

## 目录

摘要.....	1
Abstract.....	2
第 1 章 绪论.....	3
1.1 研究背景与意义.....	3
1.2 国内外研究现状.....	4
1.3 本文的主要工作.....	8
1.4 本文的组织结构.....	9
第 2 章 相关理论概述.....	10
2.1 量化交易概述.....	10
2.2 深度学习概述.....	11
2.2.1 深度学习相关概念与原理.....	11
2.2.2 循环神经网络与 LSTM 神经网络.....	12
2.3 股票技术分析指标.....	14
2.4 本章小结.....	15
第 3 章 LSTM 模型设计.....	16
3.1 LSTM 模型基本原理.....	16
3.2 模型具体设计与实现.....	16
3.3 本章小结.....	18
第 4 章 实证分析与结果.....	19
4.1 实验环境.....	19
4.2 数据收集与数据预处理.....	19
4.3 模型训练与预测结果.....	24
4.4 技术指标与具体策略.....	26
4.5 回测结果.....	27
4.6 本章小结.....	28
第 5 章 总结与展望.....	29
致谢.....	31
参考文献.....	32

附录 1.....	35
英文原文.....	35
附录 2.....	40
译文.....	40

## 基于深度学习的融资融券信息预测模型研究

### 摘要

随着国内股票市场的逐步发展与完善，越来越多的人希望通过股票为自己带来额外的收益，量化交易概念的传入，使得使用计算机技术进行股票交易成为如今的热点问题。如今在量化交易当中，越来越多的研究者选择使用深度学习来预测股票价格进而进行交易。

本文中，我们选择长短期记忆模型（LSTM）作为研究的预测模型，对模型的参数进行了一定的设置，并选择使用最近 5 年的融资融券信息等股票的交易数据作为训练数据来训练模型，将股票涨跌预测问题转变为了一个二分类问题。在本研究中，我们使用该模型对股票收盘价的涨跌进行预测，并且结合 MACD 指标选择股票择时进行买入和卖出交易。本研究使用聚宽平台进行实证研究，通过测试计算使用历史数据进行交易的收益率，并且得到了高达 8% 的年化收益率。

**关键词：**融资融券、深度学习、量化交易、LSTM、MACD

## **Abstract**

With the gradual development and improvement of the domestic stock market, more and more people hope to bring additional benefits for themselves through the stock, and the introduction of the concept of quantitative trading, making the use of computer technology for stock trading become a hot issue today. Nowadays, in quantitative trading, more and more researchers choose to use deep learning to predict the stock price and then trade.

In this paper, we choose Long Short-Term Memory (LSTM) as the prediction model of the study, and set the parameters of the model to a certain extent, and choose the trading data of the latest five years, such as the information of margin trading and securities lending, as the training data training model, to transform the stock up and down prediction problem into a two classification problem. In this study, we use this model to predict the rise and fall of the closing price of the stock, and combine MACD index to select stocks to buy and sell at the right time. In this study, we use JoinQuant platform to conduct empirical research, through testing to calculate the return rate of transactions using historical data, and get an annual return rate of up to 8%.

**Keywords:** margin trading, deep learning, quantitative trading, LSTM, MACD

## 第1章 绪论

### 1.1 研究背景与意义

近年来，我国的经济在快速繁荣地向前发展，人民的生活水平已经得到了质的提高，国民财富在不断积累增加，人们也越来越重视对自己资产的投资与管理。股票，一种既是高风险同时又是高收益的投资方式，显然会引起众多投资者的重视。股票市场的重要特征之一就是波动，中国股市是一个新兴的市场，跟许多其他国家相比较，我们的股市制度尚且不够完善，波动尤其剧烈并且频繁<sup>[1]</sup>。股市一直在波动，同时股市的波动也为通过股市获取利润提供了更大的可能。投资者们无不期望运用各种各样的手段与方法来预测股票价格的变化，通过预测的价格进行交易，进而获取收益，显然，价格预测地越准确，那么收益也就会越高。

量化投资是一种借助股票价格、股票的成交量等股票的交易数据，运用统计学、数学的相关内容与知识，使用计算机技术用编程构造模型，指导投资的现代化投资方式<sup>[2]</sup>。量化投资能够基于大量金融市场，科学高效地进行投资决策，可以在秒级时间内组合大量金融产品进行交易，量化交易可以避免交易者由于心理因素做出不理性的投资决策，进而可以为投资者带来更高的收益<sup>[2]</sup>。自从 20 世纪 80 年代开始，量化投资在美国的华尔街兴起，而近年来，随着我国股票市场体系与计算机相关技术的不断创新发展，中国金融衍生产品市场也在不断地建立和发展，与其相关的投资理论与投资策略必将会成为今后投资的研究热点<sup>[3]</sup>。

近年来，随着理论研究的不断深入和计算机性能的不断提高，人工智能得到了飞速的发展，人工智能已经成为了人类生产与生活中不可或缺的重要组成部分<sup>[4]</sup>。而深度学习神经网络又是人工智能所运用的主要预测算法，深度学习的相关算法就是传统机器学习相关算法在一定程度上的扩展和延伸<sup>[4]</sup>。深度学习算法是指人工智能领域中对大量数据进行深度学习的方法，深度学习算法是通过构建多层的神经网络，使用大量数据对神经网络进行训练以提取其特征的一种机器学习方式，由于神经网络有多层，因此称之为深度神经网络<sup>[5]</sup>。目前已有多个行业应用到了深度学习的相关算法。比如在图像识别领域中，车辆识

别、人脸识别、无人驾驶、无人机等领域使用卷积神经网络构造的系统已经取得了较好的效果；在语音识别与自然语言处理（NLP）领域中，使用长短期记忆网络（LSTM）与循环神经网络算法设计出来的语音机器人、智能问答系统等产品已取得了很大的成功；在游戏与类人脑领域，使用深度学习算法构造的各种游戏角色与人类对战甚至可以击败人类的专业选手；在数据分析以及推荐系统领域，深度学习帮助人类挖掘到了更多有用的信息，基于深度学习算法的智能推荐系统能够为人类推荐更加合适的产品等等<sup>[5,6]</sup>。

随着人们对深度学习研究的深入以及学科交叉与学科融合大趋势的发展，开始有研究人员将深度学习的内容与知识应用到金融市场上来。深度学习在金融领域的应用主要在股票、期货与外汇等几个主流方面，特别是股票价格与涨跌趋势的预测成为了近年来的投资界研究的热点之一<sup>[7]</sup>。为了预测股票市场的波动，不同领域的研究者近些年来进行的大量的理论与实证研究<sup>[8]</sup>。研究者们通过利用深度学习的相关算法与知识，研究寻找隐藏在大量金融数据背后的知识与规律，同时，研究者们也已经得到了大量有价值的研究成果<sup>[8]</sup>。

金融数据通常情况下都是一个时间序列的数据，而长短期记忆网络（LSTM）作为深度学习中一种经典的算法以其良好的记忆性著称，因此 LSTM 神经网络是一种较适用于处理金融数据的深度学习算法<sup>[4]</sup>。通常来说，股票的融资融券交易数据以及交易量、大盘指数等信息都可以在一定程度上反应股票价格的变化，因此，我们可以使用这些信息来预测股票的价格变化趋势。

综合运用历史数据来预测股票价格的未来股价涨跌情况，这是一个将股票价格变化划分为涨跌两种情况的二分类研究问题。本文使用深度学习中经典的算法之一——LSTM 与量化交易相结合，预测股票市场价格变化的涨跌情况，既能在此基础上进一步深入研究，也可以将本文的研究应用于投资当中，为投资者提供参考带来一定的收益。这既有一定的理论研究发展意义，也有一定的实践应用指导价值。

## 1.2 国内外研究现状

国外很早便开始了对量化投资的相关研究。国外量化交易的研究是经历了从最初使用传统统计学知识预测股价进行量化投资到之后的使用传统机器学习算法预测股价进行量化投资再到如今使用深度学习算法知识预测股价进行量化

投资的发展变化。在 1989 年, F. A. G. Den Butter 等人使用单变量 ARIMA 模型预测与向量 ARIMA 模型的三种不同预测相结合的方式, 对荷兰未来股票价格进行预测研究, 他们的组合权重通过回归分析来确定, 每个月的组合权重都会有所不同, 研究预测的结果是非常理想的<sup>[9]</sup>。在 1990 年, E.Schöneburg 等人使用神经网络模型来预测股票短期(10 天)内的价格, 研究了感知器、ADALINE、MADALINE 和反向传播网络的使用, 得到了较高的精确率, 取得了令人满意的效果, 并且为未来进一步的研究打下了基础<sup>[10]</sup>。在 1993 年, Gia-Shuh Jang 等人基于双模神经网络(dual-net)预测短期的股票价格, 并且使用了一种新的自适应反转识别机制, 该机制可以自动调整阈值, 以识别买入或卖出股票的最佳时机, 他们的研究得出了双网结构比单模块神经网络具有更好的泛化能力的结论<sup>[11]</sup>。在 1996 年, D. Brownstone 使用神经网络, 根据对均方差和均方根差的测量, 预测了 5 天之前和 25 天之前收盘的每日市场, 将每个单独的测试用例预测与实际的市場结果进行了比较, 并类似地计算出了整个测试集的总准确性, 研究得出了以最小均方差为基础的预测与相同的测试用例之间的关系并不是不大, 而应该用总百分比精度来衡量的结论<sup>[12]</sup>。在 1998 年, Saad E W 等人比较了 TDNN、RNN 和 PNN 三种神经网络预测结果的准确率, 讨论了不同的可预测性分析技术, 并基于每日收盘价的历史对收盘价的可预测性进行了分析, 研究结果表明所有的网络都是可行的, 首要的选择指标是方便为主<sup>[13]</sup>。在 2000 年, Kyoung-jae Kim 等人提出了一种基于遗传算法的特征离散化方法和人工神经网络(ANN)算法来预测股价指数的新方法, 研究中发现, 遗传算法不仅可以改善机器学习算法, 还可以降低特征空间的复杂度, 实验结果表明, 基于遗传算法的特征离散化模型明显优于其他的传统模型<sup>[14]</sup>。在 2002 年, Alan M. Safer 使用股票内部交易数据, 使用神经网络算法分析研究了共计 343 家公司为期 4 年半的股票数据, 得出了一些能够加强异常收益率的预测方法, 对实际投资使用具有一定的指导作用与意义<sup>[15]</sup>。在 2006 年, Hyun-jung Kim 等人研究了基于人工神经网络(ANN)和遗传算法(GA)的自适应时滞神经网络(ATNN)和时滞神经网络(TDNN)的时间序列特性混合方法在股市预测任务时间模式检测中的有效性等问题, 基于研究, 他们提出了一种基于试错和各种启发式或统计技术的通用方法, 结果表明, 本文提出的综合方法比标准的 ATNN、TDNN

