由一些单比特量子旋转门和多比特受控旋转门组合。

如图 9 所示, 每层量子神经网络包括多个量子感知机, 多层量子神经网络构成了前馈神经网络 (feedforward neural networks, ffNN)<sup>[26]</sup>。注意, 虽然量子计算的线路模型中有延迟测量 (deferred measurement) 的技巧, 但每个量子感知机的辅助量子比特的量子态必须立马测量<sup>[27]</sup>, 不能延迟到整个量子线路的末端, 否则无法提供中间层的非线性激活<sup>[28]</sup>, 再多层的量子神经网络都只相当于一层。

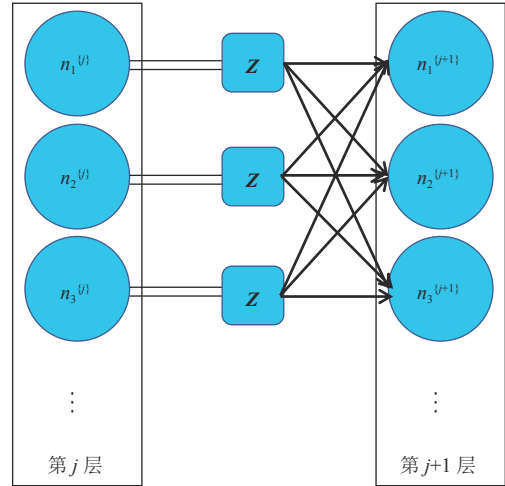


图 9 用量子感知机组成量子神经网络, 测量后的经典信息通过量子控制操作加载到量子态上

### 3.2 振幅编码

量子编码是一个将经典信息转化为量子态的过程。在使用量子算法解决经典问题的过程中, 量子编码是非常重要的一步。在量子神经网络中, 第一步就是将经典信息编码到量子态中。大多数量子编码的方法都可以看作是作用在初始化  $|0\rangle^N$  态上的参数化线路  $U(\theta_x)$ , 并且参数化电路中的参数是由经典信息  $X$  决定的。

由于量子比特天然的叠加性, 对  $m$  维经典数组能以  $\text{ceil}(\log_2 m)$  个量子比特的空间复杂度进行存储计算, 这里  $\text{ceil}$  指向上取整。本次股价预测的特征因子  $X^{(k)} = [x_0^{(k)}, x_1^{(k)}, \dots, x_{23}^{(k)}]$  的长度达到 24, 用振幅编码的方法需要 5 位量子比特就能编码。

在 PennyLane 量子计算框架中, 把经典特征数据编码到量子态中的过程叫作嵌入 (embedding), 本文采用 embeddings 类的 AmplitudeEmbedding 方法加载经典特征向量  $X^{(k)}$ 。

### 3.3 量子神经网络的训练方法

参数化量子电路 PQC 被认为是在量子计算中的一种神经网络的新范式, 不同于经典神经网络的权重矩阵, PQC 是以量子门中的相位  $\theta$  作为训练参数, 从而将量子初态变到期望的量子末态上。经过参数化量子变分线路的酉变换  $U(\theta)$  输出了一个可观测量的期望值  $f(\theta)$ 。和经典神经网络相同, 在训练 PQC 时研究者同样希望找到其相位的梯度, 从而使量子神经网络的损失函数 (cost function) 收敛。

首先需要定义损失函数。对于分类器, 经典数据集  $\{[X^{(k)}, y^{(k)}]\}$  的标签  $y^{(k)}$  对应一个量子可观测量  $Y$ 。经过编码量子线路后得到量子态  $|\psi_x\rangle$ , 再经过 ansatz 参数化量子线路, 对  $Y$  的期望  $f(\theta)$  进行观测,



测量结果  $\langle Y \rangle$  与标签  $y^{(k)}$  的均方误差 (mean-square error, MSE) 作为损失函数  $C(\theta)$ :

$$f(\theta_k) = \langle \psi_X | U^\dagger(\theta_k) Y U(\theta_k) | \psi_X \rangle$$

$$C(\theta_k) = \frac{1}{n} \sum_k |y^{(k)} - f(\theta_k)|^2 \quad (6)$$

然后用量子线路计算出梯度。基于量子线路的梯度计算与经典计算有很大的区别, 要求 PQC 满足一个重要的性质, 参数  $\theta = \{\theta_1, \theta_2, \dots, \theta_L\}$  全部是泡利旋转矩阵的旋转角, 如式 (7) 所示。这里  $\sigma$  代表泡利矩阵。

$$U(\theta_k) = e^{-i\theta_k \sigma} \quad (7)$$

这样的一块参数化量子线路  $U(\theta)$  在计算测量期望值  $f(\theta)$  的梯度时有独特的优势。结合式 (6) 和式 (7), 本文得到梯度如式 (8) 所示:

$$\nabla_{\theta_k} f = \frac{\partial f}{\partial \theta_k} = f\left(\theta_k + \frac{\pi}{4}\right) - f\left(\theta_k - \frac{\pi}{4}\right) \quad (8)$$

