

中图分类号：

密级：公 开

学科分类号：

论文编号：172115011

山东财经大学

硕士学位论文

基于长短期记忆神经网络的 股票价格预测研究

作者姓名：宋刚

学科专业：计算机应用技术

指导教师：张云峰

培养学院：计算机科学与技术学院

二〇二〇年六月

Research on stock price forecast based on Long Short-Term Memory neural network

A Dissertation Submitted for the Degree of Master

Candidate: Song Gang

Supervisor: Prof. Zhang Yunfeng

School of Computer Science and Technology
Shandong University of Finance and Economics

中图分类号：

密级：公 开

学科分类号：

论文编号：172115011

硕 士 学 位 论 文

基于长短期记忆神经网络的 股票价格预测研究

作 者 姓 名： 宋刚

申请学位级别： 工学硕士

指导教师姓名： 张云峰

职 称： 教授

学 科 专 业： 计算机应用技术

研 究 方 向： 数据处理

学 习 时 间： 自 2017 年 9 月 1 日 起至 2020 年 6 月 30 日 止

学位授予单位： 山东财经大学

学位授予日期： 2020 年 6 月

山东财经大学学位论文独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得山东财经大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：

宋刚

日期：2020 年 6 月 5 日

山东财经大学学位论文使用授权声明

本人完全同意山东财经大学有权使用本学位论文（包括但不限于其印刷版和电子版），使用方式包括但不限于：保留学位论文，按规定向国家有关部门（机构）送交学位论文，允许学位论文被查阅、借阅和复印，将学位论文的全部或部分内容编入有关数据库进行检索，采用影印、缩印或其他复制手段保存学位论文。

保密学位论文在解密后的使用授权同上。

学位论文作者签名：

宋刚

日期：2020 年 6 月 5 日

指导教师签名：

张云萍

日期：2020 年 6 月 5 日

摘 要

股票市场在国家整体经济体系中的重要地位越发突出,对股票价格的分析与预测是当前学术界最具吸引力的研究问题之一。股票价格时间序列作为一种动态的、不规则的、非线性的、非参数化的复杂数据,给股票价格预测研究工作带来挑战。

随着神经网络技术的迅速发展,长短期记忆神经网络(Long Short-Term Memory, LSTM)因其在时间序列研究中的良好表现受到广泛关注。然而,在 LSTM 网络学习训练过程中,隐藏层中神经元数目、时间窗口、批处理数据量等关键控制参数的选取严重影响模型预测性能。为了解决这个问题,本文在 LSTM 网络的基础上进行改进,提出一种基于自适应粒子群优化的 LSTM 股票价格预测模型。通过设计快速有效的智能化参数寻优算法优化 LSTM 网络结构中的控制参数,使股票数据特征与网络拓扑结构相匹配,提高神经网络结构参数的可解释性与模型预测精度。实验结果表明,与统计模型 ARIMA、机器学习模型 SVM 以及神经网络模型 MLP、RNN、LSTM 相比,本文提出的股票价格预测模型具有更高的预测精度与稳定性,在处理不同类型股票数据的预测问题中显示出一定的普遍适用性。本文主要工作如下:

第一、在传统粒子群优化算法的基础上,提出一种具有自适应学习策略的粒子群优化算法。该算法通过快速搜索聚类,将整个粒子种群划分成若干子群,不同子群采用不同的学习策略对粒子进行更新。该算法增加粒子种群多样性,提高了目标优化的准确性。

第二、在 LSTM 神经网络的基础上,与智能化参数寻优算法相融合,提出一种针对股票价格时间序列预测的 PSO-LSTM 模型。利用收敛速度快、求解精度高的自适应粒子群优化算法对关键网络结构参数寻优,实现股票数据特征与 LSTM 神经网络的有效结合。

第三、将 PSO-LSTM 股票价格预测模型应用于沪市、深市、港股股票价格预测中。实验表明,与传统统计模型和深度学习预测模型相比,该模型具有更高的预测精度与普遍适用性。

关键词: 自适应; 粒子群优化; 长短期记忆; 股票价格预测

Abstract

The stock market plays an increasingly important role in the whole economic system of our country. The analysis and forecasting of stock data is one of the most attractive research problems in the current academic circles. Due to the high-noise, dynamics, non-linear, and non-parameters characteristics of stock data, the accurate forecasting of stock prices remains a challenging task.

With the development of artificial intelligence, the Long Short-Term Memory (LSTM) has received much attention for its excellent performance in dealing with time series. However, the key control parameters such as the number of neurons in the hidden layer, the time window, and the amount of batch data seriously affect the mode forecasting performance. In order to improve the accuracy of stock price forecasting, this paper proposes an adaptive particle swarm optimization based LSTM stock price forecasting model. To enhance the interpretability of the neural network structure parameters and improve the model forecasting performance, an adaptive particle swarm optimization algorithm is used to optimize the LSTM network control parameters. The experimental results show that compared with several popular time series forecasting models, the proposed model has the best forecasting performance and universal applicability. The main contents are as following:

First, based on the traditional particle swarm optimization (PSO) algorithm, an improved particle swarm optimization algorithm with adaptive learning strategy is proposed. The algorithm divides the entire particle population into several subgroups by quickly searching the clusters, and different subgroups update the particles with different learning strategies. This algorithm increases the diversity of particle populations and improves the accuracy of target optimization.

Second, the PSO-LSTM stock price forecasting model is proposed by combining the LSTM with the adaptive particle swarm optimization algorithm. The adaptive particle swarm optimization algorithm can quickly and accurately determine the optimal hyper-parameters based on the characteristics of the stock data.

Third, the PSO-LSTM stock price forecasting model is used to forecast stock prices in Shanghai stock market, Shenzhen stock market and Hong Kong stock market. Compared with traditional statistical models and deep learning forecasting models, the proposed

model has higher forecasting accuracy and universal applicability.

Key words: Adaptive; Particle Swarm Optimization; Long Short-Term Memory;
Stock Price Forecasting

目 录

第 1 章 绪论.....	1
1.1 研究背景.....	1
1.2 国内外研究现状.....	2
1.3 论文的主要研究内容.....	5
1.4 论文的组织结构.....	5
第 2 章 相关技术介绍.....	7
2.1 股票价格预测理论.....	7
2.1.1 基本面分析.....	7
2.1.2 技术面分析.....	8
2.2 统计学模型.....	9
2.2.1 时间序列的统计特性.....	9
2.2.2 时间序列预测的分类.....	10
2.2.3 ARIMA 模型.....	11
2.2.4 指数平滑模型.....	12
2.3 机器学习模型.....	13
2.4 神经网络理论.....	13
2.4.1 多层感知机与 BP 神经网络.....	14
2.4.2 循环神经网络.....	14
2.4.3 长短期记忆神经网络.....	17
2.5 本章小结.....	20
第 3 章 智能化参数寻优算法设计.....	21
3.1 粒子群优化算法.....	21
3.2 自适应粒子群优化算法.....	22
3.2.1 自适应种群划分.....	22
3.2.2 不同种群间学习策略.....	23
3.2.3 子群间信息交互.....	24
3.2.4 种群多样性与收敛速度权衡.....	24
3.3 实验结果与分析.....	24
3.4 本章小结.....	29

第 4 章 基于自适应粒子群优化 LSTM 的股票预测模型.....	30
4.1 超参数优化过程.....	30
4.2 PSO-LSTM 模型构建过程.....	31
4.3 实验结果及分析.....	33
4.3.1 数据预处理	33
4.3.2 误差评价指标	34
4.3.3 模型参数设置	35
4.3.4 实验结果分析	35
4.4 本章小结	40
第 5 章 总结与展望	41
5.1 工作总结	41
5.2 未来工作展望	41
参考文献	43
攻读学位期间取得的学术成果	49
致谢	50

第1章 绪论

1.1 研究背景

我国金融市场不断发展壮大,金融制度体系日益健全,促使民众广泛参与到金融市场投资中,使得对金融信息服务的需求越来越迫切。随着现代社会信息化和网络化的发展,计算机技术体系愈发完善,为解决民众的金融信息需求问题提供了强有力的基础和工具。在整个国家的经济体系中,金融市场占据重要地位。特别是近年来,学术界、投资界等诸多领域越来越关注于金融市场研究,并尝试利用当前先进的或是主流的计算机技术解决金融市场中的信息需求问题,其中,预测股票的未来趋势引起研究者的广泛兴趣,对相关预测理论和方法的研究逐渐成为国内外的研究热点。

近年来,深度学习技术在人工智能领域迅猛发展。作为目前最前沿的人工智能技术,深度神经网络(Deep Neural Networks, DNN)模型已经在图像处理、检测与识别、行为追踪、自然语言处理等诸多领域得到成功应用,有效促进了人工智能技术的发展。深度学习方法能够取得广泛应用主要是由于大数据时代下,数据量爆炸式的增长为达到了深度模型学习和训练所需的量级。而且,硬件技术的日益成熟为机器提供了强有力硬件保障,有效提升了机器的运算性能。这就形成了研究人员与工业界的良性循环,硬件环境的提升必然促进研究人员的拓展深度,从而反馈给工业界以提升相关技术。深度学习方法的主要优势在于其直接基于最原始的数据特征进行学习,通过对底层特征的不断提取和组合形成高层特征的抽象表达。这种处理方式更利于发现隐藏在数据中的内在联系。而且,基于多层的网络结构和激活函数,模型可以逼近任何非线性函数,使得深度学习方法更好拟合复杂的非线性问题。影响股票市场的因素纷繁复杂,使得市场变动规律复杂、价格波动性大,仅依靠传统的股票预测方法往往不足以有效解决这些问题。深度学习方法的出现提高了对复杂模式识别问题的解决能力,其训练出的模型可以刻画数据丰富的内在信息,且具有较强的泛化能力,这为股票价格预测问题的解决提供了重要的参考价值。基于对深度学习理论和实践方法的研究,对股票历史数据预处理,消除噪声干扰,然后有针对性地提取股票价格时间序列关键性特征,最后结合先进的深度学习方法建立模型提高预测精度,进而为股票投资提供较为准确的投资策略和投资方向,具有重要的理论与现实意义。

1.2 国内外研究现状

股票价格数据是金融时间序列的一种,迄今为止对于金融时间序列的预测方法大致可以分为两类:统计模型和机器学习模型。

传统的统计模型通常基于正态分布变量之间的线性假设^[1]。其中用于处理时间序列的模型主要包括时间序列回归,指数平滑,自回归积分移动平均 (ARIMA),广义自回归条件异方差 (GARCH) 等^[2-3]。这些模型通常假设研究中的时间序列是从线性过程生成的。然而与其他类型的数据相比,金融时间序列呈现出大量的不确定性和噪声^[4],具有较强的非线性,并且在每个时间点无法观察到相同的统计特性。除此之外,独立变量和因变量之间的关系的动态变化经常发生在金融时间序列中^[5-8],这些特征使得传统的统计方法很难有效地应用于股票价格预测。因此,金融市场被认为是复杂的,进化的,噪声的,非线性的和非参数的动态系统^[9],需要更多的自适应和灵活的机制来提高预测能力。

以神经网络为首的多种机器学习模型能够反映复杂数据的非线性和不规则性,受到研究人员的广泛关注。早在 20 世纪 90 年代,white^[10]曾用简单神经网络对 IBM 股票的日报酬率进行预测,因为受到 BP 神经网络梯度爆炸的局限性,模型只收敛于局部极小值,因而没有得到准确的预测结果。Kimoto 等人^[11]利用神经网络技术开发了 TPOIX 预测系统,用以预测东京证券交易所的加权平均指数(与道琼斯指数相似)。实验结果中预测结果达到预期水平。Gen^[12]于 1996 年在人工神经网络领域中发展了正向人工神经网络模型,并利用该模型对道琼斯工业指数进行预测,其实验结果验证了神经网络模型的预测能力优于简单的移动平均统计算法。在移动平均模型和前向神经网络的基础上,Rodriguez 等人^[13]于 2000 年建立了一种对股市预测的综合模型,并在西班牙股市数据上做了实证分析。该研究表明,神经网络模型能够作为一种有效的投资策略指导方式。2004 年,Zhang 等人^[14]针对时间序列预测问题进行对比试验,通过比较神经网络模型和 ARIMA 模型的预测结果,验证了神经网络模型更适合分析处理非线性数据,ARIMA 模型的预测精度相对较低。Liu 等人^[15]采用独立成分分析方法对股票价格时间序列进行预处理,以减小股票价格数据中的较高白噪声和随机性的不利影响,然后利用神经网络建立股票趋势预测模型。实验结果与小波分析方法进行比较,神经网络模型优于小波分析。Bayogly 和 Bahadir 等人^[16]分别通过贝叶斯估计和神经网络进行股市预测,实验比较了两种方法的预测效果,结果表明,神经网络

方法取得了更为突出的结果。**Murat**^[17]对预测 TKC 股票的大盘数据采用了神经网络模型,并将趋势、模式、波动性等参数引入模型中,实验结果表明神经网络模型的预测效果优于其他预测模型。**林楠**等人^[18]对传统的 BP 神经网络和 GARCH 模型进行了实证比较。结果表明,神经网络模型在预测股票价格方面具有显著优势。**Hammd** 等人^[19]通过运用多层 BP 神经网络对约旦证券市场进行股价预测实证分析。结果表明,与统计和计量分析方法相比,多层 BP 神经网络具有更强的预测能力和较高的精度,但传统的 BP 神经网络容易陷入局部极小值的问题没有被提及。相对于统计模型和机器学习方法,神经网络模型对时间序列数据的非线性关系具有更好的描述能力^[20-23]。

与此同时,各国研究人员试图将机器学习算法、不同的预测模型相结合建立综合模型,以提高模型的预测能力。**Arnold**^[24]创新性地将遗传算法、粗糙集分类和神经网络三种算法结合起来,建立了综合预测模型。然而,文章没有单独实现三种方法的预测效果,也没有比较分析三种模型的单项预测效果与综合模型预测效果间的差异,亦没有对综合模型进行验证,这使得综合模型预测结果不客观。**Melike** 等人^[25]将 ARCH、GARCH 模型与 BP 神经网络模型相结合提出 ANN-APGARCH 综合模型,借助伊斯坦布尔股市 30 年的股票历史数据训练模型,实验结果表明,该模型有效地提高了预测结果的准确性和可靠性。**肖静**等^[26]针对容易陷入局部极小(局部最优解)的传统 BP 神经网络,结合 LM 算法建立了三层 BP 神经网络预测模型,克服了 BP 算法陷入局部最优的缺陷。同时还利用遗传算法优化了网络中神经元连接的权值和阈值,提高了模型的精度和准确性。**Hsieh** 等人^[27]突破性地将周期 BP 神经网络与小波变换及人工蚁群算法结合,数据预处理阶段采用小波变换对股票数据进行处理,再将数据输入神经网络,利用人工蚁群算法对神经元连接权值进行优化。**张坤**等人^[28]还考虑了股票价格时间序列的不确定性,结合小波技术和神经网络建立预测模型。其中,为了更好地发挥神经网络模型的预测能力,股价数据先被小波技术分解成不同的周期,再借助网络预测,最后用 BP 神经网络对不同周期的预测结果拟合。实验表明,比起单独使用小波分析或神经网络,该综合模型的预测精度更高。此外,也有学者建立基于遗传算法与神经网络相结合的上证指数预测模型,试验并没有达到预期理想效果^[29]。

随着人工智能技术的发展,以深度神经网络、卷积神经网络和循环神经网络(Recurrent Neural Network, RNN)为代表的深度神经网络模型,在自动驾驶、目标检测、语音识别、人脸识别等诸多领域取得了优异的表现。越来越多的研究人员将目

