

面向金融科技的深度学习技术综述

周帆, 陈晓蝶, 钟婷, 吴劲

引用本文

周帆, 陈晓蝶, 钟婷, 吴劲. [面向金融科技的深度学习技术综述](#) [J]. 计算机科学, 2022, 49(11A): 210900016-17.

ZHOU Fan, CHEN Xiao-die, ZHONG Ting, WU Jin. [Survey of Deep Learning Technologies for Financial Technology](#) [J]. Computer Science, 2022, 49(11A): 210900016-17.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于差分进化算法的字符对抗验证码生成方法](#)

Adversarial Character CAPTCHA Generation Method Based on Differential Evolution Algorithm
计算机科学, 2022, 49(11A): 211100074-5. <https://doi.org/10.11896/jsjcx.211100074>

[融合多层次视觉信息的人物交互动作识别](#)

Human-Object Interaction Recognition Integrating Multi-level Visual Features
计算机科学, 2022, 49(11A): 220700012-8. <https://doi.org/10.11896/jsjcx.220700012>

[基于注意力机制的手写体数字识别](#)

Handwritten Digit Recognition Based on Attention Mechanism
计算机科学, 2022, 49(11A): 211100009-5. <https://doi.org/10.11896/jsjcx.211100009>

[融合ViT卷积神经网络的木板表面缺陷识别](#)

Wood Surface Defect Recognition Based on ViT Convolutional Neural Network
计算机科学, 2022, 49(11A): 211100090-6. <https://doi.org/10.11896/jsjcx.211100090>

[基于YOLOv3与改进VGGNet的车辆多标签实时识别算法](#)

Multi-label Vehicle Real-time Recognition Algorithm Based on YOLOv3 and Improved VGGNet
计算机科学, 2022, 49(11A): 210600142-7. <https://doi.org/10.11896/jsjcx.210600142>

面向金融科技的深度学习技术综述

周 帆 陈晓蝶 钟 婷 吴 劲

电子科技大学信息与软件工程学院 成都 610054

(fan.zhou@uestc.edu.cn)

摘 要 近年来,深度学习技术被广泛应用于金融领域,并受到了国内外学术界和商业界的广泛关注。研究人员利用深度学习技术对各种金融数据进行发掘和分析,取得了大量的研究成果。深度学习在多个金融关键应用上的表现超过了传统的统计机器学习模型,包括金融市场预测、交易策略改进和金融文本信息挖掘等。为了更全面地把握深度学习技术在金融领域中的研究和应用趋势,促进它们之间的深层次融合和发展,着重梳理了近年来深度学习技术在金融科技研究中的发展脉络及前沿动态,分析和总结了深度学习模型在金融领域的主要应用和最新算法。根据金融领域中的具体应用场景将现有的深度学习金融研究进行详细分类,分析并总结各个领域的最新研究,并展望了金融科技领域未来的研究热点、技术难点和发展趋势等。

关键词: 金融科技;深度学习;价格预测;投资组合管理;趋势预测;风险评估

中图法分类号 TP311

Survey of Deep Learning Technologies for Financial Technology

ZHOU Fan, CHEN Xiao-die, ZHONG Ting and WU Jin

School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China

Abstract In recent years, deep learning techniques have been widely applied in addressing various problems in financial technology (Fintech) and have attracted increasing attention from both academia and business. Researchers utilize deep learning techniques for mining and analyzing financial data while finding the economic patterns behind tremendous data. Deep learning outperforms traditional statistical machine learning models in a range of crucial financial applications, including market movement prediction, trading strategy improvement, financial text processing, etc. To facilitate the development of Fintech and the deployment of new deep learning techniques, this paper provides a comprehensive survey of the deep learning-based Fintech studies published in recent years. Our survey focuses on the most recent advances in Fintech and provides a roadmap of financial problems as well as corresponding solutions. To this end, we investigate the widely used methodologies in finance data mining and summarize the popular deep models in Fintech data learning. Besides, we propose a taxonomy that categorizes existing Fintech research into ten well-studied applications in the literature. Subsequently, we systematically review the state-of-the-art deep learning methods and provide insights on the improvement for future endeavors. Finally, the pros and cons of existing research are summarized, followed by outlining the trend, open challenges, and opportunities in the Fintech research community.

Keywords Financial technology, Deep learning, Price prediction, Portfolio management, Trend forecast, Risk assessment

1 引言

金融的核心是在不同时间、不同空间的价值交换。金融在国家经济建设和社会发展的过程中发挥着不可或缺的作用,同时金融数据也是展现市场表现最直观的形式。20 世纪,人工智能技术开始应用于金融领域以推动金融行业的发展,技术革新成为金融创新的重要推动力,技术和金融的融合体——金融科技(Financial Technology, Fintech)应运而生,金融科技的浪潮席卷全球。金融科技的核心是利用数据科学和

人工智能技术推动经济和金融业务向智能化、个性化方向发展,从而提升金融系统的效率、客户体验、风险控制水平以及优化成本和效益等。

金融领域中的数据包括文本、视频、音频和时间序列数据等,这类数据往往具有高噪声、高维度、时间依赖等特点。早期金融市场的分析建模方法主要为传统机器学习方法^[1],主要包括马尔可夫模型(Markov Model, MM)、支持向量机(Support Vector Machine, SVM)、逻辑回归(Logistic Regression, LR)、线性回归模型、随机游走模型(Random Walk,

基金项目:国家自然科学基金(62072077, 62176043);国家重点研发计划(2019YFB1406202);四川省科技计划(2020GFW068, 2020ZHCG0058, 2021YFQ0007)

This work was supported by the National Natural Science Foundation of China(62072077, 62176043), National Key Research and Development Program of China(2019YFB1406202) and Sichuan Science and Technology Program(2020GFW068, 2020ZHCG0058, 2021YFQ0007).

通信作者:钟婷(zhongting@uestc.edu.cn)

RW)、向量自回归模型(Vector Autoregression, VAR)、自回归综合移动平均模型(Autoregressive Integrated Moving Average, ARIMA)以及神经网络(Neural Networks, NN)等方法,它们常被用作与新的学习模型进行比较的基线。

随着深度学习的蓬勃发展,深度学习技术已广泛应用于各个领域,例如图像识别、自然语言处理、医药生物以及金融等领域,并取得了瞩目的成果^[2]。深度学习在金融市场上的应用,主要是利用股票、货币、商品以及衍生品等交易数据来预测未来市场趋势。这些金融时间序列通常是非线性、不稳定和具有高噪声的数据,因此如何有效地建模金融时间序列信息以更好地预测市场规律是一个巨大的挑战^[3]。

目前有许多不同的深度学习模型已应用于金融领域,例如深度多层感知器(Deep Multilayer Perceptron, DMLP)、深层信念网络(Deep Belief Networks, DBN)^[4]、卷积神经网络(Convolutional Neural Networks, CNN)^[5]、循环神经网络(Recurrent Neural Network, RNN)、长短期记忆(Long Short-Term Memory, LSTM)^[6]、自编码器(Auto-Encoder, AE)、图神经网络(Graph Neural Network, GNN)^[7]和语言模型Transformer^[8]等。深度学习在金融市场中的应用主要分为两大类。第一类应用的研究重点在于如何设计更好的深度学习模型结构以获得更准确的预测结果。第二类应用侧重研究如何利用深度学习技术从海量文本数据(如财经新闻、在线社交媒体上的评论和公司财务报告等)中挖掘与金融相关的先验信息,从而改善金融市场的预测性能。

人工智能技术应用于金融领域是近年来一个研究热点。伴随着深度学习的蓬勃发展,出现了大量相关的综述文章,但目前关于深度学习在金融领域应用方面没有全面且系统的文献综述。之前的研究更关注深度学习在金融领域的某一特定应用,且大多数集中在股票预测^[9-10]等方面。例如,文献[9]中总结了近年来支持向量机、神经网络、深度学习等基于机器学习模型的股票预测技术,但其研究的重点不是深度学习,且缺乏对除股票预测外的其他应用的研究。文献[10]基于统计学习、模式识别、机器学习及情绪分析这4个方面回顾了股票分析和预测方面的研究成果,同样仅对

金融领域中股票预测这一应用进行分析总结。文献[3]总结了深度学习在金融领域中的市场运动预测、文本信息处理以及交易策略改进方面的应用,但在金融领域的细化程度不够且叙述较少。最新的综述^[11]详细介绍了深度学习常用的模型,并细分了金融领域的应用,但其重点是金融时间序列预测的深度学习技术。我们的综述更关注最近五年(2016—2021年)深度学习技术应用于各个金融领域的发展现状。

本文通过梳理目前关于金融科技方面的研究文献,总结了深度学习在各个金融领域的应用,具体包括:提炼了基于深度学习的金融科技研究的主要流程和实现方法;对近五年关于金融数据预测和分类任务进行了详细的整理并按具体应用分别进行阐述;分析了利用自然语言处理技术进行情感分析和文本挖掘来改善预测结果的方法;归纳出了深度学习方法之间互相结合或与其他领域的技术进行多模态融合等提高预测准确率的途径。本文的工作可以帮助相关研究人员全面地了解基于深度学习的金融领域科技研究现状,并提供研究和应用的思路和路线图。

本文第2节对深度学习在金融应用中的总体情况进行总结;第3节针对深度学习技术在金融科技中的应用进行系统的分类整理;第4节展望了深度学习在金融领域未来的一些研究方向,并对需要解决的问题进行概括;最后对全文进行总结。

2 深度学习应用于金融科技的主要框架

本文根据现有的研究总结了深度学习应用于金融科技行业的主要研究流程框架。如图1所示,该框架包括4个部分:数据集选择、数据预处理、模型构建以及模型性能评估。数据集选择为训练模型确定所需数据集,包括时间范围的确定以及训练集和测试集的分配。数据预处理阶段包括对数据进行噪声消除,降低输入变量维度,以及规范化和标准化。模型构建阶段包括深度学习模型和算法的设计,执行训练过程并调整训练参数。定义评估指标,进行模型性能的评估,并根据评估指标不断改进和优化模型。

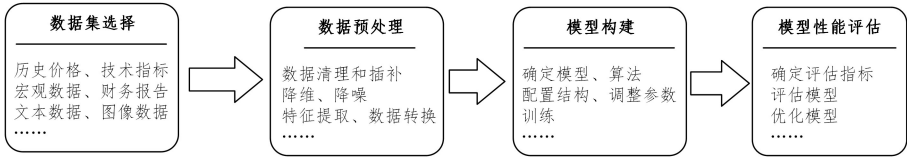


图1 金融科技研究的一般框架

Fig. 1 General framework of FinTech

2.1 数据集选择

金融市场预测的第一步是选择学习模型的输入数据,这决定了预测模型的可行性。金融数据大致分为7类,包括历史市场数据、技术指标、文本数据、财务报告、宏观经济数据、财务比率和图像数据,如表1所列。

除表1列出的7种数据类型外,还包括在预测商品价格时往往需要考虑的多种相关环境因素。例如在预测电价时,需要温度、电网负荷和天气等有意义的解释变量作为输入^[19]。金融风险评估涉及公司/个人的财务信息,其中公司

财务信息包括息差、恢复率、行业和地区等,个人财务信息包括年龄、债务比率和月收入等。检测信用卡欺诈行为时研究人员往往选择历史交易信息作为输入数据^[20],该信息包含用户基本信息、交易信息和行为记录等。

历史市场数据、技术指标、宏观经济数据、财务报告和财务比率都属于结构化输入数据,这类数据通常以表格形式呈现,其特征或属性可以描述为表的某一列。非结构化数据,如金融新闻、社交媒体、搜索引擎中的文本数据以及图像数据,须通过预处理转换为数字信息才能用作模型的输入和训练。

表 1 主要金融数据集分类
Table 1 Financial dataset classification

数据集类型	描述	常用数据集/数据类型
历史市场数据	金融资产市场中发生的所有交易活动,例如开盘价、最高价、最低价、收盘价和交易量(Open, Close, High, Low, Volume, OCHLV),股票收益及每日汇率等	美国标准普尔 500S&P500 ^[12] 、韩国综合股价指数 KO-SPI ^[13] 、中国沪深 300 指数 CSI300 ^[14] 和美国道琼斯工业指数 DJIA ^[15]
技术指标	通过数学公式计算出来的基于金融市场数据(主要是价格和数量)的指标集合	移动平均线、指数移动平均线、移动平均收敛发散、相对强弱指数
文本数据	来自于社交媒体、财经新闻和搜索引擎的海量文本数据	推特、新浪微博、路透社彭博社新闻等
财务报告	由投资银行或研究公司提供的报告,如资产负债表、收益表以及从出售或购买股票的建议中提取的数据	公司的业务模式与活动 ^[16]
宏观经济数据	反映了消费者物价指数、国内生产总值等国家或部门的经济状况	汇率 ^[17] 、商品价格 ^[17] 和居民消费价格指数 ^[18] 等
财务比率	可以从财务报表中获得,常用于信用评估和资产预测等	流动性、可解性、劳动生产率、资本生产率、利润率、投资回报率和收益率变化等
图像数据	与金融数据有关的图像	烛台图和财务报告单等