和递归神经网络（RNN）方法具有更高的精确度<sup>[16]</sup>。在 2010 年，Melek Acar Boyacioglu 等人利用自适应网络模糊推理系统（ANFIS）研究股票市场收益的可预测性问题，其目的是为了确定 ANFIS 演算法是否能够准确预测股市收益，研究使用了 6 个宏观经济变量和 3 个指数作为输入变量，模型成功地预测出了 ISE 国家 100 指数的月收益率，准确率高达 98.3%，得出了 ANFIS 可以用作经济学家和投资者进行股票价格指数收益预测的有用工具的重要结论<sup>[17]</sup>。在 2014 年，Ruxanda 等人以罗马尼亚的 BET 指数为例，对如何配置人工神经网络进行股市预测提出了一些相关的建议，并通过对克罗地亚股市数据的训练网络测试得到了证实结果，并且指出了 Broyden-Fletcher-Goldfarb-Shanno 训练算法在模型收敛性和泛化能力等诸多方面都是一个股价预测的很好的选择<sup>[18]</sup>。在 2018 年，V V Ramalingam 等人使用自然语言处理（NLP）的方法研究新闻与一般公众意见对股价的影响，发现了这些数据与股市波动之间所存在的一些相关性，得到的预测结果也具有较高的准确率<sup>[19]</sup>。

由于国内股票的引入迟于西方市场，因而国内对量化投资的研究也略晚于国外。在 1997 年，张悦华使用 BP 神经网络用从 1996 年 1 月 11 日到 1996 年 1 月 22 日的长春华联股票数据训练出模型，预测自 1996 年 1 月 23 日到 1996 年 1 月 31 日的股票收盘价，得到了预测价格与实际价格的误差不大于 8% 的优良结果<sup>[20]</sup>。在 1998 年，刘新勇等人针对股价预测所存在的一些问题，基于神经网络算法，提出了联合最小二乘法与遗传算法的综合训练网络权值与结构的新方法，通过对上证指数股票的模拟预测得到了令人满意的预测结果<sup>[21]</sup>。在 2000 年，张世英提出了一种根据国家综合经济的相关数据对证券投资时机进行指导预测的新方法，该方法采用模糊逻辑与遗传算法相结合的方式来描述综合经济数据和各种资产类之间的相关关系，并且通过实证测试，得到了较好的效果<sup>[22]</sup>。在 2003 年，韩旭里使用道氏三假说作为研究的理论基础，设计了一个三层的 BP 网络，建立了相应的数学模型，并使用上海石化共计 140 天的数据进行训练，预测 10 天的收盘价得到预测价格与实际价格的离差大小仅为 0.0146，几乎可以忽略<sup>[23]</sup>。在 2006 年，袁婷使用改进的遗传神经网络和双神经网络与传统 BP 神经网络进行了对比研究，分析了各种类型的神经网络自身的特点，为使用神经网络进行股价预测研究提供了一定的基础



[24]。在 2009 年，刘成竹基于前人对股票预测的研究，从股价指数自身的特点出发，选择影响因子，使用统计学与人工智能相结合的方法，建立了 GARCH 模型和 BP 神经网络模型来预测股票的价格，并且对上证指数进行了研究，通过对大量样本数据的学习预测，证明了所建立模型具有较好的准确率与可行性<sup>[25]</sup>。在 2012 年，张翔宇通过使用 BP 神经网络预测股价发现了其存在一些问题，进而使用了遗传算法对 BP 神经网络模型进行了进一步的优化，并且对比了两次预测结果的准确度，发现优化后的 BP 神经网络具有更好的预测性能<sup>[26]</sup>。在 2015 年，杜晓使用 MEA-BP 神经网络模型，选取了 8 个股票数据的基本指标作为预测数据，使用上证指数，与 BP 神经网络模型和 GA-BP 神经网络模型的实证分析的结果进行了比较，发现，MEA-BP 无论是收敛速度还是准确率都明显优于另外两者<sup>[27]</sup>。在 2017 年，毛景慧通过使用深度学习中 LSTM 深度神经网络模型对股票的价格进行了预测研究，使用沪深 300 成分股 2006 年到 2017 年的数据信息进行研究实证结果分析，并且对比了 BP 神经网络以及传统 RNN 预测的结果，发现由于 LSTM 模型具有良好记忆性，其预测结果明显优于 BP 神经网络和传统的 RNN 的预测结果<sup>[28]</sup>。在 2018 年，季子峥等人利用社交媒体的话题与情感信息数据，使用自然语言处理（NLP）的相关方法进行处理，使用 LDA 与 ITS 算法对股票价格进行预测研究，研究得出了较为不错的预测结果<sup>[29]</sup>。

在量化交易领域，国内与国外都已经做过了大量的研究，当然，我们可以预见无论是国内还是国外也都将会继续进行更加深入的研究。在人工智能时代，使用计算机技术进行的量化交易不可避免地成为了研究者与投资者集中选择研究的热点之一。同时，对比国内外的研究，我们可以发现，国内的量化交易研究开始的确是迟于国外，但是随着国内计算机技术与金融理论研究的不断深入，随着我国国内研究人工智能浪潮的兴起，随着国内学科交叉与融合的逐渐深入，如今国内量化交易的研究也已经到了如火如荼的地步，研究的广度与深度也已经处于国际的领先水平。我们可以相信，使用深度学习的量化交易是人类技术的一大进步，选择对其进行研究，在为投资者带来巨大的经济收益的同时也必然可以改变人类的生活，深度学习技术也必然会改变如今的证券交易方式与制

度，为人类生产方式以及生活方式带来翻天覆地的改变。

### 1.3 本文的主要工作

本文首先介绍了量化交易的发展过程与现状，并且提出了一种基于融资融券交易数据以及各股与大盘指数数据使用深度学习模型预测股票涨跌的新的量化交易方法。本研究的主要任务就是根据股票历史融资融券交易数据以及各股与大盘指数数据使用深度学习模型预测股票涨跌的趋势，同时参考股票技术指标进行预测研究，并且使用测试数据对策略进行实证分析。

由于股票的融资融券数据以及各股与大盘指数数据都是时间序列数据，故选用 LSTM 神经网络作为本研究的预测模型较为合理。备选股票池为近 5 年都存在融资融券交易的一千多只沪深 A 股的股票。训练数据从备选股票池中随机选择 100 只股票，测试数据从备选股票池中除训练数据之外的其余股票中随机选取，测试数据随机选取 150 只股票分为 3 组，每组 50 只股票，共计构成 3 个测试数据集，以备回测使用。训练数据选取的时间段为从 2015 年 1 月 1 日到 2020 年 2 月 20 日共计 1249 个交易日的数据。由于融资融券交易数据只能在交易日的第二天 10 点左右才可以获取到，因此本研究是使用前一日的融资融券数据预测第二个交易日收盘价对应当日收盘价的涨跌情况。训练数据所选取的数据列包括 7 列融资融券交易数据：融资余额、融资买入额、融资偿还额、融券余额、融券卖出额、融券偿还额、融资融券余额；以及大盘指数数据即深证成指的日收盘价；8 列各股数据：每日开盘价、每日收盘价、每日最高价、每日最低价、每日交易量、每日交易金额；训练数据共计 14 列。使用 LSTM 神经网络时，要对参数进行一定的调整，以使训练结果更加合理。根据股票数据时间序列的特点，在使用 LSTM 神经网络时应当设置适当的时间步（timestep），在本研究中 timestep 设置为 8，即是使用连续 8 个交易日的融资融券交易数据以及各股与大盘指数数据来预测一次的收盘价涨跌情况。模型预测的结果包括两种，即上涨/不上涨，当预测收盘价将会上涨时应当买入，否则应当卖出或者不买入，这也就把本研究简化成了一个二分类的问题。

本研究的收据收集以及预处理工作是在本地进行的。收集的数据来自聚宽（JoinQuant）平台提供的 JQData（本地量化数据）数据集，通过调用对应的函数即可获取所需的数据。数据预处理工作包括缺失值处理以及数据归一化，预

处理之后的数据存在 CSV 文件以备训练模型时使用。本研究的回测工作是在聚宽平台上在线进行的，首先使用随机选取出的 100 只股票 5 年多的数据即本地已经收集并做好预处理的数据来训练模型，训练之后选择使用测试数据集的股票进行预测研究，根据预测的结果进行模拟交易，预测工作是模拟历史上的每天而进行的。历史模拟运行完成之后，聚宽平台将会自动提供策略收益率、基准收益率以及换手率等参考数据的运算并将其显示出来。在聚宽平台上将 3 个测试数据集进行不同时间段的回测工作之后就可以评价训练出来的模型的性能。

#### 1.4 本文的组织结构

本文共分为 5 个章节，内各章节的基本内容如下：

第 1 章为绪论，本章主要是对本研究的相关内容进行了简单介绍与阐述，对量化交易在国内和国际上的研究历史与现状做了简要的介绍，接下来又阐述了本文的主要研究工作，最后则说明了本文的组织结构。

第 2 章为相关理论概述，详细介绍了量化交易和深度学习特别是循环神经网络的相关理论内容，通过本章理论内容的介绍，将易于对后文 LSTM 模型具体设计以及具体研究的理解。

第 3 章为 LSTM 模型设计，该部分详细地叙述了本研究所使用的 LSTM 神经网络的理论基础以及具体设计的相关内容，其中包括了数据格式、相关参数使用等诸多的问题。

第 4 章为实证分析与结果，这部分也就是使用具体的训练数据训练设计好的 LSTM 模型，并且对测试数据集的股票进行回测，并且展示了回测的相关结果，并且根据回测结果而提出了一种改进的模型，将改进的策略进行回测得到了更高的收益，最后又对设计的策略进行了一定的评价。

第 5 章为总结与展望，这部分对本次的研究工作所解决的问题与本研究的意义进行了阐述，对本研究进行了一定的总结，对研究的成果与尚未解决的问题进行了分析与探讨，并且对进一步研究的可能性进行了展望并发表了自己的观点。

## 第2章 相关理论概述

### 2.1 量化交易概述

20 世纪 80 年代开始, 国外提出了量化交易的概念。由于量化交易具有收益稳定, 操作简单等优点, 量化交易越来越被投资者所青睐, 其占据的市场份额也就越来越大。目前越来越多的研究者投入到量化交易的研究当中来, 这些研究者有着不同的学科背景, 他们原来既有从事计算机专业的也有从事数学专业的更有从事金融专业的, 对量化交易的研究明显表现出了学科交叉融合的趋势。具体的量化交易就是指通过使用股票价格、股票交易量以及其他有关股票交易所产生的数据, 作为数据分析的依据来设计出合理的数学模型, 使用计算机技术对股票价格变化等信息进行分析预测, 最后根据预测分析的结果使用程序化的指令进行股票投资的交易<sup>[30]</sup>。量化交易具有纪律、及时和系统等诸多的优良特性。