光转向非线性表达能力更强的深层神经网络^[30-31], 尝试将深度学习技术应用于股票市场预测。其中, RNN 内部网络结构中存在特有的反馈连接, 这使得过去时间的信息得以保存^[32]。具备时间表示能力的 RNN 在处理财务预测、自然语言处理和语音识别^[33]等序列分析任务中表现出天然的优势。目前, 国内外许多研究利用包括长短期记忆模型 (Long Short-Term Memory, LSTM) 在内的 RNN 对金融时间序列进行预测研究^[34-38], 并取得了一定的研究成果。Lin 等人^[39]利用 RNN 的一种变体回声状态网络 (echo state networks, ESN) 预测了下一个交易日的收盘价。Wei 和 Cheng^[40]提出了一种混合 RNN 股价预测模型, 利用综合特征选择来检测股市预测的关键技术指标。他们利用逐步回归和决策树来减少财务数据的维度, 实验结果表明了该模型的优越性。Dixon^[41]将 RNN 应用于高频交易, 从金融期货限价订购单到预测下一次事件的价格翻转, 对短期价格走势进行分类。Fischer 和 Krauss^[42]利用 LSTM 网络对标普 500 指数成份股从 1992 年到 2015 年的走势进行了预测, 并将模型预测结果与随机森林 (Random Forest, RF)、深度神经网络和逻辑回归等无记忆分类器进行了比较。实验表明 LSTM 网络相比于其他模型具有明显的优势, 验证了 LSTM 网络适用于股票市场研究。Cai 等人^[43]将全局搜索算法和进化算法与 RNN 相结合, 对时间序列数据中的缺失值进行估计。Hsieh 等人^[44]提出了一种结合人工蜂群算法 (Artificial Bee Colony, ABC) 的 RNN 的股票市场预测系统, 他们利用 ABC 对 RNN 的连接权重进行优化, 利用小波变换对市场数据进行分解, 去除噪声。Rather 等人^[45]提出了一种综合模型, 该模型由 RNN 和 ARIMA、指数平滑两种线性模型组成, 预测股票收益。Chen 等人^[46]以中国股市为背景, 建立 LSTM 网络模型预测股票价格, 实证结果表明该模型对国内股市预测的有效性。Jia 等人^[47]通过一系列的实证分析验证了 LSTM 网络模型对股价趋势预测的有效性。Siami-Namini^[48]对于金融时间序列进行预测, 验证了基于深度学习的 LSTM 网络的预测性能优于传统统计模型。借助股票历史数据和技术指标, Nelson 等人^[49]建立了 LSTM 股票波动趋势预测模型。实验表明, LSTM 网络的预测性能优于其他机器学习方法。

综上所述, 早期对于时间序列的预测主要基于统计模型, 随着人工智能的发展, 越来越多的机器学习方法得到广泛应用。特别的, 目前以 LSTM 网络为代表的循环神经网络模型取得较好的预测性能, 其预测精度优于传统统计模型与机器学习方法。

1.3 论文的主要研究内容

本文通过设计一种智能化参数寻优算法,发挥其收敛速度快、求解精度高等优势,与 LSTM 网络进行深度融合,致力于快速寻找网络模型的最优超参数,减少占用人力与计算资源的同时,达到模型预测性能最优化。本文的主要研究内容如下:

- 第一, 设计智能化参数优化算法,在经典粒子群优化算法的基础上,提出一种具有自适应学习策略的粒子群优化算法,克服传统粒子群优化算法存在的易陷入局部最优的缺点,保持算法求解精度的同时提高收敛速度。
- 第二, 将智能化参数优化算法与 LSTM 网络相融合,通过自适应粒子群优化算法对 LSTM 网络的关键参数进行寻优,使股票数据特征与网络拓扑结构相匹配,提高股票价格预测精度。
- 第三, 对于提出的参数优化算法进行数值模拟实验,与相关算法进行对比分析,验证算法的优化性能与收敛速度。
- 第四, 利用不同类型的真实股票历史数据,对提出的股价预测模型进行对比实验,验证模型的有效性和适用性。

1.4 论文的组织结构

第 1 章是绪论部分,详细介绍了股票价格时间序列预测的研究背景,讨论和分析了国内外研究现状,以及论文的主要研究内容和结构安排。

第 2 章介绍了股票价格时间序列预测的相关理论与预测模型,包括股票价格预测理论、统计模型、机器学习模型和神经网络,此部分为后续的工作提供了理论支持。

第 3 章论述了智能化参数寻优算法的设计过程。首先介绍了经典粒子群算法的基本原理。其次,对有自适应学习策略的粒子群优化算法的构造过程及性质进行分析。最后,对提出的算法进行数值模拟实验。结果表明,该算法既能保持较快的收敛速度,又能有效地避免陷入局部最优。

第 4 章将智能化参数寻优算法与 LSTM 网络相融合,提出一种基于自适应粒子群优化的 LSTM 股票预测模型。首先,分析了关键网络结构参数智能化选取的重要意义。其次,分别对关键网络结构参数的优化过程以及 PSO-LSTM 股票价格预测模型的构建流程进行了详细论述。最后,实验部分利用不同类型真实股票数据验证了该模型的有效性。本章内容已被《北京航空航天大学学报》录用。

第 5 章总结和展望,一方面总结了本文的主要工作以及取得的研究成果。另一方面分析了未来待解决的问题,并从将分形理论应用于金融时间序列预测的角度对以后的工作进行了展望。

第2章 相关技术介绍

自股票市场建立以来,研究人员一直尝试各种方法来预测股票价格时间序列。起初,时间序列预测方法主要基于统计学理论实现。随着对机器学习技术的研究,使用机器学习方法学习时间序列数据并建立模型进而预测成为该领域研究的重点。近几年,深度学习方法在自动驾驶、自然语言处理、检测与识别等诸多领域得到成功应用,引起了时间序列预测领域的广泛关注,利用各种深度神经网络模型进行股票预测的尝试也越来越多。

本章主要对股票价格时间序列预测的基本概念和相关理论进行介绍,并对该领域中的代表性工作进行分析,讨论与总结现有工作的优势及局限性,为后续章节的研究及本文预测方法引入提供预备知识。

2.1 股票价格预测理论

中国股票市场规模的不断壮大,股票交易的日益繁杂,促使股票价格的波动难以捉摸。股市投资者最关注的就是股价的变化。对股票价格的分析能够帮助投资者认识股票市场的变化规律,实现理性投资。因此,国内外学术界和投资界对股价分析方法的研究十分关注。在股票市场的发展过程中,逐渐形成两种主要的股票价格分析方法:基本面分析方法和技术面分析方法,其中,前者主要基于国家政策制度和公司基本面等信息,而后者则利用数学理论和模型对历史数据建模预测股票未来走势,强调对历史交易数据的研究。

2.1.1 基本面分析

基本面分析是指基于金融学、经济学和证券投资学等学科知识,通过分析国家相关政策、宏观经济情况、公司业绩等影响股票投资价值的基本因素,研究持股公司或企业的未来财务状况和行业动态,进而对股票的投资价值和浮动程度进行评估和判定,为投资者的下一步投资计划提供相关的决策和依据。基本面分析方法重在研究决定股票投资价值的基本要素,主要包括公司分析、宏观经济分析和行业分析。

公司分析是基本面分析方法的核心,通过借助公司财务报表、产品与市场、股利分配政策等,对公司的财务状况、经营能力、未来发展前景等方面进行综合分析,以评估股票的投资价值和未来走势。公司盈利的增长会增加投资者的收益,管理层的正

确决策能够提高市场竞争力和市场盈利,因此公司分析主要集中在财务分析和管理层分析两个方面。财务分析首先是从现金流量表、资产负债表等报表中提取利于决策的信息,然后借助财务比率等指标分析公司的经营状况和投资收益。管理层分析主要依靠公司管理人员的管理能力、工作能力、专业技术等。

宏观经济分析研究国家相关政策和经济指标对股票价格的影响,可分为宏观经济因素分析和宏观非经济因素分析。其中,前者主要分析财政政策、利率、汇率等宏观政策和宏观指标,后者主要分析政局变动等能够影响宏观经济的政策和指标。

行业分析研究行业背景对股票价格的影响。主要从行业结构、市场类型、生命周期、业绩等方面,研究上市企业在国民经济中的地位,以及横向比较在不同行业中的发展状况,为投资者提供准确的行业背景分析。

综合来看,基本面分析法能够比较全面地把握股票的投资价值,但需要较长的预测时间跨度,因此,该方法对于成熟的市场或是长期投资较为适用。

2.1.2 技术面分析

技术面分析强调价格的行为特征,主要利用历史数据预测股票价格,即,以历史股票市场行为为研究对象,基于历史价格、成交量等数据,运用统计等数学方法描述并建立股价和指数变动的历史轨迹模型,寻找股票市场交易的变化规律,对股价的未来走势进行预测。

技术面分析通常以股价、成交量和时间为主要分析要素。首先,成交价和成交量是股市交易中最基本的交易行为。通常情况下,当股价和成交量同时增长时,股价会保持上涨趋势,而当股价和成交量趋势不一致时,股价会出现相反走势。时间也是影响股票市场的重要因素,例如,一般而言,当股票趋势已形成就不会发生反转。此外,技术面分析需要基于以下基本假设:一切因素对股市的影响必须通过股价变化反映;股价会沿着已成趋势移动;投资者的交易行为会受其心理因素的影响导致历史重演。常用的技术面分析方法包括K线组合分析、价格形态分析和经典技术指标分析。

K线用于记录股票在一个交易日内的价格变化情况,是投资者常用的一种股价分析工具。K线是由实体和影线两部分组成的柱状线。实体表示一个交易日的开盘价和收盘价。影线有两条,分别位于实体的上、下部分,称为上影线和下影线,表示一个交易日的最高价和最低价。根据开盘价和收盘价之间的关系,K线可分为阳线和阴线。前者代表收盘价高于开盘价的情况,后者则表示收盘价低于开盘价。将多个交易日的

K 线连在一起可以反映某段时间内股票价格的波动情况。此时，根据时间的不同，K 线又可分为日线、周线、月线和年线。K 线将股价的波动情况清晰地反映给投资者，能够帮助投资者直观地了解股票价格的变化，提高了投资效率。

价格形态指卖出和买入操作对股票价格产生的变化。在某段交易日内，股票卖出数量增多会引起股票价格的下降，而买入数量的增多会拉动股票价格的提升。常见的价格形态有头肩顶、头肩底和矩形。头肩顶主要用于上涨尾声的看跌形态，相反，头肩底常用在下跌尾声的看涨形态。矩形形态常见于多空双方均衡导致的股价波动。

技术指标是利用数理方式对股票数据的波动以及数据之间的复杂关系加以量化。使用技术指标分析股价时，需要满足一定的前提和范围，才能发挥出指标性能。在实盘操作中，多个指标的综合应用往往能够增加投资的准确性。常用的技术指标有随机指数、移动平均线和平滑异同平均线等。随机指数指标基于最高价、最低价及收盘价计算，利用股价的波幅判断未来走势，具有快速、直观判断行情的能力，常被用于中短期趋势预测。移动平均线指标通过将某段交易日内的股票价格加以平均，再连接不同时间段的平均值形成，主要用于减少股价随机变动产生的影响。当移动平均线指标出现黄金交叉时可以积极买入，死亡交叉则预示行情下跌。平滑异同平均线指标利用快速和慢速移动平均线，加上双重平滑计量，判断买卖时机，可以减少利用移动平均线指标的高频买卖操作。

技术面分析能够直观反映出股票价格的变化，定量分析最佳买卖时机，但是该方法不适合于长线投资，且不能及时判断非常规交易或是突发事件。

2.2 统计学模型

早期的股票价格预测模型通常依托统计学习模型建立。借助统计学理论作为支撑，这类模型具有较强的可解释性强且预测效率高，因此，此类方法在当下的一些预测任务中也表现出较好的预测性能。常用于股票价格预测的统计学模型主要有整合移动平均自回归模型（Autoregressive Integrated Moving Average model, ARIMA）、指数平滑模型和移动平均模型等。下面先概述时间序列的基本概念，再对具有代表性的 ARIMA 和指数平滑模型进行介绍。

2.2.1 时间序列的统计特性

时间序列的统计特性主要表现为平稳性与随机性。

(1) 平稳性

任何的时间序列都可视为由随机过程生产。通俗来讲,在一个时间序列中如果严格消除了周期性变化,期望、方差等统计特性不会随时间产生系统的变化,那么该时间序列就是平稳的。具体地,假设某时间序列中的各个数值由某个概率分布采用产生,即任何时间点上的数值通过一个随机过程生产。假定该时间序列满足以下条件:

① 期望和方差在任何时间过程上为常数,即均与时间点无关;

② 任何两个时间间隔的协方差不依赖于这两个时间点,只与此时间间隔有关。

则认为该时间序列为平稳的。从统计意义上来看,平稳性可视为时间序列的一种可预测性表达。在诸如期望、方差和协方差等均不依赖与时间点变化时,就可以通过建立统计模型找出隐藏在时间序列中的基本数据特征,并利用其对序列的未来变化趋势进行预测。时序图观察和自相关图检验是两种常用的验证平稳时间序列的方法。

(2) 随机性

根据平稳时间序列的定义,自协方差函数只与时间间隔有关,与时间点无关。在时间间隔相同的情况下,其自协方差函数为常数。若该常数为0,则此时序列之间的相关性为0,即不相关。对于这样一种无法挖掘出可用规律的不相关序列,变称之为随机性时间序列。假设某时间序列中任何时间点上的数值通过一个随机过程生产。假定该时间序列满足以下条件:

①期望和方差任何时间过程上为常数,即均与时间点无关;

②任何两个时间间隔的协方差在任何不相等时刻均为0,则认为该时间序列为随机的。可以发现,随机性时间序列没有记忆性,即当前时间点上的数值只与该点有关,而与之之前的任意时间点无关。因此,随机性时间序列的任何两项之间都是无关的,无法从中获得有用信息并建立预测模型判断未来趋势的变化。时间序列的随机性可以通过随机性检验确定,Q统计量和LB统计量是2种常用的检验方法。

2.2.2 时间序列预测的分类

按照任务要求的不同,时间序列预测问题可分为单步预测和多步预测。其中,单步预测指的是通过时间序列的历史数据预测未来一个时间步的序列值,实现多步预测的基础。多步预测指的是通过时间序列的历史数据预测未来多个时间步的序列值,主要包括直接预测法和迭代预测法:

(1) 直接预测法：考虑到模型的预测结果由多个输出值组成，为了更方便地输出预测结果，在模型过程中采用多输入多输出的形式，这样最终的模型预测结果只含有一个输出值。

(2) 迭代预测法：该方法的主要思想是将每个时间步的预测信息视为历史数据进行学习，即，模型首先预测未来一个时间步的结果，然后将这个单步预测结果作为已知数值加入到序列中进行一个时间步的预测，通过不同的迭代获得最终的预测结果。

分析来看，直接预测方法思想简单，较为直接，能够实现对输入输出的直接建模。但对于这种复杂的多对多映射关系，若要建立性能好的预测模型，需要大量的数据作为支撑。相对而言，由于迭代预测法是一种多对一的学习模式，因此所需的数据量较小。但该方法将模型的预测值作为数据迭代预测，可能会导致预测误差的逐级放大。

2.2.3 ARIMA 模型

ARIMA 模型是对自回归滑动平均模型（Autoregressive moving average model, ARMA）的改进。ARMA 模型同样基于统计学理论提出，在时间序列预测领域也具有广泛应用。ARMA 模型主要由自回归部分和移动平均部分组成，具体表达形式如下：

$$\hat{x}_t = \sum_{i=1}^p \phi_i \cdot x_{t-i} + \sum_{j=1}^q \theta_j \cdot \varepsilon_{t-j} + \varepsilon_t,$$

其中， x_t 和 \hat{x}_t 分别为时间序列值和预测值， ϕ_i 和 θ_j 分别为自回归部分和移动平均部分的系数， p 和 q 分别为模型阶数。公式中的第一项为自回归部分，主要描述历史序列值与预测值之间的线性关系，第二项为移动平均部分，刻画预测误差与预测值之间的线性关系，第三项 ε_t 代表白噪声。ARMA 模型算法易于实现且效率高，但是必须在时间序列是平稳的这一前提假设下使用，对于具有周期性或是趋势明显的时间序列，该模型的预测效果可能会变差。为了对非平稳时间序列进行有效预测，产生了 ARIMA 模型。

ARIMA 的主要思想是对于不稳定序列先进行稳定化操作，然后再利用 ARMA 模型处理。为了稳定化序列，ARIMA 模型引入了 d 阶差分操作。这种操作是统计学常用的一种用于将不平稳序列转换为平稳序列的方法，其中，一阶差分（ $d=1$ ）和二

