

doi:10.19920/j.cnki.jmsc.2023.04.008

基于深度强化学习的自适应股指预测研究^①

卜 湛¹, 张善凡², 李雪延², 马丹丹², 曹 杰³

(1. 南京审计大学智能审计学院, 南京 211815; 2. 南京财经大学信息工程学院, 南京 210023;
3. 合肥工业大学管理学院, 合肥 230009)

摘要: 基于股指成分股基本面和技术面数据构建了时序股票关联网络, 然后利用深度图神经网络学习股票关联网络层次化表征, 以端到端的方式获得候选预测信号. 在此基础上, 提出了一种考虑动作评估反馈的深度强化学习方法 (Action Evaluation Feedback based Deep Q-Learning, AEF-DQN), 旨在将不同的候选预测信号融入智能体的动作空间, 并基于股票关联网络层次化表征、股票市场整体运行状态和历史动作评估反馈学习环境状态; 借鉴前景理论中的参照依赖特性估计奖励值函数, 从而建立状态、动作与奖励值之间的映射关系. 最后, 采用沪深 300 指数、标普 500 指数、英国富时 100 指数和日经 225 指数的成分股历史数据, 构造了股指期货交易模拟器, 在投资胜率、最大回撤率、阿尔法比率和夏普比率 4 个回测指标上对股指预测模型展开实证分析. 研究结果表明: 1) 通过层次化聚合股票关联网络的节点属性信息可以动态捕捉不同行业对股指价格波动的影响, 进而可提升预测方法的准确率; 2) 考虑动作评估反馈的深度强化学习结构可智能化选择适用于当前股票市场环境的最优模型结构, 进而可提升预测方法的鲁棒性.

关键词: 自适应股指预测; 股票关联网络; 深度图神经网络; 深度强化学习

中图分类号: F832.5 **文献标识码:** A **文章编号:** 1007-9807(2023)04-0148-27

0 引 言

股票市场是一个具有高噪声、非线性、非平稳和混沌等特征的复杂系统. 股指时间序列是股票市场内在复杂性的综合外在表现, 为投资者制定投资策略提供了重要参考. 股指预测不仅有利于更好的监测和管理与股票市场高度关联的金融市场, 还能为投资者的投资决策提供有效指导.

早期股指预测研究主要包含基本面分析和技术分析. 基本面分析聚焦于国家经济政策与上市公司的基本面等信息, 基于股指的内在价值来评估股指价格. 而技术分析更注重捕捉市场本身的运行规律, 试图将历史数据代入数学模型或机器

学习算法对市场做短期分析. 尽管有效市场假说 (Efficient Markets Hypothesis, EMH) 认为基本面分析或技术分析不会给投资者带来任何超额收益, 一些研究尝试衡量成熟和新兴股票市场的不同效率水平来揭示股票市场并非完全有效^[1], 而另一些研究基于不同的理论假设建立针对股票市场的有效预测模型. 伴随着股票市场的异常波动和越来越多的市场异象, 一些学者基于统计学方法构建了各种时间关系模型, 试图根据股指历史价格走势对未来短期变化趋势 (或收益率) 进行预测. 其中, 最具代表性的是 ARIMA 模型^[2] 和 GARCH 模型^[3]; 前者对提取金融时间序列的线

① 收稿日期: 2021-05-05; 修订日期: 2023-01-16.

基金项目: 国家重点研发计划资助项目 (2019YFB1405000); 国家自然科学基金资助项目 (71871109; 92046026); 江苏省高等学校自然科学研究重大项目 (20KJA520011); 江苏省未来网络科研基金一般项目 (FNSRFP-2021-YB-22).

作者简介: 卜 湛 (1987—), 男, 安徽铜陵人, 博士, 教授. Email: zhanbu@nau.edu.cn

性特征具备较大优势;后者对误差的方差进行了系统性建模,更加适用于股票市场波动性的分析和预测。然而,股票市场作为一种影响要素众多、各类不确定性交互的复杂系统,其价格波动不仅来自于市场参与人员的相互博弈,同时也会受到经济环境、政策干预、投资心理和交易技术等诸多要素的共同影响。传统基于统计学的时间关系模型难以有效捕捉股票市场的非线性和非平稳性,无论是在预测精度还是在稳定性方面都没有保障。近年来,股票市场中涌现的海量数据引起了众多学者的关注,以支持向量机(Support Vector Machine, SVM)^[4]、决策树(Decision Tree, DT)^[5]、人工神经网络模型(Artificial Neural Network, ANN)^[6]等为代表的因果关系模型被广泛应用于股指预测。因果关系模型认为股指价格波动与其影响要素之间在一定时间尺度内保持着固定关系,相较于直接从股指价格自身历史运动轨迹中寻找规律的时间关系模型,后者试图在原因和结果之间建立某种有效和稳定的因果联系。同时,因果关系模型在处理海量复杂数据方面展现了独特优势,较传统的时间关系模型,能够最大程度的捕捉股票市场的非线性特征,从而极大提高了股指预测精度。

面对股指预测问题,尽管近年来各类预测方法和技术不断涌现,预测效果逐步提高,仍然存在一些开放性问题的有待被更好的解决:

1) 重指数预测,轻行业及个股分析。股指时间序列是股票市场内在复杂性的综合外在表现,受共同的宏观经济或行业因素影响,股指成分股价格波动之间存在复杂的关联性。文献[7]指出通过度量股票价格波动的相关性可以揭示市场的总体风险,从微观层面上看,隶属于相同行业板块的股票价格波动趋同性较高,市场风险的增加与股票价格波动的协同率呈正相关。从中观层面上看,单一行业的价格波动可能影响其他行业的价格波动,这种“波动溢出效应”意味着“板块联动”甚至“普涨普跌”极易发生^[8]。从宏观层面上看,宏观经济波动会直接影响企业外部交易环境,进而对股票市场稳定性产生冲击^[9],例如宏观经济下行会导致股票市场系统性风险上升,而其中具有低市值、高波动和高杠杆率等特征的股票系统性风险可能会显著提升。股票市场中少数权重

(龙头)股票的价格波动,会引起行业板块内其他股价波动;受行业波动溢出影响,市场的“板块联动”现象极易发生,进而可能引致较大的规模的系统性风险。尽管部分学者试图基于股指成分股价格波动的相关性构建股票关联网络,进而提取中观尺度的拓扑结构特征(如网络聚类系数、平均最短路径长度、模块度等),最后分析这些特征与股指价格波动的相关性^[10-13]。然而,中观尺度的拓扑结构特征忽略了成分股的基本面和技术面信息,难以充分捕捉和刻画股票市场内在的复杂性。

2) 重特征构建,轻层次化特征提取。股票市场积累了大量反映股指运行趋势的原始数据,既包含大量成分股数据,也包括各类行业板块和宏观经济数据。股指波动是诸多因素的综合反映,现有研究对股票市场的内部运行机制缺乏深刻认识,很难人工判断哪些特征对股指预测有较大的影响。因此,有效地选取输入特征成为决定模型预测准确性的关键因素。股指预测领域现有的特征提取方法大多使用原始股指时序数据或宏观基本面数据,利用近邻互信息选择与股指收益率关联性较强的历史数据构造支持向量回归机的特征变量^[14]。文献[15]指出,由于股票市场中极其复杂的信息环境和投资者有限理性和非理性的交易行为,股指成分股间产生了领先滞后、价格同步、行业轮动、风险传导等多种复杂的关联性特征。因此,如何从股票关联网络的视角构建股指预测特征指标是一个值得深入思考的问题。此外,在股指运行过程中的不同阶段,不同行业板块对股指未来价格波动的影响权重也具有时变属性。如何实时捕捉不同行业板块对股指未来运行趋势的影响在现有的股指预测研究中还鲜有报道。区别于传统的基于原始股指时序数据或宏观基本面数据的特征提取方法,解决上述问题的关键是如何层次化聚合股票关联网络中的成分股特征,进而从中观尺度实时捕捉不同行业板块对股指未来价格波动的影响权重。

3) 重短期预测,轻自适应模型调整。在模型评估过程中,已有研究多将训练后的模型直接应用于股指的短期预测。伴随着预测周期的延长,预测精度总体上呈下降趋势。为了提升预测模型的鲁棒性,部分研究尝试基于滑动时间窗口技术不

断用最新的数据样本动态训练特定的股指预测模型. 由于不同时期股票市场波动特点的差异性显著, 仅仅改变模型参数并不能够完全捕捉股票市场运行规律, 甚至还需改变模型自身的结构. 股票市场中不断博弈产生的时间序列的收益率不满足正态分布, 表现出“尖峰重尾”和非对称等特征. 因此使用单一的预测模型, 都只能拟合股票市场运动规律和趋势的某一方面, 难以对股票市场的整体进行描述. 为此, 一些学者试图构建组合预测模型综合描述股指时间序列线性、非线性特征, 期望在一定程度上提高模型的预测效果^[16]. 回顾现有的组合预测模型, 绝大多数是借鉴集成学习的思想, 首先基于一组个体学习器获得候选预测结果, 然后采用加权或投票的方式获得最终的预测结果. 这种组合预测模型在要求个体学习器具备较高的预测精度同时, 还需具备一定的多样性, 在某些噪声较大的分类或回归问题上存在过拟合缺陷. 此外, 传统的组合预测模型缺乏对自身决策模式的学习完善, 无论股票市场发生何种变化, 历史预测精度如何波动, 这类方法还是按照既定模式做出决策, 缺乏对股票市场收益率可预测性的动态感知以及对自身决策模式的自动调整.

有鉴于此, 本研究试图设计可动态调整自身决策模式的自适应模型以提升股指预测的准确率和鲁棒性. 研究的边际贡献主要体现在以下方面:

1) 为了从微观层面捕捉股指价格波动的内在复杂性, 本文同时考虑股指成分股价格波动相关性以及成分股基本面和技术面指标, 构建了股指时间序列属性图模型. 具体地, 为了度量股指成分股价格波动的非线性特征, 首次采用可视图 (Visibility Graph Model, VG) 模型^[17]来分析股指成分股间的相依结构, 进而基于可视网络边集的 Jaccard 相似性来刻画任意两支成分股股价波动相关性. 区别于已有的股票关联网络研究, 本文所提方法还考虑了股指成分股的属性信息, 选取了 7 种常用的基本面指标和 14 种技术指标, 来刻画特定时间窗口内成分股的运行状态, 进而产生一系列带有标签信息的时序股票关联网络.

2) 首次提出将股指预测问题转化为一种图分类问题, 即学习一个映射函数将股票关联网络映射到一组标签的集合. 为了从中观层面捕捉不同行业板块对股指未来价格波动的影响权重, 基

于可微池化 (DIFFPOOL)^[18] 框架层次化聚合股票关联网络的节点属性信息, 试图从不同行业视角提取股指价格波动的中观尺度预测因子, 进而重构股票关联网络的粗化图. 这种特征提取模型以股票关联网络分类任务为驱动, 可有效聚合不同行业板块股指成分股的特征指标, 并基于股票关联网络的层次化聚类结果捕捉不同类簇的相互影响, 进而自动学习适用于股指短期趋势预测的预测变量. 其最终聚合得到的层次化表征向量与股票市场整体运行状态向量拼接后, 作为可微分类器的特征输入, 整个系统使用随机梯度下降进行端到端的训练. 此外, 考虑股指运行的阶段趋势特征, 引入了包括逻辑回归、多层感知器和栈式自编码器在内的 3 种可微分类模型, 通过其与 DIFFPOOL 模块的组合, 进而产生不同的股指预测信号.

3) 为了提升股指预测结果的鲁棒性, 首次提出了基于深度强化学习的自适应股指预测框架. 具体地, 考虑到股指预测中动态时变的股票市场环境以及面向动态状态下的候选预测信号寻优目标, 本文将自适应股指预测问题形式化为一种强化学习 (Reinforcement Learning) 结构, 即在动态环境 (股票市场) 与动作 (候选预测信号选择) 的交互过程获得奖励 (股指收益率), 并以最优化全局奖励为目标学习动作的选择策略. 同时, 为了在训练过程中规避因状态 - 动作空间过大而产生的“维数灾难”问题, 本研究借鉴 NDQN (Nature Deep Q-Learning)^[19] 思想, 提出了一种考虑动作评估反馈的深度强化学习方法 (Action Evaluation Feedback based Deep Q-Learning, AEF-DQN), 以在决策过程中考虑股票市场环境对智能体历史决策反馈等影响. 此外, 考虑到投资者对投资标的预期报酬越高则所能忍受的波动风险也越高, 反之则越低^[20], 本研究首次引入前景理论^[21] 中的参照依赖特性来估计奖励值函数. 基于智能体历史动作评估结果, 设计了两种预期回报函数以捕捉其在不同场景下状态 - 动作奖励值.

最后, 选取国内外 4 个最具代表性股票市场的股指成分股数据, 基于不同模型的预测结果构建了股指期货交易模拟器. 在投资胜率、最大回撤率、阿尔法比率和夏普比率 4 个回测指标上分析对比了不同模型的预测性能. 通过大规模实证

分析验证了所提方法的有效性,并分析了其在不同股票市场的表现,为未来的量化投资、算法交易等问题提供了决策依据。

1 文献综述

1.1 股指预测指标

长期以来股票市场积累了大量反映股指运行趋势的原始数据,既包含大量成分股数据,也包括各类行业板块和宏观经济数据。在早期的研究中,包含开盘价、收盘价、最高价、最低价、成交量和换手率等基础的交易数据被众多学者用于股指预测的常用指标。伴随着量化投资交易的不断发展完善,众多由金融和统计领域行业专家构造的技术指标被广泛应用于股指预测。胡熠和顾明^[22]从安全性、便宜性以及质量3个维度构造了综合性指标 B-score 来刻画巴菲特的价值投资风格,研究发现,在控制了多个横截面指标和不同市场状态后, B-score 对股票未来收益还有很好的预测能力,即使在2015年—2016年的极端市场环境中,该策略仍然可以获得正回报。Berkman等^[23]利用1918年—2006年期间的447场国际政治危机的数据库创建了一个显示出随时间发生显著变化的危机指数,研究发现该指数的变化对世界股市回报的平均水平和波动性都有很大的影响。杨晓兰等^[24]将财经博客作为投资者之间的互动媒介,利用文本挖掘技术和社会网络研究方法,构建反映投资者之间社会互动程度、情绪属性以及社会网络中心程度的变量,探讨社会互动对股市政策效应的影响。金融文本的语调与情绪含有上市公司管理层以及个体投资者表达的情感信息,并对股票市场产生影响,姚加权等^[25]通过词典重组和深度学习算法构建了适用于正式文本与非正式文本的金融领域中文情绪词典,并基于词典构建了上市公司的年报语调和社交媒体情绪指标,实证结果表明构建的年报语调指标和社交媒体情绪指标能有效地预测上市公司股票的收益率、成交量、波动率和非预期盈余等市场因素,并优于基于其他广泛使用情绪词典构建的指标。

受共同的宏观经济或行业因素影响,股指成分股价格波动之间存在复杂的关联性。从微观层

面上看,隶属于相同行业板块的股票价格波动趋同性较高;从中观层面上看,单一行业的价格波动可能影响其他行业的价格波动,这种“波动溢出效应”意味着“板块联动”甚至“普涨普跌”极易发生^[8];从宏观层面上看,宏观经济波动会直接影响到企业的外部交易环境进而对股票市场稳定性产生冲击^[9]。尤其是在经济全球化背景下,机构投资者之间互相持股关系紧密,外界任何不利因素的冲击都会导致股票市场系统性风险的不不断累积。因此,股票市场的稳定性与其自身的网络结构、行业波动溢出以及宏观经济状态密不可分^[15]。

