

中国科学技术大学

本科毕业论文



基于深度学习的股票价格水平预测

作者姓名：	马子睿
学 号：	PB20111623
专 业：	计算机科学与技术
导师姓名：	李勇 副教授
完成时间：	2024 年 6 月 7 日

摘要

深度学习是一种基于人工神经网络的机器学习方法，具有强大的数据处理和模式识别能力。近年来，深度学习在金融领域的应用日益广泛，其中一个重要应用就是股票价格预测。

股票市场的数据通常具有高度的非线性和噪声，传统的统计模型在处理这种复杂数据时可能表现不佳。深度学习模型，如长短期记忆网络（LSTM）和卷积神经网络（CNN），可以通过学习历史数据中的复杂模式和特征，提高预测的准确性。本论文主要探讨了长短期记忆网络（LSTM）以及超参数优化中的鲸鱼算法在股票预测中的应用。

首先，本文介绍了股票预测的背景和相关研究现状，以及 LSTM 神经网络和鲸鱼算法的基本原理。其次，本文介绍了 LSTM 网络在股票预测中的应用，并介绍了 LSTM 网络调参的重要性。然后，本文介绍了鲸鱼算法的基本原理和优化过程，并将鲸鱼算法应用于 LSTM 网络的超参数优化中。最后，本文通过引入混沌理论、非线性因子和最优抖动等策略对鲸鱼算法进行了改进，提高了算法的性能和稳定性。

实验结果表明，改进后的鲸鱼算法在股票预测中取得了较好的效果，验证了其在超参数优化中的有效性和可行性。

关键词：股票预测；鲸鱼算法；长短期记忆网络

ABSTRACT

Deep learning is a machine learning method based on artificial neural networks, known for its powerful data processing and pattern recognition capabilities. In recent years, deep learning has been increasingly applied in the financial sector, with one significant application being stock price prediction.

Stock market data typically exhibit high levels of nonlinearity and noise, which traditional statistical models may struggle to handle effectively. Deep learning models, such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), can enhance prediction accuracy by learning complex patterns and features from historical data.

This paper primarily explores the application of LSTM networks and the Whale Optimization Algorithm (WOA) in hyperparameter optimization for stock prediction.

First, the paper introduces the background and current state of research on stock prediction, as well as the basic principles of LSTM neural networks and the Whale Optimization Algorithm. Second, the paper discusses the application of LSTM networks in stock prediction and emphasizes the importance of tuning LSTM network parameters. Then, it explains the basic principles and optimization process of the Whale Optimization Algorithm and its application in the hyperparameter optimization of LSTM networks. Finally, the paper enhances the Whale Optimization Algorithm by incorporating strategies such as chaos theory, nonlinear factors, and optimal jitter to improve the algorithm's performance and stability.

Experimental results show that the improved Whale Optimization Algorithm achieves better results in stock prediction, validating its effectiveness and feasibility in hyperparameter optimization.

Key Words: Stock Prediction; Whale Optimization Algorithm; Long Short-Term Memory Network

目 录

第一章 引言	4
第一节 研究背景与研究意义	4
一、研究背景	4
二、研究目的	5
三、研究意义	5
第二节 领域研究现状	6
一、股票预测的有效性	6
二、股票预测的方法	6
第三节 本文的主要工作	8
第二章 传统股票预测方法	9
第一节 概述	9
第二节 基本面分析	9
一、公司财务分析	9
二、行业和竞争分析	10
三、宏观经济分析	10
第三节 股票预测的时间序列分析	10
一、移动平均模型	10
二、自回归模型（AR）	11
三、差分自回归移动平均模型（ARIMA）	11
第三章 数据选取与预处理	12
第一节 数据集简介	12
第二节 数据指标	12
第三节 数据预处理	12
一、数据清洗	12
二、数据标准化	13
三、数据集划分	14

第四章 LSTM 模型设计	15
第一节 LSTM 模型的基本原理	15
第二节 LSTM 模型的构建	16
一、LSTM 模型的工作流程	16
二、LSTM 模型的超参数	17
三、模型的优化器和调度器	17
第三节 LSTM 模型预测结果分析	18
一、训练损失收敛情况	18
二、预测结果分析	19
第四节 本章小结	20
第五章 基于鲸鱼优化算法的超参数调优	21
第一节 超参数调优概述	21
第二节 鲸鱼优化算法	21
一、启发式来源	21
二、基本原理	22
第三节 基于鲸鱼优化算法的改进 LSTM 预测	24
一、鲸鱼优化算法的优化对象	24
二、算法优化流程	24
三、实验结果分析	24
第四节 本章小结	25
第六章 鲸鱼优化算法的改进策略	27
第一节 鲸鱼算法的问题分析	27
第二节 混沌策略初始化种群选择	27
一、初始化种群选择问题	27
二、Tent 混沌序列生成初始化种群	27
第三节 非线性因子的引入	28
一、递减收敛因子的非线性调整	29
二、螺旋形状因子的非线性调整	29
第四节 自适应权重与最优局部抖动	30
一、自适应权重	30

二、最优局部抖动	31
第五节 实验结果	32
一、训练损失收敛情况	32
二、预测结果对比	32
第六节 本章小结	32
第七章 成果总结与展望	34
第一节 成果总结	34
第二节 未来展望	34
参考文献	36
致谢	38

第一章 引言

第一节 研究背景与研究意义

一、研究背景

近年来,随着我国资本市场全面深化改革的推进,股票市场越来越被我国投资者关注。2014 年,我国股民数量为 0.7 亿,而到了 2023 年,这个数字已经远超 2 亿,也就是说,我国每 7 个人中就存在一位股民。股票市场不仅是我国资本市场的重要组成部分,也是我国经济发展的重要指标之一。因此,对股票市场的研究和预测一直是金融领域的热门研究方向。