阶差分（ $d=2$ ，即在一阶差分的基础上进行第二次差分操作）在实际中广泛使用，公式如下：

$$u'_t = x_t - x_{t-1},$$

$$u''_t = x_t - 2x_{t-1} + x_{t-2},$$

其中， u'_t 和 u''_t 分别代表一阶差分和二阶差分后的项。可见， p 、 q 和 d 这 3 个超参数对 ARIMA 模型的预测性能影响较大，参数 d 主要借助平稳性检验确定。而参数 p 和 q 通常需要利用自相关图和偏自相关图中的截尾和拖尾现象确定。因此，在 ARIMA 模型中，参数 p 和 q 的确定较为困难，需要人工干涉。目前也有自动确定这 2 个参数的方法，但通常计算效率较低，效果不好。而且，考虑到实际应用中会出现易漂变的问题，参数的确定通常不利于模型对序列模式变化的刻画。此外，ARIMA 模型中没有增量更新操作，模型每次更新需要对参数值进行重新估计，降低了模型的处理效率，通常难以准确刻画序列变化趋势。

2.2.4 指数平滑模型

指数平滑模型也是一种应用广泛的统计预测模型，该模型主要利用当前预测时刻最近的几个时间序列值进行预测，并对不同时刻的序列值赋予不同的权重。下面给出一阶指数平滑的表达形式：

$$\hat{x}_t = \alpha x_{t-1} + (1-\alpha)\hat{x}_{t-1},$$

其中， x_t 和 \hat{x}_t 分别是 t 时刻的序列值和预测值， α 称为平滑系数，在 $[0,1]$ 范围内取值。较大的 α 容易使得预测结果偏向于当前预测时刻的序列值，而较小的 α 会使得预测结果偏向于之前的序列值。经过迭代计算可得：

$$\hat{x}_t = \alpha x_{t-1} + (1-\alpha)\alpha x_{t-2} + (1-\alpha)^2 \alpha x_{t-3} + \cdots + (1-\alpha)^{t-1} \hat{x}_1,$$

由于 $\alpha \in [0,1]$ 可知，当序列值离预测时刻 t 越远，其对预测结果的贡献越小，反之亦然。

除了上述的一阶指数平滑法，二阶和三阶指数平滑法也是较为常用的指数平滑模型，其中，二阶指数平滑适用于解决趋势性序列预测，而三阶指数平滑适用于解决季节性序列预测。指数平滑模型的局限性在于参数调整较为困难，平滑系数 α 需要一定的先验知识确定。而且，根据模型的建模机理，该模型在短期任务上的预测效果较好，

但不适合处理更加复杂多变的时间序列。

综上所述,作为早起时间序列预测模型的代表,ARIMA 模型和指数平滑模型等统计预测模型原理简单,易于实现,在处理传统的小规模时间序列预测任务上效率较高。但是,也因为此类模型构造较为简单,对复杂的线性时间序列预测性能欠佳,并且缺乏对时间序列特性的建模。

2.3 机器学习模型

在不考虑序列变量相关性的情况下,预测模型可看作是寻找历史序列值和预测值之间的一种映射关系,而这种映射关系可以通过机器学习模型建立。

对于时间序列长度较长或者预测序列长短不一致的情况,在模型训练过程中,一般需要选择一个大小为 w 的滑动窗口来截取多段子序列 $x_{t-w} \sim x_t$ 作为机器学习模型的输入,输出为该子序列后 h 个时间步的预测结果 $\hat{x}_{t+1} \sim \hat{x}_{t+h}$,真实值为该子序列后 h 个时间步的序列值。通常,模型中的参数通过建立真实值和预测结果之间的损失函数进行确定。平方误差损失是一种常见的损失函数,其表达式如下:

$$Loss = \frac{1}{2} \sum_{i=t+1}^{i=t+h} (x_i - \hat{x}_i)^2,$$

相对于统计学预测模型,机器学习方法对非线性的时间序列具有更好的拟合能力,而且参数的敏感性较低,模型的泛化能力相对较强^[50-53]。而且,机器学习方法直接对原始序列数据学习,建模过程更加直观简便。然而,此类方法仍然是对历史序列值和预测值之间映射关系的直接拟合,序列变量之间的相关性没有考虑在内。

2.4 神经网络理论

人工智能技术的进步推动了人工神经网络的发展,诸多神经网络模型被应用在各个领域,取得了突破性进展。同样,在时间序列预测领域,研究人员愈发关注于神经网络模型的构建。与机器学习模型的思想类似,神经网络模型也是通过数据训练得出历史序列值和预测值之间的映射关系,而且这类模型对非线性数据也具有强大的拟合能力。据相关研究表明,一个含隐藏层的三层神经网络可以任意精度逼近任何非线性函数^[54]。基于神经网络的时间序列预测的早期探索主要集中在对网络模型的构建研究上。

2.4.1 多层感知机与 BP 神经网络

多层感知机（Multilayer Perceptron, MLP）一般用来代表传统的神经网络。一个三层的神经网络，如图 2-1 所示，通常包括输入层、隐藏层和输出层。每层中的神经元可视为一种运算处理单元，主要进行权重相加和非线性激活两种运算操作。

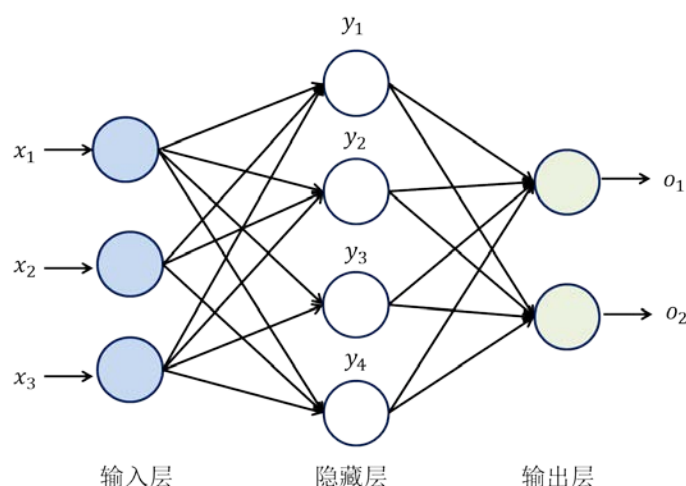


图 2-1 三层神经网络

后向传播神经网络（Back Propagation, BP）是一种特殊的多层感知机，是应用最广泛的神经网络模型之一。BP 网络利用误差反向传播算法对网络进行训练，该算法解决了网络中隐藏层连接权重的学习问题。网络以预测值和真实值的误差平方和为目标函数，采用梯度下降法，通过反向传播算法调整训练网络参数，使每次训练的目标函数值最小。BP 网络具有对大量非线性映射关系的学习和存储能力，无需提前揭示描述映射关系的数学表达。

BP 神经网络的优点包括：具有较强的非线性映射能力；能够将学习的内容存储在网络模型权值之中，具有较强的自学习和自适应能力；具有较好的泛化能力以及容错能力。基于以上优点，BP 神经网络得到广泛应用。与此同时，也暴露出许多不足，主要包括：BP 神经网络对于初始网络权重十分敏感，容易陷入局部最优；算法本质为梯度下降法，容易导致模型训练收敛速度慢。

2.4.2 循环神经网络

RNN 网络是一种专门为处理序列数据建立的深度网络模型。在 RNN 网络中，不同的时间步长具有相同的一套权重，循环地进行跨时间步的连接。由于权重的共用，RNN 网络中的参数量比 MLP 网络大大减少，且能够建立序列数据前后关系的非线性

模型。

如同人类会借助历史记忆更好地认识世界,RNN 网络最大的特点在于先前知识和当前信息共同决定预测结果,即,模型的建立不仅考虑了前一时刻的输入,还对历史序列数据具有记忆功能。这是由于模型通过跨时间点的反馈结构连接,能够存储历史时间点的序列数据信息,从而实现对时间序列的显式建模。RNN 网络的循环性体现在当前的输出与先前的输出有关,具体表现为网络对先前信息记忆,并用于当前输出的计算中。这种方式使得隐藏层之间的节点变为有连接形式,并且将前一时刻隐藏层的输出和输入层的输出同时作为隐藏层的输入。RNN 网络的结构如下图 2-2 所示:

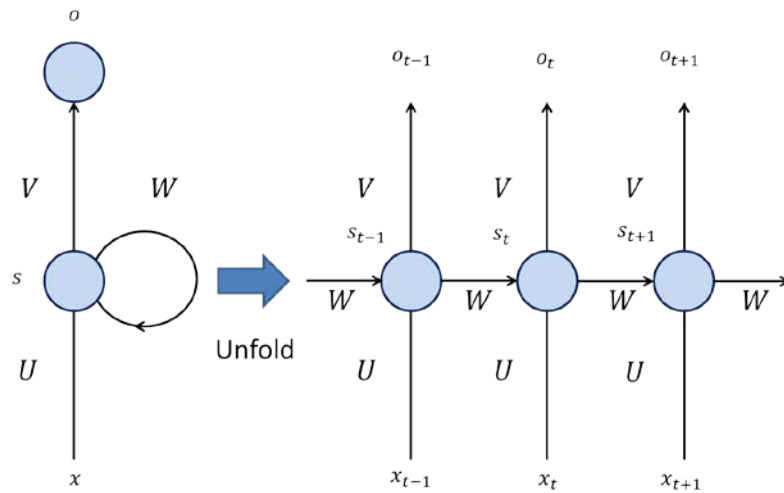


图 2-2 RNN 的神经结构图

未展开的 RNN 网络结构如图 2-2 左侧所示,可以看出,同一般的 MLP 网络相同, RNN 网络包括输入层、输出层和隐藏层,不同的是隐藏层中的神经元不仅与输入和输出有关,还存在一个与其自身的回路,即,上一时间步的隐藏层输出信息会作用于下一时间步的隐藏层神经元。图 2-2 右侧为 RNN 网络结构按时间序列的展开形式。定义时间序列 $[\cdots x_{t+1}, x_t, x_{t-1} \cdots]$ 为输入数据, $[\cdots s_{t+1}, s_t, s_{t-1} \cdots]$ 为隐藏层数据, $[\cdots o_{t+1}, o_t, o_{t-1} \cdots]$ 为输出结果, U 、 W 、 V 为输入与隐藏层之间、上一时间步的隐藏层到当前时间步的隐藏层之间以及隐藏层与输出层之间的权重矩阵。与传统网络不同,RNN 网络中不同的时间步共享同一套权重矩阵。

RNN 网络的前向传播过程由输入序列数据按照时间顺序经过层层激励得到最终的预测输出。在任意序列索引 t 时,隐藏层可表示为:

$$h(t) = \sigma(z(t)) = \sigma(U \times x(t) + W \times h(t-1) + b)$$

其中, σ 表示激活函数,一般常取 \tanh 函数, b 为线性偏倚。最终的预测输出为:

$$y(t) = \sigma(o(t)) = \sigma(V \times h(t) + c)$$

RNN 网络通常情况下用于分类问题，因此，上式中的激活函数一般取 *softmax* 函数。

RNN 网络的反向传播过程采用沿时间的反向传播算法。与 BP 网络中的反向传播算法相同，该算法也是借助梯度下降法调整网络中的权重和偏倚参数。不同的是，RNN 网络中每个时间步的输出依赖于当前时刻和上个时间步的信息，因此在反向传播时需要按时间步展开。而且，由于网络中每个时间步共享权值，所以反向传播时对相同的参数更新。下面，我们以交叉熵损失函数为例，对 RNN 网络中的反向传播过程进行说明，这里，隐藏层和输出层分别选取 *tanh* 函数和 *softmax* 函数作为激活函数。网络损失 L 为每个时间步的损失之和，即，

$$L = \sum_{t=1}^{\tau} L(t)$$

可得 V 、 c 的梯度计算如下：

$$\frac{\partial L}{\partial c} = \sum_{t=1}^{\tau} \frac{\partial L(t)}{\partial c} = \sum_{t=1}^{\tau} \hat{y}(t) - y(t)$$

$$\frac{\partial L}{\partial V} = \sum_{t=1}^{\tau} \frac{\partial L(t)}{\partial V} = \sum_{t=1}^{\tau} (\hat{y}(t) - y(t))(h(t))^T$$

序列索引 t 时的损失由当前时刻的输出损失和序列索引 $t+1$ 时的损失构成。根据前向传播算法，为求 U 和 W 的梯度，需要先求出 σ_t 的梯度，即，

$$\sigma_t = \frac{\partial L}{\partial h_t} = \frac{\partial L}{\partial \sigma_t} \times \frac{\partial \sigma_t}{\partial h_t} + \frac{\partial L}{\partial h_{t+1}} \times \frac{\partial h_{t+1}}{\partial h_t}$$

$$\sigma_t = V^T (\hat{y}(t) - y(t)) + W^T \delta_{t+1} \text{diag}(1 - h_{t+1}^2)$$

则， U 、 W 、 b 的梯度计算如下：

$$\frac{\partial L}{\partial U} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial h_t} \times \frac{\partial h_t}{\partial U} = \sum_{t=1}^{\tau} \text{diag}(1 - h_t^2) \delta_t x_t^T$$

$$\frac{\partial L}{\partial W} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial h_t} \times \frac{\partial h_t}{\partial W} = \sum_{t=1}^{\tau} \text{diag}(1 - h_t^2) \delta_t h_{t-1}^T$$

$$\frac{\partial L}{\partial b} = \sum_{t=1}^{\tau} \frac{\partial L}{\partial h_t} \times \frac{\partial h_t}{\partial b} = \sum_{t=1}^{\tau} \text{diag}(1 - h_t^2) \delta_t$$

传统的 RNN 模型存在梯度消失问题，这是由于在隐藏层激活函数选择 \tanh 函数的情况下， \tanh 函数的导数大于 0 且小于 1。在反向传播过程，连接权重也是大于 0 且小于 1 的数，则当序列索引 t 很大时，多个 \tanh 函数的倒数与连接权重乘积的累加便会趋于 0，从而导致梯度消失。为了解决这个问题，许多改进的 RNN 网络模型被提出，比如，门控 RNN 网络增加了控制网络信息传递的门控机制，Highway RNN 网络采用门函数缓解网络过深造成的梯度回流受阻问题，以及下面要介绍的长短期记忆网络。

2.4.3 长短期记忆神经网络

长短期记忆神经网络（Long Short-Term Memory, LSTM）是一种特殊的 RNN，其建立的最初目的是要解决 RNN 存在的容易导致梯度爆炸或梯度消失的缺陷。相较于传统 RNN 单元结构，LSTM 创造性的引入了控制门结构，使得网络模型能够有效的处理时间序列中长期依赖关系，历史时间信息得到有效传递。因此，对于时间序列的预测问题，LSTM 在单元结构上具有独特优势。

LSTM 的网络结构与传统 RNN 非常相似，同样具备隐藏层之间的反馈连接，主要区别在于 LSTM 单元内部结构。传统的 RNN 单元结构仅由 \tanh 激活函数构成，而 LSTM 对其单元结构进行具有突破性的升级。LSTM 单元结构如图 2-3 所示，包括最主要的两大组成部分：控制门结构和单元细胞状态。其中控制门结构主要包含了遗忘门、输入门与输出门。单元细胞状态是一条连通在所有 LSTM 单元结构中的信息流轨道，其状态具有“记忆”功能，保存着各时刻的记忆信息。