自从 Mantegna^[10]首次通过构建股票关联网络发现股票市场层级结构,进而分析其与标普500指数收益率相关性以来,越来越多的国内外学者开始运用复杂网络理论方法对股指成分股价格波动关联进行建模分析。股票关联网络的构建是将股指成分股视为节点,成分股间的价格波动关联视为连边。全连接的股票关联网络包含较多的噪声信息,为简化网络并保留网络的关键信息,已有研究通常运用最小生成树(Minimum Spanning Tree, MST)^[11]、阈值模型(Threshold Model, TM)^[12]或平面极大过滤图(Planar Maximally Filtered Graph, PMFG)^[13]来过滤关联性较弱的连边。股票关联网络领域的研究主要分为静态网络研究和时序网络研究。静态网络研究基于固定的时间窗口构建股票关联网络,侧重于研究网络的基本拓扑结构特征及其层级结构等。Chi等^[11]基于赢家通吃准则对纳斯达克指数采用阈值法构建股票关联网络,发现股指的平均日收益率仅仅和部分成分股有关。时序网络研究基于滑动窗口构建一系列在时间上连续的股票关联网络,并分析网络拓扑结构特征及其内在关联模式的动态变化规律。在此方面, Namaki等^[12]利用安曼证券交易市场价格波动指数的月度数据构建了股票关联网络并发现经济增长率对股票价格波动具有积极影响。上述研究所提取的股票关联网络特征大多属于中观尺度的拓扑结构特征,忽略了节点(成分股)的量价信息,难以充分捕捉股票市场内在的复杂性。

股票关联网络本质上隶属于图数据范畴,基于特定时间窗口生成的股票关联网络具有一个隐

含的图标签(例如下时刻股指的运动方向). 股指预测问题也可被形式化为一种图分类问题. 图分类(Graph Classification)作为一类重要的图挖掘任务,旨在对不同类型的图数据进行分类预测. 近年来,随着深度学习的兴起,图神经网络(GNN)^[26]被广泛应用于有监督图分类任务. 这类方法在首先学习得到节点嵌入向量后,进行整合转化生成图级别的表征向量,进而为后续的图分类提供数据支撑. Niepert 等^[27]提出了一个基于图结构的 CNN 网络框架(PATCHY-SAN)用于有向图和无向图的分类任务. Wang 等^[28]提出了一种计算点云特征的算法 DGCNN,提升了点云分类、分割等任务的性能. Zhang 等^[29]提出了一种用于图分类的端到端的深度学习架构 SortPool,该架构将多层图卷积神经网络输出拼接在一起,通过对节点进行排序来进行全局池化. Ying 等^[18]提出了一个用于图分类的可微池化模块 DIFFPOOL,该模块可以生成图的层次表征,并以端到端的方式与各种 GNN 模型相结合. 实证表明,该方法在图分类任务上的平均准确率可提高 5%~10%. 尽管现有图分类研究在社会网络分析、计算机视角和化合物识别等应用领域取得了显著效果,还鲜有工作研究时序股票关联网络的分类问题. 本文借鉴现有图分类算法思想,将深度学习框架应用于时序股票关联网络的特征提取和候选股指预测信号的学习.

1.2 股指预测模型

股指预测是金融时间序列分析领域的一个重要的研究问题. 上世纪 60 年代,以美国芝加哥大学 Fama 为代表的经济学家提出了有效市场假说. 该假说认为市场是理性的,信息在市場中的传递是高速有效的,任何能影响金融资产价格的公开信息都能够被及时、理性地反映到资产价格当中,所以金融市场的价格也是随机不可预测的. 然而 Abu-Mostafa 等^[30]实证分析认为,尽管金融时间序列的趋势变化从短期来看是随机的,但从长期来看确有确定性发展趋势,这也为金融时间序列的预测分析提供了理论基础. 伴随着股票市场的异常波动和越来越多的市场异象,很多学者开始重新审视有效市场假说,进而提出了众多金融时间序列模型来把控和分析金融市场的变化. 其

中,早期比较有代表性的是 ARIMA 模型^[2]和 GARCH 模型^[3];前者对提取金融时间序列的线性特征有很大的优势,后者聚焦于金融市场的波动性,并考虑了引起波动项的滞后值和扰动项条件方差的滞后值. 随着计算机科学的发展以及相关研究的逐步深入,一些研究者从不同的角度,提出了支持向量机(SVM)^[4]、决策树(DT)^[5]、人工神经网络模型(ANN)^[6]等因果关系模型,这些模型可以通过对历史金融时序数据有效处理来分析 and 挖掘金融市场的潜在规律并对市场进行概率性的预测. 为了克服传统时间关系模型分析处理收益率数据非线性因素的不足,张贵生和张信东^[14]提出了一种基于近邻互信息的股票价格预测模型(SVM-GARCH),实验表明该模型在时序数据除噪、趋势判别以及预测精确度等方面均优于传统的 ARMA-GARCH 模型.

1.3 强化学习在金融领域的应用研究

强化学习(Reinforcement Learning)^[31]能够通过和环境的交互来自我学习和更新,也是一种不断试错、通过得到的评价性信息来不断修正自己行为的机器学习算法. 近年来,部分学者将强化学习应用于金融领域的价格预测和投资组合研究. 强化学习在金融领域的最早应用可以追溯到 1969 年, Samuelson 等^[32]将动态规划算法运用于最优投资与消费选择策略的求解,给出在连续时间下两类资产的最优投资与消费问题的解决办法. 自 Q-Learning 算法在上世纪 90 年代被提出后, Moody 等^[33]使用股指收益率作为价值函数,在月度标普 500 指数上对比了 Q-Learning 和最大化即时奖励的实时循环算法(RTRL),实证分析表明强化学习在较低风险下可取得更高的收益. Lee 等^[34]提出了一种基于多代理 Q-Learning 方法,通过定义合作执行股票定价和选择决策的必要角色来有效地划分和克服股票交易问题. Deng 等^[35]尝试通过引入深度神经网络来实现实时金融信号表示和交易,该方法利用深度学习自动感知动态市场条件,然后利用强化学习模块与深层表示交互做出交易决策以在未知环境中积累最终奖励. 梁天新等^[36]以金融领域常用的强化学习模型的发展为脉络,对交易系统、自适应算法、交易策略等方面的诸多研究成果进行了综述. Wang

等^[37]基于深度强化学习提出一种能够有效在回报和风险之间取得平衡的 DeepTrader 模型来优化投资策略,该模型包括两个互补且统一的决策模块:前者以负的最大回撤作为奖励函数,通过动态调整基金间的比例捕获资产之间的时间和空间依赖关系;后者则是以股票收益率为奖励函数,通过评估单个资产从历史数据中学习动态投资模式。经典强化学习如动态规划法、蒙特卡洛法和时序差分法共同点是需要维持一个 Q 值矩阵,当状态-动作对过多时,在训练过程中会陷入“维数灾难”问题。近年来,将深度学习与强化学习相结合的深度强化学习算法受到了研究人员的广泛关注。深度强化学习(Deep Q-Learning, DQN)^[19]通过引入神经网络结构从环境中学习状态信息,然后将状态映射到相应的动作上,并基于环境反馈的回报来衡量动作的价值。从应用场景视角看,现有的基于强化学习的金融序列分析研究主要应用于投资组合,鲜有学者将强化学习应用于股指预测。本文尝试利用强化学习解决自适应股指预测问题,其主要挑战是如何高效表征复杂股票市场以学习环境状态和生成候选预测信号。从决策者视角来看,对股指的预测并不能直接影响股票市场。因此,更加需要设计股票市场环境对智能体历史决策的反馈机制,以实现智能体对自身决策模式的动态调整。

2 股票关联网络层次化表征学习

股指成分股价格波动关联是股票市场运行的基本规律。股票市场中少数权重(龙头)股票的价格波动,会引发行业板块内其他股价的波动;受行业波动溢出影响,市场的“板块联动”现象极易发生,进而可能引起较大规模的系统性风险。本节试图利用股指成分股的基本面数据和历史行情数据,基于滑动窗口技术构建时序股票关联网络;进而考虑不同行业波动溢出效应对股指运行趋势的影响,利用 DIFFPOOL^[18]深度学习框架学习时序股票关联网络的层次化表征;最后将其作为特征输入可微分类模型生成候选股指预测信号。

2.1 股指时间序列属性图模型

目前构造股票关联网络最常用的方法是基于

股票间价格波动的相关性,首先计算成分股价格序列的 Pearson 相关系数,然后通过最小生成树(MST)、阈值模型或平面极大过滤图(PMFG)得到邻接矩阵,最后提取网络的拓扑指标研究其与股指波动的相关性。由于 Pearson 相关性只能刻画时间序列间线性相关性,而不能度量股票市场中非常重要的非线性特征,以及在非常规金融风险事件发生场景下成分股间的关联性。为此,引入复杂网络领域的可视图(Visibility Graph, VG)模型^[17]来分析股指成分股间的相依结构。VG 模型常被用来描述随机、分形和混沌的单变量时间序列,以及多元时间序列,已有文献证明 VG 模型非常适合捕捉非平稳系统中的时间序列不可逆性,从而使其在金融领域的应用成为可能。Yang 等^[38]将汇率序列映射成具有层次结构的无标度网络,进而用来量化序列的 Hurst 指数。Yan 等^[39]指出时间序列可视图的连通性可以用来测量股票价格的超指数变化的幅度,进而利用 VG 中节点的度来预测金融市场价格峰值和低谷。Flanagan 和 Lacasa^[40]利用 VG 模型量化 1998 年—2012 年期间的 35 个金融指数的时间不可逆性,实验结果表明 VG 模型可以充分适应金融时间序列情景,揭示股票价格演化结构中潜在的有价值信息。Goncalves 等^[41]用信息理论概念分析了全球贸易市场可视图的连通性,通过提取市场风险的量词构建风险指标,实证表明该指标与金融不稳定时期高度相关。Zou 等^[42]利用复杂网络方法来描述基于时间序列的动力系统,将 VG 模型应用于工程学和经济学等不同领域,并展示了时间序列网络在解决现实世界科学问题方面的巨大潜力。

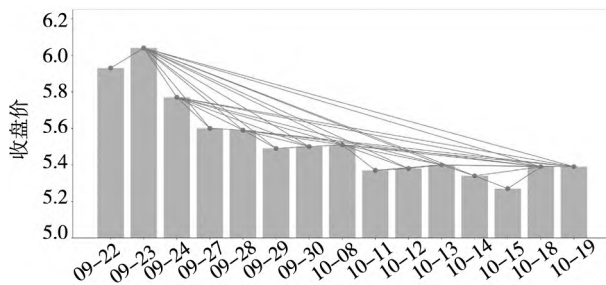
考虑到股指成分股定期调整的影响,在构建 t 时刻股票关联网络时,基于股指当前时刻最新的成分股数据,首先选取股指成分股 i 在时间范围 $\alpha \in [t-l+1, t]$ 内的收盘价序列 $P_i^{\alpha, l} = (p_i^{t-l+1}, \dots, p_i^t)$,其中 l 表示滑动时间窗口长度。VG 模型可将成分股 i 的收盘价序列 $P_i^{\alpha, l}$ 投影到如图 1 所示的二维平面。具体地, VG 采用直方图可视化序列 $P_i^{\alpha, l}$ 中的 l 个样本点,每个样本点 (α, p_i^α) 看作可视网络中的一个节点,任意两个节点 (α, p_i^α) 和 (β, p_i^β) ($\alpha < \beta$) 之间如果存在连边,当且仅当

$p_i^\gamma < p_i^\beta + (p_i^\alpha - p_i^\beta) \frac{\beta - \gamma}{\beta - \alpha}$, 即在序列 $P_i^{t,l}$ 中不存在样本点 (γ, p_i^γ) ($\alpha < \gamma < \beta$) 阻挡 (α, p_i^α) 和 (β, p_i^β) 彼此视线. 基于上述规则, VG 从收盘价序列 $P_i^{t,l}$ 中提取一个无向连通的可视网络, 并获得该网络的边集

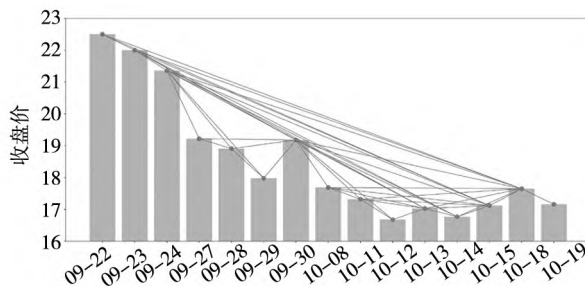
$$E_i^{t,l} = \left\{ e_{\alpha\beta}^{t,l} \mid \alpha < \gamma < \beta, p_i^\gamma < p_i^\beta + (p_i^\alpha - p_i^\beta) \frac{\beta - \gamma}{\beta - \alpha} \right\} \quad (1)$$

使用时间窗口内可视网络边集的 Jaccard 相似性来度量任意两支成分股 i 和 j 间股价波动相关性

$$\varpi_{i,j}^{t,l} = \text{Jaccard}(E_i^{t,l}, E_j^{t,l}) = \frac{|E_i^{t,l} \cap E_j^{t,l}|}{|E_i^{t,l} \cup E_j^{t,l}|} \quad (2)$$



(a) 申万宏源 2021 年 9 月 22 日到 10 月 19 日的收盘价可视图



(b) 中国宝安 2021 年 9 月 22 日到 10 月 19 日的收盘价可视图

图 1 基于可视图模型的收盘价序列投影方法

Fig. 1 The projection method of closing price sequence based on visibility graph model

如图 1 所示, 当时间窗口长度 $l = 15$, 成分股申万宏源在 2021 年 9 月 22 日到 2021 年 10 月 19 日时间范围内的收盘价序列可被投影到一个包含 42 条连边的可视网络, 成分股中国宝安则可被投影到一个包含 45 条连边的可视网络. 从图 1 不难看出, 两支成分股在相应时间窗口内都呈下降趋势, 但是它们价格波动的局部差异显著: 申万宏源在 9 月 24 日和 9 月 27 日经历连续两次较大幅度的下跌后, 其股价在后续 12 个交易日内呈现较为平稳的运行态势; 相反, 中国宝安股价波动较

为剧烈, 在 9 月 30 日出现了单日大幅上涨后, 其股价形成了一个局部新高. 基于 VG 模型对股价序列的投影, 可以清楚地观察到局部新高所对应时间节点 (如图 1(b) 中的 9 月 30 日) 的度相对较大; 而局部新低所对应时间节点 (如图 1(b) 中的 9 月 29 日) 的度相对较小. 基于式 (2), 两支成分股在相应时间窗口内可视网络边集的 Jaccard 相似度仅为 0.475, 表明它们价格波动的相似度较低. 与此相反, 价格波动的 Pearson 相关系数高达 0.958. 由此可见, 相较于 Pearson 相关性, VG 模型可以识别时间序列的局部拐点, 进而可以准确捕捉股指成分股价格波动的局部差异.

基于上述计算过程, 可计算出股指成分股在时间窗口 $[t-l+1, t]$ 内的相似矩阵. 考虑到相似性矩阵为全连接矩阵, 为了剔除冗余的连边, 采用平面极大过滤图 (PMFG) [13] 法对相似性矩阵做如下裁剪: 1) 将成分股价格波动相似度计算结果从大到小进行排序; 2) 依次从剩余的节点对中选择价格波动最相似的节点对, 在保证图的可平面性质条件下进行连接; 3) 重复上述过程直至遍历完所有节点对. 最终得到满足可平面性质的最大过滤图. 相比于最小生成树 (MST) 和阈值模型, PMFG 不仅可保留图的层次化结构, 还允许有更多的连边存在, 其过滤后的平面图中还保留环和派系等子图结构.

最后, 选取 7 种常用的基本面指标和 14 种技术指标 (如表 1 所示), 来刻画时间窗口 $[t-l+1, t]$ 内成分股的运行状态, 进而构建该窗口内股票关联网络的属性图 $G^{t,l} = \langle A^{t,l}, F^{t,l} \rangle$, 其中 $A^{t,l} \in \{0, 1\}^{n \times n}$ 表示股票关联网络的邻接矩阵, $F^{t,l} \in \mathbb{R}^{n \times k}$ 表示股票关联网络的属性矩阵, 其中每一行表示成分股 i 在 t 时刻的基本面和技术指标. 与此同时, 股票关联网络 $G^{t,l}$ 具有一个隐含的图标签 y^t 信息, 即股指在 $t+1$ 时刻相对于 t 时刻的运动方向. 令 I^t 表示股指 I 在 t 的指数价格, 如果 $I^t < I^{t+1}$, 则 $y^t = +1$; 反之, $y^t = -1$.