量化交易的策略研究主要可以分为两类, 第一种使用统计学的相关知识, 通过概率来预测股票价格变化的相关趋势。这种策略属于使用纯数学专业知识的预测方法, 目前一些技术因子的引入使用都属于这一类的策略。这种策略开始最早, 最初的量化交易基本都是采取了这种策略, 目前该策略主要是作为量化交易的辅助或者参考来使用的。第二种就是使用计算机编程的方法, 结合机器学习的相关内容对股票价格的相关趋势进行预测。第二种策略主要是在人工智能大潮到来之后才逐渐火起来的, 这种策略选取了适当的股票交易数据, 选择合适的机器学习数学模型进行训练, 使用训练好的模型预测股票的走势, 进而进行股票的交易。该策略的核心问题就是机器学习的算法问题, 选择合适的机器学习算法是该策略成功的关键因素<sup>[30]</sup>。这两种策略有一定的共同点, 主要就是它们最根本的思想都是数学, 它们都是使用数学的知识指导股票交易的典型方法。当然这两种策略的不同点也比较明显, 第一种策略是使用数学加上金融学的知识来指导交易的, 其对金融学的知识要求较高, 而第二种策略不要求使用者拥有较好的金融学知识, 主要要求研究者与使用者要具有一定的人工智能理论基础, 对使用者的学科背景要求更为宽泛。

对量化交易的策略的评测是比较简单的, 主要就是通过模拟回测来计算收

益，并且将得到的收益与基准收益进行对比，即可在一定程度上得出量化交易的好坏，当然，根据策略运行出来的收益曲线也可以评测策略的好坏，具体就是根据策略收益曲线与基准收益曲线的走势情况来评测的。

## 2.2 深度学习概述

### 2.2.1 深度学习相关概念与原理

在古代的东方，有女娲造人、鲁班造木鸢的故事；在古代的西方，有希腊神话人物创造潘多拉的故事。这些虽然都是传说中的神话故事，但这些故事无不体现着人类对创造智能体的梦想与追求。人类一直在思考自身的思维方式，同时也在思考着如何能设计出能够代替人类工作帮助人类进行决策的智能体，这始终是人类要探索的重要课题之一。在 1842 年，阿达·洛芙莱斯，作为人类社会首次进行构思可编程计算机第一人，在构思可编程计算机时就已经思考过是否可以使计算机变得智能。而如今，人工智能已经成为了一个活跃的研究课题，具有许多的实际应用和众多的研究者，拥有着众多实际的应用落地产品。如今的人工智能，就是让机器从历史数据当中学习，获取经验，根据不同层次化的概念体系来理解现实生活世界。同时层次化概念的引入，让机器的学习能力拥有了更好的性能，当然，层次化的存在也就引出了深度学习的概念。深度学习的核心思想就是：以分层网络的方式，对分层次的所有深浅层的特征信息由浅到深进行逐层地获取。

深度学习与机器学习拥有着许多共同点，最主要的就是他们都是需要训练数据进行“训练”的。经过训练数据“训练”后的深度学习模型或者机器学习模型都具备着一定的预测能力，如果训练数据比较少，则深度学习模型预测能力不会明显优于机器学习模型的预测能力，但是当训练数据特别巨大的时候，深度学习的优势就会表现得淋漓尽致，其模型将会明显好于机器学习模型。特别地，当训练数据达到千万条级别的时候，深度学习模型的性能将会明显好于人类的辨别能力。

深度神经网络之所以称之为网络，是因为他们通常都是用众多不同的函数复合在一起来表示的。那么它们为什么被称之为神经网络呢？这是因为这些网络的设计者是受到了神经学的启发而将其设计出来的。对于深度神经网络模型，可以根据模型的输出结果与模型本身二者之间是否有反馈（feedback）连接来将

神经网络分成前馈神经网络（FNN）、循环神经网络（RNN）两个大的类别。前馈神经网络的最后一层被叫做输出层（output layer），输出层之前、输入层之后的不给出每一层的输出的层就被叫做隐藏层（hidden layer）。深度神经网络有深度（depth）和宽度（width）的概念，所谓的深度就是指神经网络模型的层数；而宽度的概念则来自于隐藏层，神经网络里面的每一个隐藏层一般都是向量值的，向量值的维度也就决定了模型的宽度大小。对于前馈神经网络，其输入信息通常是仅仅从输入层单向向前通过网络到达输出层，其具体结构如图 2-1 所示：

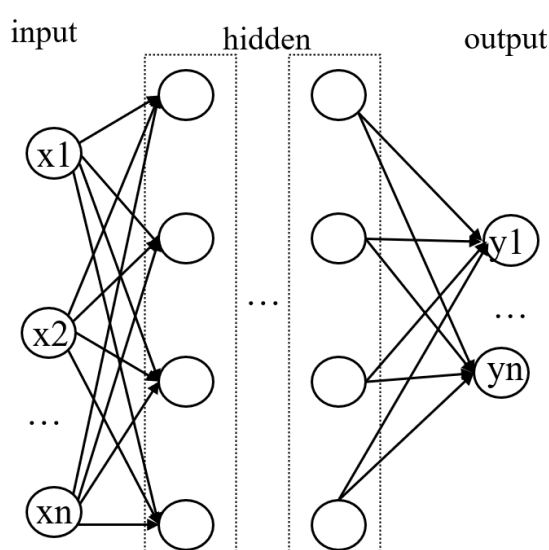


图 2-1 前馈神经网络结构图

### 2.2.2 循环神经网络与 LSTM 神经网络

由于前馈神经网络是单向向前传播的，这就要求输入前馈神经网络的假设数据都要是独立同分布的。但是在我们的现实生活中，并不是所有的数据都具有上述的特点，比如我们做自然语言处理的时候，我们的文本数据前后就是有关联的，段与段之间也是有一定的关联的，再比如我们的金融数据，不同交易日之间的数据自然就是一个序列，它们前后也是有一定的关联的，这就说明如果我们使用前馈神经网络进行训练则需要把整个序列数据都输入到模型内对其进行训练，但是这样就可能会造成维度灾难（Curse of Dimensionality），导致计算机无法运行下去。那也就说明我们需要一种能够在序列上进行操作的神经网络算法来处理这种类型的序列数据，而深度神经网络中的第 2 大类——循环神经网络（RNN）便满足这一特殊的要求。

很明显，人类在思考一个问题的时候，并不是时时刻刻都是从空白的大脑开始进行思考的，比如在阅读一篇文献的时候，我们是基于对前文的一定的理解的基础上才能够对下文进行理解的，如果是割裂了上下文而仅仅使用空白的大脑进行思考，显然就无法理解相关的内容了。很显然，传统的单向流动的神经网络就割裂了数据“上下文”的联系，使用空白的“大脑”对数据的内容进行“理解”的，所以我们在处理序列数据这种前后有联系的数据的时候需要使用一种非空白的“大脑”——循环神经网络。循环神经网络，也就是通常所说的 RNN，它是一类神经网络的总称，这一类神经网络比较适用于处理序列数据，比如文本数据、语音数据、两维的图像数据以及金融数据等等。在使用循环神经网络来处理时间数据的时候，如果网络在获取整个序列数据之前就能够观察到整个完全的序列数据的话，循环神经网络就可以有相对于时间向后的连接。

在传统的神经网络模型里面，从输入层到隐藏层再到输出层，它们的层与层之间是全连接的，但是它们每个层内部的节点与节点之间是没有连接的，这样也就导致了信息是单向流动的，导致这种网络对于像序列数据这种前后有关联的数据的问题，这种类型的神经网络就会难以去处理。循环神经网络，顾名思义，指的是该序列当前输出的信息是与前面输出的信息也是有关系的一种神经网络，具体表现出来的就是该神经网络会对前面的信息进行一定的记忆并且应用到当前输出的计算当中去，也就是说隐藏层的层内部节点与节点之间不再是没有连接的而是有一定的连接的，而且，隐藏层的输入数据不仅仅是包括输入层的输出数据，它还包括了之前的隐藏层的输出数据。这样的循环神经网络也就能够对序列数据这种前后有关联的数据表现出良好的性能来。循环神经网络的基本结构图如下图 2-2 所示：

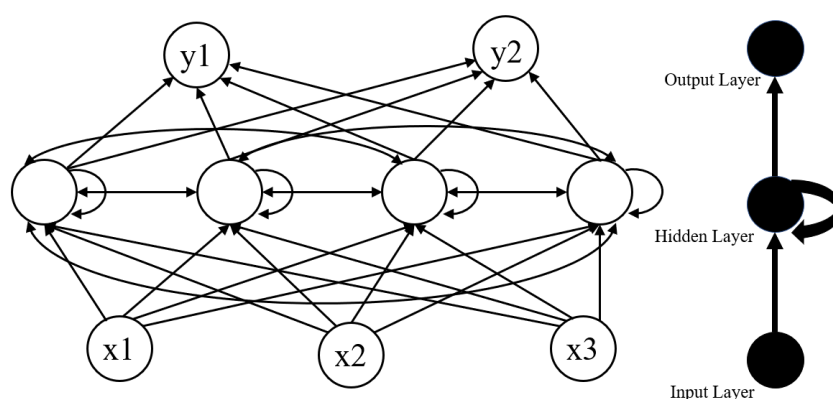


图 2-2 循环神经网络结构图



根据循环神经网络输入序列以及输出序列长度的不同，循环神经网络可以大致分为 4 中类型：一对一的、一对多的、多对一的、多对多的。

在多种类型的循环神经网络当中，长短期记忆模型（LSTM）被广泛应用于如今的工业界以及学术界研究当中。长短期记忆模型是一种特殊的循环神经网络，它是一种典型的门控循环神经网络（门控 RNN），就像渗透单元那样，门控循环神经网络也基于生成通过时间的路径，在这其中，它的导数既不会消失也不会爆炸，门控循环神经网络将这推广成为了每一个时间步（timestep）都有可能改变的连接权重。相对于渗透单元，门控循环神经网络可以自主地决定什么时候清除状态而不需要手动去决定。长短期记忆模型，顾名思义，模型将会对长短期不同时期的信息进行不同的对待处理，对于传统的循环神经网络在处理序列数据的时候会把所有之前的信息全部传递，这样就可能有冗余的信息被传递下去，而长短期记忆模型可以有选择的把长期与短期的信息进行选择性遗忘，这样就可以解决冗余信息的问题了，因此，长短期记忆模型是对传统循环神经网络的一种改进模型。长短期记忆模型是通过一定的刻意的设计来避免长期依赖问题的。在实践中，记住长期的信息是长短期记忆模型的默认行为，而并非是其需要付出很大的代价才能够获得的能力。