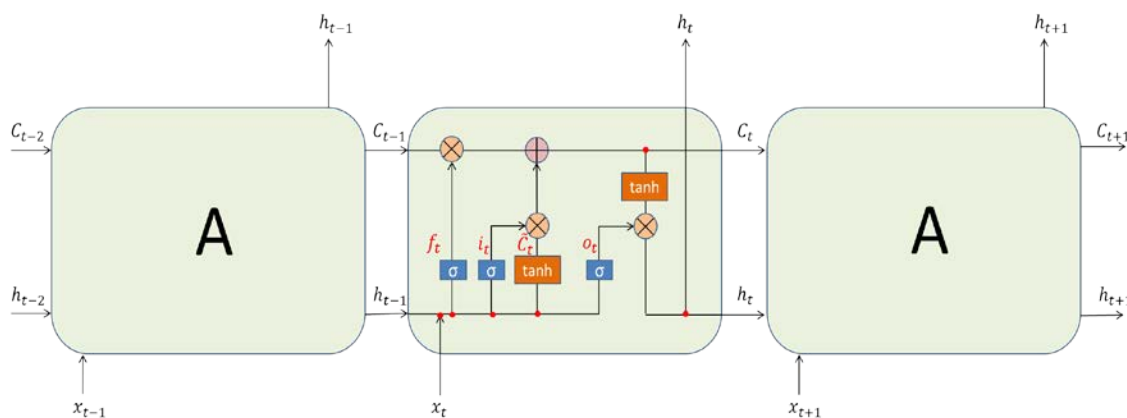


图 2-3 LSTM 单元结构图

(1) 遗忘门结构

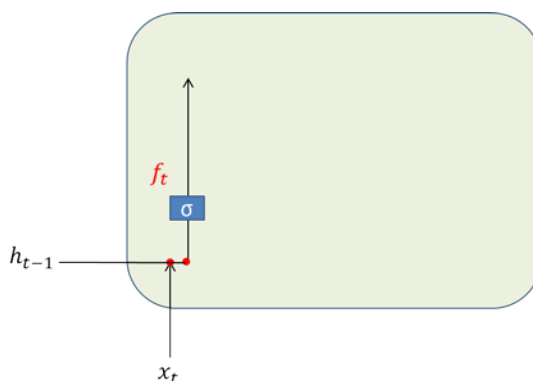


图 2-4 遗忘门结构

如图 2-4 所示,遗忘门控制着历史记忆数据有多少数据量需要遗忘或保留。上一时刻单元的输出信息与当前时刻的输入信息共同作为遗忘门的输入,经过加权处理,最后利用 *sigmoid* 激活函数计算遗忘门的输出为[0,1]之间的数值。利用该值的大小确定来自历史中的哪些信息需要遗忘,哪些信息需要保留。遗忘门结构的计算表达式如下所示:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2.1)$$

其中, x_t 表示 t 时刻单元的输入信息, h_{t-1} 表示 t-1 时刻单元的输出信息, σ 为 *sigmoid* 激活函数, W 表示权重矩阵, b 表示偏差矩阵。经过遗忘门结构,实现了去除冗余数据,保留长时间依赖关系的记忆。

(2) 输入门结构

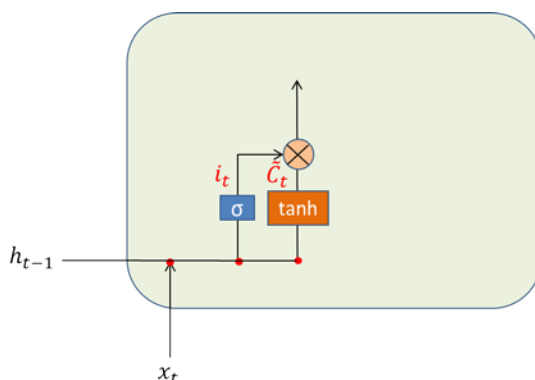


图 2-5 输入门结构

如图 2-5 所示,输入门结构控制着历史时刻与当前时刻信息有多少输入到此刻单元细胞状态中。上一时刻单元的输出信息与当前时刻的输入信息共同作为输入门的输

入，经过加权处理与 *sigmoid* 激活函数输出[0,1]之间的数值，“0”表示不允许任何信息输入到单元细胞状态，相反“1”表示允许全部记忆输入到单元细胞状态。记忆门结构的计算表达式如下所示：

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2.2)$$

经过输入门结构，实现了有多少新的信息需要记忆到单元细胞状态中。

(3) 单元细胞状态

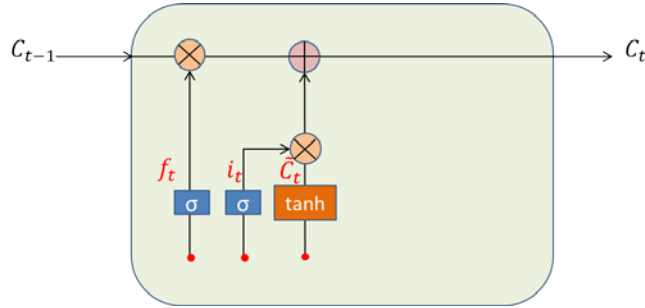


图 2-6 细胞状态结构

如图 2-6 所示，单元细胞状态用于存储历史时刻与当前时刻的记忆信息。单元细胞状态的数据信息由遗忘门结构与输入门结构进行控制。单元细胞状态的计算表达式如下所示：

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \quad (2.3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \quad (2.4)$$

其中， \tilde{C}_t 表示 t 时刻的暂时单元细胞状态， C_t 表示 t 时刻的单元细胞状态。上一时刻单元的输出信息与当前时刻的输入信息共同作为输入，经过加权操作，利用 *tanh* 激活函数将输入信息转化到[0,1]区间，其值作为 t 时刻的暂时单元细胞状态 \tilde{C}_t 。遗忘门结构与输入门结构分别作用于上一时刻单元细胞状态 C_{t-1} 与当前时刻暂时单元细胞状态 \tilde{C}_t ，从而更新当前时刻单元细胞状态。

(4) 输出门结构

如图 2-7 所示，输出门结构控制着单元细胞状态中的记忆信息哪些需要输出。上一时刻单元的输出信息与当前时刻的输入信息共同作为输入，经过加权处理与 *sigmoid* 激活函数输出[0,1]之间的数值，“0”表示不允许任何信息输出，相反“1”表

示允许全部记忆输出。记忆门结构的计算表达式如下所示：

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (2.5)$$

$$h_t = o_t \cdot \tanh(C_t), \quad (2.6)$$

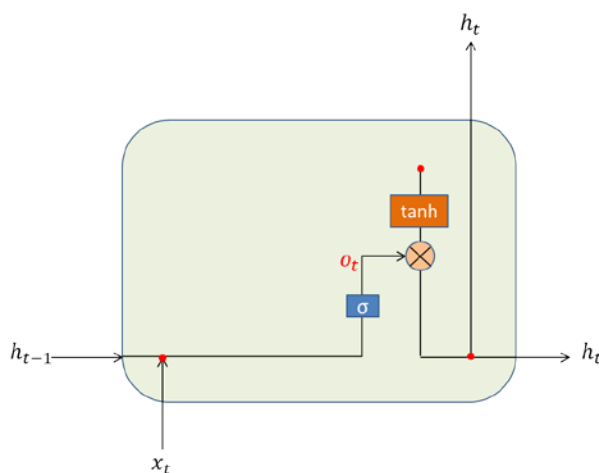


图 2-7 输出门结构

最后，对当前时刻的单元细胞状态进行 \tanh 激活函数归一化变化，利用输出门结构控制记忆信息的输出。 h_t 表示 t 时刻的 LSTM 单元输出信息。

综上所述，具有控制门结构的 LSTM 网络，能够记忆长时间的依赖关系，在时间序列处理中更具有优势。

2.5 本章小结

本章以时间序列预测方法的发展方向为主线，介绍了股票预测理论以及几种流行的时间序列预测模型的基本概念。其中包括统计模型，机器学习模型以及神经网络，主要对循环神经网络和长短期记忆神经网络进行较为详细的描述，为接下来的基于长短期记忆神经网络的股票预测研究，奠定了坚实的理论基础。

第3章 智能化参数寻优算法设计

LSTM 网络中关键结构参数直接影响着预测性能与模型稳定性,选择合适的网络结构参数是模型构建过程中非常重要的一步工作。然而目前研究显示,各国研究人员多通过人为经验或重复性计算实验对相关参数进行设置。对于多个参数选择的情况,仅依靠人工实验调参很难得到模型最优参数。因此,需要一个简单高效的智能化参数寻优算法来取代人工调参过程。本章在粒子群优化算法(Particle Swarm Optimization, PSO)的基础上,设计了智能化参数寻优算法,充分发挥粒子群优化算法收敛速度快的优势,追求更高的求解精度。首先介绍了经典的粒子群优化算法的基本概念,其次提出一种具有自适应学习策略的粒子群优化算法,最后利用数值模拟实验验证提出的优化算法的有效性,在求解精度、收敛速度方面与典型粒子群算法进行对比分析。

3.1 粒子群优化算法

PSO 算法是一种模拟种群社会行为的智能优化算法,易于实现且收敛速度快。PSO 的思想源于鸟群觅食行为研究,鸟群在捕食过程中,通过个体间的信息交互快速传递食物位置信息。PSO 将鸟群表示为粒子群形态,其中每个粒子具有速度与位置两个属性,分别表示粒子变化位置的距离大小与当前粒子位置信息。适应度函数是 PSO 中最重要的组成部分,其本质为粒子的优化目标,多以最小化形式出现。通过计算适应度函数取值,判断粒子的优劣程度。最后,各粒子通过速度更新公式与位置更新公式对目标进行迭代寻优。

传统 PSO 算法的数学描述如下:在 D 维搜索空间中,假设 t 时刻粒子 $i(i=1,2,\dots,n)$ 的各维度速度信息 $V_i^t=[v_{i1}^t, v_{i2}^t, \dots, v_{id}^t]^T$ 、各维度位置信息 $X_i^t=[x_{i1}^t, x_{i2}^t, \dots, x_{id}^t]^T$ 、个体历史最优位置信息 $p_i^t=[p_{i1}^t, p_{i2}^t, \dots, p_{id}^t]^T$ 和种群的全局最优位置信息 $p_g^t=[p_{g1}^t, p_{g2}^t, \dots, p_{gd}^t]^T$, 则粒子在 t 时刻的速度更新公式和位置更新公式如下所示:

$$v_{id}^{t+1} = wv_{id}^t + c_1r_1^t(p_{id}^t - x_{id}^t) + c_2r_2^t(p_g^t - x_{id}^t), \quad (3.1)$$

$$x_{id}^{t+1} = x_{id}^t + v_{id}^{t+1}, \quad (3.2)$$

其中, w 表示惯性权重,它控制粒子对前一时刻速度继承的程度,正常数 c_1 和 c_2 表示

学习因子, 其中 c_1 控制粒子向个体历史最优位置学习的步长, 而 c_2 则控制粒子向全局最优位置学习的步长, r_1 和 r_2 分别为在 $[0,1]$ 范围内服从均匀分布的两个随机数。一般的, 将速度和位置的取值范围分别设置为 $[-V_{\max}, V_{\max}]$ 和 $[-X_{\max}, X_{\max}]$ 。

速度更新公式(3.1)包括三个部分: 第一部分称为动量部分, 是粒子的惯性速度信息, 表示粒的初始动量, 使得粒子继续按照自身原有速度进行飞行; 第二部分称为个体认知部分, 是粒子与个体历史最优位置之间的距离信息, 表示粒子对于自身的认知, 被称为; 第三部分称为社会认知部分, 是粒子与全局最优粒子位置之间的距离信息, 表示粒子之间的合作与共享。这三个部分相互平衡与制约, 共同影响着粒子的飞行方向, 也决定了粒子群优化算法的性能。各粒子通过公式(3.2)来计算位置更新的坐标。

传统的粒子群算法具有收敛速度快(尤其是在算法的早期过程)的特点, 但也正因为算法的快速收敛性使其容易陷入局部最优, 产生早熟收敛现象, 导致收敛精度较低。

3.2 自适应粒子群优化算法

传统 PSO 具有原理简单、易于实现、收敛速度快、可调参数少、具有较强的寻优能力等独特优势, 已广泛应用于多个领域。然而传统的粒子群优化算法容易陷入局部最优, 出现早熟收敛现象, 导致优化结果精度不高。因此, 我们基于收敛速度快特点, 在传统 PSO 的基础上进行改进, 提出了一种具有自适应学习策略的粒子群优化算法(A-PSO), 致力于在保证算法收敛速度的同时获得较高的求解精度。算法根据粒子自身的分布情况, 首先利用一种聚类算法自适应地将整个种群划分成若干子群, 然后根据每个子群中粒子类型的不同, 采用不同的学习策略进行更新, 以提高种群多样。

3.2.1 自适应种群划分

PSO 在迭代寻优的过程中, 由于粒子更新公式仅与自身历史最优位置和全局最优粒子位置进行信息互换, 不可避免的降低了种群多样性, 造成算法陷入局部最优。多种群技术是一种通过划分子群操作增强种群多样性的方法, 并被广泛地应用于智能优化算法领域。为了提高子群划分的准确性, 我们受粒子具有自然聚集特性的启发, 提出了一种自适应的种群划分方法, 主要基于一种高性能的快速聚类方法^[55]实现。该聚

类方法能够自适应地确定子群个数，并能实现对任意子群形状的识别。每个类簇中心都具有如下两个特征：

一方面，簇类中心粒子被其他粒子紧密包围，即，簇类中心具有较高的局部密度 ρ_i ，定义如下：

$$\rho_i = \sum_{j \neq i} \exp(-(\frac{d_{ij}^2}{d_c})), \quad (3.3)$$

另一方面，类簇中心粒子远离其他类簇的中心粒子，即，簇类中心粒子之间具有较大的距离 δ_i ，其表达式如下所示：

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}), \quad (3.4)$$

对于局部密度最大的粒子， $\delta_i = \max_j d_{ij}$ ，即，局部密度最大的类簇中心粒子具有邻近粒子中最大的距离。基于式(3.3)和式(3.4)可知，簇类中心为 ρ 值与 δ 值均比较大的粒子。对于剩余的粒子，将其分配到与其最近的且具有高密度的粒子所在的子群，得出子群划分结果。

3.2.2 不同种群间学习策略

经过自适应地子群的划分操作，整个粒子群被分为若干个子群，与此同时每个粒子的类型也得到化分，分为普通粒子与局部最优粒子。

首先，对于普通粒子，由局部最优粒子引导社会认知部分的更新，即：

$$x_i^d = \omega x_i^d + c_1 r_1^d (pbest_i^d - x_i^d) + c_2 r_2^d (cgbest_c^d - x_i^d), \quad (3.5)$$

其中， w 是惯性权重， c_1 和 c_2 表示学习因子， r_1^d 和 r_2^d 分别为在 $[0,1]$ 区间上均匀分布的随机数， $pbest_i^d$ 表示粒子的个体历史最优位置信息， $cgbest_c^d$ 表示第 c 个子群中局部最优粒子的位置信息。

对于局部最优粒子，由所有子群中局部最优粒子的综合信息引导社会认知部分的更新，即：

$$x_i^d = \omega x_i^d + c_1 r_1^d (pbest_i^d - x_i^d) + c_2 r_2^d (\frac{1}{C} \sum_{c=1}^C cgbest_c^d - x_i^d), \quad (3.6)$$

其中， C 是子群的个数。

传统 PSO 算法的学习策略中，粒子仅依靠全局最优粒子指引搜索方向容易造成种

群多样性的缺失。采用局部最优粒子引导更新后,每个子群都能学习到不同的搜索信息,从而提高了种群的多样性。局部最优粒子除了受自身历史最优位置的影响,还可以接收到其他子群的位置信息,从而进一步增加了种群多样性,在一定程度上解决了PSO容易陷入局部最优解而无法跳出的问题。

3.2.3 子群间信息交互

第一,增强子群间的合作能够进一步提高种群多样性。在人类社会中,每个群体都有自己的优势和劣势,并通过交换外部资源来获取利益。受人类互动学习行为的启发,在学习策略中考虑了多个子群之间的信息交换。在这种情况下,可以学习更多的搜索信息来增加多样性。其次可以看到,我们对普通粒子和局部最优粒子设计了不同的学习策略,这主要是考虑到普通粒子的学习策略不能适用于局部最优粒子。对于普通粒子的学习策略,只使用子群内的信息会使粒子在局部搜索中过于集中。随着演化过程的进行,粒子越来越集中,造成多样性降低,使得全局搜索能力变弱。局部最优粒子与普通粒子不同的是,局部最优粒子指引着子群中所有粒子的位置变化。假设局部最优粒子的更新方式与普通粒子相同,一旦发生早熟收敛,局部最优粒子将会带领整个子群陷入局部最优解。因此,对局部最优粒子的学习策略进行了不同的设计,在提高种群多样性的同时,提高全局搜索能力。第三,考虑到局部最优粒子是最有可能找到最优解的粒子,这些粒子具有相对较多的最优搜索信息。因此,采用所有子群中 $cgBest_c$ 的综合信息来指导粒子的更新。