在早期的股票预测中,主要采用技术分析和基本面分析^[1]等方法,这些方法主要依赖于历史数据和财务数据等信息,对未来的股票价格走势进行预测。技术分析主要是通过研究股票的历史价格走势和成交量等信息,来预测未来的价格走势。基本面分析则是通过研究公司的财务报表、经营状况等信息,来评估公司的价值和未来的发展前景。

然而,这些传统的方法存在很大的局限性。首先,股票价格的走势会受到很多因素的影响,包括宏观经济环境、政策法规、国际形势等,这些因素很难通过技术分析和基本面分析等方法来准确预测。其次,传统的方法主要依赖于历史数据和财务数据等信息,对未来的预测能力有限,很难适应市场的快速变化和复杂性。最后,传统的股票预测很难不掺杂预测者的主观因素,很难做到客观、准确的预测。

随着量化金融分析和人工智能技术的发展,越来越多的研究者开始尝试利用深度学习等技术来预测股票价格^[2]。与传统的预测技术相比,深度学习技术具有更强的数据处理能力和模式识别能力,能够更好地挖掘数据之间的关系,提高预测的准确性和稳定性。其中,神经网络作为深度学习的代表技术,不仅可以模仿人类大脑的工作方式来实现复杂的模式识别,还可以通过大量的数据训练来提高预测的准确性,这对于纷繁复杂、噪声干扰较大的股票市场来说,具有很大的优势。

在股票预测领域,长短期记忆网络(LSTM)是一种特殊的循环神经网络,具有很强的记忆能力和时间序列建模能力,被广泛应用于股票预测、自然语言处理等领域。YanQi Feng 等人(2015)^[3]使用包括 LSTM 的深度学习对股票价格进行

预测的效果，并使用了数值和文本信息比较了不同模型的预测性能，发现 LSTM 预测的误差较小，展现了 LSTM 在股票预测方面的较大优势。

二、研究目的

股票市场的变化与动荡通常受到了多方面因素的影响，这些因素的作用通常不仅仅是简单的叠加关系，这使得传统方法预测股票价格是十分困难的。通过结合机器学习技术，特别是深度学习技术，可以更好地挖掘数据之间的关系，提高预测的准确性和稳定性。但是，由于我们无法精准的通过先验知识来指导深度学习模型的训练，因此基础深度学习模型的训练效果往往不尽如人意。本文首先探索 LSTM 模型在股票预测中的应用，然后通过引入超参数调整算法，根据股票市场特征进行超参数调整，提高 LSTM 模型的预测准确性，从而帮助投资者更好地进行股票投资。

三、研究意义

随着股票市场的发展和投资者的增多，股票市场的波动性和不确定性也越来越大，投资者需要更加准确的预测方法来指导投资决策。通过对传统机器学习模型进行自动化参数改良^[4]，可以结合股票市场的特征，提高模型的预测准确性，这在机器学习理论领域具有一定的研究意义。同时，通过提高股票预测的准确性，可以帮助投资者更好地进行投资决策，提高投资的收益率，这在实际应用中具有一定的实用价值。

此外，股票的高效预测对现实社会生活也有积极的意义。金融模型的建立和及时更新对个人投资者、政府部门和金融机构等都具有重要的意义。

1. 对于我国规模相当庞大的个人投资者来说，随着股票市场越来越复杂，传统的预测方法对于个体投资者而言成本往往过高。通过使用深度学习技术，可以更好地挖掘数据之间的关系，提高预测的准确性和稳定性，帮助投资者更好地进行投资决策，提高投资的收益率。

2. 对于政府部门来说，股票市场的稳定发展是我国经济发展的重要组成部分，股票的搞笑预测可以帮助政府更好、更及时地监管股票市场，防范金融风险，促进资本市场的稳定发展。

3. 对于金融机构来说，股票市场价格的不确定性往往会对金融机构的经营和管理带来一定的挑战，通过提高股票预测的准确性和及时性，可以帮助金融机构更好地管理风险，提高盈利能力，促进金融机构的稳健发展。

第二节 领域研究现状

一、股票预测的有效性

一直以来，人们关于股票市场是否可预测的争论一直存在两个相反的观点：一方认为股票市场是随机的，不可预测；另一方认为股票市场是有规律的，是可以预测的。而近些年，随着机器学习和人工智能技术的兴起，虽然有效市场假说一直被公认，但是越来越多的研究者开始尝试利用深度学习等技术来预测股票价格，取得了一定的成果。

众多学者和研究人员在这方面付出了很多努力，Daniel 和 Hirshleifer (1998)^[5]认为投资者情绪和行为可能导致市场价格的偏离，提供了股票市场可预测性的相关证据；D.Lo 和 A.W. (2004)^[6]提出了自适应市场假说，认为市场在演化的过程中会产生一定的可预测性，并通过进化理论解释了市场的非理性行为；Ang 等人 (2012)^[7]则是探讨了不同市场制度和环境下股票市场的行为变化，从而对股票市场的可预测性提出了一些新的思考。由此可知，市场的可预测性虽然是一个复杂的问题，但是通过深入研究和分析，我们依然可以找到一些规律和方法来对其进行预测。