2.2 数据预处理

由于金融数据包含许多历史数据点,这些数据本质上是非线性、高维和高噪声的,这些特性使模型难以对该数据进行分析 and 建模。通过数据预处理可以过滤数据集中不相关的特征和噪声,降低数据的复杂性,使金融数据能够用于建模,进而提高预测模型的准确性。金融科技领域常用的数据预处理技术如下。

(1)数据插补和纠正:将金融数据输入模型时需要考虑具有不同采样频率的市场数据和收益、利润率等基本面数据的一致性。为避免过程中遗失重要信息,可以将采样频率较低的数据按照从上一有效观察值到下一有效观察值向前传播的方式依次进行插值或补全^[21]。

(2)数据降维:金融市场的预测受到众多因素的影响,如动态随机的市场条件和各种错综复杂的环境因素等。若直接使用原始的高维数据会导致算法复杂程度高,且会干扰预测结果,降低准确性,因此降低原始数据复杂性的同时尽量保留住原始数据特性至关重要。常见的降维技术有数据分割、主成分分析、度量学习和数据聚类。

(3)数据去噪:金融市场数据中的噪声会导致预测误差较大,通过数据降噪技术可以减小噪声数据的影响,如使用聚类 etc 等常用去噪技术。

(4)数据规范化:数据规范化包括归一化和标准化。特征归一化指将所有特征值缩放到[0,1]或[-1,1]之间;特征标准化指将所有数据缩放为均值为 0、标准差等于 1 的数据。规范化后的数据作为深度学习模型和数据挖掘模型的输入,可以避免原始数据中较大值对模型训练过程的影响,从而提高模型的鲁棒性和效率。

(5)特征提取与选择:特征提取过程是从高度随机的金融数据中学习到包含数据本身的原始属性以及隐式时序依赖关系的特征。特征选择目的则是从提取到的所有特征中选择有重要意义的特征输入模型进行训练,这个过程也会降低特征的维数和算法复杂性。

2.3 基本深度学习模型

根据大量关于金融科技技术的研究,本文总结了 10 种常用的深度学习模型,分别是 DMLP、受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)^[22]、DBN^[4]、CNN^[5]、RNN^[23]、AE^[24]、生成对抗网络(Generative Adversarial Network, GAN)^[25]、GNN^[7]、Transformer^[8]和深度强化学习(Deep Reinforcement Learning, DRL)^[26]。

DMLP 本质上是全连接神经网络,通过适当的激活函数

来处理非线性问题,具有多层非线性的分层学习能力,从而能够利用较少数量的计算单元对复杂函数进行建模。

RBM 由包含对应输入参数的输入(可见单元)和对应训练结果的隐藏单元组成。DBN 是由 RBM 层层堆叠而成,以探索变量之间复杂的依赖关系。与单个 RBM 相比,堆叠模型具有更强的学习能力,并能在一定程度上缓解训练过程中的过拟合问题。

在金融科技领域中,CNN 从特定时间点的金融市场中提取特征图数据或直接将金融市场的财务图像数据作为模型输入,并从中提取相应时期金融市场的特征,进而构建学习和预测模型。CNN 常用于金融时间序列预测,如股票预测、波动率预测、欺诈检测和风险评估等。近年来,随着 CNN 日趋成熟,研究人员发现深层 CNN 存在传输丢失特征信息、网络收敛困难以及梯度消失/爆炸等问题,使得网络无法有效地训练。残差网络(Residual Network, ResNet)^[27]通过将输入信息跳跃连接到输出,在一定程度上解决了深层 CNN 信息丢失和梯度消失问题,从而降低了训练的难度。

RNN^[23]使用反馈连接来调用历史时间步长的神经状态,构建映射以保存来自过去输入的相关信息,能够通过历史信息预测下一个时间步的数据,适用于对多元时间序列数据的动态行为进行建模。特定时刻的金融数据通常包含与历史价格和趋势的某种联系,因此 RNN 非常适用于金融时序数据的处理。但随着 RNN 网络的加深和时序数据跨度的加大,梯度消失/爆炸问题开始影响预测结果。为解决这一问题,LSTM 网络被提出。LSTM 在解决 RNN 长期依赖中梯度消失和爆炸问题的同时,可以有选择地丢弃不相关的细节,保留重要的相关信息。在金融应用中,LSTM 具有能够在一段时间内保留历史数据中有效信息的能力,因此在时间序列建模过程中往往能更好地揭示延迟状态中非线性时间序列的相关性。门控循环单元(Gate Recurrent Unit, GRU)^[28]与 LSTM 非常相似,也是 RNN 的一种扩展。LSTM 与 GRU 都可以缓解标准 RNN 的梯度消失问题,它们的主要区别在于门(Gate)的应用方式。为了从文本数据中捕获上下文参数信息,每个 GRU 单元都封装了不同时刻的信息,它的更新门取代了 LSTM 中的输入门和遗忘门,直接应用于历史的隐藏状态,无需像 LSTM 单元中那样使用额外的存储单元。由于 GRU 的参数少于 LSTM,因此 GRU 训练速度会比 LSTM 快,但是当数据规模很大时,LSTM 往往会带来更好的结果。

AE 模型主要包括编码和解码两个部分。编码器将输入数据编码成潜在表征,而解码器根据该潜在表征重构输入

数据。在训练过程中通过最小化重构误差来训练模型参数。此外,由于在编码过程中可以将输入向量变换为低维的潜在表征,因此 AE 也经常被用于数据降维。变分自编码器(Variational Autoencoder, VAE)^[24]是一种广泛使用的基于似然自编码器模型,它利用变分推断对潜在空间施加限制使其拟合已知分布,解决了普通 AE 在构建输出时可能不连续的问题,已被广泛用于金融领域^[29-32]。

GAN^[25]由生成器 G 和鉴别器 D 两个不同的神经网络组成。生成器 G 负责生成尽可能真实的数据来混淆鉴别器 D;而鉴别器 D 是一个二分类的分类器,用于判断输入数据是实际数据还是生成数据并向生成器 G 提供反馈,从而形成一个动态的博弈过程。GAN 在金融时间序列建模过程中能够恢复金融时间序列的统计属性,例如线性不可预测性、波动率和杠杆效应等^[33],并被用于欺诈检测^[34]等应用。

将深度学习方法在图上进行扩展,由此产生了一个新的研究热点图神经网络(Graph Neural Networks, GNN)^[7]。GNN 从节点和图拓扑的局部信息中有效地学习全局属性,相较于其他深度学习模型能够提供更好的概括性以及关系推理能力。现有的 GNN 模型,如图卷积网络、图自动编码器、图注意力网络等,已被广泛应用于金融情绪分析、信用评分和欺诈检测等方面,并取得了较为瞩目的成果^[35]。

近年来,一种完全基于注意力机制的模型架构 Transformer^[8]在自然语言处理中取得了巨大的成功。Transformer 采用多头自注意力机制来捕获全局上下文信息,以非循环方式对时间依赖性进行建模,从而增强了学习长期依赖关系的能力。2018 年 Google 开发了一种双向 Transformer 框架 BERT (Bidirectional Encoder Representations from Transformers)^[36],该框架自提出以来便在多项任务自然语言处理任务上均取得了最佳的性能^[37-38]。目前在金融领域中,Transformer 和 BERT 常用于金融情绪分析^[39]、股票预测^[40]以及波动率预测^[37]等。

DRL 结合了深度学习和强化学习,从而能够有效与外部环境进行交互。DRL 在金融领域主要应用于投资组合优化。由于 DRL 可根据市场反应来学习交易操作和捕获最佳交易时机,在动荡的金融市场中动态地重新分配投资组合,从而最大程度地降低风险和最大化回报^[41],因此 DRL 技术也被常用于股票和外汇趋势预测等。

2.4 模型性能评估

在整个框架中,判断模型的性能表现并根据结果不断地对模型进行优化和改进不可或缺。金融科技预测模型的评估指标分为 3 种类型:回归任务指标、分类任务指标和其他指标。

(1) 回归任务常用指标

回归指标用于衡量预测的时间序列与实际时间序列之间的接近程度。常用的度量标准包括平均绝对误差(Mean Absolute Error, MAE)、平均绝对百分比误差(Mean Absolute Percent Error, MAPE)、均方误差(Mean Squared Error, MSE)、均方根误差(Root Mean Squared Error, RMSE)以及归一化均方误差(Normalized Mean Square Error, NMSE)等。这些误差指标的值越小,意味着预测结果更准确。

(2) 分类任务常用指标

分类任务的评估指标用于衡量在资产价格趋势预测上的

表现。常见的度量指标有方向预测准确性(Accuracy, ACC)、马修斯相关系数(Matthews Correlation Coefficient, MCC)、精确率(precision)、召回率(recall)、F-值以及接收者操作特征曲线(Receiver Operating Characteristic Curve, ROC)和曲线下面积(Area Under Curve, AUC)等。ACC 是正确预测在预测数据总数中所占的比例。MCC 则用于描述二分类任务中实际分类与预测分类之间的相关系数。精确率指正确预测为正的样本的占全部预测为正的样本的比例。召回率反映的是正确预测样本占正确样本的比例。F-值对精确率和召回率进行了加权平均。ROC 曲线是以真阳性率为纵坐标、假阳性率为横坐标绘制的曲线,AUC 是 ROC 曲线下的面积,通常 AUC 值越大的分类器,分类越准确。

(3) 其他指标

其他指标包括评估交易模拟结果,即采用的交易策略是否具有利润或风险。常用总收益、年化收益率、投资组合价值、夏普比率(Sharpe Ratio, SR)、最大回撤(Maximum Draw-down, MDD)等作为衡量指标。在金融领域中,总收益指自投资以来总共获得的收益,年化收益率为投资一年所获得的收益,投资组合价值即全部买入的资产价值的总和。SR 是综合考虑收益和风险的指标,通过调整投资风险来衡量投资绩效的常用方法,该方法的好处是即使风险不同,也可以比较不同的收益。MDD 代表了资产价值最大可能的跌幅,可以用于评估风险。SR 和 MDD 的计算式分别如式(1)和式(2)所示,其中 R_p 是投资组合的收益, R_f 是无风险利率,而 σ_p 是投资组合的超额收益的标准差, P_t 和 P_τ 分别是 t 时刻和 τ 时刻的价值。

$$SR = \frac{R_p - R_f}{\sigma_p} \quad (1)$$

$$MDD(T) = \max_{t, \tau \in (0, T)} [P_t - P_\tau] \quad (2)$$

3 深度学习在金融科技中的应用

根据预测目标的输出是一个或多个连续的数值还是一组有限的类,即特定价格还是价格走势,将现有的深度学习技术在金融领域中的研究分为预测任务和分类任务两大类,详细的应用分类如图 2 所示。

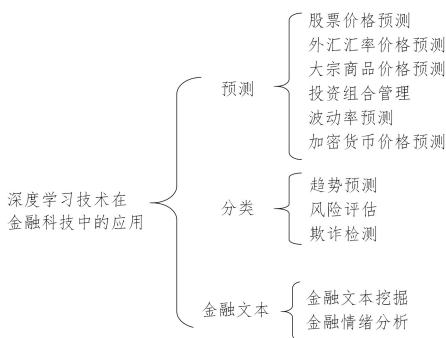


图 2 金融应用研究的分类

Fig. 2 Taxonomy of financial application research

诸多因素以复杂的非线性方式影响着金融市场中股票、加密货币、外汇和大宗商品等资产的价格,在训练过程中探索这些因素与资产价格之间的关系并将这些影响因素作为学习模型输入数据的一部分,能提供比仅使用大型历史价格数据集更好的性能。

3.1 预测和分类

3.1.1 股票价格预测

股票价格预测旨在为投资者提供正确的投资咨询服务以及投资建议,是当下金融科技领域一个重要的研究热点。股票交易涉及跨时间的价值交换,并和未来盈利或亏损状态相关联。预测股票价格的最大问题是股票原始时间序列是动态的、非线性的、高维的、非参数的和高噪声的,因此预测股票价格是极具挑战的。根据预测时间范围,需要动态选择输入参数,包括高频交易和日内价格变动到每日、每周甚至每月的