为了准确刻画股票市场的整体运行状态, 引入包括消费物价指数 (CPI)、生产者物价指数 (PPI) 和采购经理人指数 (PMI) 在内的 3 种指标刻画宏观经济整体的运行状态. 同时, 引入牛熊分界指标 (BBI) [43] 刻画股票市场的整体运行状态; 当月度股指收益率大于近 3 年月度收益率

均值时,相应指标变量取值为 1(表示当前市场为牛市环境);反之取值为 0.此外,文献[44]基于现有相关研究成果从理论上建立了考虑消费与投资者情绪的资产定价模型,然后选取 2005 年 5 月至 2018 年 4 月期间的中国 A 股上市公司数据以及消费和投资者情绪数据,研究结果表明在资产定价模型中引入消费和投资者情绪因子能够在理论上对 CAPM、Fama-French 三因子和五因子模型进行合理拓展和修正.参考文献[45]的做法,选用封闭式基金折价率、IPO 数量、IPO 首日收益和

换手率作为原始投资者情绪指标,然后进行主成分分析,选取前两个主成分作为投资者情绪的复合指标(ISI):当情绪指标高于总体均值时,相应情绪指标取值为 1(表示当前市场投资情绪良好);反之取值为 0.最后,借鉴文献[46]的思想,基于公募基金季度持仓数据,利用 Pearson 相关性、Kendall 秩相关性和 Tail 相关性构建金融机构间的关联网;然后选取各层网络的平均权重作为市场整体运行状态(见表 2)的补充.

表 1 成分股运行指标定义

Table 1 Definitions of the movement indices of constituent stocks

成分股运行指标		指标定义
基本面指标	每股收益	税后利润(净利润)与股本总数的比率
	总资产收益率	净利润与总资产的比率
	净资产收益率	净利润与净资产的比率
	市盈率	每股市价与每股盈利的比率
	市净率	每股股价于每股净资产的比率
	每股净资产	股东权益(净资产)与流通在外普通股股数的比率
	股息率	股息与股票买入价之比
技术指标	收益率	$ROC_i^{t,l} = (p_i^t - p_i^{t-l+1}) / p_i^{t-l+1}$, p_i^t 表示成分股 i 在 t 时刻的收盘价
	相对强弱指标	$RSI_i^{t,l} = 100 - 100 / (1 + RS_i^{t,l})$, $RS_i^{t,l}$ 表示成分股 i 在 $[t-l+1, t]$ 内涨幅合计与跌幅合计比值
	William's %R	William's % $R_i^{t,l} = (H_i^{t,l} - p_i^t) / (H_i^{t,l} - L_i^{t,l})$, $H_i^{t,l}$ 和 $L_i^{t,l}$ 分别表示成分股 i 在 $[t-l+1, t]$ 内的最高和最低价
	Stochastic %K	Stochastic % $K_i^{t,l} = (p_i^t - L_i^{t,l}) / (H_i^{t,l} - L_i^{t,l})$
	Stochastic %D	Stochastic % $D_i^{t,l} = (\%K_i^{t,l} + \%K_i^{t-1,l} + \%K_i^{t-2,l}) / 3$
	滑动平均	$MA_i^{t,l} = \left(\sum_{o=1}^l p_i^{t-l+o} \right) / l$
	指数滑动平均	$EMA_i^{t,l} = 2(p_i^t - EMA_i^{t-1,l}) / (l+1) + EMA_i^{t-1,l}$
	加权滑动平均	$WMA_i^{t,l} = \left(2 \sum_{o=1}^l op_i^{t-l+o} \right) / (l(l+1))$
	三角滑动平均	$TMA_i^{t,l} = \left(\sum_{o=1}^l MA_i^{t-l+o,l} \right) / l$
	平滑异同滑动平均	$MACD_i^{t,l} = EMA_i^{t,9} - EMA_i^{t,26}$
	市场偏差指标	$MBI_i^{t,l} = (p_i^t - MA_i^{t,l}) / MA_i^{t,l}$
	顺势指标	$CCF_i^{t,l} = (TP_i^t - SMATP_i^{t,l}) / MD_i^{t,l}$, $TP_i^t = (H_i^t + L_i^t + p_i^t) / 3$, $SMATP_i^{t,l}$ 表示 TP_i^t 的简单算数平均, $MD_i^{t,l}$ 表示移动窗口内 TP_i^t 和 $SMATP_i^{t,l}$ 的平均绝对差
	心理线指标_	$PSY_i^{t,l} = RDay_i^{t,l} / l$, $RDay_i^{t,l}$ 表示成分股 i 在 $[t-l+1, t]$ 内的上涨天数
	平衡交易量指标	$OBV_i^{t,l} = OBV_i^{t-1,l} + \phi_i^t v_i^t$, v_i^t 表示成分股 i 在 t 时刻的成交量,若当日股价相当于前一日收盘价上涨, $\phi_i^t = 1$,否则 $\phi_i^t = -1$

表 2 股票市场运行指标定义

Table 2 Definitions of the movement indices of stock market

市场运行指标	指标定义
消费物价指数 (CPI)	
生产者物价指数 (PPI)	
采购经理人指数 (PMI)	
牛熊分界指标 (BBI) ^[43]	当股指收益率大于近 3 年月度收益率均值时, 相应指标变量取值为 1; 反之取值为 0
投资者情绪指标 (ISI) ^[45]	当复合情绪指标高于总体均值时, 取值为 1; 反之取值为 0.
$\text{Avg. } w_{ij}^p$ ^[46]	基于 Pearson 相关性的金融机构关联网络平均权值
$\text{Avg. } w_{ij}^k$ ^[46]	基于 Kendall 秩相关性的金融机构关联网络平均权值
$\text{Avg. } w_{ij}^t$ ^[46]	基于 Tail 相关性的金融机构关联网络平均权值

2.2 股票关联网络层次化表征和图分类算法

基于上述股指时间序列属性图模型, 结合滑动窗口技术构建时序股票关联网络. 然后融合股票关联网络的标签信息, 获得一组含有标签信息

的图数据 $D = \{(G^{u,l}, y^l), (G^{u+1,l}, y^{l+1}), \dots\}$. 这样, 股指预测问题可被形式化为一种图分类问题, 即学习一个映射函数 $f: G \rightarrow Y$, 将股票关联网络映射到一组标签的集合.

区别于传统的有监督分类任务, 解决上述问题的主要挑战来自如何从股票关联网络提取有用的分类特征. 传统的研究多是基于特征工程的方法, 人工定义股指预测模型的预测因子. 尽管这种方式构建的股指预测模型具有较强的可解释性, 但是人工定义的预测因子难以充分捕捉股票市场的非线性和非平稳性, 极大限制了股指预测结果的准确性和鲁棒性. 为此, 本研究引入深度学习方法, 自动学习适用于股指预测任务的预测因子. 考虑到相同行业板块的股指成分股具有显著的价格同步特征, 不同行业板块的股指成分股存在领先滞后和轮动效应^[15]. 借助可微池化 (DIFFPOOL)^[18] 框架层次化聚合股票关联网络的节点属性信息, 试图从不同行业视角提取影响股指价格波动的中观尺度预测因子, 进而重构股票关联网络的粗化图. 上述特征提取模型以股票关联网络分类任务为驱动, 其最终聚合得到的层次化表征向量与股票市场整体运行状态向量 (如表 2 所示) 拼接后, 作为可微分类器的特征输入, 整个系统使用随机梯度下降进行端到端的训练.

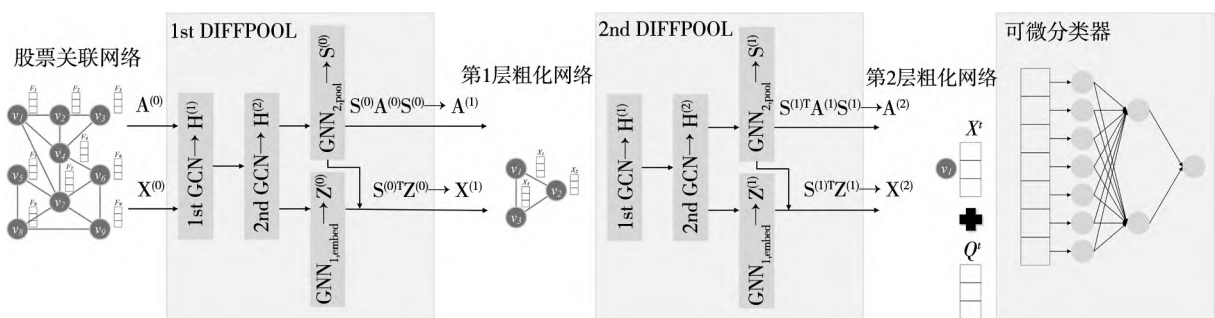


图 2 基于 DIFFPOOL 架构的股票关联网络分类模型

Fig. 2 The classification model for stock correlation network based on the DIFFPOOL framework

传统的 GNN 模型仅仅通过图的连边来传递信息, 且试图针对所有节点嵌入进行全局池化, 忽略了图中可能存在的层级结构. Ying 等^[18]提出了可微池化操作模块 DIFFPOOL, 其核心思想是通过一个微分模块分层地聚合节点信息, 然后基于

学习到的嵌入, 将节点映射为簇 (cluster) 的集合, 进而在 GNN 的每一层学习一个可微的软分配 (soft assignment), 最后通过分层的方式堆叠多个 GNN 层用于图分类任务. DIFFPOOL 的整体架构如图 2 所示. 为了学习有用的图分类表示, 本文在

单层 GNN 模块中使用图卷积神经网络 (Graph Convolution Network, GCN)^[47] 的消息传递方式. 给定一个股票关联网络 $(G^{t,l}, y^t)$, 初始化节点嵌入矩阵为 $H^{(0)} = F^{t,l}$, GCN 的消息传播函数可表示为

$$H^{(k)} = \text{ReLU}(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(k-1)} W^{(k-1)}) \quad (3)$$

其中 $\tilde{A} = A^{t,l} + I$ 表示带自环的股票关联网络邻接矩阵, \tilde{D} 为度矩阵, $W^{(k)} \in \mathbb{R}^{\kappa \times \kappa}$ 表示权重矩阵. 单层 GCN 模块迭代 K 次, 最终输出节点嵌入矩阵 $Z = H^{(K)}$.

由于 DIFFPOOL 架构不需要了解单层 GNN 模块的消息传递函数, 因此可以将经过 K 次迭代单层 GCN 模块的输出抽象表示为 $Z = \text{GCN}(A^{t,l}, F^{t,l})$. DIFFPOOL 旨在定义一种通用、端到端且可微的策略, 使得可以用一种层次化的方式堆叠多层 GNN 模型. 给定 GCN 模块的输出 $Z = \text{GCN}(A^{t,l}, F^{t,l})$ 和股票关联网络的邻接矩阵 $A \in \{0, 1\}^{n \times n}$, DIFFPOOL 试图寻找一种方式得到一个新的包含 $m < n$ 个节点的粗化图 (coarsened graph), 其中这个新图的邻接矩阵为 $A' \in \{0, 1\}^{m \times m}$, 节点嵌入矩阵为 $Z' \in \mathbb{R}^{m \times \kappa}$. 该粗化图可作为下一层 GCN 的输入, 重复 D 次可得到了一系列的粗化图.

定义第 d 层学习到的类簇隶属矩阵为 $S^{(d)} \in n_d \times n_{d+1}$ ($n_d > n_{d+1}$), n_d 表示第 d 层的节点数, n_{d+1} 表示第 $d+1$ 层的节点数, $S^{(d)}$ 可理解为第 d 层的每个节点到第 $d+1$ 层每个节点 (类簇) 的概率. 基于 $(A^{(d+1)}, X^{(d+1)}) = \text{DIFFPOOL}(A^{(d)}, Z^{(d)})$ 的方式进行图粗化, 其中 $A^{(d+1)}$ 表示第 $d+1$ 层类簇间的连接强度, 采用如下公式更新

$$A^{(d+1)} = S^{(d)T} A^{(d)} S^{(d)} \in \mathbb{R}^{n_{d+1} \times n_{d+1}} \quad (4)$$

$X^{(d+1)}$ 表示第 $d+1$ 层类簇的嵌入矩阵, 其更新策略如下

$$X^{(d+1)} = S^{(d)T} Z^{(d)} \in \mathbb{R}^{n_{d+1} \times \kappa} \quad (5)$$

DIFFPOOL 采取两种不同 GNN 模块分别学习每层的类簇隶属矩阵为 $S^{(d)}$ 和节点嵌入矩阵 $Z^{(d)}$

$$Z^{(d)} = \text{GNN}_{d, \text{embed}}(A^{(d)}, X^{(d)}) \quad (6)$$

$$S^{(d)} = \text{softmax}(\text{GNN}_{d, \text{pool}}(A, X^{(d)})) \quad (7)$$

其中 $\text{GNN}_{d, \text{pool}}$ 表示第 d 层的池化 GNN, 其输出维度对应于第 d 层预定义的最大类簇数. DIFFPOOL 框架的倒数第二层, 令 $S^{(D-1)}$ 为一个全 1 向量, 即所有节点在最后的第 D 层被划分到同一个类簇. 这样, DIFFPOOL 可生成对应于原始股票关联网络的一个嵌入向量 $X^t \in \mathbb{R}^{\kappa}$. 最后, 将 X^t 与股票市场整体运行状态向量 O^t 进行拼接后, 输入可微分类器. 整个系统使用随机梯度下降进行端到端的训练. 由于训练 $\text{GNN}_{d, \text{pool}}$ 本质上属于一个非凸优化问题, 本文使用辅助链路预测目标来训练池化 GNN, 即在第 d 层, 最小化如下的损失函数

$$L_{LP}^{(d)} = \|A^{(d)}, S^{(d)} S^{(d)T}\|_F \quad (8)$$

其中 $\|\cdot\|_F$ 表示 Frobenius 范数. 考虑到节点的类簇隶属应该接近于 one-hot 向量, 以便明确定义每个类簇或子图间的关系. 为此, 在每一层引入如下的损失函数来规范类簇隶属矩阵的熵

$$L_E^{(d)} = \frac{1}{n_d} \sum_{i=1}^{n_d} \text{entropy}(S_i^{(d)}) \quad (9)$$

其中 $\text{entropy}(\cdot)$ 表示熵函数. 在训练期间, 将每层的 $L_{LP}^{(d)}$ 和 $L_E^{(d)}$ 加入到分类损失函数中, 整个系统的损失函数定义为

$$L_{\text{total}} = L_{\text{base}} + \sum_{d=1}^D (L_{LP}^{(d)} + L_E^{(d)}) \quad (10)$$

其中 L_{base} 表示可微分类器 (如逻辑回归、多层感知器和栈式自编码器等) 的损失函数. 相较于直接从原始股指时间序列提取预测因子的传统股指预测方法, 本研究基于可微池化框架层次化提取股票关联网络的非线性特征, 有效聚合不同类簇的技术指标, 并基于股票关联网络的层次化聚类结果捕捉不同类簇的相互影响, 进而自动学习适用于股指短期趋势预测的预测变量.