而在我国，虽然量化金融分析起步较晚，但依旧有不少研究人员借助先进技术对股票市场进行了深入研究。朱倩、陈浩等人 (2015)^[8]对 GARCH 模型和 SVR 模型进行了研究，并基于中国股票市场给出了具体的预测结果和分析；孙婧等人 (2019)^[9]利用马尔可夫模型对中国股票市场的短期预测进行了研究，并分析了预测结果的准确性和稳定性；韩思远等人 (2020)^[10]更利用中国 A 股市场的交易数据对股票价格波动率进行了预测，研究发现中国 A 股存在一定的可预测性，可以通过合适的模型进行预测。以上研究均能表明，股票市场的变化并非无迹可寻，通过合适的模型和方法，我们依然可以对股票市场进行一定程度的预测。

二、股票预测的方法

1. 基本面和技术分析

基本面分析和技术分析是基于传统经济学理论的股票预测方法。基本面分析融合了微观经济学、宏观经济学、公司金融等多个学科的知识，结合 ROA（净资产收益率）、ROE（净资产收益率）、EPS（每股收益）等财务指标，通过分析公司的财务报表、经营状况等信息，来评估公司的价值和未来的发展前景^[11]。技术分析则是通过研究股票的历史价格走势和成交量等信息，来预测未来的价格

走势，主要包括 K 线图、均线、MACD 等技术指标。基本面分析和技术分析是传统的股票预测方法，虽然在一定程度上可以帮助投资者进行投资决策，但是其预测能力有限，很难适应市场的快速变化和复杂性。

2. 统计和计量方法

统计和计量方法是一种基于历史数据和数学模型的股票预测方法，主要包括时间序列分析、回归分析、协整分析等方法。其中，时间序列分析是通过分析股票的历史价格走势和成交量等信息，来预测未来的价格走势，主要包括 ARIMA 模型^[12]、GARCH 模型^[13]等。回归分析^[14]则是通过分析股票的历史价格和财务数据等信息，来评估股票的价值和未来的发展前景，主要包括线性回归、逻辑回归等。协整分析则是通过分析股票之间的协整关系，来预测未来的价格走势，主要包括协整检验、误差修正模型等。在 21 世纪初期，统计和计量方法确实在股票预测领域小有成就，但是随着股票市场的快速变化和复杂性，这些方法的预测能力逐渐受到了挑战。

3. 机器学习与深度学习

机器学习方法在 20 世纪 50 年代末被提出，它使用了大量的股票数据作为输入，并通过特定的训练算法，指导模型学习数据之间的关系，从而实现对股票价格的预测。机器学习方法主要包括支持向量机、随机森林、神经网络等。其中，神经网络作为深度学习的代表技术，不仅可以通过模仿人类大脑的工作方式来实现复杂的模式识别，还可以通过大量的数据训练来提高预测的准确性。传统的机器模型受限于模型复杂度、训练成本等因素，其挖掘数据的能力较差，不能很好地适应股票市场的快速变化和复杂性。

为了解决传统机器学习的弊端，深度学习在 21 世纪初被提出，它通过多层神经网络来实现对数据的高阶抽象和模式识别，从而提高预测的准确性和稳定性。深度学习方法主要包括卷积神经网络、循环神经网络、长短期记忆网络等。其中，长短期记忆网络（LSTM）是一种特殊的循环神经网络，具有很强的记忆能力和时间序列建模能力，被广泛应用于股票预测、自然语言处理等领域，YanQi Feng 等人（2015）^[3]使用包括 LSTM 的深度学习对股票价格进行预测的效果，并使用了数值和文本信息比较了不同模型的预测性能，发现 LSTM 预测的误差较小，展现了 LSTM 在股票预测方面的较大优势。

第三节 本文的主要工作

通过对股票市场的研究和分析,我们发现,当前对于股票价格预测的研究主要集中在如何借助深度学习的手段,提升现有模型的普适性和准确性。具体主要表现在以下几个方面:

- 对现有的预测模型进行结构、层次和参数的调优
- 引入更多的股票市场特征,提高模型的预测能力
- 结合具体的股票价格特征,进行具有针对性的短期模型训练

基于现存成果,本文主要工作内容如下:

1. 通过自行搭建的 **LSTM** 网络,对中国平安股票的价格进行预测。我们将详细介绍 **LSTM** 网络的结构和训练过程,以及预测结果的评估和分析。
2. 引入超参数调整算法,根据股票市场特征,引入混沌算法和非线性因子对算法进行调整,提高 **LSTM** 模型的预测准确性。

第二章 传统股票预测方法

第一节 概述

传统的股票预测方法主要基于历史数据和市场分析，通过数学和统计模型来预测股票价格的未来趋势。常用的时间序列分析方法包括移动平均法、自回归模型和差分自回归移动平均模型，这些方法利用过去的股票价格数据来识别趋势和预测未来价格。统计分析方法，如回归分析和多元回归分析，通过建立股票价格与其他变量（如经济指标和公司财务数据）之间的关系模型，提供预测依据。技术分析通过使用技术指标和图表分析，基于历史价格和交易量数据来识别价格模式和趋势。基本面分析则通过研究公司的财务报表和宏观经济指标，评估公司的财务健康状况和经济环境对股票价格的影响。此外，市场情绪分析通过新闻和媒体报道、社交媒体讨论等渠道，了解市场情绪和投资者心理，从而预测股票价格的短期波动。这些传统方法在实际应用中常常结合使用，以提高预测的准确性和可靠性。

第二节 基本面分析

基本面分析法是一种通过评估公司财务状况和宏观经济环境来预测股票价格的方法。以下是基本面分析法的主要内容和步骤：

一、公司财务分析

1. 财务报表分析

- **资产负债表**：分析公司的资产、负债和股东权益，评估公司的财务稳定性和偿债能力。
- **损益表**：评估公司的收入、成本和利润，了解公司的盈利能力和经营效率。
- **现金流量表**：分析公司的现金流入和流出，判断公司的现金流管理和流动性。