这样, 同一个参数线路  $U(\theta)$  不仅可以计算  $f(\theta)$ , 还能计算梯度  $\nabla_{\theta} f$ , 线路结构不变, 仅仅是线路参数  $\theta_k$  不一样<sup>[29-30]</sup>, 这就是量子变分线路梯度计算的 parameter-shift 方法。

神经网络的学习过程就是沿梯度进行参数调整的过程。量子线路输出梯度后, 经过经典优化算法对量子参数  $\theta$  进行调整改善, 再把新参数返回到量子线路, 循环迭代, 直到损失函数收敛到一定值。基于式 (8) 的梯度计算, 可以应用很多经典的梯度下降算法, 尤其是考虑到量子测量本身的随机性, 自然地适用随机梯度下降法 (stochastic gradient descent, SGD)。用变学习率的 SGD 方法训练股价预测器量子神经网络的收敛过程如图 10 所示。

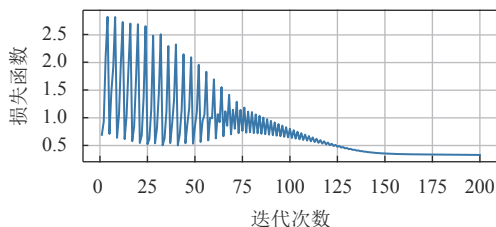


图 10 量子神经网络学习过程中损失函数的收敛

本文针对运行时间和最高累计回报率两项指标, 对比经典深度神经网络 DNN 与混合量子-经典神经网络 HQCNN 的表现。每种算法模型训练 10 次, 取平均值, 得到 DNN 运行 235.3 s, 最大回报率 25%, HQCNN 运行 217.1 s, 最大回报率 28%。结果发现量子算法的运行时间比经典算法少

了 7.7%, 且预测准确率更高导致回报率高出 3%。证明了量子算法的表达力强、鲁棒性高。

## 4 结束语

股市预测拥有海量数据, 是深度学习大展身手的应用场景。用复杂的神经网络对大量的股价样本进行训练是行之有效的, 但非常消耗算力资源。量子神经网络旨在利用量子信息处理过程, 把输入特征映射到输出分类上, 在完成经典神经网络模型的任务同时, 发挥量子优势, 开发更高效准确的算法, 设计能利用量子并行性和量子纠缠等量子效应的量子算法。

后续量子金融方法的改进, 需要加入更复杂的交易策略模型, 考虑更具体精细的指标, 如收益曲线的最大回撤率等。同时针对量子神经网络进行优化, 寻找更具表达能力的 ansatz。更重要的一点是量子变分线路的梯度大小随着量子比特数目的增长而指数衰减, 使得模型失去可训练性, 这是量子人工智能领域急需解决的问题。

本文研究工作还得到建信金融科技有限责任公司研究性课题 (KT2000050) 及成都市重点研发支撑计划重大科技应用示范项目 (2021-YF09-00114-GX) 的资助, 在此表示感谢。

## 参 考 文 献

- [1] DIXON M F, HALPERIN I, BILOKON P. Machine learning in finance: From theory to practice[M]. Cham: Springer International Publishing, 2020.
- [2] 张仕斌, 黄曦, 昌燕, 等. 大数据环境下量子机器学习的研究进展及发展趋势[J]. 电子科技大学学报, 2021, 50(6): 802-819.
- [3] ZHANG S B, HUANG X, CHANG Y, et al Research progress and perspectives of quantum machine learning in big data environment[J]. Journal of University of Electronic Science and technology, 2021, 50(6): 802-819.
- [4] JIANG W. Applications of deep learning in stock market prediction: Recent progress[J]. Expert Systems with Applications, 2021, 184: 115537.
- [5] 赵健, 陈昭昀, 庄希宁, 等. 量子态制备及其在量子机器学习中的前景[J]. 物理学报, 2021, 70(14): 140307.
- [6] ZHAO J, CHEN Z Y, ZHUANG X N, et al. Quantum state preparation and its prospects in quantum machine learning[J]. Acta Physica Sinica, 2021, 70(14): 140307.
- [7] 李楠, 程锦房, 钱富. 加权信号张量子空间拟合算法[J]. 电子科技大学学报, 2013, 42(4): 546-548, 591.
- [8] LI N, CHENG J F, QIAN F. Weighted signal tensor subspace fitting algorithm[J]. Journal of University of Electronic Science and Technology of China, 2013, 42(4):