3 自适应股指预测方法

股指价格受诸多因素共同影响, 呈现出不同的特点, 任何模型都很难将股票市场的影响因素全部囊括, 更不可能将全部因素的影响都体现在模型中, 单一的模型更是难以满足股指预测需求. 因此, 国内外一些学者试图构建组合预测模型综

合描述股指时间序列线性、非线性特征,期望在一定程度上提高预测模型的鲁棒性. 现有的组合预测模型大多是借鉴集成学习的思想,在要求个体学习器具备较高的预测精度同时,还需具备一定的多样性. 传统的组合预测模型缺乏对自身决策模式的学习完善,无论股票市场发生何种变化,历史预测精度如何波动,这类方法按照既定模式做出决策,缺乏对股票市场收益率可预测性的动态感知以及对自身决策模式的自动调整. 为此,研究试图设计可动态调整自身决策模式的自适应模型以提升股指预测的鲁棒性. 考虑到股指预测中动态时变的股票市场环境以及面向动态状态下的候选预测信号寻优目标,该问题可以被形式化一种强化学习(Reinforcement Learning)结构,即在动态环境(股票市场)与动作(候选预测信号选择)的交互过程获得奖励(股指收益率),并以最优化全局奖励为目标学习动作的选择策略. 引入 NDQN^[19]思想,提出了一种考虑动作评估反馈的深度强化学习方法 AEF-DQN(Action Evaluation Feedback Deep Q-Learning),以在决策过程中考虑股票市场环境对智能体历史决策反馈等影响. 具体地,研究将自适应股指预测问题建模为马尔科夫决策过程 $M(\mathbb{A}, \mathbb{S}, R)$, 其中 \mathbb{A} 表示动作空间, \mathbb{S} 表示状态空间, R 表示奖励函数. 为简化起见, AEF-DQN 以单个交易日为决策周期,在每个决策周期,智能体通过与股票市场环境的交互,选择动作空间的候选预测信号对股指短期波动进行预测,并基于环境反馈的回报来衡量动作的价值. 决策过程 M 中的主要元素的定义如下:

1) 动作空间

针对自适应股指预测问题,考虑将 DIFFPOOL 模型与不同分类器组合学习到的候选股指预测信号作为智能体的动作. 考虑到股指常见有 3 种运行趋势:1) 股指持续上涨形成上升走势;2) 股指持续下跌形成下跌走势;3) 股指震荡运行形成震荡走势. 当股指运行在较为明显的上升(下降)趋势中时,训练样本中股指成分股特征指标将呈现较为显著的线性分布,此时考虑在 DIFFPOOL 模块使用逻辑回归(Logistics Regression,

LR)用于捕捉股指在上升(下降)趋势的走势特征. 当股指运行在较为明显的震荡趋势中时,训练样本中股指成分股特征指标将呈现较为显著的均匀分布,考虑在 DIFFPOOL 模块中使用多层感知器(Multilayer Perceptron, MLP)用于捕捉股指在震荡趋势中的走势特征. 需要特别指出的是,当股指运行在相对稳定的上升(下降)趋势中时,市场内部(外部)环境的突发利好(利空)消息可能会引发股指出现异常波动,进而打破其原先的阶段运行趋势. 此时的股指预测可看作解决非平衡样本的离群点检测问题. 有鉴于此,基于异常检测自编码器(Outlier Detection with Autoencoder Ensembles)^[48]设计思想,尝试在 DIFFPOOL 模型中引入一个简单的栈式自编码器(Stacked Autoencoders, SA)来动态捕捉股指运行过程中离群点(趋势拐点). 具体地,根据 DIFFPOOL 模块输出的表征向量 X^t 与市场整体运行状态向量 O^t 拼接后的向量维度 $K = |X^t| + |O^t|$, 构建一个结构为

$K \times \frac{K}{2} \times \frac{K}{4} \times 1$ 的编码器和一个结构为 $1 \times \frac{K}{4} \times \frac{K}{2} \times K$ 的解码器. 在模型训练过程中,将训练集中

类别占比较高的股票关联网络视为正例样本;反之则视为负例样本. 在此基础上,首先训练一个仅针对正例样本数据的栈式自编码器,然后利用负例样本的重构误差反向传播更新 DIFFPOOL 模型参数. 综上所述,本研究在 DIFFPOOL 模块分别使用 3 种可微分类器,即逻辑回归、多层感知器和栈式自编码器,对股票关联网络进行分层聚合进而生成候选预测信号. 其中,逻辑回归和栈式自编码器激活函数采用 Sigmoid 函数,多层感知器激活函数采用 tanh 函数. 为简化起见,智能体在 t 时刻可采取的动作 $a^t \in \mathbb{A} = \{\hat{y}^{LR}, \hat{y}^{MLP}, \hat{y}^{SA}\}$.

2) 状态空间

在每个时刻 t 的状态信息包含了股票关联网络的层次化表征信息、市场总体运行状态信息(见表 2)以及最近 w 个决策周期的动作评估信息. 具体地,本研究基于 DIFFPOOL 和 3 种可微分类器学习股票关联网络的层次化表征,进而可得

到一个 k 维嵌入向量 $\mathbf{X}^t \in \mathbb{R}^k$. 同时, 考虑智能体最近 w 个决策周期的动作评估向量 $\mathbf{U}^t = (\mu_1^t, \dots, \mu_w^t) \in \mathbb{R}^w$, 其中 $\forall o \in \{1, \dots, w\}$

$$\mu_o^t = \begin{cases} 1 & y^{t-w+o-1} = a^{t-w+o-1} \\ -1 & y^{t-w+o-1} \neq a^{t-w+o-1} \end{cases} \quad (11)$$

最后, 将两者与股票市场整体运行指标向量 \mathbf{O}^t 进行拼接, t 时刻的状态可表示为 $s^t = [\mathbf{X}^t, \mathbf{O}^t, \mathbf{U}^t]$.

3) 奖励函数

假设智能体在学习过程中不断模拟股指期货投资. 考虑到投资者对投资标的预期报酬越高则所能忍受的波动风险也越高, 反之则越低^[20]. 即理性的投资者在选择投资标的与投资组合时是在固定预期报酬下寻求最低风险或在所能承受固定风险下追求最大报酬为主要目的. 本研究引入前景理论^[21]中的参照依赖特性来估计奖励值函数. 为了降低极端行情对动作-值函数更新的影响, 利用 Sigmoid 函数对智能体的奖励值进行平滑处理: 如果智能体准确预测股指移动方向 ($e_w^{t+1} = 1$), 获得 $\sigma(|I^{t+1} - I^t| / I^t)$ 数值的收益, 其中 I^t 表示 t 时刻的股指价格, $\sigma(\cdot)$ 表示 Sigmoid 函数; 反之, 则获得相应数值的惩罚. 为了鼓励智能体在纠错的同时尽可能持续采取正确动作, 基于智能体当前动作的评估反馈考虑两种投资收益参考点的设置: 如果智能体准确预测了股指移动方向, 则其投资收益的参考点 (预期回报) 设置为最近 w 个决策周期采取动作所获最低奖励值; 反之, 将考点

设置为最近 w 个决策周期采取动作所获平均奖励值. 由此, t 时刻智能体的奖励值表示为

$$r^t = \begin{cases} \sigma\left(\frac{|I^{t+1} - I^t|}{I^t}\right) - \min_{o \in \{1, \dots, w\}} \mu_o^t \sigma\left(\frac{|I^{t-w+o} - I^{t-w+o-1}|}{I^{t-w+o-1}}\right) & \mu_o^{t+1} = 1 \\ -\sigma\left(\frac{|I^{t+1} - I^t|}{I^t}\right) - \frac{1}{w} \sum_{o \in \{1, \dots, w\}} \mu_o^t \sigma\left(\frac{|I^{t-w+o} - I^{t-w+o-1}|}{I^{t-w+o-1}}\right) & \mu_o^{t+1} = -1 \end{cases} \quad (12)$$

根据 Bellman 方程, 当前状态的动作——值函数为 $Q(s^t, a^t) = r^t + \gamma Q(s^{t+1}, a^*)$, 其中 s^{t+1} 表示智能体到达的下一个状态, a^* 表示下一个状态中最好的动作, r^t 为采取动作 a^t 到达下一个状态得到的收益. 为优化 $Q^*(s^t, a^t) = E_{a^*} [r + \gamma Q(s^{t+1}, a^*) | s^t, a^t]$, 传统的强化学习方式是在状态 s^t 下通过多次迭代训练更新 Q 值. 但当状态空间过大时, 传统方法在更新 Q 值时, 无论是在时间开销还是空间开销上都力所不逮. 为此, 引入 NDQN^[19] 算法, 通过神经网络 (Q 网络) 对 Q 值进行估计, 即 $Q(s^t, a^t; \theta) \approx Q^*(s^t, a^t)$. NDQN 的核心思想是选用两个结构相同的 Q 网络, 一个作为当前 Q 网络用于选择动作, 计算当前 Q 值, 并迭代更新模型参数; 另一个作为目标 Q 网络用于计算目标 Q 值. 为了减少目标 Q 值和当前 Q 值的相关性, 目标 Q 网络参数不需要迭代更新, 而是采用延时更新策略, 即每隔一段时间直接从当前 Q 网络复制过来.

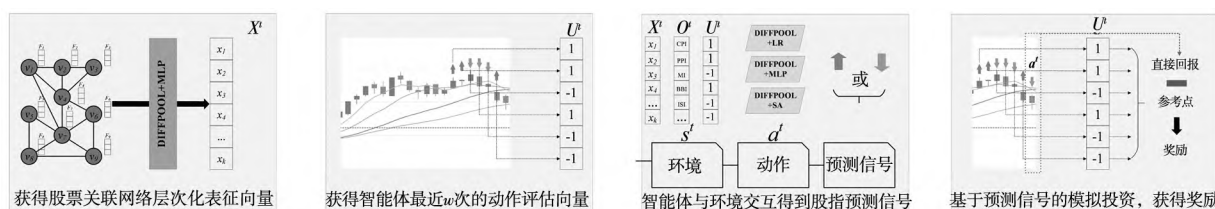


图3 自适应股指预测模型单步迭代过程

Fig. 3 The iterative process in single step of adaptive stock index prediction model

综上所述, 自适应股指预测问题可以被视为智能体在动态的股票市场中, 通过对股票关联网络进行层次化表征、当前市场运行环境更新以及历史动作评估反馈, 获得环境状态信息. 然后基于股票关联网络的分类结果执行某种预

测动作进而与环境交互, 股票市场受到动作的影响并返回智能体的奖励和下一个环境状态信息, 从而构成一个完整的单步迭代强化学习 (如图3所示). 本文所提 AEF-DQN 方法的伪代码如算法1所示.

算法 1: AEF-DQN 算法

输入: 经验回放池 Ξ 最大容量 $sample_limit$, 经验回放池 Ξ 最少元素数 $start_train_number$, 批处理大小 $batch_size$, 折扣系数 γ 、学习率 η 、更新频率 ρ 、最大迭代次数 $episode_max$ 、样本数 T

过程:

1. 随机初始化当前网络和目标网络参数: $\theta = \theta'$
2. 初始化经验回放池: $\Xi \leftarrow \phi$
3. 初始化动作评估向量: $U^t \leftarrow 0^m$
4. 初始化迭代次数: $episode \leftarrow 0$
5. While $episode \leq episode_max$:
6. $t \leftarrow 0$
7. 计算股票关联网络 $G^{t,l}$ 层次化表征 X^t
8. 计算股票市场总体运行向量 O^t
9. 计算智能体动作评估向量 U^t
10. 拼接获得系统状态: $s^t \leftarrow [X^t, O^t, U^t]$
11. While $t \leq T$:
12. 选择动作 $a^t \leftarrow \pi_{\epsilon}(s^t)$ (基于 ϵ 贪心策略(为较快收敛, 探索率 ϵ 随着迭代的进行而变小), 利用当前 Q 网络在状态 s^t 下选择动作)
13. 执行动作 a^t , 计算奖励 r^t 并更新 DIFFPOOL 网络参数和动作评估向量 U^{t+1} , 生成下一时刻的状态 s^{t+1}
14. If $|\Xi| \geq sample_limit$:
15. 删除 Ξ 中第一个元素
16. $\Xi \leftarrow \Xi \cup (s^t, a^t, r^t, s^{t+1}, t)$
If $|\Xi| \geq start_train_number$:
17. 从经验回放池中任取 $batch_size$ 个元素作为训练集, 对其中每一个元素 $(s_j, a_j, r_j, s'_j, t_j)$ 计算目标 Q 值

$$Q_j = \begin{cases} r_j & t_j = T \\ r_j + \gamma \max_{a'_j} Q(s'_j, a'_j; \theta') & t_j < T \end{cases}$$
18. 以 η 为学习率, 采用均方差损失函数 $\frac{1}{batch_size} \sum_{j=1}^{batch_size} (Q_j - Q(s_j, a_j; \theta))^2$ 训练当前 Q 网络, 并更新参数 θ
19. If $t \% \rho = 0$:
20. 更新目标 Q 网络的参数 $\theta' \leftarrow \theta$
21. $t \leftarrow t + 1$
22. End while
23. $episode \leftarrow episode + 1$
24. End while
25. 输出: 目标 Q 网络的参数 θ^*

4 实证分析

4.1 实验设置

为考察 AEF-DQN 模型在实际股票市场的适用性, 选取国内外 4 个最具代表性股票市场的股指成分股数据, 即沪深 300 指数(SHSZ)、标普 500 指数(SPX)、英国富时 100 指数(UKX)以及日经 225 指数(NKY)。实验将每个样本集分为两个部分, 前 70% 的样本作为样本内数据, 用于训练 DIFFPOOL 模型的深度图神经网络和深度强化学习的 Q 网络; 后 30% 的样本作为样本外数据, 用于股指预测效果的评价。表 3 描述了数据集的基本统计信息。从中可以看出, 4 个股指数据集中正负样本的分布较均匀。在样本外模型评估阶段, 实验基于滑动时间窗口技术对 DIFFPOOL 模型参数和深度强化学习模型中的动作-值函数进行递归估计。具体地, 为了产生第 t 交易日的候选股指预测信号, 首先用第 t 个交易日的前 T_{train} 个样本(基于 t 时刻的股指成分股数据构建股票关联网络)迭代训练 DIFFPOOL 模型, 以更新 DIFFPOOL 模型的权重矩阵; 然后利用训练后的模型生成第 t 个交易日候选股指预测信号。为了解决动作-值函数的适应性问题, 利用第 t 个交易日前 20% 样本更新 Q 网络参数, 以获得最终股指预测结果。

表 3 股指数据集统计描述

Table 3 Statistical descriptions of stock index datasets

数据集	训练集			测试集		
	时间跨度	上涨天数	下跌天数	时间跨度	上涨天数	下跌天数
SHSZ	2005-10-11— 2015-07-24	1 295	1 085	2015-07-27— 2019-09-26	529	490
SPX	2011-10-26— 2017-05-11	746	647	2017-05-12— 2019-09-24	334	263
UKX	2015-02-23— 2018-05-10	426	386	2018-05-11— 2019-09-24	174	174
NKY	2016-10-14— 2018-10-30	270	234	2018-10-31— 2019-09-24	118	98

本文构建一个简易股指期货日内交易模拟器来评估股指预测模型的性能. 为简化起见, 该模拟器只考虑交易手续费、交割手续费和平仓手续费3种期货交易成本, 且不考虑买卖期货时所需的最低保证金要求. 具体地, 根据中国金融期货交易所2022年2月9日公布的数据, 一手沪深300股指期货合约的交易手续费标准是万分之0.23, 交割手续费标准是万分之0.5, 平仓收取率标准是万分之3.45, 合约乘数为300元/点. 股指期货采用T+0的交易方式, 不仅可以做多, 也可以做空. 做多即对行情看涨, 买入股指期货, 如果当日股指价格上涨, 则平仓盈利; 做空即对行情看跌, 卖出股指期货, 如果当日股指价格下跌, 则买入归还获得盈利. 模型根据第 t 个交易日的股指预测结果(看涨或看跌), 在第 $t+1$ 个交易日开盘时将现有模拟资金全部用于做多或做空操作(期货合约价格与当日股指开盘价相同), 并在日内股指收盘时进行强制平仓. 基于上述交易和结算规则, 假设初始投资本金为10 000 RMB, 采用4种回测指标衡量不同预测模型的效果:

1) 胜率(Profitable Rate, PR): 统计在投资周期内盈利次数的比率.

2) 最大回撤率(Maximum Drawdown Rate, MDR): 最大回撤率是一个重要的风险指标, 其统计投资周期内资金走到最低点时, 收益率下降幅度的最大值.

3) 阿尔法比率(Alpha Ratio, AR): 阿尔法比率衡量投资策略相对于市场无风险利率和市场基准收益率的超额回报. 实验中基于总的交易次数计算均值和标准差, 市场无风险利率设置为4%.