2. 财务比率分析

- **财务杠杆比率**：评估公司的债务水平和财务风险。
- **流动比率和速动比率**：衡量公司的短期偿债能力。
- **资产回报率（ROA）和股本回报率（ROE）**：评估公司的盈利能力。

- **市盈率 (P/E Ratio) 和市净率 (P/B Ratio):** 通过股票价格与收益或净资产的比率, 判断股票的估值水平。

二、行业和竞争分析

1. 行业前景

研究行业的发展趋势、市场规模和增长潜力, 了解行业的整体前景和公司的市场地位。

2. 竞争分析

分析公司的主要竞争对手及其市场份额, 评估公司的竞争优势和劣势。

三、宏观经济分析

1. 经济指标

分析宏观经济指标, 如国内生产总值 (GDP)、失业率、通货膨胀率和利率, 评估经济环境对公司的影响。

2. 政策环境

研究政府的经济政策、财政政策和货币政策, 判断政策变化对行业 and 公司的影响。

第三节 股票预测的时间序列分析

时间序列分析是股票预测中常用的方法之一, 利用过去的价格数据来预测未来的价格走势。以下是时间序列分析中的一些主要方法和具体例子, 包括移动平均模型、自回归模型 (AR) 和差分自回归移动平均模型 (ARIMA)。

一、移动平均模型

移动平均模型通过计算一段时间内的平均值来平滑价格数据, 从而识别趋势。常用的移动平均模型有简单移动平均 (SMA) 和加权移动平均 (WMA)。

1. 简单移动平均 (SMA)

简单移动平均是最基本的移动平均方法, 计算公式如下:

$$SMA_t = \frac{1}{n} \sum_{i=0}^{n-1} P_{t-i}$$

其中, SMA_t 为第 t 天的简单移动平均值, P_{t-i} 为第 $t-i$ 天的股票价格, n

为移动平均的时间窗口。

2. 加权移动平均 (WMA)

加权移动平均对不同时间点的数据赋予不同的权重，最近的数据权重更高。计算公式如下：

$$WMA_t = \frac{\sum_{i=0}^{n-1} w_i P_{t-i}}{\sum_{i=0}^{n-1} w_i}$$

其中， WMA_t 为第 t 天的加权移动平均值， w_i 为权重，通常 $w_i = n - i$ ，即离当前日期越近，权重越高。

二、自回归模型 (AR)

自回归模型假设当前的股票价格是过去价格的线性组合，公式如下：

$$P_t = \alpha + \sum_{i=1}^p \beta_i P_{t-i} + \epsilon_t$$

其中， P_t 为第 t 天的股票价格， α 为常数项， β_i 为自回归系数， p 为自回归阶数， ϵ_t 为随机误差项。

三、差分自回归移动平均模型 (ARIMA)

ARIMA 模型综合了自回归 (AR)、差分 (I) 和移动平均 (MA) 三个部分，适用于处理非平稳时间序列数据。ARIMA 模型的表示为 $ARIMA(p, d, q)$ ，其中 p 为自回归阶数， d 为差分次数， q 为移动平均阶数。

1. ARIMA 模型的公式

$$\phi(B)(1 - B)^d P_t = \theta(B)\epsilon_t$$

其中， B 为滞后算子， $\phi(B)$ 为自回归多项式， $\theta(B)$ 为移动平均多项式， ϵ_t 为随机误差项。

具体展开如下：

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d P_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)\epsilon_t$$

其中， ϕ_i 为自回归系数， θ_i 为移动平均系数。

第三章 数据选取与预处理

第一节 数据集简介

本文所选取的数据集来自中国平安 A 股股票的历史交易数据，数据包括股票的开盘价、收盘价、最高价、最低价、成交量等信息，时间跨度为 2016 年 1 月 1 日至 2020 年 12 月 31 日，共计 5 年的数据，均来自 CSMAR 中国经济金融研究数据库。作为中国 A 股市场的龙头股，中国平安的股票价格受到了市场的广泛关注，其股票价格的波动性和不确定性也较大，是一个典型的股票预测问题。通过对中国平安股票的历史交易数据进行分析 and 预测，可以帮助投资者更好地进行投资决策，提高投资的收益率。

第二节 数据指标

本文所选取的数据集包括中国平安 A 股股票的历史交易数据，主要包括以下几个基础指标：

- open：股票的开盘价，即当日股票的第一笔交易价格。
- high：股票的最高价，即当日股票的最高交易价格。
- low：股票的最低价，即当日股票的最低交易价格。
- close：股票的收盘价，即当日股票的最后一笔交易价格。
- change：股票的涨跌幅，即当日股票的涨跌幅度。
- pct_chg：股票的涨跌幅百分比，即当日股票的涨跌幅度的百分比。
- vol：股票的成交量，即当日股票的成交量。
- amount：股票的成交额，即当日股票的成交额。
- pre_close：股票的前收盘价，即前一日股票的收盘价。

第三节 数据预处理

一、数据清洗

机器学习模型的训练非常依赖于知识库的“有效性”，即数据中不能出现残损、缺失、异常等情况。但即便是完善的数据库给出的数据，也有可能丢失掉某些关键数据项的关键指标，这对机器学习模型的训练过程将是致命的。因此，我