股票收盘价。此外,技术面分析、基本面分析、社交媒体反馈和投资者情绪也是用于股票预测模型的不同参数。研究者通过将预测出的股价与实际股价进行对比,并利用 MSE, RMSE, MAE 以及 NMSE 等指标来衡量模型预测股票价格的准确率。股票价格预测的核心是设计一种预测准确率高的模型。随着研究的深入,深度学习技术应用于股票预测由最初的单个模型发展到越来越多地使用集成模型并用于股市预测。表 2 列出了近年来重要的基于深度学习的股票预测方法。

表 2 深度学习方法预测股价
Table 2 Deep learning methods for stock price prediction

分类	输入数据	模型	评估指标	基线	文献
股票 预测	历史市场数据	LSTM	收益/损失, SR	RRL	[43]
		DRL	总利润	—	[12]
		DMLP	MAE, NMSE, RMSE	AR, ANN	[13]
		SFM	MSE	AR, LSTM	[46]
		BSd-RNN	RMSE, MAE, MAPE	DWT-FFNN, ARIMA 和 GARCH	[42]
		带有嵌入层的 LSTM	MSE, ACC	AELSTM, ELSTM, DBN, DBN-MLP, DMLP	[47]
		CEEMD-CNN-LSTM, EMD-CNN-LSTM	RMSE, MAE, MAPE	SVR, DTR, CNN-LSTM, LSTM	[48]
	历史市场数据+宏观经济数据	LSTM	收益, 年波动率, SR, ACC	GA-SVR	[49]
	历史市场数据+文本数据	CNN	ACC, MCC, 利润	BoW+SVM, 结构化事件元组+ANN	[50]
	历史市场数据+技术指标	GAN	ACC, RMSRE,	ARIMA-GARCH, ANN, SVM	[44]

(1)股票价格预测的单一模型

针对股票预测,RNN 及其变种成为最优先选择的深度学习模型。例如,Hajiabotorabi 等^[42]采用优化的 RNN 预测了 4 种不同的股票价格。为了解决高噪声股票市场数据没有可学习的监督信息的问题,并保证交易决策的鲁棒性,Si 等^[43]采用四层全连接网络和 LSTM 对中国股市进行预测,通过不同权重的目标对模型进行训练,获得了更好的预测效果。

AE 和 DRL 等也是研究人员常用的深度学习模型。在一项多特征研究中,Jeong 等^[12]使用 DRL 预测 S&P500 指数价格,并通过迁移学习方法防止因财务数据不足而过度拟合的问题,该方法使得在交易系统 中的交易总利润成倍增加。Chen 等^[14]利用沪深 300 期货中的高频交易数据,将 AE、RBM、径向基函数神经网络和极限学习机在大、中、小 3 种规模的数据集中进行股票价格预测,并比较了这几种不同机器学习模型的性能,发现深度学习的预测性能优于传统机器学习方法,并且不同规模的数据集实验对比结果表明增加样本量会大大提高深度学习的预测准确性。

(2)股票价格预测的集成模型

单一深度学习方法有其自身的优缺点。如 DMLP 具有解决复杂非线性模式的能力,但局部最小化会影响其训练过程;LSTM 虽然能够长时间记住信息并分析数据中的交互作用和隐藏模式,但在写入和读取数据时缺乏索引,计算费时。因此,部分研究尝试将两个或多个深度学习模型融合,以结合单一模型的优点构建新模型来改善性能。现有的许多研究已经表明集成学习方法比单一深度学习模型具有更好的性能。

Chong 等^[13]将 DMLP 应用于自回归模型来预测基于高频数据进行交易的 KOSPI 中的股票价格,并展示了 DMLP 和自回归模型集成具有良好的预测性能。文献[44]是首个将

GAN 用于价格预测的研究。该研究通过 LSTM 与 CNN 进行对抗训练获得的混合模型 GAN-FD 来预测股价,采用 13 个技术指标作为输入数据来预测中国股票市场数据,具有优于其他基准方法的预测能力。此外,该研究发现模型更新周期对预测性能有影响,较小的模型更新周期可以获得较好的预测性能。文献[45]使用堆叠式降噪自编码器(Stacked Denoise Autoencoder,SDAE)结合 LSTM 的方法来提高预测精度。Zhang 等^[46]受离散傅里叶变换的启发,设计了一种状态频率记忆(State Frequency Memory,SFM)循环网络,模拟股票价格波动的频率,然后用反傅里叶变换的方式进行非线性映射,从而预测具有多种频率交易模式的股票价格。

最近,Pang 等^[47]基于词向量的方法将数据向量化以减少向量维数,从而在低维空间进行表示学习,提出了带有嵌入层和带有自编码器的 LSTM 网络以预测中国股市价格。其中,嵌入层和自编码器用于过滤掉无关信息,从而减少输入数据的维数。文献[48]在 CNN-LSTM 混合的基础上引入频率分解算法来降低数据序列复杂性。

3.1.2 外汇汇率预测

随着经济全球化,汇率变化的不确定性很容易造成货币危机,甚至金融危机。对国家而言,汇率预测可以为指定合适的汇率政策提供指导意见,从而避免或减轻汇率波动对国内经济的不利影响。对于个人投资者而言,准确的汇率预测可以帮助投资者及时调整投资策略,降低资产风险,并获取更多利润。基于深度学习的汇率预测方面的部分研究见表 3。

文献[51]在预测欧元与美元(EUR/USD)汇率时指出,在给定数量不断增长的数据的情况下,具有自主学习能力的 LSTM 比 ARIMA 具有更好的预测性能。Shen 等^[53]通过共轭梯度算法来不断微调 DBN 模型以预测英镑(GBP)、巴西雷亚尔(BRL)、卢比(INR)与美元的每周汇率过程中的误差

权重矩阵,以最小化误差。在文献[53]的基础上,Zheng 等^[52]通过共轭梯度算法改进后的 DBN 模型预测人民币(CNY)、INR 与 USD 的汇率,使用多元方差分析来确定输入节点、

隐藏节点的敏感范围,并分析了 DBN 的节点和隐藏层数对预测误差的影响。实验表明,随着隐藏层数的增加,隐藏层节点对 DBN 误差的影响减小。

表 3 深度学习方法预测汇率

Table 3 Deep learning methods for exchange rate prediction

分类	输入数据	模型	汇率	评估指标	基线/比较模型	文献
汇率 预测	历史市场数据	LSTM	EUR/USD	RMSE	ARIMA	[51]
		DBN+CRBM	CNY/USD,INR/USD	MAPE	DMLP	[52]
		DBN	GBP/USD,BRL/USD,INR/USD	RMSE,MAE,MAPE,ACC	DBN,FFNN,RW,ARMA	[53]
		SAE+SVR	USD,GBP,EUR,JPY,AUD,CAD,CHF 的组合	MAE,MSE,RMSE	ANN,SVR	[54]
		AE+CNN	USD/GBP	SR,收益率	MLP,B&H	[55]
		CDBN-FG	EUR/USD,GBP/USD	利润	CDBN,BP-FG	[56]
	历史市场数据+宏观经济数据	DC-LSTM	USD/CNY	ACC,MAE,RMSE	ARIMA,SVR,CNN,LSTM	[57]

Shen 等^[54]使用集成堆叠式自编码器(Stacked Autoencoder,SAE)网络和支持向量回归(Support Vector Regression,SVR)的集成模型 SAE-SVR 预测美元、英镑、欧元、日元(JPY)、澳元(AUD)、加元(CAD)和法郎(CHF)等货币之间的汇率。SAE-SVR 模型由一个输入层、 K 个隐藏层和一个 SVR 输出层组成,训练过程中每个隐藏层都可以学习数据集的部分特征,并将此特征作为下一隐藏层的输入,以此类推,直到第 K 个隐藏层提取的高维抽象特征,然后将提取到的特征送入 SVR 模型中进行预测,最后对整个 SAE-SVR 训练模型进行反向传播微调来降低预测结果的误差。研究表明,SAE-SVR 模型在其中 21 个货币对中的预测性能优于传统的机器学习方法。Korczak 等^[55]在基于多主体的交易环境中结合使用 AE 和 CNN 作为预测模型,利用几种不同的输入参数

来预测 GBP/PLN 汇率,结果表明该混合模型降低了预测误差。Zhang 等^[56]将模糊理论与深度神经网络相结合,先将数据输入到模糊粒化体系结构中,再使用连续深度信念网络来预测 EUR/USD 和 GBP/USD 的汇率。Cao 等^[57]考虑到影响汇率预测的多个因素之间的相互作用即耦合,提出了深度耦合的长短期记忆结构以实现 USD/CNY 的汇率预测。该结构是由 LSTM 堆叠而成的分别捕获影响汇率预测的市场级和宏观级耦合,然后将这两部分的耦合馈入堆叠 LSTM 用于模拟市场级耦合和宏观级耦合之间的深层交互。

3.1.3 大宗商品价格预测

另外有许多研究关注一些贵金属商品如黄金、白银和铜的价格预测,还有部分研究关注能源商品如天然气、碳、电和石油的价格预测,具体研究情况如表 4 所列。

表 4 深度学习方法预测大宗商品价格

Table 4 Deep learning methods for commodity price prediction

分类	商品类型	输入数据	模型	评估指标	基线/比较	文献
商品	铜价	历史市场数据	Elman RNN	RMSE	ARIMA,MLP	[60]
	一般商品	历史市场数据+技术指标	CNN	ACC	LR,SVM	[17]
		技术指标	DMLP,RNN	ACC	—	[58]
	金价	历史市场数据+宏观经济数据	DBN	RMSE,MAPE,MAE	传统 BPNN,GA-BP,ARIMA	[59]
		历史价格	CNN-LSTM	MAE,RMSE	SVR,DMLP,LSTM	[61]
	原油	历史市场数据	ARMA+DBN,RW+LSTM	MSE	RW,ARMA,DBN,LSTM	[62]
			EEMD-LSTM	RMSE,MAPE	EEMD-ELM,EEMD-KRR,EEMD-LSSVR,EEMD-ANN	[63]
	电价	历史市场数据+环境变量	GRU-DMLP,CNN,LSTM-DMLP,DMLP	MAPE	移动平均模型,机器学习与统计方法,基于 fARX 的模型	[19]
		历史市场数据	WT-SAE-LSTM	MAE,MAPE,RMSE	LSTM,BiLSTM,BP,ELM	[64]

(1)贵金属商品

在商品价格预测中,输入深度学习网络的变量包括单变量和多变量这两种形式。单变量输入即仅将历史价格数据作为深度学习网络的输入,多变量方法即除历史价格数据外还将其他影响价格的变量作为模型的输入。实验表明在输入变量均有效的情况下,多变量模型通常比单变量模型表现更好^[17,58-59]。

现代社会中,金属与许多行业如电线、建筑和投资等紧密相关,因此预测金属价格尤为重要。如 Lasheras 等^[60]仅将纽约商品交易所 COMEX 的每日收盘价作为 RNN 网络的输入来预测铜的现货价格。另有部分研究关注多种因素对黄金价格

的影响。如 Widegren 等^[58]将外汇、索引数据集、技术指标等作为输入数据,用于预测黄金价格。Dingli 等^[17]还探讨了石油对黄金价格的影响,将黄金和石油价格的时间序列数据作为输入数据,使用 CNN 模型预测黄金的价格。Zhang 等^[59]使用 DBN 模型将原油价格、联邦基金利率和美国的消费者物价指数等数据作为输入变量对黄金价格进行预测。与传统的神经网络模型和线性模型相比,DBN 模型具有更高的准确性。Livieris 等^[61]在黄金价格的预测中将 CNN-LSTM 混合模型与 SVR,FFNN 和 LSTM 进行了比较,该模型的表现比其他模型较好。

(2)能源商品

原油和电力的价格预测是能源商品价格预测中研究最为

广泛的。Chen 等^[62]基于 DBN,LSTM,ARMA 和 RW 这 4 种模型以及它们的混合模型来预测 WTI 原油价格,构建出来的混合模型包括 RW-DBN,ARMA-DBN,RW-LSTM,AR-MA-LSTM。使用 MSE 来衡量模型预测性能,以上几种模型在预测 WTI 原油价格时,RW-DBN 的 MSE 最小,获得最佳性能。文献[62]的研究验证了 DBN 和 LSTM 这两种深度学习方法可用于对原油价格走势预测的非线性动力学建模。部分原油价格预测的研究还结合了其他时序处理方法。如 Wu 等^[63]基于 LSTM 和总体经验模态分解方法(Ensemble Empirical Mode Decomposition,EEMD)提出了 EEMD-LSTM 模型预测原油价格,其中 EEMD 将原始信号分离成多个分量并从中提取相关信息。从实验结果来看,在预测原油价格的不同模型中 EEMD-LSTM 模型表现最佳,但该研究仅使用原油价格的高频信息预测价格走势。