### 2.3 股票技术分析指标

票的技术指标是相比较于股票的基本分析而言的，技术分析是指通过图表或者技术指标的一些记录，来研究市场的行为反应，进而推测出股票价格的相对变化趋势的研究方式。股票的技术指标主要是通过股票价格、股票成交量以及股票涨跌的指数等数据来计算出的。股票技术分析指标是属于统计学的范畴，它利用股票的交易数据来论证股票的趋势进而可以指导对股票进行买卖。目前研究人员已经提出了多种股票技术指标，在本研究中主要用到了异同移动平均线（MACD）与相对强弱指数（RSI）两种，下面将对这两个指标进行介绍。

**MACD：**异同移动平均线（MACD）是移动平均线的变种之一，它利用快速移动平均线以及慢速移动平均线的聚合与分离，进而进行双重的平滑计算，根据结果发出指导买卖的信号。快速移动平均线一般是比较小的，慢速移动平均线一般是比较大的，具体选用的天数数值可以自己自由选取。MACD 的主要组成部分有正负差（DIF）以及异同平均数（DEA）这两个，它们是使用 MACD



预测趋势的核心。一般来说，当 DIF 向上交叉 DEA 的时候是买入的信号，此时就是所谓的金叉。当 DIF 向下交叉 DEA 的时候是卖出的信号，此时就是所谓的死叉<sup>[31]</sup>。

**RSI：相对强弱指数（RSI）**是根据某一时期内股票上涨点数与股票下跌点数和的比率制作出来的一种技术曲线。RSI 能够反映股票市场在一定时期内的景气程度，比较适合断线差价操作。可以设置参数计算多少日的相对强弱指标。RSI 的范围为 0 到 100，一般来说，国内股票的 RSI 分布在 20 到 80 之间，其值为 80 到 100 时表现为极强应当卖出，其值为 50 到 80 时表现为强，应当买入，其值为 20 到 50 时表现为弱，适合观望，其值为 0 到 20 时表现为极弱，适合买入<sup>[32]</sup>。

本研究设计的量化交易策略是主要使用深度学习预测指导交易，同时结合股票技术指标加以买卖限制进行交易的，通过深度学习与技术指标相结合可以实现很不错的效果。

## 2.4 本章小结

本章主要介绍了与本研究相关的一些理论，包括量化交易的相关理论、深度学习的相关理论、股票的相关技术指标等等，这些理论的介绍将会有助于读者对本研究、及本论文的理解。这一部分最主要的内容是对深度学习中深度神经网络进行了较为系统的阐述，介绍了几种典型的深度神经网络模型，通过上述的具体阐述可以更好地理解本研究中使用的长短期记忆模型（LSTM）的具体设计，对本实验的核心预测问题可以有更好地了解。下面的章节将会就本研究中使用的 LSTM 模型进行详细的介绍，将会介绍包括 LSTM 模型的设计、实验的具体进行与实验结果等相关的一些内容。

### 第3章 LSTM 模型设计

#### 3.1 LSTM 模型基本原理

长短期记忆模型（LSTM）第一次提出是在 1997 年，通常而言，LSTM 是指含有 LSTM 区块的深度神经网络，LSTM 区块（block），也就是部分书籍与文献所称呼的智能网络单元，LSTM 区块能够记忆不同时间长度的数值，区块中的门控（gate）能够决定输入（input）的内容是否很重要以至于能够被记住以及它应不应该被输出（output）。LSTM 一般具有 3 个门控，分别就是遗忘门（forget gate）、输入门（input gate）和输出门（output gate）。输入门用来决定 input 的值是否应该进入到下一层，若此时产生的值近似接近于 0，输入门就会把它挡住而不能进入下层；遗忘门是用来遗忘的，如果此时的值接近于 0，它就会把区块中的内容遗忘掉；输出门用来决定区块中的值是否可以输出，如果输出满足条件则输出，不然将会将其拒绝输出。类似地，我们可以把这 3 个门控——输入门、输出门、遗忘门各自的功能类比为通常所说的读、写、重设，这样就可以进行更好的理解。根据研究表明，LSTM 门控中最重要的就是遗忘门（forget gate），其次重要的就是输入门（input gate），最不重要的就是输出门（output gate）。LSTM 区块的基本结构图如下图 3-1 所示：

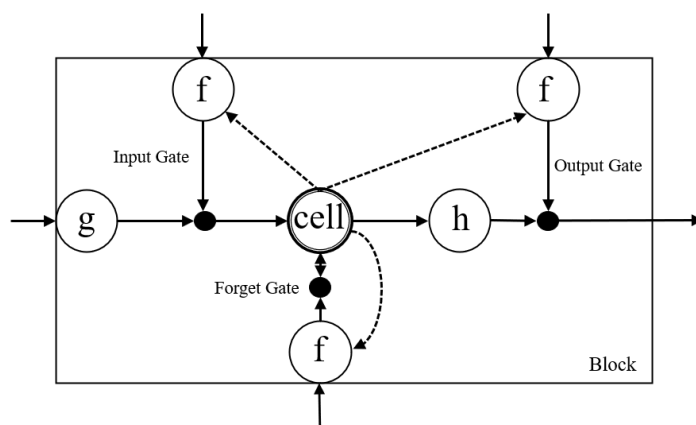


图 3-1 LSTM 区块基本结构图

#### 3.2 模型具体设计与实现

在本研究中，我们使用股票的融资融券交易信息等股票交易产生的数据预测股票的涨跌趋势，由于金融数据具有连续性，我们使用的是多天的股票交易数据来预测之后相邻两天之间的股票收盘价涨跌的趋势。假设我们使用的预测

指标共有  $N$  个，那么我们可以构造向量  $X$ ，其中  $X$  为一个  $N$  维向量， $X$  可用来表示一天的股票交易数据的  $N$  个预测指标，其中  $X(i)$  代表的是对应交易日的股票交易数据的第  $i$  项的数值。由于股票的交易数据为序列数据，那么我们需要用  $T$  天的数据来预测一天的涨跌趋势，也就是说我们模型中的时间步 (timestep) 设置的大小为  $T$ 。这样就构造出来了一个  $T*N$  大小的矩阵  $V$ ，其中  $V(i,j)$  就代表该对应的时间段内第  $i$  天的第  $j$  项股票交易数据的数值， $V(i,)$  则代表该时间段内第  $i$  天的股票交易数据，而  $V(,j)$  代表该时间段内第  $j$  项交易数据的所有交易日数据。一个矩阵  $V$  就是我们 LSTM 模型的一个输入值 (input)。而我们要预测的内容则是股票的收盘价涨跌趋势，这里涨用 1 表示，不涨用 0 来表示，我们构造向量  $Y$ ， $Y$  为 2 维向量，若  $Y=(1,0)$  则代表涨，若  $Y=(0,1)$  则代表不涨，也就是说我们的 LSTM 模型的一个输出值 (output) 就是一个  $Y$  向量。这里我们构造的输入数据就类似于计算机网络中的滑动窗口，输入值的滑动学习窗口构造如图 3-2 所示：

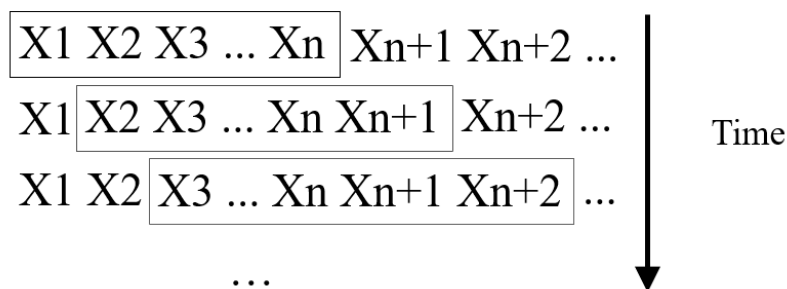


图 3-2 滑动学习窗口

在本研究中，我们使用的股票交易数据共 14 项，分别为：7 列融资融券交易数据——融资余额、融资买入额、融资偿还额、融券余额、融券卖出额、融券偿还额、融资融券余额；大盘指数数据——深证成指的收盘价；8 列各股数据——每日开盘价、每日收盘价、每日最高价、每日最低价、每日交易量、每日交易金额。设置的时间步 (timestep) 为 8，也就是说我们用 8 天的 14 项股票交易数据来预测之后两天相对的股票涨跌趋势。由于前一天融资融券交易数据在第二个交易日的上午 10 时左右更新，因此，本研究使用的是前 8 天的融资融券等股票交易数据来预测下一个交易日的收盘价相对于当日收盘价的涨跌的趋势。

在本次研究当中，我们设计的这个模型是将 3 个 LSTM 神经网络层叠放在

一起，使得该模型能够学习到更高层次的时间表示。其中前两个 LSTM 神经网络返回的是完整的输出序列，但最后一个 LSTM 神经网络只返回出输出序列的最后一步，这样就可以降低时间维度，即将输入序列转换成了单个的 2 维向量。LSTM 神经网络具体的设计如图 3-3 的 LSTM 结构设计图所示：

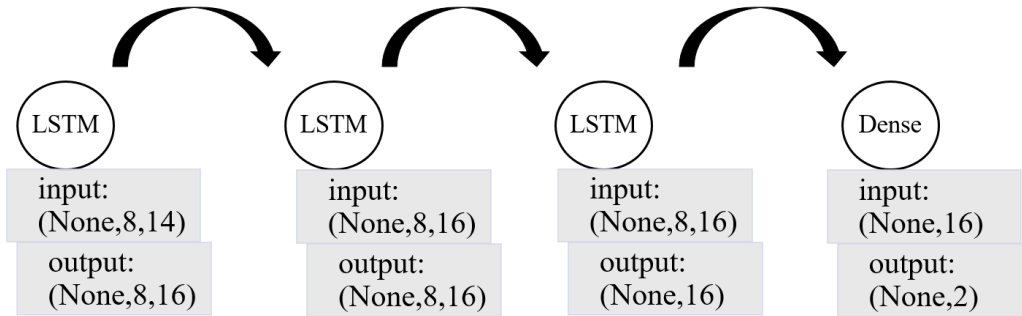


图 3-3 LSTM 结构设计图

至此，LSTM 神经网络模型的理论设计部分基本完成，数据格式的介绍也基本完成了，已经具备了进行模型具体训练的基础，使用 python 编程语言对设计好的模型编写并使用数据进行训练就可以了。