们在进行模型构建和模型训练之前，应当对数据进行清洗，以确保数据的完整性和有效性。在机器学习中，数据清洗的方法一般有以下几种：

1. 删除无效值法：若某一数据项的值为无效值，则直接删除该数据项。这种方法的优点是简单易行，但若源数据集中存在大量无效值，将会导致数据集的大量减少，从而影响模型的训练效果。
2. 特殊值替换法：将无效值替换为特殊值，如 0、平均值、中位数等。这种方法的优点是能够保留数据集的完整性，但若特殊值替换不当，将会影响模型的训练效果。
3. 线性插值法：通过已知数据点的线性插值来估计无效值。这种方法的优点是能够保留数据集的完整性，但若数据点之间的关系不是线性关系，将会影响模型的训练效果。

经过调查，我们发现由该数据库给出的数据集中无效值极少，因此我们选择删除无效值的方法进行数据清洗。

二、数据标准化

由于数据集存在多个指标，而每个指标的取值范围和量纲不同，如果将这些数据用于训练，则很有可能影响模型的收敛速度和学习效果。因此，我们需要对数据进行标准化处理，将数据的范围限定在一定范围内，以便于模型的训练。

在本文中，我们采用了 **Min-Max** 标准化^[15]方法对数据进行标准化处理，具体公式如下：

$$x_{\text{new}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (3.1)$$

其中， x_{new} 为标准化后的数据， x 为原始数据， x_{\min} 和 x_{\max} 分别为原始数据的最小值和最大值。

经过这种处理，我们可以将数据的范围限定在 $[0, 1]$ 之间，以便于模型的训练。当然，在模型的预测过程中，我们也需要对预测结果进行反标准化处理，其公式如下：

$$x = x_{\text{new}} \times (x_{\max} - x_{\min}) + x_{\min} \quad (3.2)$$

这就要求我们在训练模型时，需要保存原始数据的最大值和最小值，以便在预测过程中进行反标准化处理。

三、数据集划分

在机器学习中，我们通常将数据集划分为训练集和测试集，以便于评估模型的泛化能力。在本文中，我们将数据集按照 7 : 3 的比例划分为训练集和测试集，其中训练集用于模型的训练，测试集用于模型的评估。具体而言，我们将数据集的前 70%（2016 年 1 月 1 日至 2019 年 8 月 31 日）作为训练集，后 30%（2019 年 9 月 1 日至 2020 年 12 月 31 日）作为测试集。

本数据集中一共有 8 个维度的数据，我们将每 99 个临近数据组成一组，作为一个样本，令模型对由 99 个数据组成的样本进行学习，以预测第 100 个数据。

第四章 LSTM 模型设计

本章主要介绍了 LSTM 模型的设计和实现过程，包括 LSTM 模型的基本原理、网络结构、训练过程和预测过程等。首先，介绍了 LSTM 模型的基本原理，包括 LSTM 模型的结构、LSTM 模型的训练过程和预测过程等。然后，介绍了 LSTM 模型的网络结构，包括 LSTM 模型的输入层、隐藏层和输出层等。最后，介绍了 LSTM 模型的训练过程和预测过程，包括 LSTM 模型的损失函数、优化器和初始超参数选取思想等。

第一节 LSTM 模型的基本原理

LSTM (Long Short-Term Memory) 是一种特殊的循环神经网络 (RNN)，具有很强的记忆能力和时间序列建模能力，被广泛应用于股票预测、自然语言处理等领域。LSTM 模型的基本原理是通过多层神经网络来实现对数据的高阶抽象和模式识别，从而提高预测的准确性和稳定性。其主要数据流动图如图4.1所示。

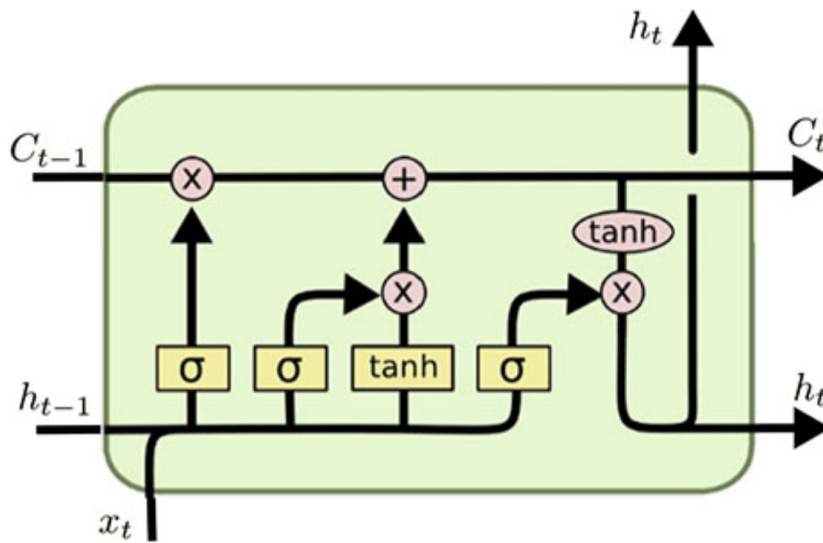


图 4.1 LSTM 模型的数据流动图

LSTM 的提出主要是为了解决长期依赖问题，即传统的 RNN 模型在处理长序列数据时，往往会出现梯度消失或梯度爆炸的问题，导致模型的训练效果不佳。为了解决这一问题，LSTM 引入了三个门控单元，包括输入门、遗忘门和输出门，通过控制信息的输入、遗忘和输出，实现对长序列数据的建模和预测。

1. 输入门：决定在细胞状态里存储的内容，并将新的信息选择性的记录到细胞状态中。它主要由 sigmoid 层^[16]和 tanh 层组成，sigmoid 层输出值为 01，