通过深度学习技术提升电价预测准确度,发电公司不仅可以获得经济利益,还可以确保电网的稳定性。Lago 等^[19]将 4 种不同的深度学习模型即 DMLP,CNN,GRU-DMLP 和

LSTM-DMLP 模型用于预测电价,并与其他电价预测文献中提出的 23 种传统的统计模型(例如随机森林和梯度增强方法)进行比较。最近,作为处理各种规模数据的有效方法 WT 被应用于电价预测中。Qiao 等^[64]将 WT、SAE 和 LSTM 模型集成,构造出一个新的预测模型 WT-SAE-LSTM。实验结果表明,WT-SAE-LSTM 模型在预测准确度方面优于其他单一预测模型(如 LSTM),这表明该混合模型在一定程度上弥补了 LSTM 模型的弊端。

3.1.4 投资组合管理

在金融市场上,个人投资者通常希望了解他们投资资产收益率的变化、未来收益率的发展趋势以及应该采取哪些措施来获取最好的投资组合。投资组合管理是一个决策过程,往往与股票预测和货币预测相结合,需要将一定数量的资金分配给多个金融资产,通过不断调整分配权重实现最大化回报和最小风险。因此,在金融投资中将预测模型纳入投资组合管理具有极大的应用前景。表 5 列出了深度学习在投资组合管理方面的应用,这里我们同样按照单一方法和混合方法的方式进行总结。

表 5 深度学习方法预测投资组合
Table 5 Deep learning methods for portfolio forecasting

分类	输入数据	模型	评估指标	基线	文献
投资 组合	历史市场数据	CNN+RL	SR,总收益,MDD	B&H,Best Stock,Uniform Constant, 再平衡投资组合	[65]
		CNN,RNN,LSTM	投资组合价值,MDD,SR	Best Stock,UCRP,UBAH	[66]
		DDPG,PPO	SR,MDD	UCRP	[67]
		SRNN,LSTM,GRU	ACC,月收益	EWP	[68]
		HC-CNN,HC-LSTM	年化收益率,SR	B&H,布林带方法	[71]
	历史市场数据+ 技术指标	LSTM+MV	MSE,RMSE,MAPE,MAE	SVM,RAF,DMLP,ARIMA	[70]

(1)单一深度学习模型进行投资组合管理

投资组合管理需要模型能够与外部环境交互、自学习并具有决策能力。深度学习模型中,DRL 能够有效与外部环境进行交互,基于外部环境而行动以取得最大化的预期利益,同时具有深度学习的自学习能力和强化学习的决策能力,因此其是目前最适合应用于投资组合管理的模型。Jiang 等人在加密货币投资组合管理方面进行了许多研究,例如在文献[65]中利用 CNN 模型来预测一年内 12 种加密货币的价格,在选定的加密货币上应用了确定性策略梯度的 DRL 模型来解决权重的分配和调整问题;在训练过程中使用投资组合权重作为模型的输出,以实现加密货币管理的最大长期回报。文献[66]提出了可扩展的 DRL 框架,并基于 RNN,LSTM 和 CNN 3 种模型以及该框架进行加密货币投资组合管理,结果显示该框架的盈利能力超过传统投资组合选择方法。Liang 等^[67]采用了不同的 DRL 模型,包括深度确定性策略梯度 DDPG、近端策略优化 PPO 和策略梯度 PG 在内的最新算法,用于金融投资组合的相关决策,并评估了这些模型在中国股票市场中的表现。研究结果表明,PG 在股票交易方面比其他两种模型更优,DDPG 次之。

此外,部分研究人员试图在投资组合中应用 LSTM 对组合收益进行预测。Lee 等^[68]使用简单 RNN,LSTM 和 GRU 对未来股票收益进行预测,实验结果表明 3 种 RNN 模型中 LSTM 网络表现最好。此外,该研究还对 LSTM 网络的股票收益预测结果设定阈值,从而构建了基于阈值

的投资组合,获得了较好的性能。

(2)混合深度学习模型进行投资组合管理

许多研究表明,在优化投资组合管理方面,混合模型往往优于单一深度学习模型^[69-71]。为解决金融环境描述困难和交易执行动态性的问题,Deng 等^[69]基于深度神经网络和强化学习提出了一种新的结构增强深度神经网络 RDNN,利用深度神经网络自动感知动态的金融市场情况以进行特征学习,RL 与深度表示交互并作出交易决策。该研究还引入了模糊学习以减少原始时间序列中的不确定性,结果表明该结构能同时进行市场情况总结和最优行为学习。为了优化投资组合结构,Wang 等^[70]首先深入探讨了最优投资组合形成之前的资产预选过程,从而保证了最优投资组合的高质量输入,然后结合了 LSTM 网络和 Markowitz MV 方法来构建投资决策模型。该模型考虑了金融市场波动的长期依赖性,并从时间序列数据中捕获了股票的长期变化模式,以找到最佳模型来预测资产的累计收益。

3.1.5 波动率预测

金融资产的波动性预测是在目标时间段内预测金融资产价格的波动程度并衡量不确定性或风险的程度,在风险评估和资产定价方面至关重要。通常,较高的波动性表示较高的市场风险,直接影响金融市场的稳定和宏观经济。预测波动性的方法主要包括深度学习方法和传统机器学习方法,其中传统机器学习方法中基于 GARCH^[72]的方法使用得较多。

表 6 列出了常用于波动率预测的深度学习模型。

(1)单一模型的波动率预测

LSTM是深度学习方法中用于预测波动率的常用模型。Zhou 等^[73]利用股市数据和百度每日搜索中 28 个与金融市场和家庭消费相关的关键词数据作为 LSTM 的输入来预测 CSI300 的波动率。文献^[74]为解决在预测过程中容易引入误差以及计量经济学模型与深度学习模型优化目标不同等问题,引入负对数似然函数作为 LSTM 的损失函数,仅使用指数收益作为模型的输入,更公平地比较了深度学习模型和基于 GARCH 的模型。最终结果表明深度学习模型能更准确地预测波动率。此外,文献^[37]将新模型 BERT 应用于波动率预测,结合了文本和音频特征,

显著地改善了预测结果。

(2)混合模型的波动率预测

一些文献将 LSTM 和计量经济学方法混合来提高波动率预测的准确性。与以往仅将单个计量经济模型与单个神经网络模型相结合的方法不同, Kim 等^[75]开发了一种将 LSTM 模型与多个 GARCH 类型模型相集成的混合方法,使用 22 个交易日和解释变量作为输入数据预测韩国综合股票价格指数(KOSPI200)股票指数第二天的波动性。实验结果显示与结合单一 GARCH 类型模型的混合神经网络模型相比,结合两个或多个 GARCH 模型的神经网络模型的预测性能显著提升。文献^[76]则将 LSTM 和 GARCH 混合,用于改善对铜价波动性的预测。

表 6 深度学习方法预测波动率
Table 6 Deep learning methods for volatility prediction

分类	输入数据	模型	评估指标	基线/比较	文献
波动率	历史市场数据+文本数据	LSTM	MSE,MAPE	GARCH	[73]
		BERT	MSE	GARCH,基于文本的模型	[37]
		RNN-Boost	ACC,MAE,MAPE,RMSE	单一 RNN,DMLP,ANN,SVR	[77]
	历史市场数据	LSTM,DMLP	损失	ARMA-GARCH	[74]
		LSTM+GARCH	MAE,MSE,HMAE,HMSE	LSTM,GARCH,以及它们的混合模型	[75]
	历史市场数据+图像数据	CNN-LSTM	MSE	CNN,LSTM,SVR,GARCH	[78]

Chen 等^[77]提出的 RNN-Boost 模型在 RNN 模型中加入了自适应集成 Adaboost 技术,并通过潜在 Dirichlet 结合主题建模以及使用混合后验概率计算的情感特征提高预测模型的准确性。实验结果表明,该模型优于其他流行方法并在预测指数波动性方面有显著提升。最近, Vidal 等^[78]通过在多层 CNN 网络中加入 LSTM 构建了一种全新的方法,使用时间序列到图像的转换,同时能够捕获静态和动态序列数据,通过马尔可夫转换模型来预测黄金价格的波动性。

3.1.1.6 加密货币价格预测

加密货币价格飙升和暴跌之间的高波动性和快速过渡对预测加密货币价格构成了巨大挑战。随着加密货币市场的蓬勃发展,泡沫的形成及其急剧爆发可能会影响全球金融市场的稳定性,因此在最近几年吸引了众多研究者的关注。对加密货币的研究大多集中在价格预测和交易策略方面。与股票价格预测相比,关于加密货币价格预测的研究较少,而且这些研究中大多没有考虑文本信息对预测的影响。

深度学习方法预测加密货币价格主要依赖于两个组件:输入特征和预测模型。我们根据输入的来源分为仅使用加密货币价格、数量等技术指标的研究,以及在技术

指标的基础上加入如 Google 趋势、Twitter 文本等社会指标的研究。

文献^[79-81]仅使用历史市场数据作为模型输入。其中文献^[79]比较了贝叶斯优化的 RNN,LSTM 和 ARIMA 在预测比特币价格时的准确性。文献^[80]将深度学习方法用于预测比特币、数字现金和瑞波币的价格,并比较了 LSTM 和加密货币市场中常用的广义回归神经网络方法(General Regression Neural Network,GRNN)的预测性能,发现在信号长度和样本规模不是非常大时,与 GRNN 相比,LSTM 模型具有更高的准确率但训练更加耗时。同样,文献^[81]利用 CoinMarket 和 CoinDesk 存储库中每天、每小时和每分钟的数据,基于 DMLP 和 LSTM 来预测以太币的价格。结果表明,相比于 DMLP 模型,LSTM 在性能上略胜一筹,LSTM 模型对于长期依赖关系更加健壮和精确。文献^[82]首次集成深度学习、分解方法和基于元启发式优化算法来预测加密货币价格。该方法具有较好的可靠性,可以捕获加密货币时间序列的非线性特征。Lopes^[83]使用社会指标加技术指标作为 LSTM 模型的输入来预测比特币 BTC 和莱特币 LTC 这两种货币的价格,实验结果表明,使用了社会指标的预测模型效果更好。表 7 列出了目前深度学习技术用于加密货币预测的研究。

表 7 深度学习方法预测加密货币价格
Table 7 Deep learning methods for cryptocurrency prediction

分类	输入数据	模型	评估指标	基线	文献
加密货币	历史市场数据	贝叶斯优化的 RNN 和 LSTM	精确率,ACC, RMSE	ARIMA	[79]
		LSTM	RMSE	GRNN	[80]
		DMLP,LSTM	MAE,MSE, RMSE,MAPE	—	[81]
		EW-T-LSTM-CS	MAE, RMSE,MAPE	单一 LSTM	[82]
	历史市场数据+技术指标+文本数据	CNN,LSTM,SFM	MSE	—	[83]

3.1.7 趋势预测

预测资产的未来状态一直是金融市场投资者的兴趣所在。资产价格波动较为频繁且具有非线性特点,因此部分研究工作将预测资产价格问题转换为二分类问题,通过

价格变化是否超过阈值来判断价格是上涨还是下跌,这比预测具体资产价格更容易实现。CNN^[84-85],LSTM^[86],RNN^[86],DBN^[87],DMLP^[88]等是预测资产价格趋势的常用深度模型,详见表 8。

表 8 深度学习用于趋势预测
Table 8 Deep learning methods for trend prediction

分类	输入数据	模型	评估指标	基线/比较模型	文献
资产趋势预测	历史市场数据	DMLP,GBT,RF	回报,SR,MDD	简单 DMLP	[88]
		小波+CNN	ACC,MSE	MLP,CNN,LSTM	[84]
		Transformer	ACC,MCC	CNN,LSTM,ALSTM	[40]
	历史市场数据+技术指标	RNN,LSTM	ACC,F-score	决策树,RF,SVM,Adaboost,XGBoost,朴素贝叶斯,KNN,LR,ANN	[86]
		DBN	MSE,召回率,精确率	SVM,LR	[87]
		CNN-LSTM	精确率,召回率,ACC,F1-分数	PCA+ANN,SVR,CNN-cor,CNNpred	[90]
	历史市场数据+文本数据	VAE+GRU	ACC,MCC	HAN,ARIMA,TSLLDA,RAND	[29]
	历史市场数据+技术指标+图像数据	CNN-TA	精确率,召回率,F-分数,年回报	LSTM,MLP 回归	[85]
	历史市场数据+技术指标+宏观经济数据	CNNpred	SR,回报	PCA+ANN,ANN,CNN-cor	[89]
	历史市场数据+图像数据	DSPNN	ACC	RF,AdaBoost,LSTM	[91]