4) 夏普比率(Sharp Ratio, SR): 夏普比率是常用的衡量风险与收益的指标, 衡量每增加一单位的风险相应所获得的额外回报. 本研究实验中夏普比率同样是基于总的交易次数计算均值和标准差.

参考股指预测领域文献^[49]的实验设置, 本研究文构建了一系列用于股指预测的神经网络模型, 包括基于多层感知机(Multi-Layer Perceptrons, MLP)、卷积神经网络(Convolution Neural

Network, CNN)、递归神经网络(Recurrent Neural Network, RNN)、长短期记忆网络(Long Short-Term Memory, LSTM)和门控神经单元(Gated Recurrent Unit, GRU)的模型. 同时, 针对RNN、LSTM和GRU等网络模型难以充分利用时间维度信息的问题, 引入注意力机制(Attention Mechanism)来区分不同时间维度信息对预测的重要程度. 实验使用股指前70%的样本数据作为训练集, 剩余30%的样本数据作为测试集. 区别于研究所提的股票关联网络层次化表征的特征提取方法, 上述基于神经网络的股指预测模型的特征指标仅包含股指开盘价、收盘价、最高价、最低价、当日成交量、涨跌幅和换手率在内的7种基础交易指标以及8种股票市场运行指标(见表2), 对应于模型中每一交易日下特征维度. 实验使用当天所有15项指标来预测下一交易日股指收盘价. 训练过程中, 采取最近30个交易日的的数据来预测最新的股指收盘价, 并使用均方根误差作为损失函数使其输出结果不断逼近真实数据. 同时, 实验将考察和对比AEF-DQN模型与基础的股票关联网络分类模型的性能差异, 考虑在DIFFPOOL模块使用3种可微分类器, 即逻辑回归、多层感知器和栈式自编码器, 以观察其对股指预测任务的实用性. 其相应预测模型分别表示为DIFFPOOL-LR、DIFFPOOL-MLP和DIFFPOOL-SA. 实验的基准方法还引入基于XGBoost的组合预测模型^[16], 该模型的输入特征同基于神经网络的预测模型一致. 最后, 为了考察模型的自适应性, 实验还将对比AEF-DQN模型与经典Q-learning和Sarsa模型的性能. 即在强化学习部分, 选择基础的Q-learning和Sarsa训练动作-值函数(相应模型本别表示为DIFFPOOL-QL和DIFFPOOL-Sarsa). 区别于研究所提方法, DIFFPOOL-QL和DIFFPOOL-Sarsa模型在训练和预测过程中的状态空间仅仅是由二值化股票关联网络表征向量和股票股票市场整体运行指标向量拼接而成.

4.2 参数分析

研究所提的自适应股指预测模型存在诸多超参, 既包含用于股票关联网络层次化表征的DIF-

FPOOL 模型参数,还包含对动作-值函数进行预测的 Q 网络参数. 本节利用网格搜索交叉验证 (GridSearchCV) 方法对模型参数进行整定.

首先,将股指预测问题看作是有监督图分类问题来确定 DIFFPOOL-MLP 模型的最优参数. 由于股票关联网络的节点数较少,根据文献^[18]的建议, DIFFPOOL 模型中 GCN 的迭代数和池化层数都设置为 2. 实验选取 2005 年 10 月 11 日到 2015 年 07 月 24 日 SHSZ 股指样本,创建一个超参网络: 测试集大小 T_{test} : [5, 6, 7, 8, 9, 10, 11, 12], 训练集大小 T_{train} : [20, 25, 30, 35, 40, 45, 50, 55], 隐藏层维度 d : [20, 40, 60, 80, 100, 120, 140, 160], 节点粗化率 δ : [0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45], 迭代次数 τ : [5, 10, 15, 20, 25, 30, 35]. 然后将训练样本等分为 5 个部分,使用 5 折的交叉验证并运行 200 个超参组合以得到最佳的超参

数. 图 4 展示了不同参数组合 200 次交叉验证结果. 不难发现在动态训练股票关联网络层次化表征模型过程中,随着测试集和训练集规模的增加,模型的投资胜率呈先升后降趋势. 当隐藏层维度 d 持续增加时,预测模型的投资胜率表现出较为稳定的上升趋势,而当该参数超过 140 时,投资胜率的提升较为平缓. 进一步分析节点粗化率 δ 对模型预测性能的影响,该参数直接影响股票关联网络层次化聚合后类簇结构的分辨率: 较低的粗化率会产生众多规模较小的类簇;反之,则会产生少数规模较大的类簇. 基于图 4(d) 的参数分析结果,将节点粗化率 δ 设置为 30%. 最后,观察图 4(e) 可以发现模型迭代次数 τ 的增加并非能显著提升模型的预测精度,较高的迭代次数可能会造成过拟合现象.

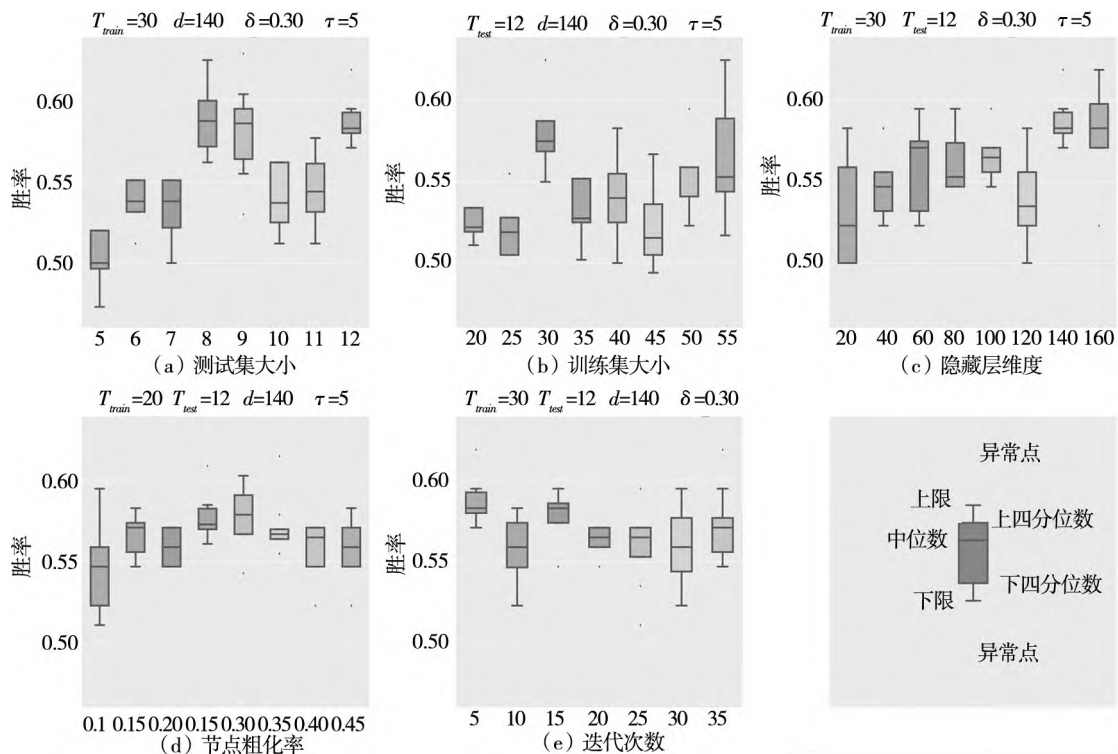


图 4 DIFFPOOL-MLP 模型参数分析

Fig. 4 Parametric analysis of DIFFPOOL-MLP model

然后,采取相同方式对深度强化学习模型进行参数分析. 实验创建一个超参网络: Q 网络隐藏层维度 d' : [5, 10, 15, 20, 25], 目标 Q 网络更新

频率 ρ : [5, 10, 15, 20, 25, 30], 随机数 ε : [0.9, 0.95, 0.99], Q 网络学习率 η : [0.01, 0.05, 0.1], 折扣系数 γ : [0.5, 0.6, 0.7, 0.8, 0.9]. 图 5 汇报

了不同参数组合 200 次交叉验证结果. 随着 Q 网络隐藏层维度 d' 、目标 Q 网络更新频率 ρ 和随机数 ε 的增加, 基于深度强化学习模型的投资胜率

呈先升后降趋势. 随着 Q 网络学习率 η 的增加, 模型预测性能持续下降; 相反, 模型预测性能随着折扣系数 γ 的增加而提升.

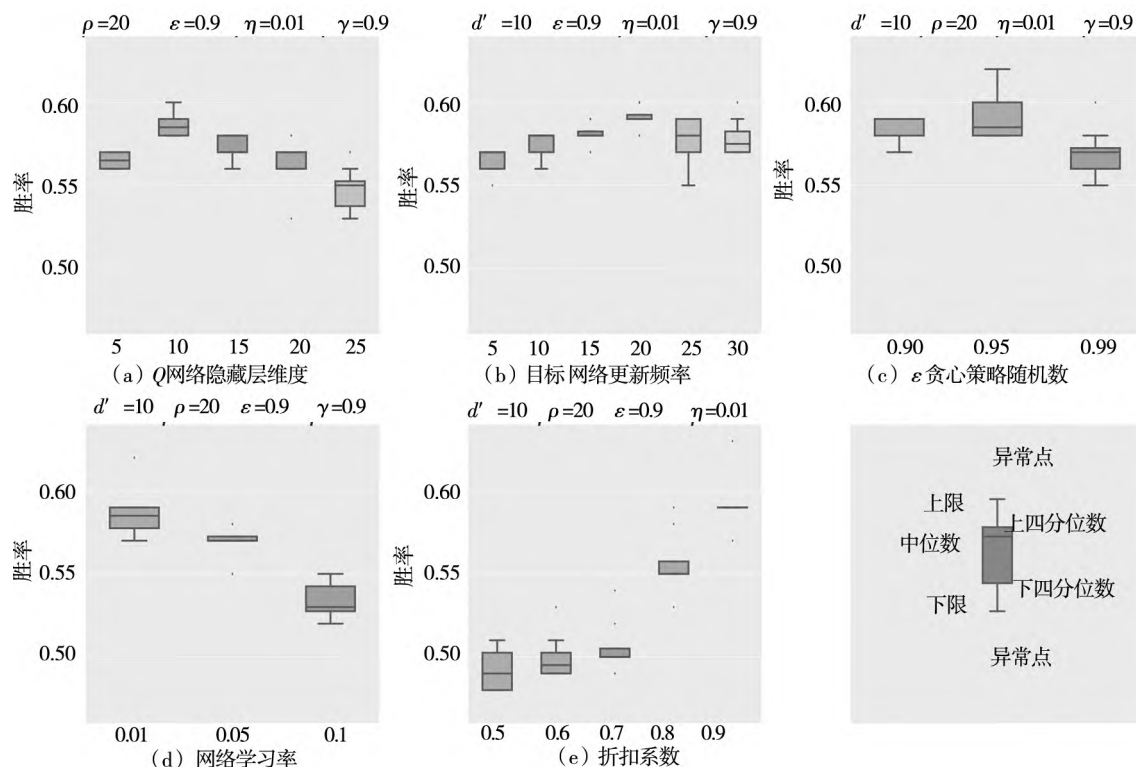


图5 深度强化学习模型参数分析

Fig. 5 Parametric analysis of DQN model

综上所述, AEF-DQN 算法的参数设置如表 4 所示. 在构建股票关联网络时, 滑动窗口长度 l 设置为 15, 任意两支成分股间的股价波动相关性用相应可视网络边集的 Jaccard 相似性来度量, 最后采用平面极大过滤图 (PMFG) 方法对相似性矩阵进行裁剪, 生成股票关联网络. 在训练股票关联网络层次化表征模型时, 用于 DIFFPOOL 的 GNN 模型建立在图卷积神经网络 GCN 架构之上, 其迭代次数 K 设置为 2, 并在每两个 GCN 层之后应用一个 DIFFPOOL 层. 在每个股指数据集上共使用 $D=2$ 个 DIFFPOOL 层. 同时, 采用两种不同的稀疏 GraphSAGE 模型^[50]分别计算每层的类簇隶属矩阵为 $S^{(d)}$ 和节点嵌入矩阵 $Z^{(d)}$. 在样本外测试过程中, 对 DIFFPOOL 模型采用动态训练的方式生成候选股指预测信号: 测试集大小 T_{test} 设置为 8, 训练集大小 T_{train} 设置为 30, 隐藏层维度 d 设置为 140, 节点粗化率 δ 设置为 30%, 迭代次数 τ 设

置为 5. 在训练基于强化学习的自适应股指预测模型时, 分别基于 3 种可微分类模型与 DIFFPOOL 模型的组合, 学习每个决策周期的动作 (候选股指预测信号), 即动作空间的规模为 $|A|=3$. 动作评估向量 U' 的维度 w 设置为 5, 通过与股票关联网络层次化表征向量和市场总体运行状态向量的拼接, 获得最终的状态向量 s' . 针对每个股指训练集, 采用 1 000 个 epoch 训练 Q 网络: Q 网络隐藏层维度 d' 设置为 10, 目标 Q 网络更新频率 ρ 设置为 20, 随机数 ε 设置为 0.9, Q 网络学习率 η 设置为 0.01, 折扣系数 γ 设置为 0.9.

4.3 行业波动效应分析

长期以来, 股指价格波动是否收到相关行业指数价格波动的影响? 行业指数价格波动的基本特征如何? 各行业指数波动是否具有协同性? 这些问题一直是金融领域的研究热点, 同样也是投资者在进行资产配置和价值投资时关注的重点.

为了探究股票市场剧烈波动时各行业板块对股指价格波动的影响,选取 2015 年 6 月至 9 月 SHSZ 指数为研究对象. 指数在这段期间经历了两阶段大幅回撤,包含 5 个重要的时间节点:6 月 19 日的指数表现疲软,跌幅超过 6%,近千支股票跌停,行业板块全线飘绿;6 月 26 日的指数大幅低开,随后开启放量杀跌模式,逾两千只个股跌停;7 月 27 日上午的指数低开后维持在 5 日线与 10 日线之间窄幅震荡,午后由于股指期货快速跳水导致

股指下跌;8 月 24 日的指数大幅低开,行业板块全线暴跌,午后股指几近跌停,临近尾盘虽有小幅上冲,但当其跌幅收窄至 6% 后表现乏力,再次下探;8 月 25 日指数大幅跳空低开,早盘股指一度跌幅收窄,但午后沪深两市均杀跌不止,股指跌破 3 000 点,创年内新低. 为了分析 SHSZ 指数大幅调整前后股票关联网络的拓扑结构特征,本文选取 6 月 15 日作为初始观察时间节点,此时市场整体运行情况相对平稳.