### 3.3 本章小结

本部分首先介绍了 LSTM 神经网络的基本内容，介绍了 LSTM 区块的相关知识，详细介绍了有关 3 个门控，即输入门、遗忘门、输出门 3 个门控的相关作用与相关内容。接下来，又介绍了有关本研究中数据格式的相关问题，介绍了 LSTM 模型的输入值与输出值的内容与构成，介绍了滑动学习窗口的构造过程，介绍了用前 8 天的股票交易数据来预测第二天相对当天的股票收盘价的变化趋势的原因与具体内容。在这部分的最后，又介绍了使用的 LSTM 层数以及 LSTM 层累积叠加的问题等等。这一章节介绍了逻辑上 LSTM 模型的实现，下面的章节我们将会介绍使用我们设计的 LSTM 模型对测试股票实证分析的结果与相关的分析内容，以及对本实验的总结与思考展望的相关内容。

## 第4章 实证分析与结果

### 4.1 实验环境

本实验分为线上线下两部分，线上部分的实验环境为聚宽（JoinQuant）平台在线运行，在线平台运行时将会使用 GPU 运行深度学习模型；线下部分为本地的计算机，实验环境为：编程环境为 Spyder3；处理器为 Intel(R) Core(TM) i7-6500U CPU @ 2.50GHz, 2592Mhz, 2 个内核，4 个逻辑处理器；内存为 8GB；操作系统为 64 位 win10 操作系统。本实验使用的所有的编程实现语言都是 Python3，本地 python 版本为 python 3.7.4。使用的深度学习编写框架是 Keras，版本号为 Keras 2.3.1。而 Keras 使用的后台使用的是 Theano，版本号为 Theano 1.0.4。

此外，本实验使用到的 python 扩展程序库包括 numpy、pandas、Keras 以及 jqdatasdk 等等。

### 4.2 数据收集与数据预处理

本实验的数据来源于聚宽（JoinQuant）平台所提供的数据集 API——jqdatasdk，在本地登录之后，通过在本地调用数据集的对应函数，可以将在线数据集中的股票交易数据等信息读取到 pandas 的 DataFrame 数据结构当中来，然后把需要的所有信息挑选整合到一起，并进行进一步处理，最后存到 CSV 文件当中。

本实验的数据收集工作主要是收集训练数据，训练数据从备选股票池当中选择，备选股票池为自 2015 年 1 月 1 日以来都存在融资融券交易的一千多只沪深 A 股股票，选择的输入数据为股票的融资融券交易信息等股票交易数据。训练数据共包括 100 只自从 2015 年 1 月 1 日之后存在融资融券交易数据的股票，首先是把自从 2015 年以来都有融资融券交易数据的股票放到备选股票池当中，训练数据从备选股票池中随机选取 100 只，而测试数据同样是从备选股票池中选择，测试数据从备选股票池中除训练数据之外的其余股票中随机选取，随机选取 150 只股票共计分为 3 组，每组 50 只，构成 3 个测试数据集，以备回测之用。训练数据选取的时间段为 2015 年 1 月 1 日到 2020 年 2 月 20 日共计 1249 个交易日的的数据。而训练数据共有 14 列，分别为：7 列融资融券交易数据——

融资余额、融资买入额、融资偿还额、融券余额、融券卖出额、融券偿还额、融资融券余额；大盘指数数据——深证成指的收盘价；8 列各股数据——每日开盘价、每日收盘价、每日最高价、每日最低价、每日交易量、每日交易金额。分析至此，数据收集工作就已经基本完成了，我们就是要将上述的 14 列 1249 个交易日的数据选择出来并且整合到一起，按股票的不同按顺序存放到 DataFrame 当中去，其中每一支股票的数据交易都由 1249 个交易日的数据构成，DataFrame 当中共计 100 个 1249 行的数据。

接下来要对 DataFrame 当中的数据进行数据预处理，数据预处理的工作主要有两个，一个就是处理缺失值，另一个为数据归一化。首先来看看缺失值的处理，缺失值的处理采取的方式是直接删除对应行，不过我们通过 jqdatasdk 获取到的数据不存在任何缺失值，这也为我们后续的试验带来了便利。接下来说数据归一化，数据归一化是为了统一量纲，使得模型输入更加合理，训练与预测结果更加准确，同时还可以提高模型收敛的速度，归一化的目标就是把数据变为 (0,1) 之间的小数。我们采取的归一化的方法是 min-max 归一化的方式，具体公式如下公式 4-1 所示：

$$x^* = \frac{x - \min}{\max - \min} \quad (4.1)$$

图 4-1 min-max 归一化公式

数据归一化之后，就可以把数据转存到 CSV 文件当中去了，本地的 CSV 文件将会留作下一步来使用。

上述的数据是训练数据的输入部分，输出部分亦即所谓的标签是通过将选择的每一支股票连续两天的收盘价进行比较运算得出的。若第一天的收盘价低于第二天，说明收盘价涨了，对应为“1”，不然说明没涨，对应为“0”。通过比较运算，我们把得出来的 0/1 标注就可以写入到 CSV 文件当中了，存放的顺序要与训练数据输入部分对应的顺序相同。当然，这里存在一个需要注意到的小问题，那就是每一支股票开始时也就是每一支股票 2015 年 1 月 1 日对应的标签的问题，这里为了统一，将其一律设置为“0”。其实，由于时间步（timestep）的存在，这里第一行数据在最后训练的时候并不能用到。

由于聚宽平台程序运行时间的限制，应当把训练数据的输入部分进行进一步的处理，具体就是把数据处理为带有时间步（timestep）的数据，也就是把数

据的每 8 行放在一起，当然这样也就产生了很多的重复行，当为了加速在线平台的运行速度却不得不这么做。然后再把数据整合到一行，也就是每 14\*8 列的数据是一个组，最后在聚宽平台上我们调用一个重新塑形的函数就可以将数据整合回带时间步（timestep）的 2 维数据以用作训练。

对于测试数据，在数据收集阶段要将其取出并且存到 CSV 文件当中去，所以我们在备选股票池除了训练数据以外的股票当中随机选择 3 组 50 只股票的股票代码存到 CSV 文件中以备回测时使用。

下面的表 4-1 将会展示本次研究所使用到的训练数据的 100 支股票的相关信息：

表 4-1 训练数据股票信息

股票代码	股票名称
000001.XSHE	平安银行
000002.XSHE	万科 A
000006.XSHE	深振业 A
000009.XSHE	中国宝安
000012.XSHE	南玻 A
000027.XSHE	深圳能源
000028.XSHE	国药一致
000031.XSHE	大悦城
000039.XSHE	中集集团
000046.XSHE	泛海控股
000049.XSHE	德赛电池
000050.XSHE	深天马 A
000060.XSHE	中金岭南
000061.XSHE	农产品
000062.XSHE	深圳华强
000063.XSHE	中兴通讯
000069.XSHE	华侨城 A
000078.XSHE	海王生物

000088.XSHE	盐田港
000089.XSHE	深圳机场
000099.XSHE	中信海直
000100.XSHE	TCL 科技
000151.XSHE	中成股份
000156.XSHE	华数传媒
000157.XSHE	中联重科
000333.XSHE	美的集团
000338.XSHE	潍柴动力
000400.XSHE	许继电气
000401.XSHE	冀东水泥
000402.XSHE	金融街
000413.XSHE	东旭光电
000415.XSHE	渤海租赁
000417.XSHE	合肥百货
000423.XSHE	东阿阿胶
000425.XSHE	徐工机械
000501.XSHE	鄂武商 A
000503.XSHE	国新健康
000513.XSHE	丽珠集团
000516.XSHE	国际医学
000525.XSHE	红太阳
000528.XSHE	柳工
000536.XSHE	*ST 华映
000537.XSHE	广宇发展
000538.XSHE	云南白药
000539.XSHE	粤电力 A
000540.XSHE	中天金融
000541.XSHE	佛山照明



000543.XSHE	皖能电力
000550.XSHE	江铃汽车
000551.XSHE	创元科技
000552.XSHE	靖远煤电
000554.XSHE	泰山石油
000559.XSHE	万向钱潮
000563.XSHE	陕国投 A
000568.XSHE	泸州老窖
000581.XSHE	威孚高科
000592.XSHE	平潭发展
000596.XSHE	古井贡酒
000598.XSHE	兴蓉环境
000607.XSHE	华媒控股
000616.XSHE	海航投资
000623.XSHE	吉林敖东
000625.XSHE	长安汽车
000630.XSHE	铜陵有色
000631.XSHE	顺发恒业
000650.XSHE	仁和药业
000651.XSHE	格力电器
000661.XSHE	长春高新
000667.XSHE	美好置业
000671.XSHE	阳光城
000680.XSHE	山推股份
000685.XSHE	中山公用
000686.XSHE	东北证券
000690.XSHE	宝新能源
000709.XSHE	河钢股份
000712.XSHE	锦龙股份

000718.XSHE	苏宁环球
000725.XSHE	京东方 A
000728.XSHE	国元证券
000729.XSHE	燕京啤酒
000731.XSHE	四川美丰
000732.XSHE	泰禾集团
000738.XSHE	航发控制
000739.XSHE	普洛药业
000750.XSHE	国海证券
000758.XSHE	中色股份
000761.XSHE	本钢板材
000762.XSHE	西藏矿业
000768.XSHE	中航飞机
000776.XSHE	广发证券
000777.XSHE	中核科技
000778.XSHE	新兴铸管
000783.XSHE	长江证券
000786.XSHE	北新建材
000789.XSHE	万年青
000790.XSHE	华神科技
000793.XSHE	华闻集团
000800.XSHE	一汽轿车
000801.XSHE	四川九洲
000807.XSHE	云铝股份

### 4.3 模型训练与预测结果

模型的训练工作是在聚宽平台上进行的,使用聚宽平台上提供的 GPU 可以加速模型训练的进度。训练模型之前需要把本地获取到的数据集上传上去,上传到平台上的数据集包括训练数据集的输入部分、训练数据集的输出部分和 3 个测试数据的股票代码等 CSV 文件。然后要读取文件,由于训练数据的输入部

分在 CSV 文件当中为一列数据,因此在使用数据对 LSTM 模型进行训练之前还需要进行重塑的操作,重塑工作使用 pandas 库自带的 reshape 函数即可完成。

接下来看看我们的 LSTM 模型所使用的一些参数,上一章已经提到了,我们设计的这个模型是将 3 个 LSTM 神经网络层叠放在一起的,而每一个 LSTM 神经网络层都是由 16 层构成的,最后一层输出预测的结果,而模型的优化器 (optimizer) 使用的是 rmsprop 函数,损失函数 (loss) 使用的是 binary\_crossentropy 函数,激活函数 (activation) 使用的是 softmax 函数。根据使用的参数以及函数我们可以知道,我们预测输出的结果将会是二维向量两个值对应为 0/1 的概率。