- 决定哪些信息需要保留， \tanh 层输出值为 -1 到 1 ，生成一个新的候选值向量。
2. 遗忘门：决定从细胞状态里扔掉什么信息，也就是将上一步细胞状态中的信息选择性的遗忘。它由 sigmoid 层组成，输出值为 0 到 1 ，决定哪些信息需要保留。
 3. 输出门：决定了最终要做出的预测结果。它由 sigmoid 层和 \tanh 层组成， sigmoid 层输出值为 0 到 1 ，决定哪些信息需要输出， \tanh 层输出值为 -1 到 1 ，生成最终的输出值。

第二节 LSTM 模型的构建

一、LSTM 模型的工作流程

本文使用的 LSTM 模型工作流程如图4.2所示。测试集数据首先会进入 LSTM 模型的输入层，然后通过隐藏层进行特征提取和抽象，最后通过输出层进行预测和评估。随后，LSTM 的输出会经过两个全连接层，第一个全连接层负责特征提取，第二个全连接层负责预测和评估，并将结果重新反向传播到 LSTM 模型中，进行参数更新和优化。

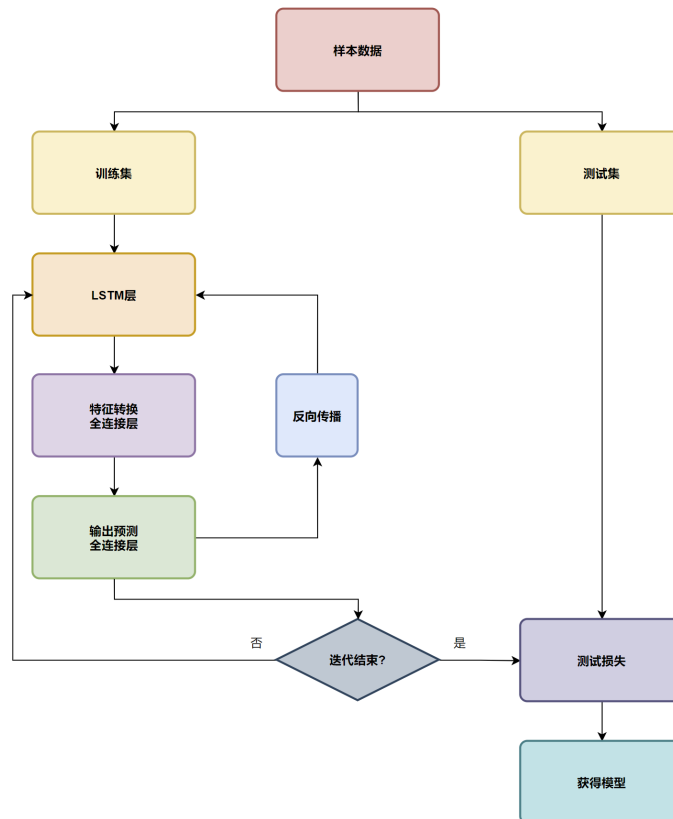


图 4.2 LSTM 模型的工作流程

在 LSTM 层中，我们加入了 Dropout 层^[17]，以减少过拟合的风险。Dropout 层是一种正则化技术，通过在训练过程中随机丢弃一部分神经元，来减少神经元之间的依赖关系，从而提高模型的泛化能力。

二、LSTM 模型的超参数

在这一部分，我们主要介绍在研究中期手动调整的超参数，包括批大小、隐藏层大小、迭代次数等。各个部分的参数选取如下：

1. LSTM 输入层神经元个数：由于本文对股票数据项的 8 个特征进行预测，因此 LSTM 输入层神经元个数为 8。
2. LSTM 隐藏层神经元个数：尝试设置 LSTM 隐藏层神经元个数为 64、128、256。
3. LSTM 网络层数：设置 LSTM 网络层数为 3。
4. LSTM 的 Dropout 层：尝试设置 Dropout 层的丢弃率为 0.1、0.3、0.5。
5. 特征转换全连接层：设置全连接层的神经元个数为 64。
6. 输出预测全连接层：设置全连接层的神经元个数为 1。

经过多次试验，我们通过不停训练在各个参数上进行调整，最终确定了上述参数的取值，以获得最佳的预测效果。参数选取如表4.1所示。

表 4.1 LSTM 模型的超参数选取

超参数	取值
LSTM 输入层神经元个数	8
LSTM 隐藏层神经元个数	256
LSTM 网络层数	3
LSTM 的 Dropout 层丢弃率	0.3
特征转换全连接层神经元个数	64
输出预测全连接层神经元个数	1

三、模型的优化器和调度器

1. 模型优化器

对于股票预测的时间序列分析问题，一般存在以下几种优化器：

- Adam 优化器：Adam 优化器是一种自适应学习率的优化器，它可以根据梯度的一阶矩估计和二阶矩估计自动调整学习率，从而提高模型的训练速度和收敛性。
- SGD 优化器：SGD 优化器是一种随机梯度下降的优化器，它通过计算梯度的均值来更新模型的参数，从而提高模型的训练速度和收敛性。

- Nadam 优化器: Nadam 优化器是 Adam 优化器和 Nesterov 动量的结合体, 它可以根据梯度的一阶矩估计和二阶矩估计自动调整学习率, 并结合 Nesterov 动量来提高模型的训练速度和收敛性。

由于股票市场价格预测是一个复杂的时间序列分析问题, 极容易出现局部最优解问题, 这时候需要借助动量的思想, 冲出局部最优的“谷底”, 从而找到全局最优解。因此, 我们选择了 Nadam 优化器作为 LSTM 模型的优化器。