表 4 实验参数设置

Table 4 Experimental parameter setting

股票关联网络 构建参数	窗口长度 l	相关性度量		矩阵裁剪			
	15	VG + Jaccard 相似度		PMFG			
股票关联网络 表征模型参数	GCN 迭代数 K	池化层数 D	测试集大小 T_{train}	训练集大小 T_{train}	隐藏层维度 d	节点粗化率 δ	迭代次数 τ
	2	2	8	30	140	30%	5
基于深度强化 学习的自适应 模型参数	评估向量长度 w	隐藏层维度 d'	网络更新率 ρ	随机数 ε	学习率 η	折扣系数 γ	迭代次数
	5	10	20	0.9	0.01	0.9	1 000

DIFFPOOL-MLP 模型对股票关联网络层次化聚合后,最终将成分股划分到 9 个不同类簇. 图 6 展示了不同时间节点股票关联网络的类簇结构. 其中,节点亮度表示成分股类簇隶属信息(即 DIFFPOOL-MLP 模型层次化聚合结果);节点形状反映成分股行业归属信息:构成 SHSZ 指数的 300 支成分股分别隶属于科技(砖石形)、房地产(圆形)、工业(六边形)、公用事业(八边形)、金融(平行四边形)、能源(正方形)、消费(圆角正方形)、医药(三角形)原材料(V 形)在内的 9 种行业板块;节点规模大小由成分股最近 15 个交易日累计收益率的 Sigmoid 值确定;最近 15 个交易日内成分股间价格波动相似性决定了节点间连边的粗细. 从图 6 可观测得出,在股指运行的不同阶段,科技、工业、金融、能源行业成分股间连接较为紧密,表明这些行业成分股价格波动存在较高相似性. 当股指运行相对平稳时,不同行业成分股价格波动的趋同性较高,相应股票关联网络的类簇结构不够显著;例如,在图 6(a)中,公用事业、金融、消费和能源行业的成分股连接较为紧密,且处于网络中相对中心的位置. 当股指经历第一轮大幅回撤

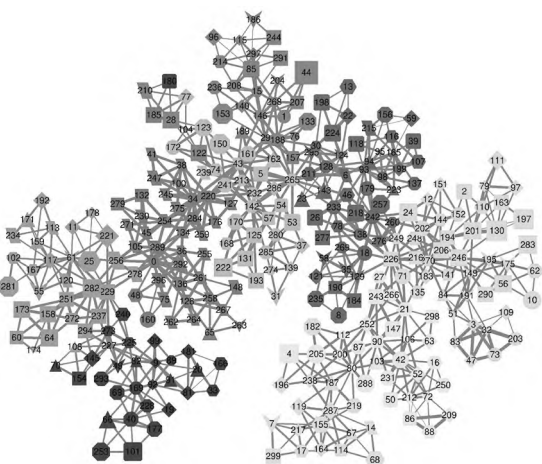
时(6 月 19 日—6 月 26 日),同一行业内成分股价格走势开始迅速分化,如图 6(b)~图 6(c)所示,股票关联网络呈现较为显著的类簇结构,随后,在 7 月 9 日—7 月 26 日期间,SHSZ 直接经历了一轮短期反弹,如图 6(d)所示,相同行业成分股走势又开始趋同,股票关联网络类簇结构显著性开始降低,表明这一轮股指反弹具有普涨效应. 随后,在 8 月 18 日—8 月 26 日期间,股指经历了第二轮大幅回撤,如图 6(e)~图 6(f)所示,股票关联网络类簇结构显著性大幅提升.

为了进一步分析个股和行业异动对股票市场的影响,实验选取 2015 年下半年 SHSZ 指数成分股中具有代表性的 9 家上市公司进行中心性分析. 如表 5 所示,所选取的 9 支股票分别隶属于 9 个不同的行业板块. 通过分析这 9 支成分股的度中心性和介数中心性在股指运行过程中不同时间节点(从 6 月 15 日—8 月 25 日)的变化趋势,可以发现当市场运行状况平稳时(5 月 16 日—6 月 15 日),以洪都航空和平安银行为代表的公用事业和金融板块与大盘阶段走势趋同. 如图 6(a)所示,这两支成分股具有较高的度中心性和

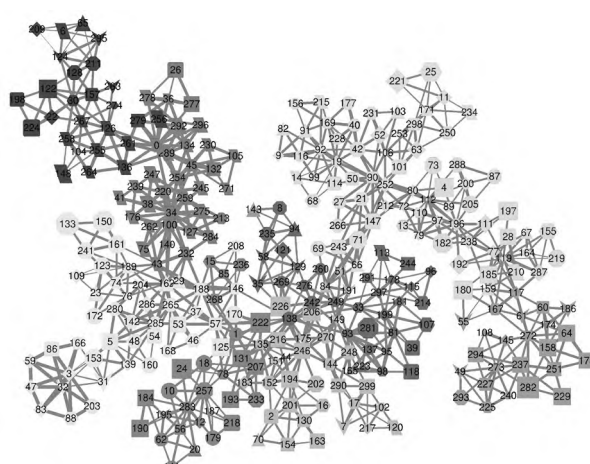
介数中心性,表明它们与其他成分股的走势相同。当 SHSZ 指数在 6 月 9 日成功站上 5 380 的历史相对高点后,平安银行在随后的 5 个交易日内经历了近 9% 的大幅回撤,同一行业的其他成分股受此影响也进行了不同程度的回撤。此后,金融行业异动迅速传导至公用事业和消费行业,进而引发股指经历了 2.5% 的小幅回撤。这一阶段的股指回撤并未全面波及到其他行业,以海信视像、国电电力和海思科为代表的科技、能源和医药行业甚至还创了阶段新高。然而在 5 月 15 日至 6 月 19 日期间,金融和消费行业的权重股继续大幅杀跌,科技、能源和医药行业出现了集体熄火,导致股指跌破 5 000 点大关。金融和消费行业权重股所引发的指数调整一直持续到 7 月 8 日,随后以金融街、三一重工和海思科为代表的房地产、工业和医药行业率先带领指数经历了为期 12 个交易日的反弹,其累计涨幅超过 14%。从图 6(d) 观察可得,金融街和三一重工的度中心性分别升至 14 和 10,后者的介数中心性甚至达到了 0.341 0。较高的度中心性和介数中心表明该成分股与其他个股走势趋同,可一定程度代表指数的阶段走势。这表明经历了股指的大幅回撤后,市场资金更倾向于进入房地产和工业等相关行业避险。8 月 18 日—8 月 26 日,受多重利空消息的影响,以海信视像、国电电力、海思科和中金岭南为代表的科技、能源、医药和原材料行业带领指数开启了第二轮

大幅杀跌,前期反弹无力的金融和消费行业继续跟跌,SHSZ 指数在 7 个交易日内跌幅超过 25%。

回顾 SHSZ 指数从 2015 年 6 月 15 日—8 月 25 日的调整过程,不难发现以平安银行为代表的权重股的异动首先引发金融行业其他成分股的短期调整。金融行业异动迅速传导至公用事业和消费行业,带动股指经历了第一轮大幅杀跌;随后房地产、工业和医药行业率先带领指数进行了反弹,但是受科技、能源、医药和原材料等行业的拖累,股指随后又进行了第二轮大幅回撤。DIFFPOOL-MLP 模型对股指成分股进行层次化聚类,围绕中心性节点(具有较高度中心性或介数中心性的成分股)而形成的类簇包含的成分股价格波动趋同性较高。当股指运行平稳时,行业轮动助推股指进行横盘整理或阶段性上涨,相同行业成分股价格波动相似性较高,而股票关联网络类簇结构的显著性有所下降;当市场恐慌情绪上升,股指进行大幅回撤时,相同行业成分股价格波动开始分化,但股票关联网络类簇结构显著性提升。以股票关联网络图分类任务为驱动的 DIFFPOOL 模型通过迭代训练深度图神经网络的权重矩阵对节点属性信息(成分股/类簇基本面和技术面指标)进行聚合,可动态捕捉不同类簇对股指波动的影响权重。相较于传统直接应用股指时序数据的股指预测模型,本文所提方法能够最大程度的捕捉股票市场的非线性特征,从而极大提高了股指预测精度。



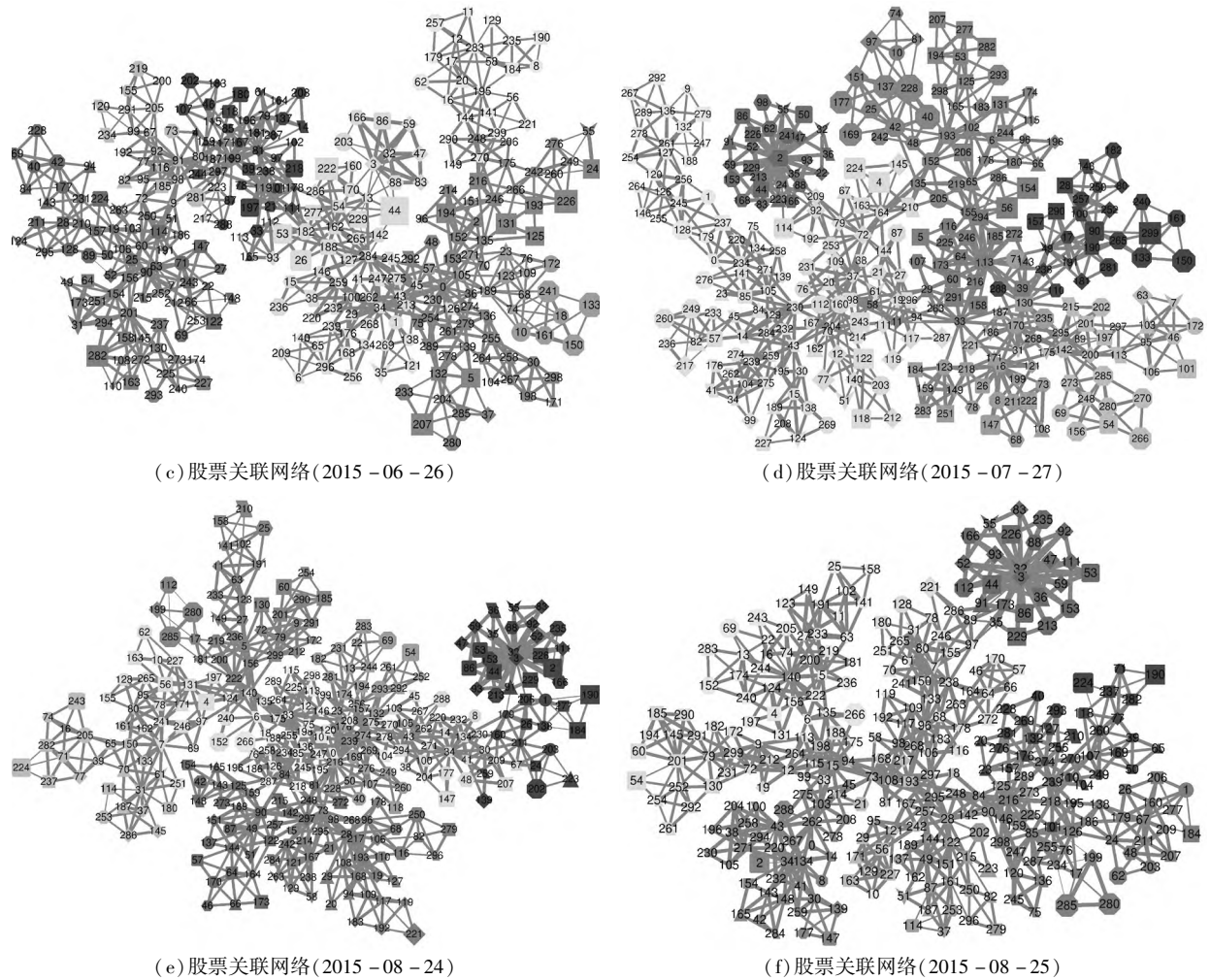
(a) 股票关联网络(2015-06-15)



(b) 股票关联网络(2015-06-19)

图 6 SHSZ 指数股票关联网络(2015-06-15—2015-08-25)

Fig. 6 The stock correlation networks of SHSZ index from June 15th to August 25th 2015



续图 6

Fig. 6 Continues

表 5 行业代表性成分股中心性分析

Table 5 Centrality analysis of industry-representative component stocks

股票名称	海信视像	金融街	三一重工	洪都航空	平安银行	国电电力	贵州茅台	海思科	中金岭南
股票代码	600060	000402	600031	600316	000001	600795	600519	002653	000060
所属行业	科技	房地产	工业	公用事业	金融	能源	消费	医药	原材料
股票关联网络节点 ID	140	18	135	169	0	216	188	103	7
2015-06-15	度中心性	5	8	3	13	19	6	9	3
	介数中心性	0.002 2	0.007 3	0.000 0	0.051 5	0.401 7	0.002 4	0.021 7	0.000 0
2015-06-19	度中心性	3	6	12	9	15	8	13	3
	介数中心性	0.002 4	0.010 6	0.147 6	0.005 6	0.195 4	0.018 6	0.329 9	0.000 0
2015-06-26	度中心性	5	3	10	4	17	9	12	4
	介数中心性	0.011 8	0.000 0	0.206 8	0.000 0	0.306 1	0.135 4	0.055 2	0.000 0
2015-07-27	度中心性	6	14	10	3	6	4	3	7
	介数中心性	0.053 9	0.059 7	0.341 0	0.000 0	0.109 0	0.001 7	0.000 0	0.004 4
2015-08-24	度中心性	15	3	3	4	3	23	4	6
	介数中心性	0.320 6	0.000 0	0.000 0	0.000 2	0.000 0	0.204 3	0.000 0	0.002 2
2015-08-25	度中心性	20	3	7	6	3	13	3	13
	介数中心性	0.152 7	0.000 0	0.023 0	0.002 4	0.000 0	0.163 2	0.000 0	0.144 9

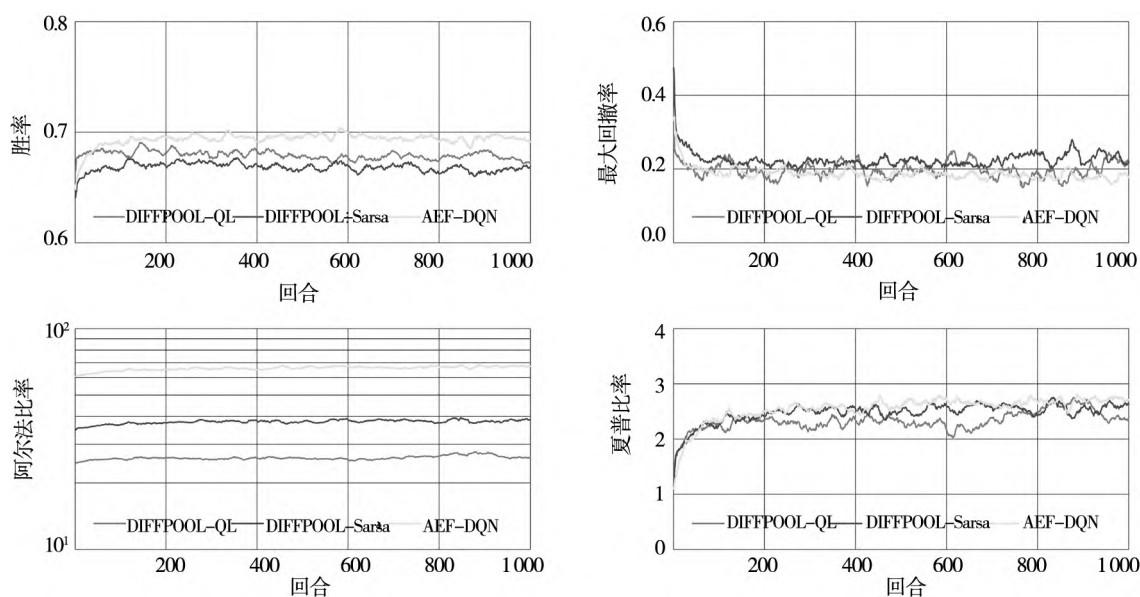


图7 不同自适应预测模型在SHSZ指数上的离线训练结果

Fig. 7 The offline training results of different adaptive prediction models on SHSZ index

4.4 股指预测性能分析

本实验首先考察基于深度强化学习结构的自适应模型对股指移动趋势的样本内拟合结果,将DIFFPOOL-LR、DIFFPOOL-MLP和DIFFPOOL-SA得到的候选股指预测信号融入智能体的动作空间,图7报告了不同强化学习算法在SHSZ指数2005-10-11至2015-07-24期间1000次epoches迭代训练过程中4种回测指标的变化趋势。由于不同的强化学习算法在SPX、UKX和NKY指数训练集上具有相同的表现,为了节约篇幅,本实验仅展示DIFFPOOL-QL、DIFFPOOL-Sarsa和AEF-DQN算法在SHSZ指数上的离线训练结果。从图7可发现,相比于DIFFPOOL-QL和DIFFPOOL-Sarsa算法,AEF-DQN算法对SHSZ指数的离线训练效果更好。AEF-DQN算法的状态空间中融入了最近动作序列的评估反馈,其可在环境与动作的交互过程中动态调整候选预测信号的选择模式,进而有效提升股指预测结果的准确性和鲁棒性。经过1000次迭代训练,AEF-DQN在样本内的投资胜率达到了69.5%,比次优的DIFFPOOL-Sarsa算法在PR指标上提升了约2%。DIFFPOOL-QL算法在最大回撤率指标上的表现相对于AEF-DQN和DIFFPOOL-Sarsa来说鲁棒性较低。经过1000次离线迭代训练,DIFFPOOL-QL取得的MDR指标仍在22%上下波动。进一步考察AR指标和SR指标,不难发现两者的移动趋势