在程序读取了测试数据集的股票代码之后,很显然需要再次构造出来 14 列的输入数据,这 14 列数据与之前所提到的测试数据使用的股票交易数据是相同的。特别注意的是,这 14 列数据列的顺序要与训练数据列的顺序也相同。构造好了 14 列的数据之后,我们要把相邻 8 个交易日的数据作为一组,把这个 8\*14 矩阵的测试数据作为输入,输入到训练好的 LSTM 模型当中去,这样就可以得出来二维向量两个值对应为 0/1 的概率。我们可以认为若预测出的结果第一个小数(概率)大于第二个则是 (1,0),此预测表示将会有上涨的趋势,同样,若测出的结果第一个小数(概率)小于第二个则是 (0,1),此预测表示将会有下跌的趋势。这样,使用 LSTM 模型预测股票涨跌的趋势的工作就完成了。由于交易手续费需要考虑,购买股票越频繁交易的手续费也将会需要越多,因此应该选择最有可能增长的股票进行购买,可以限制每天最多可买入 5 支股票,而限制的条件就是使用前一交易日的交易量或者前一交易日融资余额增长的比例,具体为购买满足预测上涨的股票中前一个交易日交易量最大的 5 支股票或者购买满足预测上涨的股票中前一个交易日融资余额增加比例最大的 5 支股票。另外需要提及的是,虽然我们使用的训练数据是基于股票日收盘价的变化,但是,实际操作中我们无法真正使用收盘时间的价格进行交易,我们需要选择接近每天收盘的一个时间进行股票交易,本实验中,我设置的交易时间为 14:50,稍提前于收盘时间。

对于上述的 3 个测试数据集,仅仅使用 LSTM 深度学习模型预测基于上述两种选股策略进行交易的 5 年多(自 2015 年 1 月 12 日到 2020 年 5 月 8 日)的

运行回测的收益曲线结果分别如下表 4-2 与表 4-3 所示：

表 4-2 使用交易量选股的纯 LSTM 预测交易收益

	基准收益率	总收益率	年化收益率
测试集 1	9.97%	21.02%	3.89%
测试集 2	9.97%	22.55%	4.17%
测试集 3	9.97%	20.77%	3.85%
平均结果	9.97%	21.47%	3.97%

表 4-3 使用融资余额增长比例选股的纯 LSTM 预测交易收益

	基准收益率	总收益率	年化收益率
测试集 1	9.97%	20.32%	3.76%
测试集 2	9.97%	21.66%	4.01%
测试集 3	9.97%	19.17%	3.55%
平均结果	9.97%	20.38%	3.77%

很明显，根据上面表格中的回测结果，该策略的收益是胜过基准收益的，可以认为该模型的设计是具有一定的意义的，该策略是一种可以使用的策略，但是收益还是不够高，可以考虑进一步进行一定的改进。

#### 4.4 技术指标与具体策略

由上述实验结果可以得知，仅仅使用 LSTM 模型来预测交易股票所带来的收益并非特别的高，虽然收益胜过了基准收益，但是其收益还比不上很多的稳健型基金的收益，因此在这里考虑要将交易策略进行一定的改进操作。首先就是考虑每天都会有闲置资金这一问题，我们可以考虑把闲置的资金全部购买稳定收益的基金，银华日利（511880.XSHG）就是一只稳定型收益的基金，同时，该基金是一种 ETF 基金，其交易并没有特殊的手续费。因此，我们考虑每天 14:50 交易之后用全部闲置的资金来购买这个基金，第二天 14:50 之前再把该基金全部卖出，这样就可以实现使用闲置资金来赚取收益了。另外，由于上述单纯使用深度学习的选股策略收益不够高，我们可以考虑在深度学习预测的基础上增加技术指标加以限制，可以选择的技术指标有 MACD 与 RSI 两种，通过实验分析表明使用 MACD 的效果会更好一些。使用 MACD 限制深度学习模

型选定的股票，具体是当模型预测需要买入并且 MACD 判定为金叉时才将其买入（满足两者才买入），在决定卖出时，如果模型预测为应该卖出或者 MACD 处于死叉就将其直接卖出即可（满足一者即卖出）。除此以外，根据大量经验表明，一只股票在一天中刚开盘时的价格通常都会高于即将收盘时候的价格，因此，我们可以考虑改变股票的卖出时间，由于还要考虑融资融券交易信息更新的时间，因此我们可以考虑在下午刚开盘的时候将策略预测要跌的股票卖出，因此这里设定股票的卖出交易时间为 13:05，而股票的买入交易时间依旧不变，为 14:50。这样本研究的具体策略就设置好了。这里有一点需要提及，那就是由于 LSTM 模型预测上涨与 MACD 判定为金叉同时满足的情况不是特别多，因此不需要使用融资余额增长比例或者交易量来限制股票数量了，因为每天购买的股票都不会超过 10 支，甚至一般不超过 5 只。

#### 4.5 回测结果

使用上述设计好的策略，回测时间设定为自 2015 年 1 月 12 日到 2020 年 5 月 8 日，使用测试数据集 1 运行的收益曲线如图 4-1 所示：

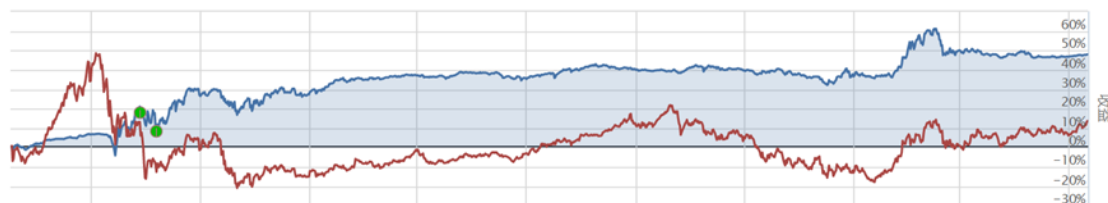


图 4-2 测试集 1 的收益曲线

使用 3 个数据集运行出来的结果分别如下表 4-4 所示：

表 4-4 使用 LSTM 模型预测与 MACD 判定的融合策略交易收益

	基准收益率	总收益率	年化收益率
测试集 1	9.97%	41.21%	7.63%
测试集 2	9.97%	43.10%	7.98%
测试集 3	9.97%	39.77%	7.36%
平均结果	9.97%	41.36%	7.70%

由此可见，改进后的新策略的收益已经可以超过大部分基金（如银华日利）的收益了，改进后的策略收益达到了改进前策略收益的几乎 2 倍，可以说该策略还算得上是比较成功的了。同时，我们也可以认为，使用股票的融资融券交易数据等股票交易数据是可以在一定程度上预测股票涨跌趋势的，也就是股票

的融资融券交易信息是有用处的。不得不说的是，本研究在实证阶段使用的是历史回测的方式，回测使用的是历史数据。为进一步探讨策略当前有效性等问题进行模拟交易是很有必要的，但是模拟交易的结果需要逐步获得，因此在本文中也就无法对其进行讨论了。聚宽平台提供延时的当前模拟交易，该策略的模拟交易已经在聚宽平台上进行运行，不过结果还尚未得到很多，所以本文在这里就不对其结果进行具体的讨论了。

#### 4.6 本章小结

在本章当中，首先讨论了有关本研究的相关内容步骤，首先是本研究使用的环境问题，研究分两部分进行，分别为线下的收据收集与处理和线上的历史回测运行。接下来叙述了有关 LSTM 模型的具体内容，展示了单独使用 LSTM 模型的预测结果进行交易的相关收益结果，并且根据交易结果对策略进行了改进，引入了技术指标对交易加以限制，使用 MACD 判断是否要买入股票，然后又展示了改进后的策略，并根据新策略测试数据，运行结果。改进后的策略实验结果表现相当不错，甚至可以优于多数的稳健型基金的收益，可以认为本策略还是一个比较成功的策略，同时也可以说明融资融券信息等股票交易数据在一定程度上可以预测出股票之后的涨跌趋势。

## 第5章 总结与展望

本研究是以量化交易以及深度学习的相关内容为基础进行的，本研究选择了融资融券交易数据等与股票交易有关的数据对股票收盘价的走势进行预测，其中使用到的预测数据包含了融资融券信息以及大盘指数、各股数据等相关的交易数据。同时为了限制每天购买的股票数量并且提高收益率而选择使用MACD来限制购买的股票数量。本研究选用了100支自2015年1月1日到2020年2月20日共计1249个交易日都存在融资融券交易数据的股票作为训练数据，又选择了3组不在训练数据当中的测试数据，每组由50支股票构成，测试时间段选择为2015年1月12日到2020年5月8日。通过本实验，我设计了深度学习与股票技术指标相结合的量化交易策略，通过在聚宽平台上进行回测，可以看出该量化交易策略的收益还是不错的，其收益可以胜过很多的稳健性基金。使用深度学习的量化交易策略是目前量化交易发展的一大趋势，同时这也是人工智能走进各行各业的一大表现。本研究所设计的量化交易策略可以考虑应用到实际股票交易当中，这个年化10%的收益率用于理财还是一个很不错的选择。当然通过本研究，我们也可以看出来根据股票的相关交易数据对股票价格变化进行预测也并非不可能，股票的交易数据在一定程度上是可以反映股票之后的价格走势的。本研究通过历史回测分析与比较，验证了本研究的思路在真实的股票市场的投资中是可以具有指导意义的。

本研究所设计的策略算得上是比较成功的，但是，在本研究当中还是存在着一些值得进一步探讨的问题，下面将其一一列出：

- 1、首先就是在本研究当中数据标注时的的问题，在实验中仅仅通过对比相邻两天收盘价的涨跌就确定0/1标签，那么，如果在标注时为两天之间的收盘价增长幅度设置一个阈值，那么预测效果是否可以更好一些呢？比如我们设置增长幅度达到30%以上的才可以买入，那么收益会不会更高一些呢？这是一个值得进一步深入研究的问题。
- 2、另外，在模型设计的时候，是否可以将模型进行进一步的优化呢？比如，如果对LSTM神经网络的层数设计以及所使用的函数等内容进一步优化会不会使得预测效果更好一些呢？这也是一个值得深入研究的问题。

- 3、除此以外，对于技术指标选择的问题，是否存在除了MACD之外的其他指标可以与LSTM模型结合使用，这种结合指标对股价涨跌进行预测会不会使收益更高一些呢？或者结合多个技术指标与LSTM模型一起对股价涨跌进行预测会不会得到更好的收益呢？这个问题也值得进行进一步的深入研究的。
- 4、另外，深度学习训练数据可以使用的内容除了那14列指标数据之外，是否还有其他交易数据也会对预测有一定的影响呢？或者把股票分行业加入行业指标会不会效果更好一些呢？这也值得进一步的研究。
- 5、有关训练数据量的问题，如果进一步扩大数据量，比如若选择500支股票的数据作为训练数据，那么模型的预测效果是否可以更好一些呢？如果选用1000支股票的数据又会如何呢？
- 6、关于时间步（timestep）的选择的问题，本研究中使用的时间步设置的是8，但是更大的时间步能不能够提高预测的准确率呢？或者说，时间步设置为多大能够实现最大的收益呢？所有的股票都会拥有相同的最优时间步吗？
- 7、除了上述问题以外，本研究所做的回测都是在与训练数据选择的时间相重叠的时间段上进行的，那如果将模型用于之后的时间段，那么该策略的预测效果又会如何呢？也就是该策略的当前可用性如何呢？所以使用本研究的策略进行模拟交易进一步探讨策略的有效性是很有必要的。