2. 模型调度器

经过我们的测试发现, 股票预测的 LSTM 模型在训练初期和后期学习率的调整是非常重要的。如果学习率设置过大, 模型在后期的损失率会急剧震荡, 导致模型收敛极为困难; 如果学习率设置过小, 模型在训练初期的收敛速度会非常慢, 导致模型收敛时间过长。因此, 我们选择了学习率指数衰减调度器, 即每隔一定的步数, 将学习率乘一个比率使其减小。经过我们的测试, 每训练 500 个数据时令学习率衰减的效果最好, 因此, 在第 i 次迭代时, 学习率的计算公式如下:

$$lr_i = lr_0 \times \text{decay}^{\lfloor \frac{i}{500} \rfloor} \quad (4.1)$$

其中, lr_i 为第 i 次迭代的学习率, lr_0 为初始学习率, decay 为衰减率。

第三节 LSTM 模型预测结果分析

一、训练损失收敛情况

在前文所述超参数组合的情况下, 我们训练了 LSTM 模型, 并记录了训练过程中的损失值。如图4.3所示, LSTM 模型的训练误差和测试误差都在不断下降, 说明 LSTM 模型在训练过程中逐渐收敛, 模型的泛化能力逐渐增强。同时, 得益于学习率调度器的设置, 虽然在前期训练时出现了一定程度的震荡, 但是在后期训练时, 模型的收敛速度明显加快, 训练效果明显提升。

在训练过程中, 我们使用 MSE (Mean Squared Error) 作为损失函数, 用于评估模型的预测效果。LSTM 的 MSE 损失值如表4.2所示。

表 4.2 LSTM 模型的 MSE 损失值

训练集 MSE	测试集 MSE
0.0429145023226738	0.01695946608408598

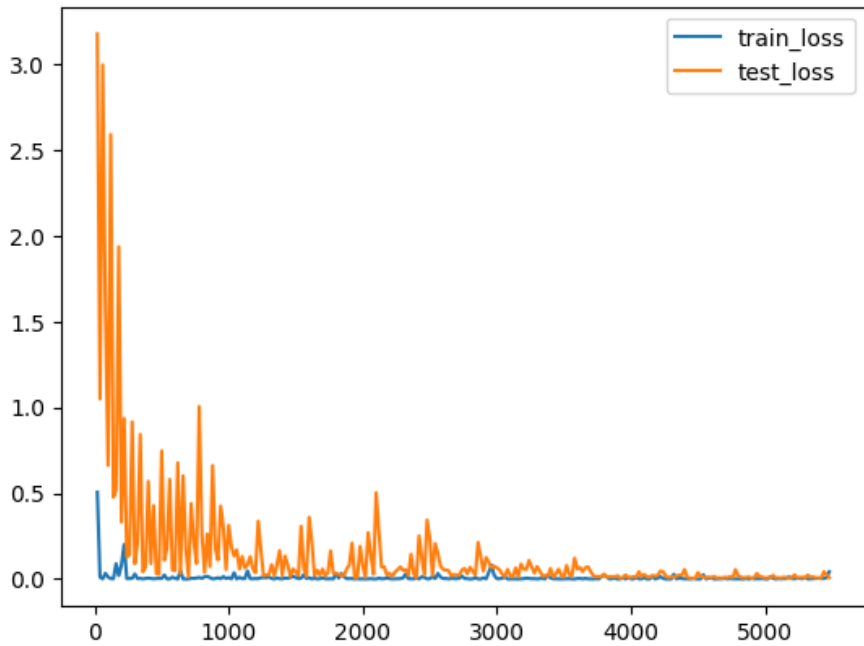


图 4.3 LSTM 模型的训练损失收敛情况

二、预测结果分析

LSTM 模型对中国平安股票的预测结果如图4.4所示。从图中可以看出,LSTM 模型的预测结果与真实值较为接近,预测的趋势与真实值基本一致,预测的波动也与真实值较为吻合。不过,在极端值附近,该模型给出的结果变化程度不够大,导致预测的准确性有所下降。

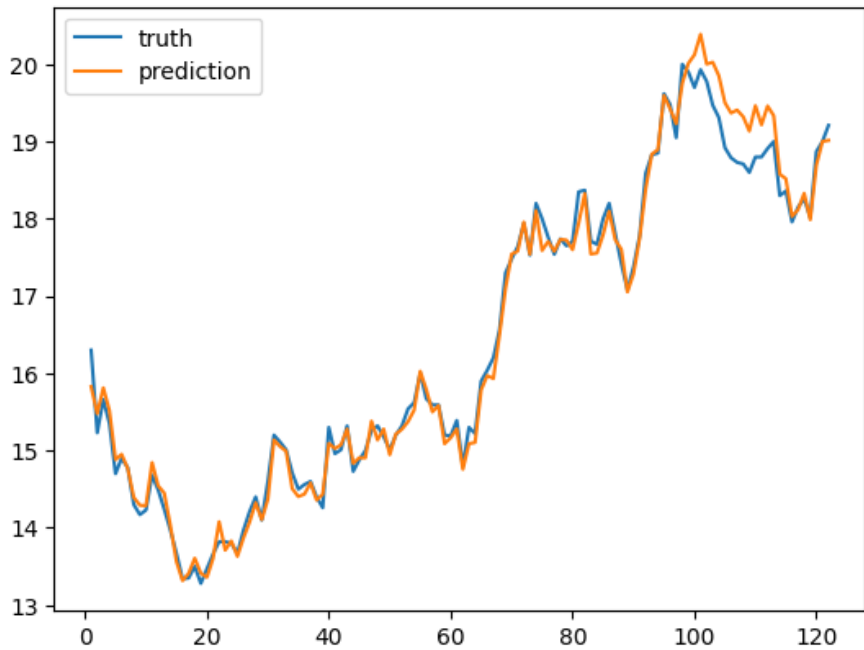


图 4.4 LSTM 模型的预测结果

第四节 本章小结

本章主要介绍了 LSTM 模型的设计和实现过程，包括 LSTM 模型的基本原理、网络结构、训练过程和预测过程等。首先，介绍了 LSTM 模型的基本原理，包括 LSTM 模型的结构、LSTM 模型的训练过程和预测过程等。然后，介绍了 LSTM 模型的网络结构，包括 LSTM 模型的输入层、隐藏层和输出层等。最后，介绍了 LSTM 模型的训练过程和预测过程，包括 LSTM 模型的损失函数、优化器和初始超参数选取思想等。