比较接近,AEF-DQN经过约400次迭代训练,其阿尔法比率(AR)和夏普比率(SR)分别保持在[60,70]和[2.3,2.8]的数值区间波动,说明其相对于市场无风险利率和市场基准收益率有出色的超额回报,同时具备理想的抗风险能力。

为了进一步分析不同股指预测算法对SHSZ股指的模拟投资表现,图8报告了应用不同股指预测模型进行股指期货日内交易模拟时账户资金的移动趋势。在实验训练周期和测试周期的初始阶段,都预设10000 RMB的投资本金。其中,DIFFPOOL-QL、DIFFPOOL-Sarsa和AEF-DQN3个模型在股指训练集上的模拟投资结果是经过离线迭代训练后的第1000次模拟投资结果。通过观察不同预测模型在训练周期和测试周期资金的变化情况,可以发现AEF-DQN对SHSZ指数样本内数据的趋势拟合效果要明显好于其他基线模型。不同预测模型在训练周期对SHSZ指数的趋势拟合效果差异性显著,尤其是在经历了2007年“股灾”后,不同预测模型的投资表现开始迅速分化。基于深度强化学习结构的自适应股指预测模型经过不断的试错学习,能够及时的捕捉市场规律,在2008年—2015年训练周期能够稳定盈利。受模拟投资复利影响,AEF-DQN在训练周期结束时模拟资金盈利近650倍。考虑到基于强化学习结构的自适应股指预测模型(DIFFPOOL-QL、DIFFPOOL-Sarsa和AEF-DQN)的模拟投资表现受随

机数 ε 的影响. 为此, 图 8 展示了 3 种自适应模型在测试周期模拟投资 50 次的平均回测结果. 考虑到在单次模拟投资中, 3 种预测模型都会对动作-值函数进行更新. 实验中, 3 种自适应预测模型在测试周期的每轮模拟投资中都分别输入相同的 Q 值配置文件(即经过 1 000 轮离线训

练后的 Q 表或 Q 网络). 在测试周期, AEF-DQN 相较于其他基线模型仍有较出色的表现. 尤其是在经历了 2015 年 SHSZ 指数的大幅调整后, AEF-DQN 通过动态调整候选预测信号的选择模式, 较其他预测模型能够获得更稳健的投资回报.

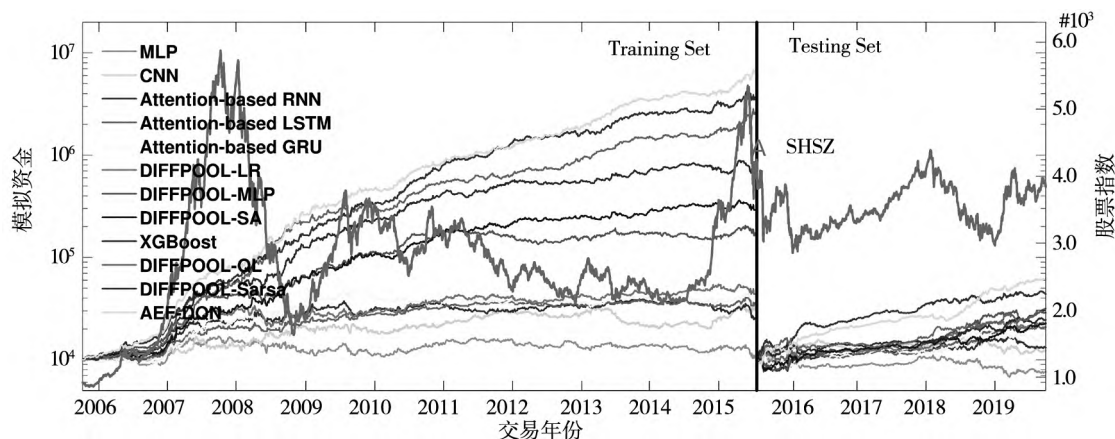


图 8 不同预测模型在 SHSZ 指数上的模拟投资结果

Fig. 8 The simulated investment results of different prediction models on SHSZ index

利用各种预测模型在测试集(样本外)的回测结果如表 6 至表 9 所示. 在 SHSZ 指数的模拟投资实验中, 相比于其他基线模型, 本文所提出的 AEF-DQN 模型在 PR、MDR、AR 和 SR 4 种指标上均表现最好. 具体地, 在近 4 年的测试周期中, AEF-DQN 模型在测试集上的 50 次模拟投资中的平均胜率达到了 61.37%, 比次优的 Sarsa 模型提升了 3.6%, 同时 AEF-DQN 模型的胜率标准差仅为 0.61%, 显著低于其它模型; 平均最大回撤率仅为 17.30%, 远低于沪深 300 指数 32.46% 的最大回撤水平; 取得的阿尔法比率和夏普比率都高于其它基线模型, 比次优的 DIFFPOOL-Sarsa 模型分别提升了 91.79% 和 38.41%, 这说明其具备更好的盈利能力.

进一步分析不同预测模型对 SPX、UKX 和 NKY 指数的回测结果, 可以得到类似结论. 如表 7 所示, 在 SPX 指数的模拟投资实验中, DIFFPOOL-MLP 模型在 MDR 指标上优势显著. 以 DIFFPOOL-QL、DIFFPOOL-Sarsa 和 AEF-DQN 为代表的基于强化学习的自适应股指预测模型在模拟投资过程中可以根据股票关联网络的层次化表征状态调整自身的决策模式, 相较于其他基线模型, 在 PR、AR 和 SR 指标上具有显著优势. 在测试周期

的 50 次模拟实验中, AEF-DQN 模型在各项回测指标上均优于 DIFFPOOL-QL 和 DIFFPOOL-Sarsa 模型, 由于智能体的状态空间考虑了其最近 w 个决策周期的动作评估向量, AEF-DQN 模型对决策模式的调整更为有效, 各项回测指标的标准差较其他两个基于强化学习结构的自适应预测模型更小, 说明其对短期行情的预测效果更加稳健. 经过 50 次在线模拟, AEF-DQN 模型取得的平均夏普比率达到 2.505 0, 远高于基线模型; 相较于 SPX 指数本身 0.460 0 的夏普比率, 提升了 444%, 说明其在标普 500 指数上可以获得市场超额收益. 表 8 对比了不同模型在 UKX 指数上的回测结果, AEF-DQN 模型得到的投资收益走势强劲, 其各项回测指标均显著优于其他基线模型. 尽管 UKX 指数在测试周期波动性较高, AEF-DQN 模型在 50 次模拟投资实验中的平均投资胜率高达 60.63%, 比次优的 DIFFPOOL-Sarsa 模型提升了 5.49%. UKX 指数在测试周期的最大回撤率为 16.41%, 以 DIFFPOOL-QL、DIFFPOOL-Sarsa 和 AEF-DQN 为代表的基于强化学习结构的自适应预测模型在 50 次模拟投资实验中的平均最大回撤率均低于 10%, 这说明基于股票关联网络层次化表征的决策模式调整机制对提升股指预测的鲁

棒性具有重要作用. 在 50 次模拟投资过程中, AEF-DQN 模型取得的平均阿尔法比率和平均夏普比率分别达到了 44.54% 和 2.893 0, 远高于次优的 DIFFPOOL-Sarsa 模型, 这说明考虑动作评估反馈的候选信号寻优策略可显著提升模型的盈利能力. 最后, 表 9 对比了不同预测模型在 NKY 指数上的回测表现. NKY 指数在测试周期表现为典型的震荡下行趋势, 相较于其他单一模型, DIFFPOOL-LR 模型在 PR、AR 和 SR 三个回测指标上表现最优; DIFFPOOL-MLP 的表现略逊于 DIFFPOOL-LR, 但在 MDR 指标上略优于前者. 除了 Attention-based LSTM 和 Attention-based GRU, 仅考虑原始股指时序数据的神经网络模型在 AR 和 SR 指标上均未跑赢 NKY 指数自身走势. 这表明股票关联网络分类模型更适用于在震荡下行市场环境中捕捉短线行情趋势. 尽管 NKY 指数在 2019 年波动性较大, AEF-DQN 模型的平均最大回撤率仅为 6.46%, 低于指数本身 21.14% 的最大回撤率, 说明其具有较强的抗风险能力. AEF-DQN 模型在阿尔法比率和夏普比率上均跑赢了市场指数和其他基线模型, 在 50 次模拟投资过程中取得平均阿尔法比率和平均夏普比率分别为 45.64% 和 2.258 5, 远高于次优的 DIFFPOOL-Sarsa 模型.

与现有的基于神经网络结构的单一股指预测模型和组合预测模型 XGBoost 相比, AEF-DQN 借助于可微池化框架层次化聚合股票关联网络的节点属性信息, 进而从不同行业视角提取股指价格

波动的中观尺度预测因子; 在此基础上重构股票关联网络的粗化图, 基于类簇隶属度矩阵捕捉不同行业板块的相互影响和价格传导信息, 进而可层次化聚合不同行业板块的属性信息. 同时, 股票关联网络在很大程度上减轻了异常行情导致的数据波动, 基于深度学习的图分类框架可以提取股票市场的非线性特征, 因此在预测过程中得到了比股指时序数据更好的预测结果. 与之形成鲜明对比的是测试集中的 SHSZ 指数分别在 2015 年和 2018 年经历了两次大幅回撤, MLP 和 CNN 模型的 MDR 指标都超过了 30%, 说明其应对系统性金融风险的能力不足. 与 DIFFPOOL-LR、DIFFPOOL-MLP 和 DIFFPOOL-SA 等非强化学习结构股票关联网络分类模型相比, 考虑动作评估反馈的 AEF-DQN 模型在决策过程中考虑股票市场环境对智能体历史决策反馈等影响, 更能准确捕捉市场运行规律, 因此表现更好. 与自适应模型 DIFFPOOL-QL 和 DIFFPOOL-Sarsa 相比, AEF-DQN 中经验回放池的设计不仅可以产生大量样本供监督学习使用, 还可解决训练样本间的相关性及非静态分布问题, 并且使其具有短时记忆能力. 在迭代训练次数较少时, 与 DIFFPOOL-QL 和 DIFFPOOL-Sarsa 模型相比, AEF-DQN 的预测效果并无明显优势, 甚至会劣于这两种模型; 但随着迭代次数的增加, AEF-DQN 的优势会逐步显示出来并最终收敛到一个具有最高胜率和最大总奖励值的状态.

表 6 不同模型在 SHSZ 上的回测结果对比

Table 6 Comparison of back-testing results of different models on SHSZ

	PR	MDR	AR	SR
SHSZ 指数		0.324 6	0.040 7	-0.064 9
MLP	0.526 5	0.412 4	-0.114 1	-0.469 3
CNN	0.523 5	0.309 9	0.003 3	0.073 5
Attention-based RNN	0.534 3	0.236 4	0.071 5	0.340 4
Attention-based GRU	0.532 4	0.296 9	0.187 8	0.665 2
Attention-based LSTM	0.540 2	0.270 1	0.196 7	0.686 1
DIFFPOOL-LR	0.540 2	0.296 0	0.229 4	0.747 7
DIFFPOOL-MLP	0.583 3	0.290 0	0.378 5	1.036 0
DIFFPOOL-SA	0.574 5	0.283 4	0.241 8	0.765 9
XGBoost	0.545 1	0.254 8	0.273 6	0.846 5
DIFFPOOL-QL	0.577 4 ± 0.012 1	0.223 3 ± 0.065 6	0.429 7 ± 0.031 2	1.132 7 ± 0.042 1
DIFFPOOL-Sarsa	0.592 1 ± 0.009 2	0.243 6 ± 0.052 1	0.712 5 ± 0.025 3	1.510 5 ± 0.033 3
AEF-DQN	0.613 7 ± 0.006 1	0.173 0 ± 0.045 3	1.371 7 ± 0.019 3	2.090 8 ± 0.026
Avg.	0.556 9	0.321 4	0.331 9	0.785 5

表 7 不同模型在 SPX 上的回测结果对比

Table 7 Comparison of back-testing results of different models on SPX

	PR	MDR	AR	SR
SPX 指数		0.164 5	0.037 7	0.460 0
MLP	0.536 3	0.215 8	0.002 1	0.034 4
CNN	0.510 6	0.190 0	0.005 3	0.145 4
Attention-based RNN	0.550 0	0.162 1	0.029 4	0.278 2
Attention-based GRU	0.577 2	0.106 5	0.196 8	1.175 2
Attention-based LSTM	0.581 8	0.085 7	0.226 2	1.290 9
DIFFPOOL-LR	0.554 5	0.083 9	0.234 5	1.353 0
DIFFPOOL-MLP	0.583 3	0.063 7	0.238 8	1.372 1
DIFFPOOL-SA	0.574 2	0.122 9	0.298 0	1.517 9
XGBoost	0.568 1	0.092 6	0.350 0	1.795 1
DIFFPOOL-QL	0.595 4 ± 0.014 1	0.156 0 ± 0.024 4	0.349 8 ± 0.036 6	1.791 2 ± 0.052 1
DIFFPOOL-Sarsa	0.601 0 ± 0.011 2	0.119 2 ± 0.016 1	0.332 7 ± 0.027 2	1.788 2 ± 0.042 3
AEF-DQN	0.609 1 ± 0.009 2	0.117 0 ± 0.007 3	0.554 0 ± 0.015 4	2.505 0 ± 0.039 3
Avg.	0.570 8	0.125 7	0.235 8	1.255 1

表 8 不同模型在 UKX 上的回测结果对比

Table 8 Comparison of back-testing results of different models on UKX

	PR	MDR	AR	SR
UKX 指数		0.164 1	0.040 0	-0.613 3
MLP	0.500 0	0.248 2	-0.146 8	-1.194 4
CNN	0.508 6	0.184 0	-0.049 8	-0.274 5
Attention-based RNN	0.543 1	0.095 4	-0.022 3	-0.073 7
Attention-based GRU	0.511 4	0.104 5	-0.015 9	0.088 3
Attention-based LSTM	0.580 4	0.097 1	0.075 8	0.525 3
DIFFPOOL-LR	0.548 8	0.095 5	0.109 7	0.789 9
DIFFPOOL-MLP	0.560 3	0.082 4	0.176 7	1.316 5
DIFFPOOL-SA	0.586 2	0.106 2	0.221 8	1.533 8
XGBoost	0.566 0	0.075 1	0.214 4	1.547 9
DIFFPOOL-QL	0.548 8 ± 0.017 2	0.093 1 ± 0.034 3	0.227 8 ± 0.027 2	1.641 0 ± 0.055 4
DIFFPOOL-Sarsa	0.574 7 ± 0.012 2	0.075 0 ± 0.025 2	0.376 5 ± 0.023 3	2.526 9 ± 0.044 6
AEF-DQN	0.606 3 ± 0.009 1	0.054 5 ± 0.021 3	0.445 4 ± 0.021 5	2.893 0 ± 0.034 4
Avg.	0.552 9	0.111 2	0.134 4	0.943 3

表 9 不同模型在 NKY 上的回测结果对比

Table 9 Comparison of back-testing results of different models on NKY

	PR	MDR	AR	SR
NKY 指数		0.211 4	0.040 5	0.117 2
MLP	0.513 9	0.171 1	-0.097 2	-0.491 6
CNN	0.518 5	0.159 7	-0.072 1	-0.304 4
Attention-based RNN	0.495 3	0.123 2	-0.041 4	-0.127 8
Attention-based GRU	0.509 2	0.205 9	0.022 8	0.197 8
Attention-based LSTM	0.541 6	0.145 7	0.130 7	0.809 3
DIFFPOOL-LR	0.578 7	0.121 2	0.281 9	1.526 9
DIFFPOOL-MLP	0.569 4	0.114 2	0.181 9	1.037 5
DIFFPOOL-SA	0.550 9	0.076 2	0.173 2	1.030 2
XGBoost	0.587 9	0.062 4	0.406 2	2.097 0
DIFFPOOL-QL	0.601 8 ± 0.015 3	0.145 3 ± 0.021 6	0.422 4 ± 0.027 3	2.111 3 ± 0.036 2
DIFFPOOL-Sarsa	0.625 0 ± 0.011 2	0.085 5 ± 0.016 4	0.424 4 ± 0.021 2	2.170 4 ± 0.026 1
AEF-DQN	0.634 2 ± 0.008 1	0.064 6 ± 0.018 3	0.456 4 ± 0.013 7	2.258 5 ± 0.014 3
Avg.	0.560 5	0.116 0	0.190 7	1.025 7

5 结束语

由于股票市场中极其复杂的信息环境和投资者有限理性和非理性的交易行为, 股指成分股间产生了领先滞后、价格同步、行业轮动、风险传染等多种复杂的关联性特征^[15], 这种关联性相互交织构成了股指成分股间的关联网络. 本研究基于可视图和平面极大过滤图模型构建时序股票关联网络, 将股指预测问题形式化为一种面向股票关联网络的图分类问题, 然后利用深度图神经网络学习股票关联网络的层次化表征, 并结合不同的可微分类模型以一种端到端的方式生成候选预测

信号. 然后提出了一种考虑动作评估反馈的深度强化学习模型(AEF-DQN), 可融合不同的候选预测信号, 并基于环境状态动态调整自身决策模式. 最后, 利用 AEF-DQN 模型对沪深 300 指数(SHSZ)、标普 500 指数(SPX)、英国富时 100 指数(UKX)和日经 225 指数(NKY)进行了实证分析. 主要实证结果表明: 1) 通过层次化聚合股票关联网络的节点属性信息可以动态捕捉不同行业对股指价格波动的影响, 进而可提升预测方法的准确率; 2) 考虑动作评估反馈的深度强化学习结构可智能化选择适用于当前股票市场环境的最优模型结构, 进而可提升预测方法的鲁棒性.