(1)单一模型预测资产趋势

文献[88]通过 DMLP、梯度增强树(GBT)和 RF 方法对次日 S&P 500 股票指数中的股票趋势进行判断,并将这 3 种不同的机器学习模型进行比较,指出深度学习模型没有表现出最佳预测性能,因为其没有进行详细的超参数优化。该工作还将这几种模型等权重组合在一起进行实验,结果显示等权重组合模型的性能优于单一模型。

对于金融时间序列趋势预测,许多研究者更倾向于使用 RNN 和 LSTM 模型。Nabipour 等^[86]通过对德黑兰股市的股票进行预测,比较了 9 种机器学习模型以及 RNN 和 LSTM 两种深度学习模型在预测股票走势方面的性能。研究者将近十年的股票历史数据中的 10 项技术指标作为输入数据并将其转换为二进制数据和连续数据,分别进行实验。结果表明,在使用二进制数据作为输入的评估中,深度学习方法是性能最好的,但与其他预测模型的差异不大。在使用连续数据作为输入数据时,RNN 和 LSTM 明显优于其他预测模型。将输入数据转换为二进制数据后,预测模型性能均有提高,这表明合适的预处理方法可以提高股票趋势预测准确率。为了捕捉金融时间序列中的极长依赖关系,例如几个月内的依赖关系,文献[40]提出了多尺度高斯先验改进的 Transformer 进行股票走势预测的方法,该方法能够捕捉金融时间序列的长期、短期和层次依赖性。

(2)混合模型预测资产趋势

由于非线性时间序列预测较为复杂,部分研究人员将预测分为两阶段进行,即先提取特征,然后将其作为模型的输入进行预测。文献[87]结合了 DBN 和 MLP 来构建预测模型,使用历史价格数据以及技术指标作为输入。与基线模型相比,混合模型均呈现出更好的结果。CNN 模型适用于多变量环境,因此也常被用于两阶段预测中的特征提取。为了解决一维输入张量^[84]和二维输入张量^[85]作为 CNN 的输入在提取特征过程中往往会忽略相关股票市场的潜在影响问题,文献[89]设计了一种三维输入张量构造方法,该方法能够从包含相关股票市场历史信息的多源数据中提取特征,从而提高

模型的预测性能。基于此,文献[90]做出了一些改进,虽然依然沿用三维输入张量构造方法进行特征提取,但结合了 LSTM 组成混合模型,最终实验结果表明三维输入张量和 CNN-LSTM 的组合可以明显改善模型的性能。类似地,Long 等^[91]的研究结合真实交易数据,基于 CNN 和 Bi-LSTM 构建了一个新的模型深层趋势预测神经网络 DSPNN,并建立了公司知识图谱来获得股票之间的相关性,从而预测股价趋势,并且将交易数据形成的矩阵作为 CNN 网络的输入进行特征提取,然后将 CNN 网络的输出与股票信息结合输入到 Bi-LSTM 中来预测中国的三只股票(中信证券、广发证券和中国平安)在下一个交易日的价格走势。该研究中不仅使用了投资者真实交易记录和公共市场数据融合构建知识图谱,而且使用 CNN 网络和基于注意力的 Bi-LSTM 集成来提高模型预测性能,实验结果表明 DSPNN 趋势预测的准确率已达到 70%甚至更高。

另有部分研究将集成模型作为整体进行训练,而不是单纯地将训练分为两阶段。Buczowski 等^[92]将 GRU 和 LSTM 集成成为一个新的模型,结合金融专家的建议来预测股票价格趋势,该集成模型在 ISMIS 2017 数据挖掘竞赛中获得了第五名的成绩。将深度学习模型与其他先进技术结合起来提高预测性能是常用方法。Xu 等^[29]受 VAE 的启发,提出了一种具有变分架构的新型解码器模型 StockNet,使用推特文本和价格信号作为输入,使用 VAE 对其进行编码并引入时间辅助,以捕获随机性,从而对股票走势作出预测。StockNet 尽管在技术上可行,但由于股票市场的高度随机性,此方法也具有泛化能力不足等问题。

3.1.8 风险评估和欺诈检测

风险评估包括破产预测、信用评分、信用评估、贷款/保险违约预测、债券评级、贷款申请、消费者信用确定、企业信用评级、抵押贷款选择决策、财务困境预测、逃税漏税预测等。由于企业、银行的安全和稳健性评定高度依赖于这些风险评估措施,因此正确识别风险状态至关重要。许多研究人员将深度学习应用于风险评估以提高准确性。

大多数风险评估的深度学习研究集中于信用评分和银行困境分类,可用于检测是否存在欺诈行为。但是,也有一些论文涉及抵押违约的可能性,风险交易检测或危机预测等,详见表 9。

表 9 深度学习用于风险评估和欺诈检测

Table 9 Deep learning for risk assessment and fraud detection

分类	具体分类	输入数据	模型	评估指标	基线/比较	文献
风险评估和欺诈检测	破产预测	财务报告+财务比率+图像数据	CNN	F1-分数, AUC	CART, LDA, SVM, MLP, AdaBoost	[94]
	信用评分	个人财务信息	DCNN	ACC, 误报/漏报	DMLP	[95]
		历史交易记录+个人财务信息	CNN+Relief	AUC, ACC	LR, RF	[99]
	欺诈检测	历史交易信息	LSTM	AUC	RF	[96]
			CNN, SLSTM, CNN-LSTM	ACC, AUC, ROC	—	[97]
			DAE	召回率, 精确率, ACC, F-分数	LR	[98]
			DMLP+RF	召回率, 精确率, ACC	LR	[103]
			历史交易信息+其它相关变量	LSTM, GRU	ACC	ANN, RNN
	违约预测	个人财务信息	CNN+RF	ACC, AUC	CNN, RF	[100]
	股市危机预测	历史价格数据+宏观经济数据	LR, CART, RF, SVM, NN, XGBoost, DMLP	AUC	—	[102]

(1)单一模型进行风险评估和欺诈检测

在 2008 年的金融危机之后,公司违约和财务困境预测在金融中变得越来越重要。Luo 等^[93]将 DBN 用于信用评分,将公司违约掉期(CDS)数据用于公司信用评级并输出相应的信用分级。在测试时与其他常用的信用评分模型如逻辑回归、多层感知器和支持向量机进行比较,结果显示 DBN 能够更有效、更准确地预测信用风险。Hosaka 等^[94]将 2 164 家公司的财务报表中得出的财务比率数据转换为灰度图像,使用 CNN 模型对图像进行训练和测试,从而进行破产预测。

金融机构可以利用消费者信用评分评估借贷人是否有违约风险,还有部分研究关注利用深度学习评价消费者信用。Neagoe 等^[95]分别使用 DMLP 和 CNN 网络,根据客户特征例如年龄、收入、就业、婚姻状况等,对客户信用进行了分类,两种模型之间的性能差异非常明显,深度 CNN 模型优于 DMLP 及其多层变体模型。

还有一些研究致力于识别交易中的欺诈行为。Roy 等^[20]使用 LSTM 网络对包含 8000 万笔信用卡交易的数据集进行分类,以检测是否存在信用卡欺诈。Jurgovsky 等^[96]使用 LSTM 从信用卡交易序列中检测信用卡欺诈,结果显示 LSTM 对线下交易检测的准确性较高。Heryadi 等^[97]探索了 3 种用于印度尼西亚银行信用卡欺诈检测的深度学习模型,分别是 CNN、堆叠 LSTM 和 CNN-LSTM 混合模型,分析了欺诈和非欺诈数据之间数据不平衡问题对性能指标的影响。之后,Al-Shabi 等^[98]发现 AE 模型对处理不平衡数据集更有效。

(2)混合模型进行风险评估

许多集成方法也被用于信用风险评估。2018 年,Zhu 等^[99]首次尝试将 CNN 应用于信用评分,并且基于 CNN 提出了 Relief-CNN 的混合模型。

消费者信用评分分类是通过将输入的消费者数据转换为二维像素矩阵,将得到的矩阵作为图像输入 CNN 进行训练,最后利用 CNN 对消费者数据转换得到的图像进行分类。在一个抵押贷款风险评估应用中,Kvamme 等^[100]引入了一种新的信用评分模型,该模型根据消费者交易数据使用 CNN 和随机森林分类器 RF 组合的模型来预测客户是否会违约。Yu 等^[101]研究了用于客户信用风险分类的 DBN 和 SVM 级联

混合模型,还利用重采样技术应对不平衡数据,最终结果显示该方法能够提高分类性能。Chatzis 等^[102]将深度学习技术和统计机器学习方法(如 SVM, RF, XGBoost 等梯度提升方法)结合来预测股票市场危机,该研究结果表明深度学习技术显著提高了分类准确度。

3.2 金融文本挖掘和情绪分析

研究者们发现除了历史市场数据外,投资者情绪、公司丑闻和金融新闻对金融市场的走向也有一定的影响。社交媒体、实时新闻、网络和微博的迅速发展,使用户能够在在线平台上表达对资产、服务或任何其他要素的态度或观点,但由于这些观点都是用自然语言表达的,机器系统很难理解新闻和网络消息的含义,更难于将其与资本市场联系起来^[104]。近年来,自然语言处理技术开始被用于金融科技研究,部分研究通过文本挖掘技术对新闻报道进行情绪分析、观点挖掘或事件提取,推动非结构化文本数据成为金融市场预测的重要信息来源。

3.2.1 金融文本挖掘

在线社交网络作为信息共享平台具有大量信息,但与金融市场相关的在线内容的质量、可靠性和全面性差异很大,其中很大一部分由低质量的新闻、评论甚至谣言构成。因此想要利用这些信息对金融市场进行预测需要提取其中相关且可靠的高质量信息,而使用文本挖掘技术结合深度学习进行是一个改善预测结果的重要途径。本节根据研究所采用的文本粒度(词、句和段落、文档)进行总结分析。

(1)词嵌入

Shi 等^[105]利用词向量技术,从路透社和彭博社的财经新闻中提取有效信息作为模型输入。该研究表明将从新闻文本中获得的特征输入到模型中可以提高分类准确率。Pinheiro 等^[106]Ding 等^[50]的模型和评估框架使用由字符级嵌入语言模型预训练的 LSTM 神经网络来探索金融新闻对公司股票价格的影响。结果表明,基于 LSTM 字符级别的模型与其他模型相比,其体系结构复杂度更低,性能更好。文献^[107]结合卷积、词嵌入、序列建模等方法从金融新闻中提取信息,然后加入技术分析指标,利用 CNN 和 LSTM 的组合模型 RCN 对股票价格进行预测,结果表明 RCN 的价格预测性能优于 LSTM。该研究还显示可以通过卷积网络从复杂的金融文本

数据中提取有效信息。另有部分研究将词向量技术和深度学习应用在风险预测中,如 Rawte 等^[108]通过 GloVe 词嵌入来表示单词,使用 CNN 网络对文本文档进行分类,从而对 219 家银行的风险进行分类和预测。

(2)句子和段落嵌入

词嵌入捕获单个词的含义,而不能捕获上下文的信息,并且文本的含义并不总是单词含义的总和。继词嵌入研究后,学者们开始尝试利用较大文本块的含义。句子向量技术常常用于银行和公司的财务风险分析中。为了简化在更广泛的上下文中检索定位相关部分和特定段落的过程,Rönnqvist 等^[109]采用句子级别的语义建模来进行银行困境预测,具体方法是从金融新闻文章中提取数据,通过句子嵌入提取固定长度语义句子向量,将其作为深度前馈网络的输入来预测是否存在银行困境。Cerchiello 等^[110]基于文献^[109]中开发的框架进行扩展,使用全连接神经网络,通过从财经新闻中挖掘金融信号对银行困境进行分类,证明了新闻数据在预测银行困境过程中的重要性。

在股价走势预测中,Akita 等^[111]同时使用数字信息和文本信息作为输入数据,其中文本信息通过段落向量技术将可变长度的新闻文章映射到固定长度的向量来获得分布式表示,然后使用 LSTM 对 10 家公司的收盘价进行预测。实验结果表明,文本信息的分布式表示优于单纯的数字数据方法和基于词袋的方法。Matsubara 等^[112]使用段落向量嵌入方法从路透社、彭博社的金融新闻文章中提取信息创建输入向量,通过深度神经生成模型(Deep Neural Generative Model, DGM)来预测 S&P500 股票指数每日股票价格走势。