这些都是值得进一步深入研究探讨的问题，当然这些问题也可以为有关使用融资融券信息等交易数据训练深度神经网络模型来预测股票价格趋势提供更多的新思路，毕竟提出问题是解决问题的关键，提出以上的问题是为了对这个课题进一步更好地进行研究。



## 致谢

写到这里，我的毕业论文就要接近尾声了，我的毕业设计课程也基本就要完成了，我的四年本科生涯也就马上要结束了。

在这里，我首先要感谢的就是我的指导老师——王华老师。正是在王老师的指导之下，我本科期间的最后一个作业，我的毕业论文终于要完成了。由于今年疫情的特殊情况，我们都无法回到学校，但是在我做毕设的期间，王老师还是通过微信语音电话对我进行着认真的指导。王老师高超的学术造诣让我仰慕，王老师认真负责的态度让我钦佩。王老师一直在督促着我的进度，把握着我的方向，帮我明确着下一步的任务，正是王老师的指导，让我避免走弯路，让我顺利完成毕业设计。

我还要感谢我的父母与家人，从小到大，在我成长的过程中，他们给我的从来都是关心与支持，我的成长必须要感谢他们。我还要感谢我大学期间的同学，我们都相遇于山东大学，在这里，我们一起学习，一起进步，一起蜕变，一起成长，而如今却又到了要各奔东西的时候，在这里我要特别感谢他们。我还要感谢那些在我的成长过程中出现过的所有的人，我的好友，我的对手，我的师长，以及我曾爱过的与爱过我的人，我曾伤害过的与伤害过我的人，甚至那些曾在我身边不经意出现过的陌生人，我要感谢他们，正是他们的出现，他们的作用，让我成为今天的我。

我还要感谢的是我的母校——山东大学。四年时光，转瞬即逝，在山大的日子我是快乐的，我是成长的，我是进步的。我即将从山东大学毕业，但是山大人这一名号永远是我的荣耀。“凡我在处，便是山大。有你在是，那便是家。”今后我要以作为山大人作为荣耀，努力为母校添光彩，为母校增荣耀。今生既入山大门，今生不枉山大人。

## 参考文献

- [1] 王娟, 李锐. 中国股市的时变波动性——基于长记忆性、杠杆效应视角[J]. 北京航空航天大学学报(社会科学版). 2019, 32(3): 57-65, 79.
- [2] 刘晶晶, 古晨. 量化投资交易策略研究[J]. 中国市场. 2017(2): 201, 223.
- [3] 谢雨晴. 浅析金融衍生品和量化投资[J]. 北方经贸. 2019(7): 102-103, 147.
- [4] 范凯隆. 深度学习对金融市场预测的影响研究进展[J]. 现代商业. 2020(07): 118-119.
- [5] 刘毅铭. 深度学习研究与应用综述[J]. 绿色科技. 2019(11): 281-283.
- [6] 万里鹏, 兰旭光, 张翰博. 深度强化学习理论及其应用综述[J]. 模式识别与人工智能. 2019, 32(01): 67-81.
- [7] 梁天新, 杨小平, 王良. 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.
- [8] Tashiro D, Matsushima H, Izumi K, Sakaji H. Encoding of high-frequency order information and prediction of short-term stock price by deep learning[J]. Quantitative Finance. 2019, 19(9): 1499-1506.
- [9] Butter F A G, Gevel F M S. Prediction of the Netherlands' money stock[J]. De Economist. 1989, 137(2).
- [10] E. S. Stock price prediction using neural networks: A project report[J]. Elsevier. 1990, 2(1).
- [11] Jang G, Lai F, Jiang B, Parng T, Chien L. Intelligent stock trading system with price trend prediction and reversal recognition using dual-module neural networks[J]. Applied Intelligence. 1993, 3(3).
- [12] Brownstone D. Using percentage accuracy to measure neural network predictions in Stock Market movements[J]. Neurocomputing. 1996, 10(3).
- [13] W S E, V P D, C W D. Comparative study of stock trend prediction using time delay, recurrent and probabilistic neural networks.[J]. IEEE transactions on neural networks. 1998, 9(6).
- [14] Kim K, Han I. Genetic algorithms approach to feature discretization in artificial

- neural networks for the prediction of stock price index[J]. *Expert Systems With Applications*. 2000, 19(2).
- [15] Safer A M. The application of neural networks to predict abnormal stock returns using insider trading data[J]. *Applied Stochastic Models in Business and Industry*. 2002, 18(4).
- [16] Kim H, Shin K. A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets[J]. *Applied Soft Computing Journal*. 2006, 7(2).
- [17] Boyacioglu M A, Avci D. An Adaptive Network-Based Fuzzy Inference System (ANFIS) for the prediction of stock market return: The case of the Istanbul Stock Exchange[J]. *Expert Systems With Applications*. 2010, 37(12).
- [18] Ruxanda, Badea. Configuring Artificial Neural Networks for stock market predictions[J]. *Technological and Economic Development of Economy*. 2014, 20(1).
- [19] Ramalingam V V, Pandian A, Dwivedi S, Bhatt J P. Analysing News for Stock Market Prediction[J]. *Journal of Physics: Conference Series*. 2018, 1000(1).
- [20] 张悦华. BP网络在股市预测方面的应用[J]. *吉林大学自然科学学报*. 1997(01): 33-34.
- [21] 刘新勇, 贺江峰, 孟祥泽. 基于神经网络的股市预测[J]. *南开大学学报(自然科学版)*. 1998(03): 39-44.
- [22] 李波, 张世英. 基于模糊-遗传杂合算法对证券投资市场时机预测的研究(英文)[J]. *Transactions of Tianjin University*. 2000(01): 28-31.
- [23] 侯木舟, 韩旭里. 基于MATLAB的神经网络在股市预测中的应用[J]. *系统工程*. 2003(02): 112-115.
- [24] 袁婷. 神经网络在股票市场预测中的应用[J]. *软件导刊*. 2006(05): 18-19.
- [25] 刘成竹. 基于时间序列和神经网络的股票指数预测研究[D]. 华中师范大学, 2009.
- [26] 张翔宇. 基于优化BP神经网络和粒计算的股指预测研究[D]. 南京大学, 2012.

- [27] 杜晓. 基于MEA-BP模型的股票预测研究[D]. 兰州大学, 2015.
- [28] 毛景慧. 基于LSTM深度神经网络的股市时间序列预测精度的影响因素研究[D]. 暨南大学, 2017.
- [29] 季子峥, 沈婷婷, 张孝. 利用社交媒体情感分析的短期股价趋势预测方法[J]. 北京理工大学学报. 2020, 40(01): 83-89.
- [30] 郭笑宇. 量化投资交易策略研究[J]. 财经界. 2019(3): 16-17.
- [31] 贾宇杰. MACD指标在股市实战中的应用及其存在的问题研究[J]. 中国商贸. 2018(12): 34-36.
- [32] Gurrib I, Kamalov F. The implementation of an adjusted relative strength index model in foreign currency and energy markets of emerging and developed economies[J]. Macroeconomics and Finance in Emerging Market Economies. 2019, 12(2): 105-123.

## 附录 1

### 英文原文

#### LSTM with Wavelet Transform Based Data Preprocessing for Stock Price Prediction

Xiaodan Liang, Zhaodi Ge, Liling Sun, Maowei He, Hanning Chen

School of Computer Science and Technology, Tianjin Polytechnic University,  
Tianjin, 300387, China

*{sll198257@163.com}*

### Abstract

For profit maximization, the model-based stock price prediction can give valuable guidance to the investors. However, due to the existence of the high noise in financial data, it is inevitable that the deep neural networks trained by the original data fail to accurately predict the stock price. To address the problem, the wavelet threshold-denoising method, which has been widely applied in signal denoising, is adopted to preprocess the training data. The data preprocessing with the soft/hard threshold method can obviously restrain noise, and a new multi-optimal combination wavelet transform (MOCWT) method is proposed. In this method, a novel threshold-denoising function is presented to reduce the degree of distortion in signal reconstruction. The experimental results clearly showed that the proposed MOCWT outperforms the traditional methods in the term of prediction accuracy.

### Introduction

Stock price prediction is a typical problem based on time series forecasting, and various stock forecasting methods emerge in an endless stream. Stock price prediction means to predict the stock price after a certain time, so as to help investors realize the maximum benefit. The methods proposed in the literatures can be roughly divided into two categories containing traditional mathematical methods and economic methods. Many previous works are based on traditional statistical methods. Typically, Kalman filter and autoregressive model are very classic statistical methods,

which are often used for financial sequence prediction. With the development of artificial intelligence, the deep learning methods are increasingly applied to the field of stock price prediction.

The deep learning method has a very superior performance compared with the traditional statistical method. One of the main reasons is that with the direct analysis, deep learning can map the original data to a nonlinear model thereby giving a better fitting effect through the multilayer neural network. In addition, deep learning has the advantage of self-selection in the application of the financial field. Most financial data is highly noisy and unstable. With deep learning, various events that have a significant impact on finance can be expressed by knowledge maps. And then features are selected automatically through deep networks to adjust parameters and weights. The results obtained in this way may be more accurate and objective. Recurrent neural networks(RNN) have been widely used in the field of natural language processing and have achieved great success. With the Advantages of neural networks, it is possible to grasp the public sentiment more accurately. In this paper, the more powerful Long Short-Term Memory (LSTM) neural network model based on the concept of RNN is used for stock prediction applications. It can save long-term memory more effectively. The “forget gate” and “input gate” are the core of its structure. With the “gate recurrent” structure, the model can remember the effective information while forgetting the useless information. In addition, due to the huge depth of the neural network in RNN, the gradient is calculated from the higher power of the matrix, sometimes accompanied by the problem of gradient explosion or gradient disappearance. However, the LSTM model with a special “gate recurrent” structure is very powerful to avoid this phenomenon.

In this paper, the wavelet denoising method is introduced into data preprocessing. The data with wavelet preprocessed were used as the training data. Our main contributions are as follows. First, we have improved the traditional wavelet denoising method and its performance is better than the traditional wavelet denoising method. Secondly, we propose a new multioptimal combination wavelet transform (MOCWT) method. Compared with the traditional wavelet method and the

improved wavelet method, MOCWT has the best performance.