参考文献:

- [1] Page L, Siemroth C. How much information is incorporated into financial asset prices? Experimental evidence[J]. *Review of Financial Studies*, 2021, 34(9): 4412–4449.
- [2] Chen Y, Eaton G W, Paye B S. Micro(structure) before macro? The predictive power of aggregate illiquidity for stock returns and economic activity[J]. *Journal of Financial Economics*, 2018, 130(1): 48–73.
- [3] 李木易, 方颖. 动态混合 HGARCH 模型的估计和预测[J]. *管理科学学报*, 2020, 23(5): 1–12.
Li MUYI, Fang Ying. Estimation and forecasting of mixture HGARCH model[J]. *Journal of Management Sciences in China*, 2020, 23(5): 1–12. (in Chinese)
- [4] 林宇, 黄讯, 淳伟德, 等. 基于 ODR-DASYN-SVM 的极端金融风险预警研究[J]. *管理科学学报*, 2016, 19(5): 87–101.
Lin Yu, Huang Xun, Chun Weide, et al. Early warning for extremely financial risks based on ODR-ADASYN-SVM[J]. *Journal of Management Sciences in China*, 2016, 19(5): 87–101. (in Chinese)
- [5] Chen Y C, Misis V V. Decision forest: A nonparametric approach to modeling irrational choice[J]. *Marketing Science*, 2022, Early Access, DOI: 10.1287/mnsc.2021.4256.
- [6] Flori A, Regoli D. Revealing pairs-trading opportunities with long short-term memory networks[J]. *European Journal of Operational Research*, 2021, 295(2): 772–791.
- [7] Pollet J M, Wilson M I. Average correlation and stock market returns[J]. *Journal of Financial Economics*, 2008, 96(3): 364–380.
- [8] 裴茜, 朱书尚. 中国股票市场金融传染及渠道——基于行业数据的实证研究[J]. *管理科学学报*, 2019, 22(3): 90–112.
Pei Xi, Zhu Shushang. Financial contagion in China's stock market: A study based on industry-level data[J]. *Journal of Management Sciences in China*, 2019, 22(3): 90–112. (in Chinese)
- [9] 姜富伟, 马甜, 张宏伟. 高风险低收益? 基于机器学习的动态 CAPM 模型解释[J]. *管理科学学报*, 2021, 24(1): 109–126.
Jiang Fuwei, Ma Tian, Zhang Hongwei. High risk low return? Explanation from machine learning based conditional CAPM model[J]. *Journal of Management Sciences in China*, 2019, 24(1): 109–126. (in Chinese)
- [10] Mantegna R N. Hierarchical structure in financial markets[J]. *The European Physical Journal B*, 1999, 11(1):

193 – 197.

- [11] Chi K T, Jing L, Francis C M L. A network perspective of the stock market[J]. *Journal of Empirical Finance*, 2010, 17 (4): 659 – 667.
- [12] Namaki A, Shirazi A H, Raei R, et al. Network analysis of a financial market based on genuine correlation and threshold method[J]. *Physica A: Statistical Mechanics and its Applications*, 2011, 390 (21 – 22): 3835 – 3841.
- [13] Tumminello M, Aste T, Di Matteo T, et al. A tool for filtering information in complex systems[J]. *Proceedings of the National Academy of Sciences*, 2005, 102(30): 10421 – 10426.
- [14] 张贵生, 张信东. 基于近邻互信息的 SVM-GARCH 股票价格预测模型研究[J]. *中国管理科学*, 2016, 24(9): 11 – 20.
- Zhang Guisheng, Zhang Xindong. A SVM-GARCH model for stock price forecasting based on neighborhood mutual information[J]. *Chinese Journal of Management Science*, 2016, 24(9): 11 – 20. (in Chinese)
- [15] 张伟平, 庄新田, 吴冬梅. 我国股市网络稳定性及其宏观影响因素——基于三种网络拓扑结构、鲁棒性的实证研究[J]. *运筹与管理*, 2020, 29(1): 185 – 194.
- Zhang Weiping, Zhuang Xintian, Wu Dongmei. The stability of China stock market network and its macroeconomic factors: An empirical study based on three network topologies and Robustness[J]. *Operations Research and Management Science*, 2020, 29(1): 185 – 194. (in Chinese)
- [16] Yun K K, Yoon S W, Won D. Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process[J]. *Expert Systems with Applications*, 2021, (186): 115716.
- [17] Lacasa L, Luque B, Ballesteros F, et al. From time series to complex networks: The visibility graph[J]. *Proceedings of the National Academy of Sciences*, 2008, 105(13): 4972 – 4975.
- [18] Ying R, You J, Morris C, et al. Hierarchical Graph Representation Learning with Differentiable Pooling[C]. *Montreal: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2018, 4805 – 4815.
- [19] Volodymyr M, Koray K, David S, et al. Human-level control through deep reinforcement learning[J]. *Nature*, 2015, (518): 529 – 533.
- [20] Bayar O, Chemmanur T J, Liu M H. A theory of equity carve-outs and negative stub values under heterogeneous beliefs [J]. *Journal of Financial Economics*, 2011, 100(3): 616 – 638.
- [21] Tan W Q, Zhang J. Good days, bad Days: Stock market fluctuation and taxi tipping decisions[J]. *Marketing Science*, 2021, 67(6): 3965 – 3984.
- [22] 胡 熠, 顾 明. 巴菲特的阿尔法: 来自中国股票市场的实证研究[J]. *管理世界*, 2018, 34(8): 41 – 54.
- Hu Yi, Gu Ming. Warren Buffett's Alpha: Empirical research from the Chinese stock market[J]. *Management World*, 2018, 34(8): 41 – 54. (in Chinese)
- [23] Berkman H, Jacobsen B, Lee J B. Time-varying rare disaster risk and stock returns[J]. *Journal of Financial Economics*, 2010, 101(2): 313 – 332.
- [24] 杨晓兰, 王伟超, 高 媚. 股市政策对股票市场的影响——基于投资者社会互动的视角[J]. *管理科学学报*, 2020, 23(1): 15 – 32.
- Yang Xiaolan, Wang Weichao, Gao Mei. The impact of stock market policies on stock market: From the perspective of investor social interaction[J]. *Journal of Management Sciences in China*, 2020, 23(1): 15 – 32. (in Chinese)
- [25] 姚加权, 冯 绪, 王赞钧, 等. 语调、情绪及市场影响: 基于金融情绪词典[J]. *管理科学学报*, 2021, 24(5): 26 – 46.
- Yao Jiaquan, Feng Xu, Wang Zanjuan, et al. Tone, sentiment and market impacts: The construction of Chinese sentiment dictionary in finance[J]. *Journal of Management Sciences in China*, 2021, 24(5): 26 – 46. (in Chinese)
- [26] Wu Z, Pan S, Chen F, et al. A comprehensive survey on graph neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4 – 24.

- [27] Niepert M, Ahmed M, Kutzkov K. Learning Convolutional Neural Networks for Graphs[C]. Proceedings of the 33rd International Conference on Machine Learning (ICML), New York: USA, 2016, 2014–2023.
- [28] Wang Y, Sun Y, Liu Z, et al. Dynamic graph CNN for learning on point clouds[J]. ACM Transactions on Graphics, 2018, 38(5): 1–12.
- [29] Zhang M, Cui Z, Neumann M, et al. An End-to-End Deep Learning Architecture for Graph Classification[C]. Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI), New Orleans, Louisiana: USA, 2018, 4438–4445.
- [30] Abu-Mostafa Y S, Atiya A F. Introduction to financial forecasting[J]. Applied Intelligence, 1996, 6(3): 205–213.
- [31] 李凯文, 张 涛, 王 锐, 等. 基于深度强化学习的组合优化研究进展[J]. 自动化学报, 2021, 47(11): 2521–2537.
Li Kaiwen, Zhang Tao, Wang Rui, et al. Research reviews of combinatorial optimization methods based on deep reinforcement learning[J], 2021, 47(11): 2521–2537. (in Chinese)
- [32] Samuelson P, Merton R C. A complete model of warrant-pricing that maximizes utility[J]. Industry Management Review, 1969, 10(2): 17–46.
- [33] Moody J, Wu L Z. Optimization of Trading Systems and Portfolios[C]. Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFER), New York: USA, 1997, 300–307.
- [34] Lee J W, Park J, Jangmin O, et al. A multiagent approach to Q-learning for daily stock trading[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part A, 2007, 37(6): 864–877.
- [35] Deng Y, Bao F, Kong Y, et al. Deep direct reinforcement learning for financial signal representation and trading[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 28(3): 653–664.
- [36] 梁天新, 杨小平, 王 良, 等. 基于强化学习的金融交易系统研究与发展[J]. 软件学报, 2019, 30(3): 845–864.
Liang Tianxin, Yang Xiaoping, Wang Liang, et al. Review on financial trading system based on reinforcement learning[J]. Journal of Software, 2019, 30(3): 845–864. (in Chinese)
- [37] Wang Z C, Huang B W, Tu S K, et al. Deep Trader: A Deep Reinforcement Learning Approach for Risk-Return Balanced Portfolio Management with Market Conditions Embedding[C]. Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI), A Virtual Conference, 2021, 643–650.
- [38] Yang Y, Wang J B, Yang H J, et al. Visibility graph approach to exchange rate series[J]. Physica A-statistical Mechanics and Its Applications, 2009, 388(20): 4431–4437.
- [39] Yan W F, Serooskerken E V. Forecasting financial extremes: A network degree measure of super-exponential growth[J]. PLOS ONE, 2015, 10(9): e0128908.
- [40] Flanagan R, Lacasa L. Irreversibility of financial time series: A graph-theoretical approach[J]. Physics Letter A, 2016, 380(20): 1689–1697.
- [41] Goncalves B A, Carpi L, Rosso O A, et al. Quantifying instabilities in financial markets[J]. Physica A-statistical Mechanics and Its Applications, 2019, (525): 606–615.
- [42] Zou Y, Donner R V, Marwan N, et al. Complex network approaches to nonlinear time series analysis[J]. Physics Reports-Review Section of Physics Letters, 2021, (787): 1–97.
- [43] Cooper K B, Steffen M, McDermott R, et al. Observation of quantum oscillations between a Josephson phase qubit and a microscopic resonator using fast readout[J]. Physical Review Letters, 2004, 93(18): 180401.
- [44] 李双琦, 陈其安, 朱 沙. 考虑消费与投资者情绪的股票市场资产定价[J]. 管理科学学报, 2021, 24(4): 86–108.
Li Shuangqi, Chen Qian, Zhu Sha. Asset pricing in the stock market with consumption and investor sentiment[J]. Journal of Management Sciences in China, 2021, 24(4): 86–108. (in Chinese)
- [45] 高大良, 刘志峰, 杨晓光. 投资者情绪、平均相关性与股市收益[J]. 中国管理科学, 2015, 23(2): 10–20.
Gao Daliang, Liu Zhifeng, Yang Xiaoguang. Investor sentiment, average correlation and stock market returns[J]. Chinese

- Journal of Management Science, 2015, 23(2): 10–20. (in Chinese)
- [46] 李守伟, 文世航, 王 磊, 等. 多层网络视角下金融机构关联性的演化特征研究[J]. 中国管理科学, 2020, 28(12): 35–43.
- Li Shouwei, Wen Shihang, Wang Lei, et al. Evolution characteristics of financial institutions' interrelationships from the perspective of multilayer network[J]. Chinese Journal of Management Science, 2020, 28(12): 35–43. (in Chinese)
- [47] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[C]. Proceedings of the 5th International Conference on Learning Representations (ICLR), Palais des Congrès Neptune, Toulon: France, 2017, 1–14.
- [48] Chen J H, Sathe S, Aggarwal C, et al. Outlier Detection with Autoencoder Ensembles[C]. Proceedings of the 2017 SIAM International Conference on Data Mining (SDM), Houston: Texas, USA, 2017, 90–98.
- [49] Zhang J, Cui S C, Xu Y, et al. A novel data-driven stock price trend prediction system[J]. Expert Systems with Applications, 2018, (97): 60–69.
- [50] Hamilton W L, Ying R, Leskovec J. Inductive Representation Learning on Large Graphs[C]. Annual Conference on Neural Information Processing Systems (NeurIPS), Long Beach, CA: USA, 2017, 1024–1034.

Adaptive stock index prediction based on deep reinforcement learning

BU Zhan¹, ZHANG Shan-fan², LI Xue-yan², MA Dan-dan², CAO Jie³

1. School of Intelligence Audit, Nanjing Audit University, Nanjing 211815, China;

2. College of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210023, China;

3. School of Management, Hefei University of Technology, Hefei 230009, China

Abstract: Based on the fundamental and technical data of constituent stocks, this paper constructs the sequential stock correlation networks, and utilizes the deep graph neural network to learn the hierarchical representations of stock correlation networks for obtaining the candidate prediction signals in an end-to-end fashion. On this basis, an Action Evaluation Feedback based Deep Q-Learning (AEF-DQN) is proposed to integrate different candidate prediction signals into the agent's action space. Specifically, AEF-DQN learns the environmental state using the hierarchical representation of stock correlation network, overall movement state of stock market and evaluation feedback of historical actions. By employing the reference dependency property in the prospect theory to estimate the reward function, AEF-DQN establishes the mapping relationship between state, action, and reward value. Finally, using the historical constituent stock data of SHSZ 300 index, S & P 500 index, FTSE 100 index and Nikkei 225 index, this paper constructs a stock index futures trading simulator to empirically analyze the proposed stock index prediction model from the profitable rate, maximum draw-down rate, Alpha ratio and Sharp ratio measures. The results show that: 1) by hierarchically aggregating the node attributes of stock correlation networks, the proposed network representation model can dynamically capture the potential impacts of different industry sectors on the price fluctuation of stock index, so as to improve the accuracy of the prediction model; 2) by considering the evaluation feedback of historical actions, the proposed deep reinforcement learning structure can intelligently select the optimal model structure suitable for the current stock market environment, so as to improve the robustness of the prediction method.

Key words: adaptive stock index prediction; stock correlation network; deep graph neural network; deep reinforcement learning