(3)文档级

在文档层面,大多研究人员将金融文章作为一个整体,进行主题和事件提取分析。金融文章的语义被分解为多个主题和相关系数,或根据定义的事件类型和角色对象来提取有关事件的复杂结构化信息,以便系统捕获新闻和网络消息的含义。Wang 等^[113]先使用基于 LDA 的文本主题模型来提取索赔中出现的事故文本描述特征,再将该文本特征、传统数值特征和文本数据的处理经验作为 DMLP 模型的输入数据进行训练,以检测汽车保险索赔是否为欺诈。François-Lavet 等^[114]参考文献^[30]中的结构,同样直接从具有分层注意力机制的推特文本中挖掘新闻向量,结合基于 VAE 的 Stock-Net 深度生成模型预测股票趋势。

有关新闻驱动的股票预测任务中,词袋和名词短语这些浅层特征无法捕获结构化的实体关系信息,从而无法准确表示完整的事件。因此 Ding 等^[50]采用神经张量网络从路透社和彭博社新闻文本中提取事件并表示为向量,其中事件由“角色-动作-对象-时间”的四元组组成,例如“Google 于 2006 年 10 月 9 日收购了 YouTube”。然后将提取到的信息用于 CNN 网络对长期和短期事件的综合影响进行建模,从而进行价格预测和指导股票交易。实验结果表明,从新闻中提取事件信息作为输入可以提高预测性能。Liu 等^[115]提出了一种联合特征提取方法,该方法使用由事件元组和财经新闻构建的特征向量,并把所有提取到的特征向量作为 LSTM 网络的输入用于股票预测。除上述方法外,Hu 等^[30]考虑到每则金融新闻对股票市场的影响程度是不同的,提出了一种新的混合注意力网络。通过使用新闻嵌入和基于金融新闻的混合

注意力网络对具有不同影响力的金融新闻进行区分,模仿人类面对混乱的在线内容时的学习过程,最后将 DMLP 网络作为最终的判别网络对股票趋势做出更可靠的预测。

3.2.2 金融情绪分析

金融情绪分析也是金融科技领域的重要研究内容,大多通过统计社交媒体、金融新闻和评论等文本数据中情感词出现频率或对不同类型的情感词赋予不同的权重来计算情感值得到文本的情感倾向性,从而挖掘公众情绪和用户观点(尤其是负面观点)。正面评价增加表示该公司状况良好,股价可能会上涨,反之亦然。金融情绪分析最初被定义为二分类问题,但是随着研究的深入,衍生了细粒度的情感分类,即将用户情绪分为多种类型。金融情绪分析有利于个人在股票交易时做出明智的决定,同时相关机构可以利用这些信息来预测未来公司收益和股票波动情况。金融文本情绪分析方法大致可以分为基于词典的方法和基于机器学习的方法。

(1)基于词典的情感分析方法

基于词典的情感分析方法一般先将文本中特定的意见词或短语提取出来,然后在该领域的意见词典中查找其代表的情绪,最后通过平均或求和相关的分数来组合它们各自的语义方向,从而决定文本的整体情感。这种方法的优点是不需要收集数据和训练数据,而它的缺点是需要该领域的意见词典且不具有处理上下文相关词的机制。

近年来,研究人员密切关注在线社交平台(如 Twitter 和新浪微博等)文本,通过情绪词典分析公众情绪来提高金融市场预测的准确率。基于 Twitter 数据的情绪分析研究很多,例如文献^[116]将微博数据作为情绪分析来源,设计了一个用于预测金融市场趋势的深度随机子空间集成分类方法。该方法首先在与金融相关的社交媒体情感计算中采用基于词典的方法进行情感分析,然后结合得到的情感特征和价格特征来预测股票的价格。此外,部分研究使用其他金融文本数据如路透社和彭博社的新闻进行情绪分析。例如,Dang 等^[117]利用金融文本挖掘和情绪分析对每日财经新闻和股票价格数据提取特征后输入双流式 GRU(Two-stream Gated Recurrent Unit, TGRU)模型,以预测标准普尔 500 指数的价格趋势。结果显示所提出的词嵌入方法比普通的嵌入方法(如 Glove 和 Word2Vec)更有效,因为它充分考虑了单词的情感价值。TGRU 与 GRU 和 LSTM 相比具有更高的准确性,但其网络较复杂,需要较长训练时间和大量的计算资源。

尽管有许多研究检验了将情绪分类加入股票预测有助于改善预测性能,但很少有研究涉及不同文本数据来源。Day 等^[118]在对 4 个不同来源的金融新闻财务情感分析中发现新闻来源不同导致模型的预测能力不同,再一次验证了结合情感分析的深度模型能够提高股票价格趋势预测的准确性。

(2)基于机器学习的情感分析方法

基于机器学习的情感分析方法通常是把文本情感分析作为分类任务进行处理,通过从文本数据中提取特征构建分类器并进行训练和实现情感分类。基于机器学习的方法比基于词典的方法更耗时且成本更高,但分类准确率通常高于基于词典的方法。

在股票市场预测中,Das 等^[119]使用了 LSTM 对 Twitter 帖子进行了情绪分析,以预测 Google、Microsoft 和 Apple 等股票价格变化。实验结果显示 LSTM 的性能优于基准模型,

加入情感特征使准确性显著提高。Zhuge 等^[120]利用朴素贝叶斯方法学习在线论坛评论中的情绪来改进 LSTM 网络的预测性能,与仅输入历史股票价格数据的模型相比较表现更好。Proskey 等^[15]使用 LSTM 网络和 CNN 对路透社的新闻文本进行情绪分析,并将这些情绪分析数据用于股市预测,以验证情感与其他各种随机事件的关系。Wang 等^[121]考虑到情感和事件之间的联系,使用分层神经网络对情感进行分类,使用了 CNN 来从转换后的事件表示中提取显著特征。

除了股票市场预测外,基于机器学习的情绪分析方法也被用于加密货币的价格预测、寻找金融专家以及欺诈识别等应用中。例如,Lopes^[83]使用来自 StockTwits 的加密货币微博数据源作为情感挖掘模型的输入,获得带标签的情感值,并将该情感值与该时间段的相应价格数据作为价格预测模型的输入以对加密货币价格进行预测。其中情感挖掘模型采用

CNN 和 LSTM 的组合模型提取文本挖掘中的情感,预测模型则使用具有解决爆炸性和梯度递减问题能力的 LSTM 网络来预测比特币 BTC 和莱特币 LTC 这两种货币的价格。实验结果表明,使用了情感数据的预测模型比没有使用情感数据的模型效果更好。Sohangir 等^[122]使用 CNN,LSTM 和 doc2vec 等几种神经网络,在金融社交媒体平台 StockTwits 中对专家发布的股市意见进行情绪分析,以判断这几种方法是否能提高情感分析的准确性。研究结果表明,CNN 在这几个方法中预测金融专家情绪的准确性最高。该研究同时还在 StockTwits 消息中使用 CNN 来查找专家^[123]。

以上研究表明,主观情绪和客观事件、事实的结合及相互利用将产生更好的预测结果。文献[15,83,120,124-125]在研究投资者情绪对股票价格走势的影响中均使用了 LSTM 网络,但采用的文本数据来源、输入数据以及情绪分类有所不同,具体情况如表 10 所列。

表 10 金融文本数据挖掘
Table 10 Financial text data mining

文献	模型	数据集	输入数据	情绪分类	结论
[83]	CNN,LSTM	比特币,莱特币,StockTwits	OCHLV,技术指标,情绪分析	3 类	在加密货币价格预测中,使用了情绪数据的预测模型比没有情绪数据的模型效果好
[119]	LSTM	Yahoo(谷歌、苹果、微软股价),Twitter 数据集	Twitter 情绪,历史股价	7 类	引入基于情绪分析和股价趋势的预测模型,将提高其预测的可靠性和可信度
[120]	朴素贝叶斯方法+LSTM	东方财富的股票帖子,网易交易数据	论坛评论,OCHL 的变化率和价格	2 类	采用机器学习的情绪分类方法,结合 LSTM 获得了出色的预测性能
[15]	LSTM,CNN,word2vec	DJIA 30 种股票,S&P500,DJI,路透社新闻	新闻标题,价格数据	3 类	使用 CNN 没有达到比基线模型更好的结果,因此尚不清楚情绪分析是否有助于股票预测
[125]	LSTM	东方财富中的在线评论,CSMAR 的 125 只股票	在线评论,股票交易数据	3 类	投资者的情绪对交易量有积极影响

3.3 深度学习在金融领域的其他应用

有部分关于金融科技的研究论文不属于前面列出的主题,包括社会保障、银行营销等。这些应用领域的研究相对较少,存在较大的研究空间。比如,社会保险包括社会保障、医疗保险等,也属于金融产品,它们在保护社会功能方面起到了不可忽视的作用^[126]。另外,已有部分研究利用深度学习技术来分析银行营销策略的有效性^[127-128]。我们认为这些方面的研究亟待提高,也是从事金融科技的研究者未来可以重点关注的领域。

4 深度学习在金融科技中的研究趋势及展望

4.1 金融数据的数据增强以及多模态学习

根据我们的综述可以知道深度学习模型在金融领域的应用可以大大提高预测的准确性,但是深度学习模型需要依赖大量数据来学习和训练,然后做出决策。因此可以通过数据增强技术,确定更多数据是否有助于解决金融业务中的数据复杂性问题,带来更好的数据挖掘和预测性能。从前面总结的表格中可以看出,文献的输入数据集中在历史时间序列数据和新闻文本上,未来的工作可以侧重于多模态学习(Multi-modal Learning)和多视图学习(Multi-view Learning)等方面。

4.2 大规模金融数据预处理和预训练

不少研究验证了数据集中类别平衡性和数据规范化的重要性^[17,97,129]。数据规范化能够提高深度学习模型输入数据

的质量,有助于构建稳健且可操作的金融科技。其次,选择输入数据中与金融相关且最具代表性的特征作为输入以减少输入数据维数,从而提高训练速度和鲁棒性。

在金融领域中,标签数据稀少,以至于大多数金融文本挖掘模型往往不能直接利用深度学习技术。针对此问题,预训练模型能够更有效地捕获金融文本数据中的语言知识和语义信息,并且与深度学习技术结合可以较为简单并显著提高整体预测性能^[130]。更重要的是,预训练技术可以借助于大量额外的数据进行表征学习,从而提高下游金融任务的学习表现。因此,使用预训练模型也是金融科技领域未来的一个重要研究方向。

4.3 模型和算法改进

深度学习技术在大数据分析中已取得了巨大成功,但在训练数据有限时,需要更强大的模型来实现更强的学习能力,因此探索新的深度学习算法在金融领域的应用至关重要。最近几年,一些研究者尝试将新的深度学习模型应用于金融数据挖掘。例如,将在数据增强方面取得成功的 GAN 模型应用于金融领域^[34];利用迁移学习解决跨市场分析^[12];探索 DRL 在算法交易系统中的应用^[67];将图神经网络框架应用于多元时间序列预测^[131];探索对未知网络类型的泛化能力^[132];将 Transformer 应用于金融文本情绪分析^[39]等。可以预见的是未来将会有更多新的深度学习模型应用于金融领域,探索优化组合的集成深度学习模型,以及将其他领域技术

发展应用于金融数据分析。

其次,改进现有深度学习模型,例如解决模型训练过程中过拟合和欠拟合问题。一般情况下更深的神经网络模型或神经元可以减小误差,迭代次数的增加有助于解决欠拟合问题;更大的输入数据以及规则化数据有助于缓解过拟合问题,未来还需要进一步研究如何在复杂金融市场数据分析中平衡过拟合和欠拟合及其负面影响。

4.4 可解释性

深度学习模型作为预测模型性能是非常优秀的,但由于难以解释预测结果,它通常被作为“黑箱”模型。如果能对深度学习模型的预测结果和模型的运行规则给出合理的解释,随后有方向性地进行优化,将可以减少大量不必要的工作并提升模型的可信度,这对其在金融领域的应用会有很大的帮助^[133]。

4.5 通用金融科技框架

在各深度学习模型结构中,提炼出适用于金融市场分析与预测的通用框架或规律性方法,总结各种模型和方法的普遍性,从而进行算法和模型的模块化应用。目前,还缺少适合于金融科技的通用深度学习框架和大规模应用的开源平台。这个方向的研究将有助于从整体上推动金融科技的发展,也是值得深入研究的方向。

4.6 金融科技应用创新

在可预见的未来,投资组合、算法交易、风险评估以及加密货币和区块链技术可能会继续占据金融科技领域的主导地位。另外一个值得思考和研究的方向是探索深度学习在金融领域的新型应用。例如,采用计算机视觉等方面的深度学习技术来对金融市场中诸如图表之类的类型更丰富的数据进行分析 and 处理;尝试利用深度学习提取深层文本信息的能力自动生成金融研报;利用深度学习捕捉信息(特别是尾部信息)的能力在风险管理方面进行应用探索;利用深度学习技术完善金融体系结构和监管体系;加快提高资金配置的能力和金融创新能力等。

结束语 金融市场对投资者和研究者有巨大的吸引力,深度学习是推动金融科技发展的关键技术。目前,研究人员和机构做了大量的工作来探索如何利用深度学习模型在金融大数据上进行可靠的数据挖掘工作。

本文对深度学习在金融领域的应用进行了全面详细的综述,归纳整理了金融领域内深度学习的经典方法以及近年来为提高金融市场预测准确率所提出的新模型。更具体说来,本文讨论了将深度学习技术应用于金融领域实现有效预测的一般框架,包括数据集选择、数据预处理、基本模型构建以及模型性能评估 4 个部分;将相关研究工作根据具体应用场景分为股票价格预测、外汇汇率预测、大宗商品价格预测、投资组合管理、波动率预测、加密货币价格预测、趋势预测、风险评估以及金融文本挖掘和情绪分析 10 个应用领域进行归纳整理;并且讨论了基于深度学习的金融科技未来可能的发展方向。笔者希望本文能为金融科技的从业人员提供有价值的参考和技术路线图。

参 考 文 献

[1] WU J J. Financial distress prediction: The comparison and appli-

cation of data mining models[J]. Journal of Tsinghua University (Philosophy and Social Sciences), 2006(S1): 45-53.