3.2.4 种群多样性与收敛速度权衡

在保持高多样性的同时提高收敛速度是一个权衡问题。由式(3.5)和式(3.6)可以看出,与式(3.1)和式(3.2)所示的传统PSO算法不同,我们的学习策略只使用位置项描述粒子。这是由于提出的学习策略虽然具有较高的多样性,但频繁的子群划分操作导致全局搜索能力不足。为了解决这一问题,受变量替换可以简化问题且降低计算复杂度的启发,我们对速度项和位置项合并,将学习策略设计成一种不含显式速度项的形式来提高收敛速度。

3.3 实验结果与分析

本节中,我们为了评估所提出的自适应粒子群算法在求解精度和收敛速度上的性

能,进行数值模拟实验。该算法在 6 个常用的基准函数上进行验证,并与典型的、最新的粒子群算法进行了比较。

表 3-1 给出了实验中采用的基准函数,这些函数可分为 3 类:单模态函数 F_1 ^[56]、多模态函数 $F_2 - F_4$ ^[56]以及移位和旋转函数 $F_5 - F_6$ ^[57]。单模态函数只含有一个全局最优解,考虑到现有的很多粒子群算法以收敛速度慢为代价保持种群多样性,因此借助单模态函数可以测试算法能否保持较快的收敛速度。多模态函数存在多个局部最优值,常用于检验算法在避免早熟收敛时的全局搜索能力。移位和旋转函数是在原始基准函数的基础上进行了移位或旋转操作,即,移动基准函数的全局最优值至新的位置,或是对基准函数通过正交矩阵旋转。因此,移位和旋转函数属于更复杂的优化问题,借助这些函数可以更系统地比较算法性能。

表 3-1 基准函数

基准函数		维数	搜索空间	最优解	可接受值
F ₁	$F_1(x)=\sum_{i=1}^D x_i +\prod_{i=1}^D x_i $	30	$[-10,10]^D$	0	0.01
F ₂	$F_2(x)=-20exp(-0.2\sqrt{1/D\sum_{i=1}^Dx_i^2})-exp(1/D\sum_{i=1}^Dcos2\pi x_i)+20+e$	30	$[-32,32]^D$	0	0.01
F ₃	$F_3(x)=1/4000\sum_{i=1}^Dx_i^2-\prod_{i=1}^Dcos(x_i/\sqrt{i})+1$	30	$[-600,600]^D$	0	0.01
F ₄	$F_4(x)=\sum_{i=1}^D[x_i^2-10cos(2\pi x_i)+10]$	30	$[-5.12,5.12]^D$	0	50
F ₅	$F_5(x)=\sum_{i=1}^D(\sum_{j=1}^iz_j)^2-450,z=x-o$	30	$[-100,100]^D$	-450	100
F ₆	$F_6(x)=\sum_{i=1}^D[z_i^2-10cos(2\pi z_i)+10]-330,$ $z=(x-o)\times M$	30	$[-5,5]^D$	-330	200

选取 3 种具有代表性的 PSO 算法与本文提出的 A-PSO 算法进行对比试验,包括如式(3.1)和式(3.2)所示的传统 PSO 算法,综合学习 PSO 算法(Comprehensive Learning PSO, CLPSO)^[58]与集成 PSO 算法(Ensemble PSO, EPSO)^[59]。其中,CLPSO 算法是一种经典的对学习策略改进的 PSO 算法,该算法利用其它所有粒子的历史最优信息对速度项更新,通过从邻域中学习更多知识增强种群的多样性,在解决多模态问题上具有良好性能。EPSO 算法是近年提出的一种集成算法,将若干表现较好的 PSO 算法集成用于解决复杂优化问题,综合能力较强。

实验中 A-PSO 算法的参数设置如下:根据经典 PSO 算法中的常用参数设置^[59,60],将惯性权重 w 设为从 0.9 到 0.4 线性减少的形式,加速因子 $c_1 = c_2 = 2.0$ 。3 种对比算

法中的参数设置与原文献保持一致。种群规模选用 20 个粒子，每个基准函数的迭代停止条件为 2.0×10^5 个评估函数^[57]。所有算法通过 MATLAB 2014b 运行，最大运算精度为 1E-323，即当值小于 1E-323 时，结果视为 0。每个算法取 30 次独立实验的平均值作为实验结果。此外，为了从统计意义上验证算法性能，实验中采用显著性水平 $\alpha = 0.05$ 的 Wilcoxon 秩和检验^[61]评估本文算法与其他算法的结果差异性。

表 3-2 给出 4 种算法在 6 个基准函数上的求解精度比较结果，包括平均误差值、误差值的标准差和 Wilcoxon 秩和检验的 P 值，这里的误差值指算法得出的最优解与全局最优解之间的误差。从表中可以看出，与其他算法相比，在所有的单模态函数和多模态函数上，本文提出的 A-PSO 算法均取得了最高的求解精度，特别地，本文算法找到了这 4 个函数的全局最优解，说明本文算法具有较好的全局搜索能力，特别是对于多模态问题，能够很好地避免局部最优。这主要源于提出的自适应种群划分策略能够自适应地确定种群规模，为学习策略多样性的发挥提供了精确的先决条件。另外，函数 F_4 (Rastrigin 函数) 是一种具有大量局部最优值干扰的复杂多模态问题，要求算法具有较高的多样性才能保证较好的求解精度。本文算法在该函数上的表现明显优于其他算法，这说明本文提出的学习策略能够很好地保持种群多样性。对于移位和旋转函数，A-PSO 算法也具有较好的表现。EPSO 算法是一种对很多性能较好的算法的集成，对于复杂优化问题具有较好的性能。因此在函数 F_5 上，EPSO 算法的求解精度要略高，但 A-PSO 算法在 6 个基准函数中的 5 个函数上求解精度最高，综合能力更强。此外，Wilcoxon 秩和检验的结果也能很好地验证 A-PSO 算法的性能。在显著性水平为 0.05 的情况下，当 P 值小于 0.05 时，认为 A-PSO 算法与其他算法的求解结果差异显著。从表 3-2 中的 P 值结果可以看出，A-PSO 算法在求解精度方面明显优于其他算法。

表 3-3 给出了 4 种算法在 6 个基准函数上的收敛速度比较结果，包括找到可接受解所需的评估函数个数，30 次独立实验中算法成功找到可接受解的概率和算法收敛速度性能的排名，这里的排名根据成功率降序和评估函数个数升序的规则给出^[62,63]。

“Inf”表示在给定的迭代停止条件下，算法没有找到可接受解。从表 3-3 的实验结果可以看出本文提出的 A-PSO 算法在 6 个基准函数上都取得了最快的收敛速度，且表现明显优于其他算法。

表 3-2 基准函数求解精度统计结果对比

基准函数	评价指标	PSO	CLPSO	EPSO	A-PSO
F ₁	平均误差	2.51E-34	1.01E-13	9.16E-27	0
	误差标准差	5.84E-34	6.51E-14	8.31E-27	0
	p 值	0.00000	0.00000	0.00000	
F ₂	平均误差	1.15E-14	2.01E-12	1.78E-14	0
	误差标准差	2.27E-15	9.22E-13	8.42E-15	0
	p 值	0.00000	0.00000	0.00000	
F ₃	平均误差	2.37E-02	6.45E-13	4.43E-03	0
	误差标准差	2.57E-02	2.07E-12	8.86E-03	0
	p 值	0.00000	0.00000	0.00000	
F ₄	平均误差	3.07E+01	2.57E-11	0	0
	误差标准差	8.68E+00	6.64E-11	0	0
	p 值	0.00000	0.00000	0.00000	
F ₅	平均误差	7.30E-02	2.98E+02	2.77E-11	5.02E-11
	误差标准差	7.56E-02	9.97E+02	3.37E-11	9.58E-13
	p 值	0.00000	0.00000	1.00000	
F ₆	平均误差	1.21E+02	1.00E+02	8.49E+01	5.11E+01
	误差标准差	5.58E+01	1.54E+01	9.19E+00	1.72E+01
	p 值	0.00021	0.00000	0.00000	

表 3-3 基准函数收敛速度统计结果对比

基准函数	评价指标	PSO	CLPSO	EPSO	A-PSO
F ₁	评估函数个数	103,077	66,525	14,635	1,552
	成功率	100	100	100	100
	排名	4	3	2	1
F ₂	评估函数个数	110,844	76,646	15,584	1,328
	成功率	100	100	100	100
	排名	4	3	2	1
F ₃	评估函数个数	99,541	81,422	12,577	1,360
	成功率	90	100	80	100
	排名	3	2	4	1
F ₄	评估函数个数	94,379	53,416	13,108	1,855
	成功率	96	100	100	100
	排名	4	3	2	1
F ₅	评估函数个数	143,175	Inf	24,064	16,692
	成功率	100	0	100	100
	排序名	3	4	3	1
F ₆	评估函数个数	93,384	23,861	18,859	13,785
	成功率	20	100	83	100
	排名	4	2	3	1
综合排名		4	3	2	1

这说明学习策略的简化能够有效加快算法的收敛,另一方面,学习策略多样性的提高增强了算法的全局搜索能力,也能够促进算法的尽快收敛。此外,从成功率的结果来看,A-PSO 算法在所有函数上均实现了 100%的成功率,说明本文算法具有较高的稳定性。A-PSO 算法的快收敛特性也能从图 3-1 所示的收敛曲线中看出。与其他 PSO 算相比,A-PSO 算法的收敛曲线下降趋势明显,能够在较少的评估函数个数下找到较优的解。综上实验结果表明,本文算法在求解精度和收敛速度方面都有很好的性能,能够有效地避免局部最优。

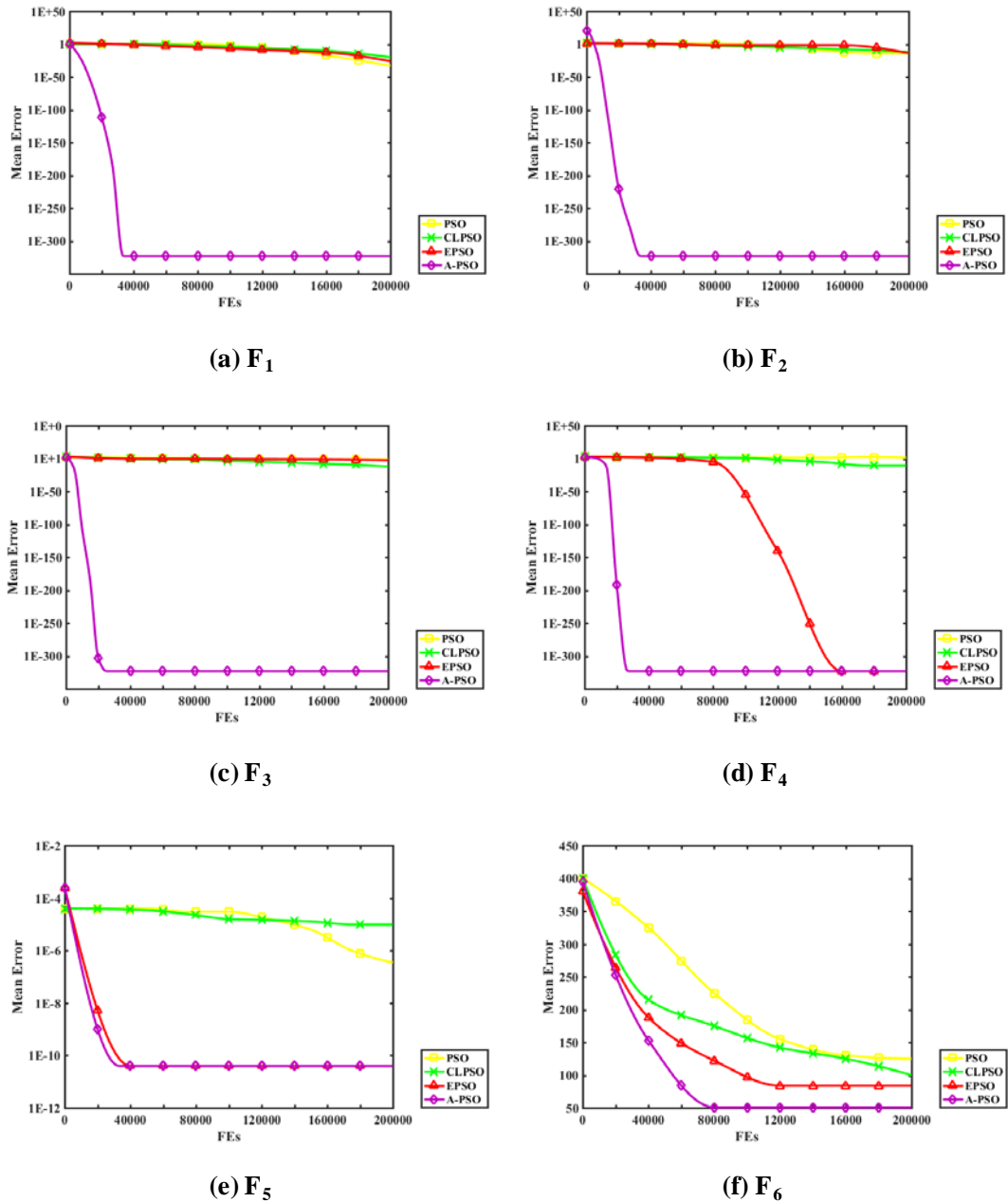


图 3-1 基准函数收敛曲线图对比

3.4 本章小结

为了设计一种收敛速度快、求解精度高的智能化参数寻优算法,本章基于传统的粒子群优化算法,提出一种具有自适应学习策略的粒子群优化算法。该算法在进化过程中,能够根据粒子的分布特性自适应地确定粒子群的大小。与传统的基于固定或参数辅助种群规模的学习策略相比,自适应子群划分成功地结合了粒子的自然演化规律,提高了学习策略的智能性。数值模拟实验表明,提出的算法在收敛速度和求解精度方面具有一定的优势。

第4章 基于自适应粒子群优化 LSTM 的股票预测模型

股票价格时间序列是一种特殊的金融时间序列。众所周知，金融时间序列具有高噪声、非线性、复杂周期性、不规则性等特征。因此，传统的时间序列模型很难对股票价格变化进行精确的刻画。LSTM 网络是处理时间序列最为出色的模型之一，本章在充分发挥 LSTM 网络的技术优势的基础上，对其进行深度优化和改进，将兼具收敛速度和求解精度的 A-PSO 算法与 LSTM 网络深度融合，提出 PSO-LSTM 股票价格预测模型，致力于获得更高的模型预测精度。本章分为三个部分，包括超参数调优过程，PSO-LSTM 网络构建过程，以及模型在真实股票数据中的预测实验。

4.1 超参数优化过程

与其他的神经网络模型类似，LSTM 网络中存在各式各样的结构参数，其中主要包括：隐藏层中神经元数目大小，时间窗口大小，批处理数据量大小等。关键的网络结构参数能够影响模型的内部结构，并与模型的训练过程具有密不可分的联系。某些参数的选取，可直接控制网络模型的性能。例如，时间滑动窗口大小设置不合适，就会破坏掉时间序列数据间隐藏的周期性关联，影响模型预测精度。再者，批处理数据量大小的设置，如果批处理数据量设置过大，则会导致不同批次之间梯度变化减小，使得模型容易陷入局部最优；如果批处理数据量设置过小，则会导致梯度振荡严重，耗费大量训练时间，并且造成模型的不易收敛。因此，对于模型关键结构参数的选取对于模型的建立过程就显得格外的重要。就目前来看，在基于 LSTM 网络的股票预测模型中，对网络结构参数最优选取的研究并不广泛。在实际的模型训练与检测过程中，这些参数绝大多数就取决于人为经验选取和大量重复性计算工作。这无疑浪费了大量的人力与计算资源，因此，我们将前一章节所提出的 A-PSO 优化算法与 LSTM 网络相融合，试图快速、有效的对相关网络结构参数进行寻优，提高模型的预测效率和预测准确性。