The remainder of this paper is organized as follows. Some related works are reviewed in Section 2. In Section 3, the LSTM model is utilized to handle the stock price forecasting task and a new method named multioptimal combination wavelet transform (MOCWT) is proposed for the aim of data denoising. Section 4 demonstrates the performance of the proposed MOCWT with other canonical methods. Finally, the conclusion is given in Section 5.

### Related Works

The issue of stock forecasting has been widely concerned by researchers. Various stock forecasting methods emerge in an endless stream. For example, some researchers use a news-oriented approach to predict stock trends.

Ziniu Hu et al. proposed a Hybrid Attention Networks (HAN), which takes into account the use of analytical news factors to predict a stock trend in learning framework. Xi Zhang et al. proposed that the investor social networks is also an emotion-oriented forecasting method, which utilizes emotional factors and stock correlation characteristics to model. In their further article, they considered the relationship between the single source and the multisource data and then employed the coupling matrix and tensor decomposition framework to study the impact of online news and user sentiment on stock price changes. In the research method of Huicheng Liu et al., an Attention-Based RNN was used to accomplish this task and a Bidirectional-LSTM is used to capture the characteristics of the information in the news text. Based on factors such as online news and investor sentiment to prediction, Jieyun Huang et al. designed tensors to capture the intrinsic connections of different sources of information and then used this method to solve the problem of data sparsity and finally proposed an improved submode coordinate algorithm (SMC) to match the use of tensors for improving the prediction accuracy. Marcelo Sardelich et al. also studied the relationship between news and stock prices and then predicted the volatility of the day. Some researchers adopted mathematical methods to make stock market forecasts. In the literature of Mahsa Ghorbani et al., the covariance information based on principal component analysis of dimensionality reduction

method was designed for stock forecasting, and company stocks from different industries were used to illustrate this method.

Currently, the deep learning methods to predict stock prices have become the most widely used method. Yue-Gang Song et al. compared the predictive performance of five neural network models for stock price forecasting. Hyeong Kyu Choi proposed the ARIMA-LSTM hybrid model to predict the stock price correlation coefficient. This method first employed the ARIMA model to filter out the linear trend in the data and then passed the residual value to the LSTM. FULI FENG et al. proposed a deep learning program for Relational Stock Ranking (RSR). The Temporal Graph Convolution method was brought in to simulate the temporal evolution of the stock and the relational network, and the solution was implemented using the LSTM network. In their other paper, a confrontational training method was proposed to simulate the randomness of stock price by increasing the perturbation. Its advantage is to prevent overfitting of data. Linyi Yang et al. proposed a dual-layer attention-based neural network with GRU, which adopted an input attention mechanism to reduce the noisy news and an output attention mechanism to allocate different weights to different days. Experiments showed that their method is effective.

In the case of stock forecasting, unlike the deep learning methods that have been widely used, there are a few researchers using convolutional neural network (CNN). CNN is originally used in image processing and has excellent performance. Ehsan Hoseinzade et al. used CNN to extract the correlation of multisource data to achieve stock market forecasting. Jinho Lee et al. used stock chart images as input to the model and used the Deep Q-Network and convolutional neural networks for global stock market forecasting. There are also some researchers who use other methods to study the stock market. Arjun R et al. proposed a triangulation qualitative reasoning to explore the strength of causal relationship between stock price search interest and real stock market outcomes on worldwide equity market indices.

#### Conclusion and Future Work

In this paper, we improved the original wavelet denoising method. The



performance of improved method is better than traditional methods.

In addition, we propose a new multioptimal combination wavelet transform (MOCWT) method, the experimental results of which show that its performance is the best compared with the traditional wavelet method. The proposed MOCWT is obviously superior to the traditional method and the improved method. For the original data method without wavelet processing, the prediction results will always have a large oscillation, and the fitting effect of the real data is poor. The overall performance of the model is affected. The experimental results illustrate that the data characteristics are of great significance to the performance of the whole model.

In this work, there still are some key experimental points worth exploring in the future. For example, the optimization method and structure of the neural network and the loss function, as well as the parameter variables in the experiment, are all worth further optimizing.

## 附录 2

### 译文

基于小波变换的 LSTM 股价预测数据预处理

Xiaodan Liang, Zhaodi Ge, Liling Sun, Maowei He, Hanning Chen

天津工业大学计算机科学与技术学院, 天津, 300387

{sll198257@163.com}

### 摘要

为了实现利润最大化, 基于模型的股票价格预测可以为投资者提供有价值的指导。然而, 由于财务数据的高噪声存在, 由原始数据训练的深层神经网络无法准确预测股票价格是不可避免的。针对这一问题, 采用小波阈值去噪方法对训练数据进行预处理, 并在信号去噪中得到了广泛的应用。采用软/硬阈值方法对数据进行预处理, 能明显抑制噪声, 提出了一种新的多最优组合小波变换 (MOCWT) 方法。在该方法中, 提出了一种新的阈值去噪函数来降低信号重建中的失真度。实验结果表明, 该方法在预测精度方面优于传统方法。

### 介绍

股票价格预测是基于时间序列预测的典型问题, 各种股票预测方法层出不穷。股价预测是指在特定时间对股价进行预测, 以帮助投资者实现利益最大化。文献中提出的方法大致可分为传统数学方法和经济方法两大类。以往的许多工作都是基于传统的统计方法。通常, 卡尔曼滤波和自回归模型是非常经典的统计方法, 常用于金融序列预测。随着人工智能技术的发展, 深度学习方法越来越多地应用于股票价格预测领域。

与传统的统计方法相比, 深度学习方法具有非常优越的性能。其中一个主要原因是通过直接分析, 深度学习可以将原始数据映射到非线性模型, 从而通过多层神经网络给出更好的拟合效果。此外, 深度学习在金融领域的应用具有自我选择的优势。大多数金融数据都是非常嘈杂和不稳定的。通过深入学习, 可以用知识地图来表达对股票有重大影响的各种事件。然后通过深度网络自动选择特征, 调整参数和权重。用这种方法得到的结果可能更加准确和客观。递归神经网络 (RNN) 在自然语言处理领域得到了广泛的应用, 并取得了巨大的

成功。利用神经网络的优势，可以更准确地把握公众情绪。本文将基于 RNN 概念的更强大的长短期记忆（LSTM）神经网络模型用于股票预测应用。它能更有效地保存长期记忆。“遗忘门”和“输入门”是其结构的核心。该模型采用“门递归”结构，能够在遗忘无用信息的同时，记忆有效信息。另外，由于神经网络在 RNN 中的深度很大，因此梯度是由矩阵的高阶次计算出来的，有时会伴随着梯度爆炸或梯度消失的问题。然而，具有特殊“门递归”结构的 LSTM 模型对于避免这种现象是非常有效的。

本文将小波去噪方法引入到数据预处理中。将小波预处理后的数据作为训练数据。我们的主要贡献如下。首先，对传统的小波去噪方法进行了改进，其性能优于传统的小波去噪方法。其次，提出了一种新的多最优组合小波变换（MOCWT）方法。与传统小波方法和改进小波方法相比，MOCWT 具有最好的性能。

本文的其余部分安排如下。第 2 部分对相关工作进行了回顾。第 3 部分利用 LSTM 模型处理股票价格预测问题，提出了一种基于多最优组合小波变换（MOCWT）的数据去噪方法。第 4 部分用其他规范方法演示了所提出的 MOCWT 的性能。最后，第 5 部分给出了结论。

## 相关工作

股票预测问题一直受到研究者的广泛关注。各种股票预测方法层出不穷。例如，一些研究人员使用面向新闻的方法来预测股票趋势。

Ziniu Hu 等人提出了一种混合注意网络（HAN），它在学习框架中考虑了利用分析性新闻因素来预测股票走势。Xi Zhang 等人提出了投资者社交网络也是一种以情绪为导向的预测方法，它利用情绪因素和股票相关特性进行建模。在他们的后续文章中，他们考虑了单一来源和多源数据之间的关系，然后利用耦合矩阵和张量分解框架来研究在线新闻和用户情绪对股价变化的影响。在刘慧成等人的研究方法中，采用基于注意的 RNN 来完成这一任务，并采用双向 LSTM 来捕捉新闻文本中信息的特征。基于网络新闻和投资者情绪等因素进行预测，黄洁云等人设计张量来捕捉不同信息源之间的内在联系，然后利用这种方法解决数据稀疏的问题，最后提出一种改进的子模式坐标算法（SMC）来匹配张量的使用以提高预测精度。Marcelo Sardelich 等人同时研究了新闻与股价的

关系，并预测了当天的波动性。一些研究人员采用数学方法对股市进行预测。在 Mahsa-Ghorbani 等人的文献中，设计了基于主成分分析的降维协方差信息用于股票预测，并用不同行业的公司股票对该方法进行了说明。

目前，利用深度学习方法预测股票价格已成为应用最为广泛的方法。宋岳刚等人比较了五种神经网络股价预测模型的预测性能。崔贤圭提出 ARIMA-LSTM 混合模型来预测股价相关系数。该方法首先利用 ARIMA 模型滤除数据中的线性趋势，然后将残值传递给 LSTM。冯富力等人提出了一个用于关系型股票排序（RSR）的深度学习程序。引入时间图卷积方法模拟股票和关系网络的时间演化，并利用 LSTM 网络实现求解。在他们的另一篇论文中，提出了一种对抗性训练方法，通过增加扰动来模拟股票价格的随机性。它的优点是防止数据过度拟合。Linyi Yang 等人提出了一种基于 GRU 的双层注意神经网络，该网络采用输入注意机制减少噪声，输出注意机制对不同的时间分配不同的权重。实验表明，该方法是有效的。

就股票预测而言，与广泛应用的深度学习方法不同，目前有少数研究者使用卷积神经网络（CNN）。CNN 最初用于图像处理，具有优异的性能。Ehsan Hoseinzade 等人利用 CNN 提取多源数据的相关性，实现股市预测。Jinho Lee 等人以股价图图像为输入，采用深度 Q 网络和卷积神经网络进行全球股市预测。也有一些研究者用其他方法来研究股票市场。Arjun R 等人提出了一种三角定性推理方法，以探讨全球股票市场指数上股价搜索兴趣与实际股市结果之间因果关系的强度。

## 结论和今后的工作

本文对原有的小波去噪方法进行了改进。改进后的方法性能优于传统方法。

此外，我们还提出了一种新的多最优组合小波变换（MOCWT）方法，实验结果表明其性能优于传统小波方法。该方法明显优于传统方法和改进方法。对于未经小波处理的原始数据方法，预测结果往往会有较大的振荡，且实际数据拟合效果较差。模型的整体性能受到影响。实验结果表明，数据特征对整个模型的性能具有重要意义。

在这项工作中，还有一些关键的实验点值得我们进一步探索。例如，神经网络的优化方法、结构和损失函数等，以及实验中的参数变量，都是值得进一

步优化的。