[2] SUN Z J, XUE L, XU Y M, et al. Overview of deep learning[J]. Application Research of Computers, 2012, 29(8): 2806-2810.

[3] SU Z, LU M, LI D X. Deep learning in Financial Empirical Applications: Dynamics, Contributions and Prospects[J]. Journal of Financial Research, 2017(5): 111-126.

[4] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.

[5] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[6] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[7] WU Z, PAN S, CHEN F, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 4-24.

[8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv:1706.03762, 2017.

[9] XU H R, XU B, XU K W. Analysis on application of machine learning in stock forecasting[J]. Computer Engineering and Applications, 2020, 56(12): 19-24.

[10] SHAH D, ISAH H, ZULKERNINE F. Stock market analysis: A review and taxonomy of prediction techniques[J]. International Journal of Financial Studies, 2019, 7(2): 26.

[11] SEZER O B, GUDELEK M U, OZBAYOGLU A M. Financial time series forecasting with deep learning: A systematic literature review: 2005-2019[J]. Applied Soft Computing, 2020, 90: 106181.

[12] JEONG G, KIM H Y. Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning[J]. Expert Systems with Applications, 2019, 117: 125-138.

[13] CHONG E, HAN C, PARK F C. Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies[J]. Expert Systems with Applications, 2017, 83: 187-205.

[14] CHEN L, QIAO Z L, WANG M G, et al. Which artificial intelligence algorithm better predicts the Chinese stock market? [J]. IEEE Access, 2018, 6: 48625-48633.

[15] PROSKY J, SONG X Y, TAN A, et al. Sentiment predictability for stocks[J]. arXiv:1712.05785, 2017.

[16] ZHANG J, MARINGER D. Using a genetic algorithm to improve recurrent reinforcement learning for equity trading[J]. Computational Economics, 2016, 47(4): 551-567.

[17] DINGLI A, FOURNIER K S. Financial time series forecasting-a deep learning approach [J]. International Journal of Machine Learning and Computing, 2017, 7(5): 118-122.

[18] TÜFEKCI P. Predicting the direction of movement for stock price index using machine learning methods[C]// Second International Afro-European Conference for Industrial Advancement AECIA. Springer International Publishing, 2016: 477-492.

[19] LAGO J, DE RIDDER F, DE SCHUTTER B. Forecasting spot electricity prices: Deep learning approaches and empirical com-

- parison of traditional algorithms[J]. *Applied Energy*, 2018, 221: 386-405.
- [20] ROY A, SUN J Y, MAHONEY R, et al. Deep learning detecting fraud in credit card transactions[C]// *Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 2018: 129-134.
- [21] JIANG W. Applications of deep learning in stock market prediction: recent progress[J]. *Expert Systems with Applications*, 2021, 184: 115537.
- [22] SMOLENSKY P. Information processing in dynamical systems: Foundations of harmony theory[R]. Massachusetts: Colorado Univ at Boulder Dept of Computer Science, 1986.
- [23] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436-444.
- [24] KINGMA D P, WELING M. Auto-encoding variational bayes [J]. *arXiv*: 1312. 6114, 2013.
- [25] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *arXiv*: 1406. 2661, 2014.
- [26] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning[J]. *arXiv*: 1312. 5602, 2013.
- [27] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2016: 770-778.
- [28] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. *arXiv*: 1406. 1078, 2014.
- [29] XU Y M, COHEN S B. Stock movement prediction from tweets and historical prices[C]// *56th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2018: 1970-1979.
- [30] HU Z N, LIU W Q, BIAN J, et al. Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction[C]// *11th ACM International Conference on Web Search and Data Mining*. ACM Press, 2018: 261-269.
- [31] ZHOU F, ZHANG S, YANG Y. Interpretable operational risk classification with semi-supervised variational autoencoder[C]// *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020: 846-852.
- [32] ZHOU F, QI X, XIAO C, et al. MetaRisk: Semi-supervised few-shot operational risk classification in banking industry[J]. *Information Sciences*, 2021, 552: 1-16.
- [33] TAKAHASHI S, CHEN Y, TANAKA-ISHII K. Modeling financial time-series with generative adversarial networks[J]. *Physica A: Statistical Mechanics and its Applications*, 2019, 527: 121261.
- [34] BA H. Improving detection of credit card fraudulent transactions using generative adversarial networks[J]. *arXiv*: 1907. 03355, 2019.
- [35] XU K, HU W, LESKOVEC J, et al. How powerful are graph neural networks? [J]. *arXiv*: 1810. 00826, 2018.
- [36] DEVLIN J, CHANG M W, LEE K, et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. *arXiv*: 1810. 04805, 2018.
- [37] YANG L Y, NG T L J, SMYTH B, et al. Htm1: Hierarchical transformer-based multi-task learning for volatility prediction [C]// *The Web Conference 2020*. Assoc Comp Machinery, 2020: 441-451.
- [38] KESKAR N S, MCCANN B, VARSHNEY L R, et al. Ctrl: A conditional transformer language model for controllable generation[J]. *arXiv*: 1909. 05858, 2019.
- [39] MISHEV K, GJORGJEVIKJ A, VODENSKA I, et al. Evaluation of sentiment analysis in finance: from lexicons to transformers[J]. *IEEE Access*, 2020, 8: 131662-131682.
- [40] DING Q, WU S, SUN H, et al. Hierarchical Multi-Scale Gaussian Transformer for Stock Movement Prediction[C]// *IJCAI*. 2020: 4640-4646.
- [41] LIANG T X, YANG X P, WANG L, et al. Review on financial trading system based on reinforcement learning[J]. *Ruan Jian Xue Bao/Journal of Software*, 2019, 30(3): 845-864.
- [42] HAJIABOTORABI Z, KAZEMI A, SAMAVATI F F, et al. Improving DWT-RNN model via B-spline wavelet multiresolution to forecast a high-frequency time series [J]. *Expert Systems With Applications*, 2019, 138: 112842.
- [43] SI W Y, LI J K, DING P, et al. A multi-objective deep reinforcement learning approach for stock index future's intraday trading [C]// *10th International Symposium on Computational Intelligence and Design (ISCID)*. IEEE, 2017: 431-436.
- [44] ZHOU X Y, PAN Z S, HU G Y, et al. Stock market prediction on high-frequency data using generative adversarial nets[J]. *Mathematical Problems in Engineering*, 2018, 2018.
- [45] HUANG T T, YU L. Application of SDAE-LSTM Model on Financial Time Series Forecasting[J]. *Computer Engineering and Applications*, 2019, 55(1): 142-148.
- [46] ZHANG L H, AGGARWAL C, QI G J. Stock price prediction via discovering multi-frequency trading patterns [C] // *23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, 2017: 2141-2149.
- [47] PANG X W, ZHOU Y Q, WANG P, et al. An innovative neural network approach for stock market prediction[J]. *The Journal of Supercomputing*, 2020, 76(3): 2098-2118.
- [48] REZAEI H, FAALJOU H, MANSOURFAR G. Stock price prediction using deep learning and frequency decomposition[J]. *Expert Systems with Applications*, 2021, 169: 114332.
- [49] MOURELATOS M, ALEXAKOS C, AMORGIANIOTIS T, et al. Financial indices modelling and trading utilizing deep learning techniques: The ATHENS SE FTSE/ASE Large Cap Use Case[C]// *Innovations in Intelligent Systems and Applications (INISTA)*. IEEE, 2018: 1-7.
- [50] DING X, ZHANG Y, LIU T, et al. Deep learning for event-driven stock prediction[C]// *24th International Joint Conference on Artificial Intelligence*. AAAI Press, 2015.
- [51] PUTRI K S, HALIM S. Currency movement forecasting using time series analysis and long short-term memory[J]. *International Journal of Industrial Optimization*, 2020, 1(2): 71-80.
- [52] ZHENG J, FU X, ZHANG G J. Research on exchange rate forecasting based on deep belief network[J]. *Neural Computing and Applications*, 2019, 31(1): 573-582.
- [53] SHEN F R, CHAO J, ZHAO J X. Forecasting exchange rate using deep belief networks and conjugate gradient method[J]. *Neurocomputing*, 2015, 167: 243-253.