网络结构超参数可以表示为一种数值向量表达形式，如 $P=[p_1, p_2, p_3, p_4]$ ；其中 p_1 表示第一层隐藏层中神经元数目， p_2 表示第二层隐藏层中神经元数目， p_3 表示时间窗口大小， p_4 表示批处理数据量大小。将网络结构超参数 P 作为 A-PSO 优化算法

的优化对象，以适应度函数 f 作为算法优化目标，其表达式如下所示：

$$f = \frac{1}{K} \sum_{i=1}^K \frac{|\hat{y}_i - y_i|}{y_i}, \quad (4.1)$$

其中 K 表示数据集中的样本数据总量， \hat{y}_i 表示第 i 个样本数据的预测结果， y_i 表示为第 i 个样本数据的真实信息。该适应度函数可表示为百分比绝对误差函数，代表了预测结果与真实信息间的差异情况。

超参数优化过程首先进行初始化操作，包括初始化粒子数量，设置的粒子搜索范围，以及初始化粒子位置信息。其次，设置迭代停止条件。最后，利用 A-PSO 算法对超参数进行迭代寻优。

4.2 PSO-LSTM 模型构建过程

以自适应粒子群算法指导的超参数寻优过程与 LSTM 网络进行深度融合，构建了 PSO-LSTM 股票价格预测模型。该模型充分发挥了 LSTM 网络善于处理时间序列数据的优势，并结合 A-PSO 优化算法的收敛速度快、求解精度高的优点，对 LSTM 网络结构参数进行深度优化。PSO-LSTM 网络架构如图 4-1 所示：

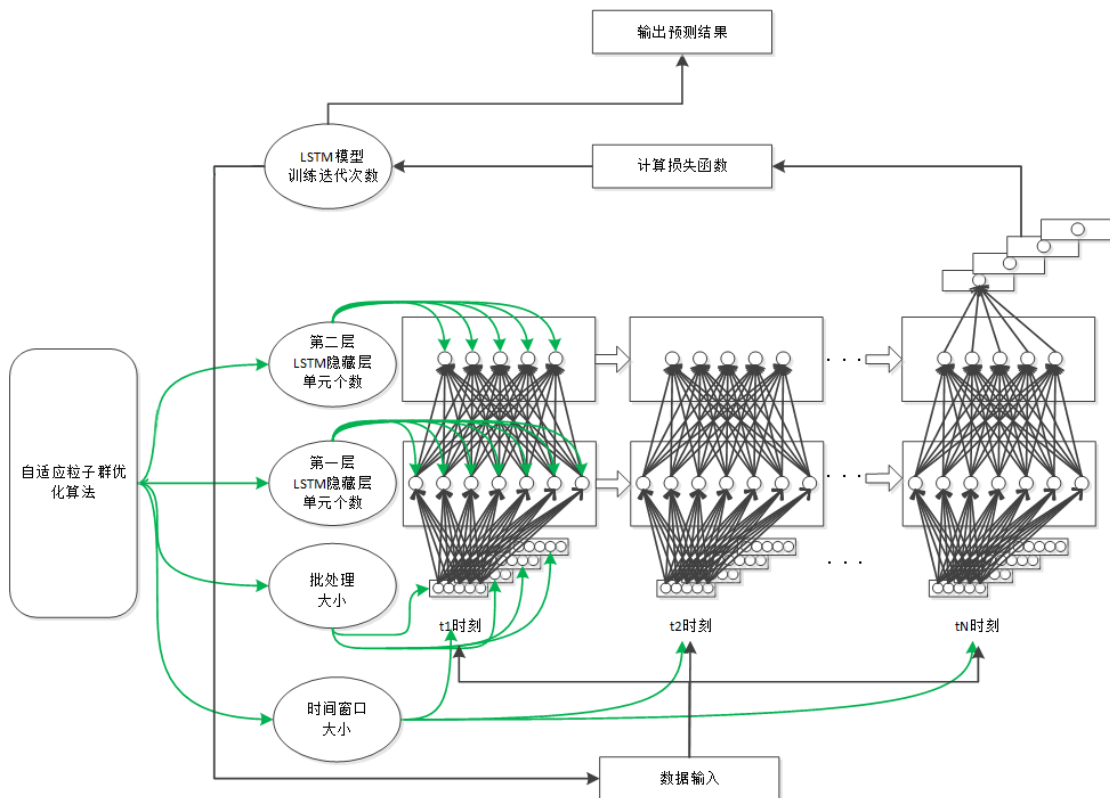


图 4-1 PSO-LSTM 网络架构图

A-PSO 优化算法与 LSTM 网络深度融合为一体,对网络结构超参数进行共同优化,PSO-LSTM 股票价格预测模型的算法流程如下所示:

Step 1: 数据预处理阶段,将股票价格历史数据按比例划分为训练数据集、验证数据集与测试数据集。股票数据首先进行数据预处理操作,去除序列数据中的噪声干扰,并对于股票数据每个维度信息进行归一化。

Step2: 结合网络结构超参数特征,初始化 A-PSO 优化算法。自适应粒子群优化算法中所有的优化对象对应于各粒子的不同维度信息,对于不同的优化对象,设置合适的粒子搜索范围。A-PSO 优化算法在初始化阶段,每个粒子不同维度的信息在搜索范围内随机生成,并对粒子数量进行初始化设置。

Step3: 算法利用公式(3.3)、(3.4)分别计算局部密度 ρ 与距离 δ ,按照自适应种群划分策略,种群中的所有粒子完成子群归属划分与粒子类型划分,划分的子群数量由算法自适应选择。

Step4: 自适应子群划分完成之后,需要对每个粒子对应的适应度值进行计算,适应度值的大小表示粒子的优劣情况。适应度值的计算步骤如下:第一步,以粒子位置信息对应的网络结构参数取值情况,利用训练数据集建立 LSTM 网络。第二步,利用训练完成的网络模型对验证数据集预测。第三步,根据公式(4.1)计算模型在验证数据集上的 MAPE 误差,得到该粒子适应度值。

Step5: 根据每个粒子适应度值的优劣情况:适应度值大小,粒子的子群归属情况:粒子属于哪个子群,以及类型归属情况:粒子是否为子群最优粒子,确定粒子历史最优位置信息 $pbest$,确定子群最优粒子最优位置信息 $cgbest_c$ 。

Step6: 不同类别的粒子采取不同的位置更新方式,普通例子根据公式(3.5)更新位置信息,子群最优粒子根据公式(3.6)更新位置信息。普通粒子仅受到自身历史最优位置与所属子群中最优粒子位置的指引;子群最优粒子除了受到自身历史最优位置的指引,还能够与其他子群最优粒子进行信息交流。这是保证粒子避免陷入局部最优解、加速收敛的关键。

Step7: 判断 A-PSO 优化算法终止条件,如果达到终止条件,将此时最优粒子对应的超参数取值情况传递到 Step8;否则,算法流程返回至 Step3,对超参数寻优过程进行迭代。

Step8: 根据网络结构超参数取值情况构建 LSTM 股票价格预测模型。

Step9: 利用训练数据集和验证数据集共同训练 LSTM 股票价格预测模型, 通过测试数据集进行预测, 获得股票价格预测结果。

4.3 实验结果及分析

为了验证本文提出的 PSO-LSTM 股票价格预测模型的预测性能, 本节选取了沪市、深市、港股的个股真实股票交易数据进行实验, 实验结果与其他 5 种时间预测模型进行了比较分析。对比方法包括了经典的统计模型 ARIMA^[64], 性能表现良好的机器学习模型支持向量机 (Support Vector Machine, SVM)^[65], 以及神经网络模型 MLP^[66]、RNN^[67]与 LSTM^[68], 上述对比实验方法已在上文中进行了详细说明。本节将从数据预处理, 误差评价指标, 模型参数设置, 实验结果分析方面进行详细说明。

4.3.1 数据预处理

实验数据均从同花顺交易软件中获取, 同花顺是一款为广大股票爱好者提供实时股票价格信息并支持股票买入卖出的交易软件, 其反应的股票数据具有真实可靠性。另外, 国内个股股票价格变化受到诸多因素的影响, 因此我们在不同的股票交易市场选取了公司规模较大, 流通性较好的三支股票价格历史数据进行研究, 以减小外部环境因素的影响。三支个股股票分别为沪市 A 股浦发银行 (股票代码 600000), 深市 A 股五粮液 (股票代码 000858) 和港股恒隆集团 (股票代码)。选取了股票数据十个属性进行分析, 包括了开盘价格、当日最高价格、当日最低价格、收盘价格、涨幅比例、振幅比例、当日成交量、当日成交额、换手率和与日成交次数。

在实验之前需要对股票价格数据进行预处理。由于在股票价格数据中不同属性之间的数量级差异巨大, 例如股票收盘价格的数量级要远小于当日成交量, 涨幅比例与股票收盘价格也不在同一数量级中。单种属性数据数量级过大或不同属性数据存在数量级差异都会导致模型无法收敛。因此, 需要对实验数据的不同属性数据进行统一量纲的操作, 常用的方法为进行归一化处理。归一化处理将数据控制在[0,1]范围内, 可避免网络模型梯度爆炸, 提升模型训练速度。本文选取归一化函数如下:

$$x_i = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)}$$

其中, x_i 为归一化处理后的数据, X_i 为原始样本数据, $\min(X_i)$ 与 $\max(X_i)$ 分别表示同一纬度中样本数据的最小值与最大值。对于预测模型的输出, 为了保证与原始数据

数量级一致，还需要进行反归一化处理，反归一化函数如下：

$$Y_i = y_i \times (\max(X_i) - \min(X_i)) + \min(X_i)$$

其中， y_i 为模型预测初始数据， Y_i 为经过反归一化处理最终输出数据。

4.3.2 误差评价指标

为了衡量不同预测模型的预测性能，对预测的准确性进行数值化描述，实验设置了 5 个误差评价指标，其中包括：RMSE、MAE、MAPE、MSE 以及 R^2 。不同的误差评价标准采用不同的计算方法，以不同的分析层面来反映预测值与真实值之间的差异程度，具体的计算公式如下所示：

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2},$$

$$MAE = \frac{1}{N} \sum_{n=1}^N |\hat{y}_n - y_n|,$$

$$MAPE = \frac{1}{N} \sum_{n=1}^N \frac{|\hat{y}_n - y_n|}{y_n},$$

$$MSE = \frac{1}{N} \sum_{n=1}^N (\hat{y}_n - y_n)^2,$$

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{n=1}^N (\hat{y}_n - \bar{y})^2}{\sum_{n=1}^N (y_n - \bar{y})^2},$$

其中， N 表示实验数据的样本量大小； \hat{y}_n 表示经过模型测试获得的预测值， y_n 表示实验数据的真实值， \bar{y} 为实验数据真实值的平均值。特别的，在 R^2 评价指标中， SSR 表示回归平方和，反映预测值与真实值之间的偏差平方和； SST 表示总离差平方和，反映真实值与平均值之间的偏离程度； R^2 评价指标的取值范围在 $[0,1]$ 之间，取值越接近 1，表示模型误差越小，模型预测性能越好。其他的 4 个评价指标，取值越是接近 0，模型预测效果越好。实验根据这 5 类评价指标，共同衡量不同预测模型的预测性能。

4.3.3 模型参数设置

本文实验构建的 PSO-LSTM 股票价格预测模型是由双层的 LSTM 网络构成，模型的损失函数选取了均方误差 MSE，使用 ADAM 优化算法对梯度进行更新训练。该模型搭建在 Python3.6，Tensorflow 与 Keras 实验环境中，由两块 2080TI 计算显卡提供计算支撑。实验对 LSTM 网络结构超参数 P 进行优化，根据股票价格数据的特征，设置超参数的粒子搜索范围。其中，设置网络中隐藏层神经元数目 p_1 、 p_2 的取值范围[10,30]，设置时间窗口大小 p_3 的取值范围[1,20]，设置批处理数据量大小 p_4 的取值范围[1,60]。另外规定粒子群的粒子数量为 30，A-PSO 优化算法的停止条件设置最大迭代次数为 500。

4.3.4 实验结果分析

浦发银行股票、五粮液股票、恒隆集团股票分别对应了沪市个股、深市个股与港股个股。实验分别建立在三种不同类型股票价格数据中进行预测研究，目的是验证本文提出的模型的有效性以及对于不同实验数据的适用性。

(1) 浦发银行股票价格预测

选取浦发银行 2016 年至 2018 年共计 708 条股票交易数据，并将其按比例划分为训练数据集，验证数据集与测试数据集。通过训练数据集与验证数据集建立 PSO-LSTM 股票价格预测模型，经 A-PSO 算法优化得到的网络结构超参数 $P = [12, 22, 10, 58]$ 。基于训练好的 PSO-LSTM 模型预测浦发银行第二日收盘价。

图 4-2 展示了不同预测模型在浦发银行股票价格数据上的预测结果，其中绿色曲线代表股票价格真实值曲线，红色曲线表示预测值曲线。为了准确地对各模型之间的预测性能进行比较，不同模型的误差评价指标客观数据结果在表 4-1 中被给出。

表 4-1 模型误差评价指标

	RMSE	MAE	MAPE(%)	MSE	R^2
ARIMA	0.230	0.357	2.436	0.048	0.945
SVM	0.185	0.148	1.415	0.034	0.967
MLP	0.192	0.155	1.467	0.037	0.983
RNN	0.208	0.160	1.515	0.043	1.342
LSTM	0.180	0.145	1.379	0.032	1.152
PSO-LSTM	0.142	0.107	1.037	0.020	1.012

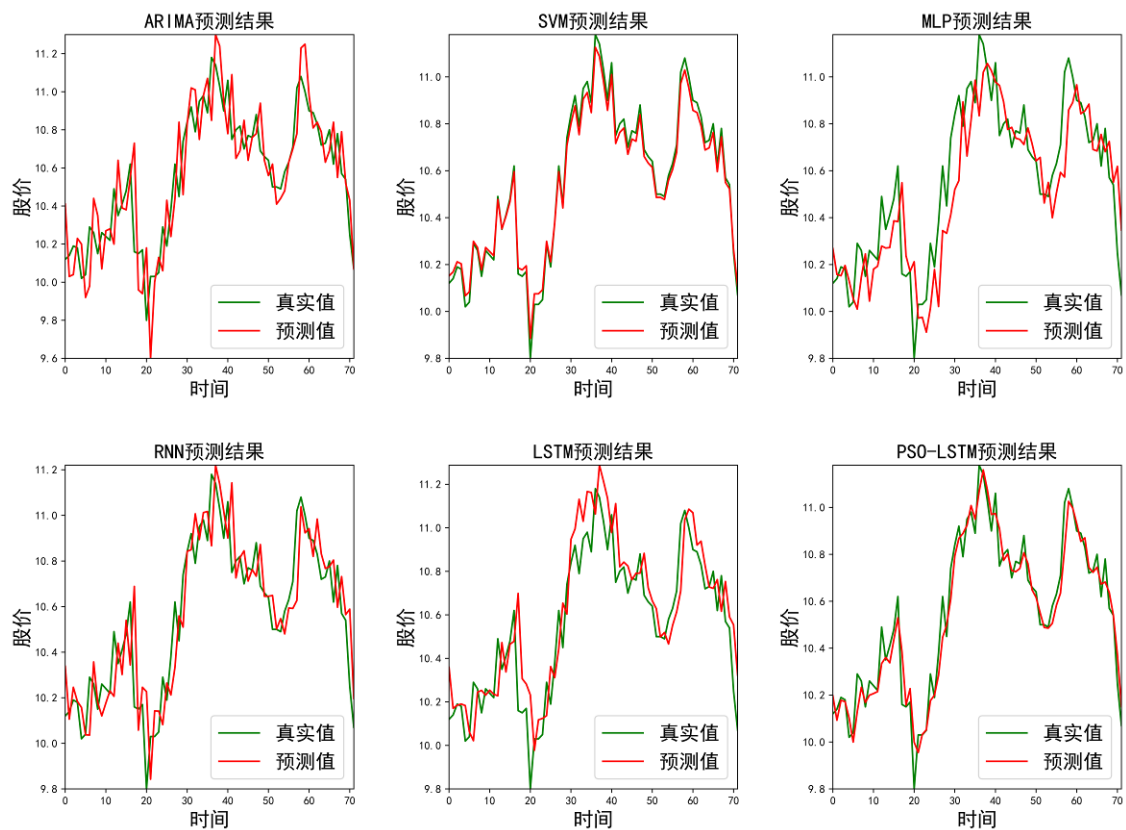


图 4-2 浦发银行股票价格预测结果

从视觉效果上来看，ARIMA 模型、MLP 模型与 RNN 模型的预测值曲线波动频繁，预测结果较差；SVM 模型、LSTM 模型与 PSO-LSTM 模型均能够较好的反映出股票价格的变化趋势，PSO-LSTM 模型在价格突变前后处的预测结果优与另外两种模型。PSO-LSTM 模型在不同的评价指标中均取得最优的结果，特别的，该模型较 ARIMA 模型的预测精度具有明显的优势。为进一步衡量该模型与 LSTM 模型的预测性能差异，实验增加了浦发银行 2019 年 1 月至 3 月的股票价格数据作为附加测试数据集，将其均匀分成五组测试样本，分别利用 LSTM 模型与 PSO-LSTM 模型对每组数据进行实验，预测结果如下。