- 546-548, 591.
- [6] 吴永飞, 王彦博. FinTech 时代量子科技在金融领域的应用展望[J]. 中国金融电脑, 2020(12): 41-45.  
WU Y F, WANG Y B Application prospect of quantum technology in financial field in fintech era[J] China Financial Computer, 2020(12): 41-45.
- [7] 刘聪. 量子计算在金融行业中的应用[J]. 投资与合作, 2021(4): 11-12, 57.  
LIU C. Application of quantum computing in financial industry[J]. Investment and Cooperation, 2021(4): 11-12, 57.
- [8] 於学松. 量子计算在智能金融发展中的应用前景分析[J]. 国际金融, 2020(2): 35-41.  
YU X S. Analysis on the application prospect of quantum computing in the development of intelligent Finance[J]. International Finance, 2020(2): 35-41.
- [9] LI Z, YANG D, ZHAO L, et al. Individualized indicator for all: Stock-wise technical indicator optimization with stock embedding[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, AK: ACM: 894-902.
- [10] NIELSEN M A. Neural networks and deep learning[M]. San Francisco, CA: Determination press, 2015.
- [11] 王磊, 刘金国. 微分万物: 深度学习的启示[J]. 物理, 2021, 50(2): 69-75.  
WANG L, LIU J G. Differentiate everything: A lesson from deep learning[J]. Physics, 2021, 50(2): 69-75.
- [12] GERLEIN E A, MCGINNITY M, BELATRECHE A, et al. Evaluating machine learning classification for financial trading: An empirical approach[J]. Expert Systems with Applications, 2016, 54: 193-207.
- [13] LI Z, TAM V. A Machine learning view on momentum and reversal trading[J]. Algorithms, 2018, 11(11): 170.
- [14] MCCLEAN J R, ROMERO J, BABBUSH R, et al. The theory of variational hybrid quantum-classical algorithms[J]. New Journal of Physics, 2016, 18(2): 023023.
- [15] FARHI E, GOLDSTONE J, GUTMANN S. A Quantum approximate optimization algorithm[EB/OL]. (2014-11-14). <https://arxiv.org/abs/1411.4028>.
- [16] PERUZZO A, MCCLEAN J, SHADBOLT P, et al. A variational eigenvalue solver on a quantum processor[J]. Nature Communications, 2014, 5(1): 4213.
- [17] SCHULD M, BOCHAROV A, SVORE K M, et al. Circuit-centric quantum classifiers[J]. Physical Review A, 2020, 101(3): 032308.
- [18] HUANG H L, DU Y, GONG M, et al. Experimental quantum generative adversarial networks for image generation[J]. Physical Review Applied, 2021, 16(2): 024051.
- [19] DALLAIRE-DEMERS P-L, KILLORAN N. Quantum generative adversarial networks[J]. Physical Review A, 2018, 98(1): 012324.
- [20] KILLORAN N, BROMLEY T R, ARRAZOLA J M, et al. Continuous-variable quantum neural networks[J]. Physical Review Research, 2019, 1(3): 033063.
- [21] ZHAO C, GAO X S. QDNN: Deep neural networks with quantum layers[J]. Quantum Machine Intelligence, 2021, 3(1): 15.
- [22] BEER K, BONDARENKO D, FARRELLY T, et al. Training deep quantum neural networks[J]. Nature Communications, 2020, 11(1): 808.
- [23] WAN K H, DAHLSTEN O, KRISTJÁNSSON H, et al. Quantum generalisation of feedforward neural networks[J]. NPJ Quantum Information, 2017, 3(1): 36.
- [24] 田宇, 林子栋, 王翔宇, 等. 基于自旋体系的量子机器学习实验进展[J]. 物理学报, 2021, 70(14): 140305.  
TIAN Y, LIN Z D, WANG X Y, et al. Experimental progress of quantum machine learning based on spin systems[J]. Acta Physica Sinica, 2021, 70(14): 140305.
- [25] BHARTI K, CERVERA-LIERTA A, KYAW T H, et al. Noisy intermediate-scale quantum (NISQ) algorithms [EB/OL]. (2021-01-21). <https://arxiv.org/abs/2101.08448>.
- [26] TACCHINO F, BARKOUTSOS P, MACCHIAVELLO C, et al. Quantum implementation of an artificial feed-forward neural network[J]. Quantum Science and Technology, 2020, 5(4): 044010.
- [27] TACCHINO F, MACCHIAVELLO C, GERACE D, et al. An artificial neuron implemented on an actual quantum processor[J]. NPJ Quantum Information, 2019, 5(1): 1-8.
- [28] SCHULD M, SWEKE R, MEYER J J. Effect of data encoding on the expressive power of variational quantum-machine-learning models[J]. Physical Review A, 2021, 103(3): 032430.
- [29] SCHULD M, BERGHOLM V, GOGOLIN C, et al. Evaluating analytic gradients on quantum hardware[J]. Physical Review A, 2019, 99(3): 032331.
- [30] MITARAI K, NEGORO M, KITAGAWA M, et al. Quantum circuit learning[J]. Physical Review A, 2018, 98(3): 032309.

编辑 蒋晓