- [54] SHEN H, LIANG X. A time series forecasting model based on deep learning integrated algorithm with stacked autoencoders and svr for fx prediction[C]//International Conference on Artificial Neural Networks. Springer, 2016: 326-335.
- [55] KORCZAK J, HEMES M. Deep learning for financial time series forecasting in A-Trader system[C]//Federated Conf on Computer Science and Information Systems(FedCSIS). IEEE, 2017: 905-912.
- [56] ZHANG R, SHEN F R, ZHAO J X. A model with fuzzy granulation and deep belief networks for exchange rate forecasting [C]//International Joint Conference on Neural Networks(IJCNN). IEEE, 2014: 366-373.
- [57] CAO W, ZHU W D, WANG W J, et al. A deep coupled LSTM approach for USD/CNY exchange rate forecasting[J]. IEEE Intelligent Systems, 2020, 35(2): 43-53.
- [58] WIDEGREN P. Deep learning-based forecasting of financial assets[D]. KTH Royal Institution Technology, 2017.
- [59] ZHANG P Y, CI B C. Deep belief network for gold price forecasting[J]. Resources Policy, 2020, 69: 101806.
- [60] LASHERAS F S, DE COS JUEZ F J, SÁNCHEZ A S, et al. Forecasting the COMEX copper spot price by means of neural networks and ARIMA models[J]. Resources Policy, 2015, 45: 37-43.
- [61] LIVIERIS I E, PINTELAS E, PINTELAS P. A CNN-LSTM model for gold price time-series forecasting[J]. Neural Computing and Applications, 2020, 32(23): 17351-17360.
- [62] CHEN Y H, HE K J, TSO G K. Forecasting crude oil prices: a deep learning based model [J]. Procedia Computer Science, 2017, 122: 300-307.
- [63] WU Y X, WU Q B, ZHU J Q. Improved EEMD-based crude oil price forecasting using LSTM networks[J]. Physica A: Statistical Mechanics and its Applications, 2019, 516: 114-124.
- [64] QIAO W B, YANG Z. Forecast the electricity price of US using a wavelet transform-based hybrid model[J]. Energy, 2020, 193: 116704.
- [65] JIANG Z Y, LIANG J J. Cryptocurrency portfolio management with deep reinforcement learning[C]//Intelligent Systems Conference(IntelliSys). IEEE, 2017: 905-913.
- [66] JIANG Z Y, XU D X, LIANG J Y. A deep reinforcement learning framework for the financial portfolio management problem [J]. arXiv:1706.10059, 2017.
- [67] LIANG Z P, CHEN H, ZHU J H, et al. Adversarial deep reinforcement learning in portfolio management [J]. arXiv: 1808.09940, 2018.
- [68] LEE S I, YOO S J. Threshold-based portfolio: the role of the threshold and its applications[J]. The Journal of Supercomputing, 2018: 1-18.
- [69] DENG Y, BAO F, KONG Y Y, et al. Deep Direct Reinforcement Learning for Financial Signal Representation and Trading[J]. IEEE Trans. Neural Netw Learn Syst, 2017, 28(3): 653-664.
- [70] WANG W Y, LI W Z, ZHANG N, et al. Portfolio formation with preselection using deep learning from long-term financial data[J]. Expert Systems with Applications, 2020, 143: 113042.
- [71] STOEAN C, PAJA W, STOEAN R, et al. Deep architectures for long-term stock price prediction with a heuristic-based strategy for trading simulations[J]. PloS one, 2019, 14(10): e0223593.
- [72] NGUNYI A, MUNDIA S, OMARI C. Modelling volatility dynamics of cryptocurrencies using GARCH models[J]. Journal of Mathematical Finance, 2019, 9: 591-615.
- [73] ZHOU Y L, HAN R J, XU Q, et al. Long shortterm memory networks for CSI300 volatility prediction with Baidu search volume[J]. Concurrency and Computation: Practice and Experience, 2019, 31(10): e4721.
- [74] JIA F, YANG B L. Forecasting Volatility of Stock Index: Deep Learning Model with Likelihood-Based Loss Function[J]. Complexity, 2021, 2021.
- [75] KIM H Y, WON C H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models[J]. Expert Systems with Applications, 2018, 103: 25-37.
- [76] HU Y, NI J, WEN L. A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction[J]. Physica A: Statistical Mechanics and its Applications, 2020, 557: 124907.
- [77] CHEN W L, YEO C K, LAU C T, et al. Leveraging social media news to predict stock index movement using RNN-boost[J]. Data & Knowledge Engineering, 2018, 118: 14-24.
- [78] VIDAL A, KRISTJANPOLLER W. Gold Volatility Prediction using a CNN-LSTM approach[J]. Expert Systems with Applications, 2020: 113481.
- [79] MCNALLY S, ROCHE J, CATON S. Predicting the price of bitcoin using machine learning[C]//26th Euromicro International Conference on Parallel, Distributed and Network-based Processing(PDP). IEEE, 2018: 339-343.
- [80] LAHMIRI S, BEKIROS S. Cryptocurrency forecasting with deep learning chaotic neural networks[J]. Chaos, Solitons & Fractals, 2019, 118: 35-40.
- [81] KUMAR D, RATH S. Predicting the Trends of Price for Ethereum Using Deep Learning Techniques [M]. Singapore: Springer, 2020: 103-114.
- [82] ALTAN A, KARASU S, BEKIROS S. Digital currency forecasting with chaotic meta-heuristic bio-inspired signal processing techniques[J]. Chaos, Solitons & Fractals, 2019, 126: 325-336.
- [83] LOPES G D L F. Deep Learning for Market Forecasts[D]. University of Porto, 2018.
- [84] DI PERSIO L, HONCHAR O. Artificial neural networks architectures for stock price prediction: Comparisons and applications [J]. International Journal of Circuits, Systems and Signal Processing, 2016, 10(2016): 403-413.
- [85] SEZER O B, OZBAYOGLU A M. Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach[J]. Applied Soft Computing, 2018, 70: 525-538.
- [86] NABIPOUR M, NAYYERI P, JABANI H, et al. Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data: a Comparative Analysis[J]. IEEE Access, 2020, 8: 150199-150212.
- [87] SHARANG A, RAO C. Using machine learning for medium frequency derivative portfolio trading [J]. arXiv:1512.06228, 2015.

- [88] KRAUSS C, DO X A, HUCK N. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500[J]. *European Journal of Operational Research*, 2017, 259(2): 689-702.
- [89] HOSEINZADE E, HARATIZADEH S. CNNpred: CNN-based stock market prediction using a diverse set of variables[J]. *Expert Systems with Applications*, 2019, 129: 273-285.
- [90] YANG C, ZHAI J J, TAO G H. Deep Learning for Price Movement Prediction Using Convolutional Neural Network and Long Short-Term Memory[J]. *Mathematical Problems in Engineering*, 2020, 2020.
- [91] LONG J W, CHEN Z P, HE W B, et al. An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market[J]. *Applied Soft Computing*, 2020, 91: 106205.
- [92] BUCZKOWSKI P. Predicting stock trends based on expert recommendations using gru/lstm neural networks[C] // *International Symposium on Methodologies for Intelligent Systems*. Springer, 2017: 708-717.
- [93] LUO C C, WU D S, WU D X. A deep learning approach for credit scoring using credit default swaps[J]. *Engineering Applications of Artificial Intelligence*, 2017, 65: 465-470.
- [94] HOSAKA T. Bankruptcy prediction using imaged financial ratios and convolutional neural networks[J]. *Expert Systems with Applications*, 2019, 117: 287-299.
- [95] NEAGOE V E, CIOTEC A D, CUCU G S. Deep convolutional neural networks versus multilayer perceptron for financial prediction [C] // *International Conference on Communications (COMM)*. IEEE, 2018: 201-206.
- [96] JURGOVSKY J, GRANITZER M, ZIEGLER K, et al. Sequence classification for credit-card fraud detection[J]. *Expert Systems with Applications*, 2018, 100: 234-245.
- [97] HERYADI Y, WARNARS H L H S. Learning temporal representation of transaction amount for fraudulent transaction recognition using cnn, stacked lstm, and cnn-lstm[C] // *IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*. IEEE, 2017: 84-89.
- [98] AL-SHABI M. Credit card fraud detection using autoencoder model in unbalanced datasets[J]. *Journal of Advances in Mathematics and Computer Science*, 2019, 33(5): 1-16.
- [99] ZHU B, YANG W C, WANG H X, et al. A hybrid deep learning model for consumer credit scoring[C] // *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)*. IEEE, 2018: 205-208.
- [100] KVAMME H, SELLEREITE N, AAS K, et al. Predicting mortgage default using convolutional neural networks[J]. *Expert Systems with Applications*, 2018, 102: 207-217.
- [101] YU L, ZHOU R, TANG L, et al. A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data[J]. *Applied Soft Computing*, 2018, 69: 192-202.
- [102] CHATZIS S P, SIAKOULIS V, PETROPOULOS A, et al. Forecasting stock market crisis events using deep and statistical machine learning techniques[J]. *Expert Systems with Applications*, 2018, 112: 353-371.
- [103] SOHONY I, PRATAP R, NAMBIAR U. Ensemble learning for credit card fraud detection[C] // *ACM India Joint International Conference on Data Science and Management of Data (CoDS-CO-MAD'18)*. New York, 2018: 289-294.
- [104] ZHAO C, YE Y W, YAO M H. Stock volatility forecast based on financial text emotion[J]. *Computer Science*, 2020, 47(5): 79-83.
- [105] SHI L, TENG Z Y, WANG L, et al. DeepClue: Visual Interpretation of Text-Based Deep Stock Prediction[J]. *IEEE Trans on Knowledge and Data Engineering*, 2019, 31(6): 1094-1108.
- [106] DOS SANTOS P L, DRAS M. Stock market prediction with deep learning: A character-based neural language model for event-based trading [C] // *Australasian Language Technology Association Workshop 2017*. 2017: 6-15.
- [107] LEE C Y, SOO V W. Predict stock price with financial news based on recurrent convolutional neural networks[C] // *Conference on Technologies and Applications of Artificial Intelligence (TAAD)*. IEEE, 2017: 160-165.
- [108] RAWTE V, GUPTA A, ZAKI M J. Analysis of year-over-year changes in Risk Factors Disclosure in 10-K filings[C] // *Proceedings of the Fourth International Workshop on Data Science for Macro-Modeling with Financial and Economic Datasets (DSMM'18)*. 2018: 1-4.
- [109] RÖNNQVIST S, SARLIN P. Bank distress in the news: Describing events through deep learning[J]. *Neurocomputing*, 2017, 264: 57-70.
- [110] CERCHIELLO P, NICOLA G, RÖNNQVIST S, et al. Deep learning bank distress from news and numerical financial data [J]. *arXiv:1706.09627*, 2017.
- [111] AKITA R, YOSHIHARA A, MATSUBARA T, et al. Deep learning for stock prediction using numerical and textual information [C] // *IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*. IEEE, 2016: 1-6.
- [112] MATSUBARA T, AKITA R, UEHARA K. Stock Price Prediction by Deep Neural Generative Model of News Articles[J]. *IEEE Transactions on Information and Systems*, 2018, E101. D (4): 901-908.
- [113] WANG Y B, XU W. Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud[J]. *Decision Support Systems*, 2018, 105: 87-95.
- [114] FRANÇOIS-LAVET V, HENDERSON P, ISLAM R, et al. An introduction to deep reinforcement learning [J]. *arXiv:1811.12560*, 2018.
- [115] LIU Y, ZENG Q G, YANG H R, et al. Stock price movement prediction from financial news with deep learning and knowledge graph embedding [C] // *Pacific Rim Knowledge Acquisition Workshop*. Springer, 2018: 102-113.
- [116] WANG Q L, XU W, ZHENG H. Combining the wisdom of crowds and technical analysis for financial market prediction using deep random subspace ensembles[J]. *Neurocomputing*, 2018, 299: 51-61.
- [117] LIEN MINH D, SADEGHI-NIARAKI A, HUY H D, et al. Deep Learning Approach for Short-Term Stock Trends Prediction Based on Two-Stream Gated Recurrent Unit Network[J]. *IEEE Access*, 2018, 6: 55392-55404.
- [118] DAY M Y, LEE C C. Deep learning for financial sentiment ana-

lysis on finance news providers[C]//IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). IEEE,2016:1127-1134.

[119]DAS S,BEHERA R K,RATH S K. Real-time sentiment analysis of twitter streaming data for stock prediction[J]. Procedia Computer Science,2018,132:956-964.

[120]ZHUGE Q,XU L Y,ZHANG G W. LSTM Neural Network with Emotional Analysis for prediction of stock price[J]. Engineering Letters,2017,25(2).

[121]WANG Y W,LI Q,HUANG Z X,et al. EAN:Event attention network for stock price trend prediction based on sentimental embedding[C]//10th ACM Conference on Web Science. Association for Computing Machinery,2019:311-320.

[122]SOHANGIR S,WANG D D,POMERANETS A,et al. Big Data:Deep Learning for financial sentiment analysis[J]. Journal of Big Data,2018,5(1):1-25.

[123]SOHANGIR S,WANG D D. Finding Expert Authors in Financial Forum Using Deep Learning Methods[C]//Second IEEE International Conference on Robotic Computing (IRC). IEEE,2018:399-402.

[124]IWASAKI H,CHEN Y. Topic sentiment asset pricing with dnn supervised learning[J/OL]. <http://dx.doi.org/10.2139/ssrn.3228485>.

[125]WANG G S,YU G J,SHEN X H. The Effect of Online Investor Sentiment on Stock Movements: An LSTM Approach[J/OL]. <http://dx.doi.org/10.2139/ssrn.3228485>.

[126]YING J J C,HUANG P Y,CHANG C K,et al. A preliminary study on deep learning for predicting social insurance payment behavior[C]//IEEE International Conference on Big Data (Big Data). IEEE,2017:1866-1875.

[127]KIM K H,LEE C S,JO S M,et al. Predicting the success of bank telemarketing using deep convolutional neural network [C]//7th International Conference of Soft Computing and Pattern Recognition (SoCPaR). IEEE,2015:314-317.

[128]ŁADY ZYŃSKI P,ZBIKOWSKI K,GAWRYSIAK P. Direct marketing campaigns in retail banking with the use of deep learning and random forests[J]. Expert Systems with Applications,2019,134:28-35.

[129]LI Y,LIN X H,WANG X W,et al. Credit risk assessment algorithm using deep neural networks with clustering and merging [C]//13th International Conference on Computational Intelligence and Security(CIS). IEEE,2017:173-176.

[130]LIU Z,HUANG D,HUANG K,et al. FinBERT: A Pre-trained Financial Language Representation Model for Financial Text Mining[C]//IJCAI. 2020:4513-4519.

[131]WU Z,PAN S, LONG G,et al. Connecting the dots: Multivariate time series forecasting with graph neural networks[C]//26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Association for Computing Machinery,2020:753-763.

[132]WANG X,JI H Y,SHI C,et al. Heterogeneous graph attention network[C]//the World Wide Web Conference. ACM Press,2019:2022-2032.

[133]KRITTANAWONG C,JOHNSON K W,ROSENSEN R S,et al. Deep learning for cardiovascular medicine: a practical primer[J]. European Heart Journal,2019,40(25):2058-2073.



ZHOU Fan, born in 1981, Ph.D, associate professor, is a member of China Computer Federation. His main research interests include big data post analysis and data mining, machine learning and deep learning.



ZHONG Ting, born in 1977, Ph.D, associate professor. Her main research interests include cloud computing security and machine learning.