如图 4-3 所示，绿色曲线代表股票价格真实值，红色曲线代表 PSO-LSTM 模型预测值，蓝色曲线代表 LSTM 模型预测值。表 4-2 给出两种模型的误差评价指标比较结果，在 RMSE 误差评价指标中，PSO-LSTM 模型在第二组、第三组、第四组、第五组的数据实验中，在模型预测精度方面较 LSTM 模型有了较大幅度的提升。

表 4-2 LSTM 与 PSO-LSTM 误差评价指标

RMSE	第一组	第二组	第三组	第四组	第五组
LSTM	0.107	0.106	0.144	0.163	0.153
PSO-LSTM	0.117	0.091	0.136	0.124	0.115

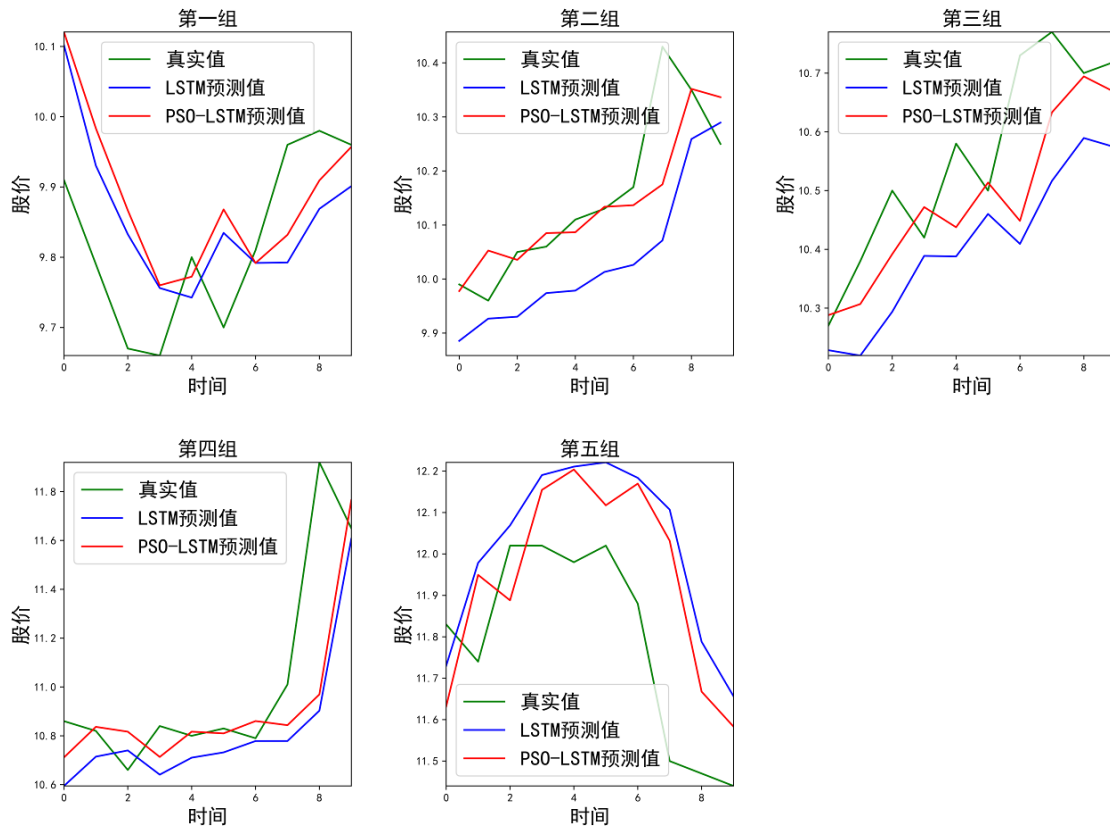


图 4-3 LSTM 与 PSO-LSTM 预测结果

实验结果显示,对于同一只股票不同时间段价格数据,LSTM 模型与 PSO-LSTM 模型均具有有效的预测能力。相较于 LSTM 模型预测结果,PSO-LSTM 模型在第二组、第三组与第五组数据中的预测效果具有明显的提升。综上所述,在浦发银行股票价格时间序列中,与 ARIMA、SVM、MLP、RNN 以及 LSTM 模型相比,PSO-LSTM 模型在预测精度与稳定性方面具有一定的优势。

(2) 五粮液股票价格预测

选取五粮液 2016 年至 2018 年共计 726 条股票交易数据,通过训练数据集与验证数据集建立 PSO-LSTM 股票价格预测模型,经 A-PSO 算法优化得到的网络结构超参数 $P=[16,32,12,60]$ 。基于训练好的 PSO-LSTM 模型预测五粮液第二日收盘价,所得结果如下。

如图 4-4 所示,对于五粮液股票价格的预测,ARIMA 模型在股票价格突变前后出的预测能力较差,这主要是因为 ARIMA 模型作为一种统计模型,预测值很大程度上依赖于前面时刻的历史数据信息,进行加权求和得到。因此,在股票价格持续上升或持续下跌时 ARIMA 模型预测值具有一定的参考价值,但在股票价格突变出的预测效果很差。SVM 模型在一段时间内的预测结果非常出色,但是在部分时间段的预测

值与真实值之间的差异较大,说明模型的稳定性不足。MLP 模型、RNN 模型与 LSTM 模型的预测结果能够大致反映出股票价格变化的趋势,但是预测结果对比真实值曲线具有不同程度的滞后性。PSO-LSTM 模型较其他模型相比能够取得出色的预测结果。如表 4-3 所示,该模型在数值化的模型评价标准中均取得良好的结果,说明 PSO-LSTM 模型在五粮液股票价格数据上具有最优异的预测能力。

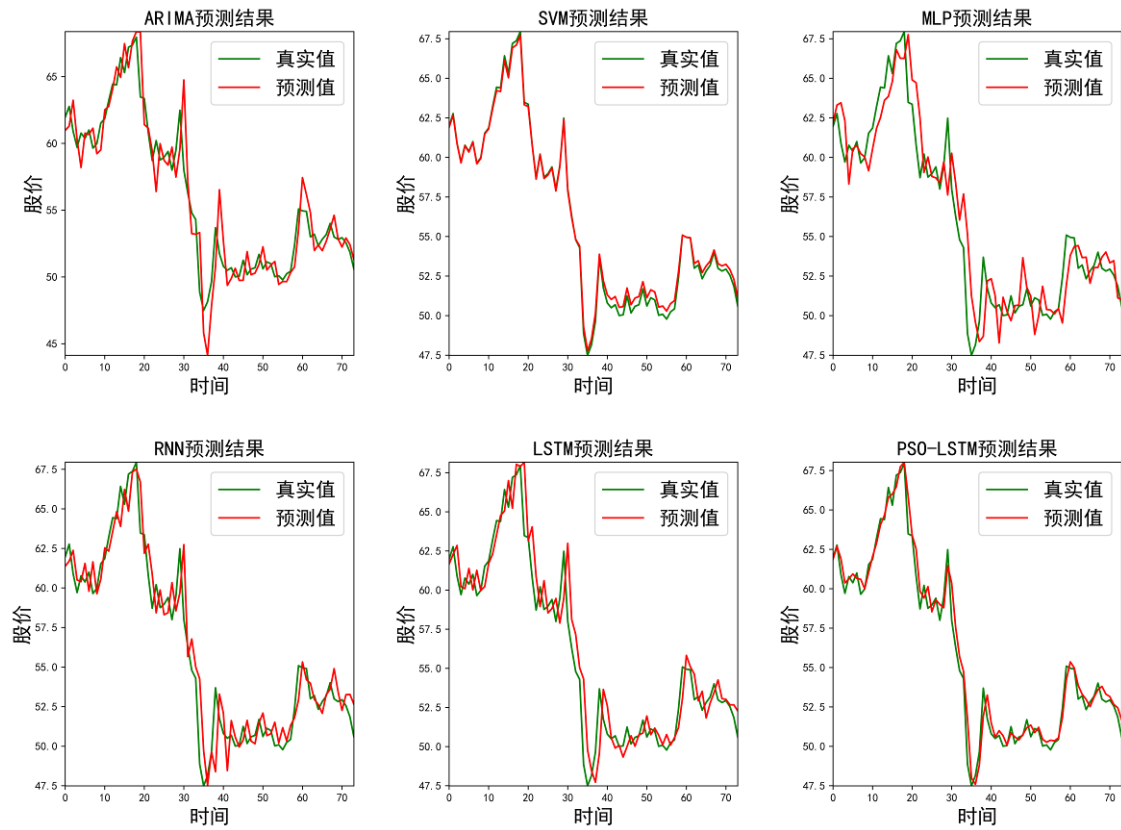


图 4-4 五粮液股票价格预测结果

表 4-3 模型误差评价指标

	RMSE	MAE	MAPE(%)	MSE	R^2
ARIMA	1.781	6.273	11.175	3.285	1.018
SVM	1.690	1.376	2.487	2.855	0.773
MLP	1.889	1.388	2.683	3.488	1.061
RNN	1.655	1.325	2.348	2.720	1.034
LSTM	1.601	1.192	2.125	2.573	1.057
PSO-LSTM	1.343	1.034	1.781	1.803	1.012

(3) 恒隆集团股票价格预测

选取恒隆集团 2016 年至 2018 年共计 733 条股票交易数据,通过训练数据集与验证数据集建立 PSO-LSTM 股票价格预测模型,经 A-PSO 算法优化得到的网络结构超

参数 $P=[16,32,12,60]$ 。基于训练好的 PSO-LSTM 模型预测恒隆集团第二日收盘价，所得结果如下所示：

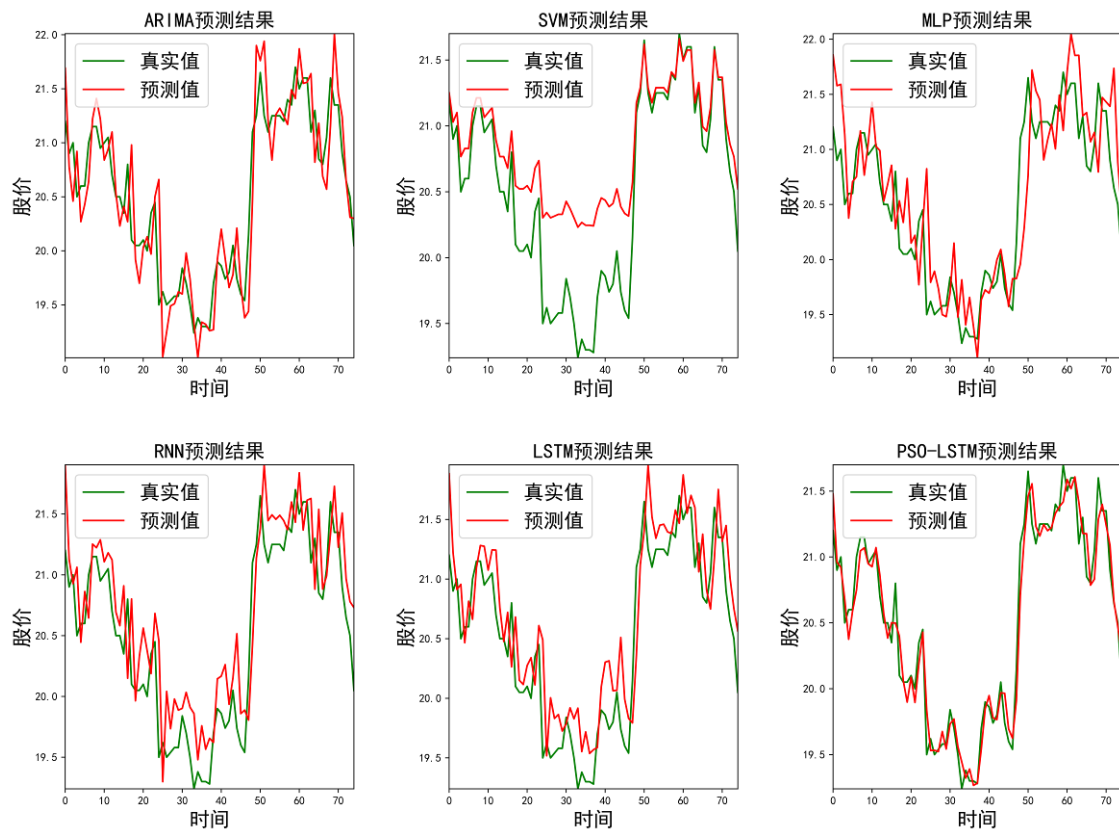


图 4-5 恒隆集团股票价格预测结果

表 4-4 模型误差评价指标

	RMSE	MAE	MAPE(%)	MSE	R^2
ARIMA	0.358	0.892	4.363	0.128	1.239
SVM	0.522	0.425	2.120	0.272	0.580
MLP	0.344	0.253	1.230	0.118	0.970
RNN	0.384	0.324	1.590	0.148	0.957
LSTM	0.365	0.303	1.483	0.133	0.957
PSO-LSTM	0.285	0.237	1.047	0.081	1.004

如图 4-5 所示，ARIMA 模型的预测曲线在价格突变处的预测值与真实值之间存在很大的差异，预测结果不理想。可以明显观察到，在很长一部分的测试时间内 SVM 模型的预测结果失去了指导意义，无法反映真实股票价格情况。同为神经网络的 MLP 模型、RNN 模型与 LSTM 模型预测结果很相似，LSTM 模型的预测结果较优于另外两个模型，但也存在滞后性和预测误差较大的问题。PSO-LSTM 模型的预测结果较其他 5 种预测模型更准确，但是仍可以观察到，在股票价格真实曲线的波峰、波谷处的预测值与真实值之间有差距。表 4-4 给出各模型的误差评价指标比较，说明了本文提

出的 PSO-LSTM 股票价格预测模型取得了最高的预测精度,在稳定性和预测精度方面展现出具有竞争力的优势。

4.4 本章小结

本节详细介绍了 PSO-LSTM 股票价格预测模型的构建过程,以及实验环境配置情况。通过分别对沪市、深市、港股中的个股股票价格进行预测实验,验证了提出模型的有效性,展现出该模型具有较好的预测精度与稳定性。对于不同类型的股票数据,本文模型都表现出优秀的预测能力,表明该模型还具有一定的普遍适用性。

第5章 总结与展望

为了更准确地预测复杂的股票价格时间序列,本文对国内外研究人员的相关研究进展进行了总结和说明,并介绍了常见的时间序列预测模型相关理论。在目前处理时间序列表现最出色的 LSTM 网络的基础上,设计了智能化参数寻优算法,与 LSTM 网络进行深度融合提出了 PSO-LSTM 股票价格预测模型。与统计模型 ARIMA、机器学习模型 SVM 以及神经网络模型 MLP、RNN、LSTM 相比,该模型具有更高的预测精度,更好的稳定性,并具有一定的普遍适用性,在金融时间序列的研究领域具有很好的应用价值。

5.1 工作总结

本文研究了基于 LSTM 网络的股票价格预测,针对现有时间序列预测模型的不足,设计了收敛速度快、求解精度高的智能化参数寻优算法,充分发挥 LSTM 网络的优势,建立了一种基于自适应粒子群优化的 LSTM 股票价格预测模型,实验结果验证了模型的有效性和适用性。本文的主要工作体现在以下方面:

1、自适应学习策略的粒子群优化算法

为了增加粒子种群多样性,提高了目标优化的准确性,本文提出一种具有自适应学习策略的粒子群优化算法。通过快速搜索聚类,整个粒子群被自适应地划分成若干子群,并自动完成所有粒子的种群归属与类别归属任务。对于不同类别粒子利用不同的学习策略,实现了位置信息更新寻优。实验结果显示,自适应学习策略的粒子群优化算法具有更快的收敛速度。数值模拟实验表明,与其他典型的 PSO 算法相比,该算法在求解精度和收敛速度上具有一定的优势。

2、基于自适应粒子群优化的 LSTM 股票预测模型

将自适应学习策略的粒子群优化算法与 LSTM 网络进行深度融合,使模型网络结构得到快速有效的优化。实验通过与相关预测模型进行比较分析,表明本文提出的 PSO-LSTM 股票价格预测模型具有更高的预测精度,对于不同类型的股票数据具有一定的普遍适用性。

5.2 未来工作展望

与传统的时间序列预测模型相比,基于深度学习模型对股票价格时间序列进行预

测研究具有更大的优势。就目前我们取得的研究进展上来看,不同模型对于股票价格的预测结果总是存在一定的滞后性,这同样也是时间序列预测中的难题,到目前为止,仍然没有得到有效解决,下一步工作将针对这一问题展开研究。

拟在优化 LSTM 网络结构的基础上,引入分形神经网络思想,深入研究有理分形函数性质,提出一种全新的分形神经网络。尝试将其研究成果与改进模型进行融合,并应用于股票价格时间序列分析与预测工作。

参考文献

- [1] Wang J. Z., Wang J. J., Zhang Z. G., etc. Forecasting stock indices with back propagation neural network[J]. Expert Systems with Applications, 2011, 38(11): 14346–14355
- [2] Lin C. S., Chiu S. H., Lin T. Y. Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting[J]. Economic Modelling, 2012, 29(6): 2583–2590
- [3] Wang B., Huang H., Wang X. A novel text mining approach to financial time series forecasting[J]. Neurocomputing, 2012, 83: 136–145
- [4] Vanstone B., Finnie G. An empirical methodology for developing stock- market trading systems using artificial neural networks[J]. Expert Systems with Applications, 2009, 36(3): 6668–6680
- [5] Hsu S. H., Hsieh J., Chih T. C., Hsu K. C. A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression[J]. Expert Systems with Applications, 2009, 36 (4): 7947–7951
- [6] Atsalakis G. S., Valavanis K. P. Surveying stock market forecasting techniques–Part II: Soft computing methods[J]. Expert Systems with Applications, 2009, 36 (3): 5932–5941
- [7] Tung W. L., Quek C. Financial volatility trading using a self-organising neural-fuzzy semantic network and option straddle-based approach[J]. Expert Systems with Applications, 2011, 38(5): 4668–4688
- [8] Kara Y., Acar B. M., Baykan Ö. K. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange[J]. Expert Systems with Applications, 2011, 38(5): 5311–5319
- [9] Huang C. L., Tsai C. Y. A hybrid SOFM-SVR with a filter-based feature selection for stock market forecasting[J]. Expert Systems with Applications, 2009, 36(2): 1529–1539
- [10] White H. Economic prediction using neural networks: the case of IBM daily stock returns[C]// IEEE 1988 International Conference on Neural Networks, 2(6): 451–458
- [11] Kimoto T., Asakawa K., Yoda M., etc. Stock Market Prediction System with Modular Neural Networks[C]// Neural Networks International Joint Conference on IJCNN, 1990, 1:

1-6

- [12] Gen C. R. Non-linear prediction of security returns with moving average rules[J]. Journal of Forecasting, 1996, 15(3): 43-46
- [13] Rodriguez F., Martel C. On the profitability of technical trading rules based on artificial neural networks: Evidence form the madrid stock market[J]. Economics Letters, 2000, 69(1): 35-37
- [14] Zhang G. P. Time series forecasting using a hybrid ARIMA and neural network model[J]. Neurocomputing, 2003, 50: 159-175
- [15] Liu H., Wang J. Integrating independent component analysis and principal component analysis with neural network to predict Chinese stock market[J]. Mathematical Problems in Engineering, 2011: 1-15
- [16] Murat A. Bahadir. Comparison of bayesian estimation and neural network model in stock market trading[J]. Intelligent Engineering Systems through Artificial Neural Networks, 2011(20): 74-81
- [17] Murat A. Neural based techical analysis in stock market forecasting[M]// Intelligent Engineering Systems through Artificial Neural Networks, ASME Press, 2008(18): 261-265
- [18] 林楠. 基于 BP 神经网络和 GARCH 模型的中国银行股票价格预测实证分析[D]. 兰州: 兰州大学, 2014.
- [19] Hammad A., Alhajali M. Forecasting the jordanian stock prices using artificial neural network[M]// Intelligent Engineering Systems through Artificial Neural Networks, Digital Collection of The American Society of Mechanical Engineers, 2012(17): 273-275
- [20] Celik A. E., Karatepe Y. Evaluating and forecasting banking crises through neural network models: An application for turkish banking sector[J]. Expert Systems with Applications, 2007, 33(4): 809-815
- [21] Zhang G. P. Time series forecasting using a hybrid ARIMA and neural network model[J]. Neurocomputing, 2003: 159-175
- [22] Zhang G. P., Qi M. Neural network forecasting for seasonal and trend time series[J]. European Journal of Operational Research, 2005, 160(2): 501-514
- [23] Tang Z., De A. C., Fishwick P. A. Time series forecasting using neural networks vs. Box- Jenkins methodology[J]. SIMULATION, 1991, 57(5): 303-310

- [24] Arnold F. Shapiro. The merging of neural networks, fuzzy logic and genetic algorithms[J]. Mathematics and economics, 2012(32): 115-131
- [25] Melike B. Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul Stock Exchange[J]. Expert Systems with Applications, 2011(36): 7355-7362
- [26] 肖菁, 潘中亮. 股票价格短期预测的 LM 遗传神经网络算法[J]. 计算机应用, 2012, 32(A01): 144-146
- [27] Hsieh T. J., Hsiao H. F., Yeh W. C. Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm[J]. Applied Soft Computing, 2011(11): 510-525
- [28] 张坤, 郁湧, 李彤. 基于小波和神经网络相结合的股票价格模型[J]. 计算机工程与设计, 2009, 30(23): 5497-5498
- [29] 闫东. 神经网络技术在股票价格短期预测中的应用研究[D]. 重庆: 重庆交通大学, 2013.
- [30] Shen F., Chao J., Zhao J. Forecasting exchange rate using deep belief networks and conjugate gradient method[J]. Neurocomputing, 2015, 167: 243-253
- [31] Zhang R., Shen F., Zhao J. 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
- [32] Yoshihara A., Fujikawa K., Seki K., etc. Predicting stock market trends by recurrent deep neural networks[C]// In Proceedings of the Pacific Rim International Conference on Artificial Intelligence, 2014: 759-769
- [33] Saad E. W., Prokhorov D. V., Wunsch D. C. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks[J]. IEEE Transactions on Neural Networks, 1998, 9(6): 1456-1470
- [34] Lai G., Chang W.C., Yang Y., etc. Modeling long-and short-term temporal patterns with deep neural networks[C]// The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2017: 95-104
- [35] Wang J., Wang Z., Li J., etc. Multilevel wavelet decomposition network for

- interpretable time series analysis[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Minings, 2018: 2437-2446
- [36] Qin Y., Song D., Chen H., etc. A dual-stage attention-based recurrent neural network for time series prediction[J]. arXiv preprint, arXiv: 1704.02971
- [37] Shi X., Chen Z., Wang H., etc. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting[J]. Computer Science, 2015, arXiv: 1506.04214
- [38] Lipton Z. C., Kale D. C., Elkan C., etc. Learning to diagnose with lstm recurrent neural networks[J]. Computer Science, 2015, arXiv: 1511.03677
- [39] Lin X., Yang Z., Song Y. Short-term stock price prediction based on echo state networks[J]. Expert Systems with Applications, 2009, 36(3-part-P2): 7313-7317
- [40] Wei L. Y., Cheng C. H. A hybrid recurrent neural networks model based on synthesis features to forecast the Taiwan stock market[J]. International Journal of Innovative Computing Information & Control Ijicic, 2012, 8(8): 5559-5571
- [41] Dixon M. F. Sequence classification of the limit order book using recurrent neural networks[J]. Social Science Electronic Publishing, 2017
- [42] Fischer T., Krauss C. Deep learning with long short-term memory networks for financial market predictions[J]. European Journal of Operational Research, 2018, 270: 654–669
- [43] Cai X., Zhang N., Venayagamoorthy G. K., etc. Time series prediction with recurrent neural networks trained by a hybrid PSO–EA algorithm[J]. Neurocomputing, 2007, 70: 2342–2353
- [44] Hsieh T., Hsiao H., Yeh W. Forecasting Stock Markets Using Wavelet Transforms And Recurrent Neural Networks: An Integrated System Based On Artificial Bee Colony Algorithm[J]. Applied Soft Computing Journal, 2011, 11(2): 2510-2525
- [45] Rather A. M., Agarwal A., Sastry V. N. Recurrent neural network and a hybrid model for prediction of stock returns[J]. Expert Systems with Applications, 2015, 42(6): 3234-3241
- [46] Chen K, Zhou Y, Dai F. A LSTM-based method for stock returns prediction: A case study of China stock market[C]// IEEE International Conference on Big Data, 2015: 2823-2824

- [47] Jia H. Investigation into the effectiveness of long short term memory networks for stock price prediction[J]. arXiv preprint: 1603.07893, 2016.
- [48] Siami-Namini S., Namin A. S. Forecasting economics and financial time series: Arima vs lstm[J]. arXiv preprint, 2018, arXiv: 1803.06386
- [49] Nelson D. M. Q., Pereira A. C. M., Oliveira R. A. Stock market's price movement prediction with LSTM neural networks[C]// International Joint Conference on Neural Networks, IEEE, 2017: 1419-1426
- [50] Kim K. j. Financial time series forecasting using support vector machines[J]. Neurocomputing, 2003, 55: 307-319
- [51] Tay F. E., Cao L. Application of support vector machines in financial time series forecasting[J]. omega, 2001, 29(4): 309-317
- [52] Lahouar A., Slama. Day-ahead load forecast using random forest and expert input selection[J]. Energy Conversion and Management, 2015, 103: 1040-1051
- [53] Rynkiewicz J. 1999. Hybrid HMM/MLP models for times series prediction[C]// ESANN, 7th European Symposium on Artificial Neural Networks, 1999: 21-23
- [54] Hornik K. Multilayer feedforward networks are universal approximators[J]. Neural Networks, 1989, 2(5): 359-366
- [55] Rodriguez A., Laio A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496
- [56] Yao X., Liu Y., Lin G. M. Evolutionary programming made faster[J]. IEEE Transactions on Evolutionary Computation, 1999, 3(2): 82-102
- [57] Suganthan P. N., Hansen N., Liang J. J., etc. Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization[R]. Nanyang Technological University, Singapore, Technical Report 2005005, 2005
- [58] Liang J. J., Qin A. K., Suganthan P. N., etc. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(3): 281-295
- [59] Lynn N., Suganthan P. N. Ensemble particle swarm optimizer[J]. Applied Soft Computing, 2017, 55: 533-548
- [60] Qin Q. D., Cheng S., Zhang Q. Y., etc. Particle swarm optimization with interswarm

- interactive learning strategy[J]. IEEE Transactions on Cybernetics, 2016, 46(10): 2238-2251
- [61] Derrac J., Garcia S., Molina D., Herrera F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms[J]. Swarm and Evolutionary Computation, 2011, 1(1): 3-18
- [62] Zhan Z. H., Zhang J., Li Y., etc. Adaptive particle swarm optimization[J]. IEEE Transactions on Systems, Man, and Cybernetics. Part B, 2009, 39(6): 1362-1381
- [63] Chen W. N., Zhang J., Lin Y., etc. Particle swarm optimization with an aging leader and challengers[J]. IEEE Transactions on Evolutionary Computation, 2013, 17(2): 241-258
- [64] Adebisi A. A., Adewumi A. O., Ayo C. K. Stock price prediction using the ARIMA model[C]// International Conference on Computer Modelling & Simulation. Piscataway, IEEE, 2015: 106-112
- [65] Lin Y., Guo H., Hu J. An SVM-based approach for stock market trend prediction[C]// The International Joint Conference on Neural Networks(IJCNN), IEEE, 2013
- [66] Usmani M., Adil S. H., Raza K., etc. Stock market prediction using machine learning techniques[C]// International Conference on Computer & Information Sciences, IEEE, 2016
- [67] Rather A. M., Agarwal A., Sastry V. Recurrent neural network and a hybrid model for prediction of stock returns[J]. Expert Systems with Applications, 2015, 42(6): 3234-3241
- [68] Selvin S., Vinayakumar R., Gopalakrishnan E. A., etc. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]// The International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2017

攻读学位期间取得的学术成果

攻读学位期间发表的论文:

- [1] 基于粒子群优化 LSTM 的股票预测模型[J]. EI 检索期刊: 北京航空航天大学学报, 2019, 45(12): 2533-2542. **第一作者**
- [2] 纹理细节保持的图像插值算法[J]. 计算机科学, 2019, 46(6A): 169-176. **第一作者**
- [3] 基于视觉感知的图像放大. 计算机工程与应用, 在线发表. **第三作者**

攻读学位期间参与的项目:

- [1] 基于分形分析的复杂纹理图像三维表面重建与特征计算关键技术研究. 国家自然科学基金面上项目, **第六完成人**
- [2] 复杂纹理的几何特征分析与重建技术研究, 山东省自然科学基金, **第七完成人**
- [3] 粒子群优化算法研究, 山东财经大学研究生创新项目, **第五完成人**
- [4] 复杂环境下的图像复原技术, 山东财经大学研究生创新项目, **第八完成人**

致谢

时光转瞬即逝，仿佛昨日才懵懵懂懂的踏入研究生的校园，今天就到了即将说再见的时刻。不禁脑海里像放电影一样播放着三年时光的点点滴滴，有开心，有悲伤，有痛苦，也有欢乐，这就像是人生的一个小缩影，我又走过了一场为期三年的人生经历。感谢陪伴着我的老师们同学们和我的家人，因为有他们在我身边，我就有了坚强的后盾，同时也承担起守护他人的责任，这种力量一直支持着我顺利完成硕士阶段的学习和研究。

首先，我最想要感谢的就是我的研究生导师，张云峰教授。导师用自己豁达的处事风格，严谨的教学态度以及独特的个人魅力深深的影响着我。每个学生都应该是自己导师的影子，传承导师的思想和风格。导师以身作则，对待科研问题一丝不苟。学习中，对于我们提出的问题，无论多忙总要抽出时间耐心仔细的分析其中的本质联系，解答我们的疑惑。生活中，导师总是宽以待人，悉心照顾到我们方方面面。时常感觉自己如此的幸运，能够在导师的教导下成长，在做人做事做学问方面都受益匪浅。再次感谢导师，谢谢您！

其次，我要感谢陪伴了我三年时光的全体 2017 级研究生，跟大家在一起经历了许许多多欢乐的时刻，也祝福大家都拥有一个美丽前程。感谢我的师姐范清兰，王平，师哥杜宏伟，师弟秦超，邵凯，师妹孙紫怡，还有我可爱的搭档刘新新，大家一起经历了太多的风风雨雨，在我陷入困难时给予我很大的帮助，相信不论时间怎么变换，你我都是最值得信赖的朋友。

最后，感谢我的家人，正是你们的支持、鼓励和温暖，我才能如此坚定的前进。心之所向，永远都是家的方向。

最后的最后，感谢自己不在该放弃的时刻没有放弃，感谢自己在需要挺住的时刻挺住！保持饥饿与谦逊，永远盯着最重要的目标走。

宋刚

2020 年 6 月 1 日