在实际结果分析中，我们看到了 LSTM 模型在股票预测过程中的良好表现，但也从极端值附近的变化预测趋势不够明显的问题中看到了 LSTM 模型的局限性。通过分析我们可以看出，虽然 LSTM 模型的泛化能力已经很强，但其超参数的选取过于依赖经验，很容易陷入局部最优解，导致模型的预测效果不佳。因此，我们需要进一步优化 LSTM 模型的超参数选取，提高模型的泛化能力和预测准确性。

第五章 基于鲸鱼优化算法的超参数调优

第一节 超参数调优概述

在深度学习模型中，超参数调整是一个非常重要的问题。超参数是指在模型训练之前需要设置的参数，如学习率、批大小、隐藏层大小等。这些参数的设置直接影响了模型的训练效果，但是由于深度学习模型的复杂性和不确定性，很难通过经验或者直觉来确定超参数的取值。因此，如何有效地调整超参数，提高模型的训练效果，一直是深度学习领域的一个难点问题。具体来说，超参数调整的难点主要有以下几个方面：

- 搜索空间巨大：超参数可以有很多不同的值和组合，搜索空间可能是庞大的，尤其是在复杂的模型中。这使得找到最优超参数组合变得非常困难。
- 计算成本高昂：调整超参数需要训练和评估大量的模型，这可能需要大量的计算资源和时间。在大规模数据集或复杂模型上，这一挑战尤为明显。
- 过拟合：在调整超参数时，存在过拟合训练集的风险，因为可能会使用过多的训练数据和评估模型的性能来选择超参数，而不是真正的泛化性能。
- 领域知识的需求：对于一些复杂的模型和任务，超参数调整可能需要对领域有深入的了解，以便能够更好地选择超参数。
- 过于依赖经验：有时候超参数的选择可能会依赖于经验或者直觉，这可能会导致选择不够准确或者不够全面。

一般而言，超参数调整被认为是一个组合优化问题，而组合优化问题往往是 NP-hard 问题，很难在多项式时间内找到最优解。一直以来，研究人员都希望通过启发式搜索算法来解决这一问题。在本文中，我们将介绍一种基于鲸鱼算法的超参数调整方法，通过模拟鲸鱼觅食的过程，来搜索超参数空间，并尝试找到最优的超参数组合。

第二节 鲸鱼优化算法

一、启发式来源

鲸鱼优化算法是一种应用于搜索问题的启发式搜索算法，由 Seyedali Mirjalili 和 Andrew Lewis^[18] 最早在 2016 年提出，主要用于解决优化问题。

WOA 源自自然界中座头鲸使用气泡网进行捕食的行为。据观察，这种觅食

是通过在螺旋路径上制造独特的气泡来完成的，如图 5.1 所示。WOA 正是受到了这种高度智能化的觅食行为的启发而提出的。

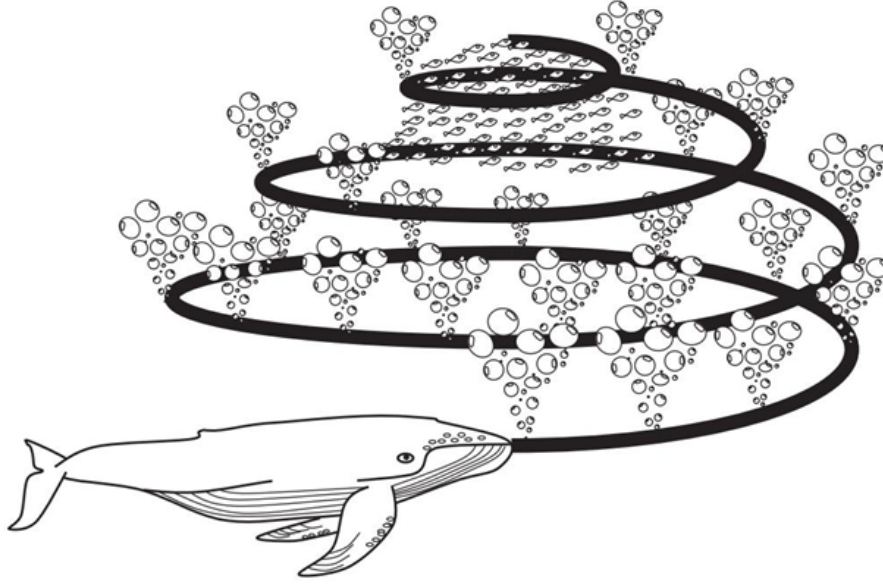


图 5.1 座头鲸的气泡网觅食法

二、基本原理

1. 初始化种群

WOA 首先需要明确种群大小和种群个体维度。种群大小是参与调优的超参数组合的个数，而种群个体维度是超参数组合的维度。设种群大小为 N ，种群个体维度为 D ，每个个体的超参数组合为 $\mathbf{X}_i (i \in [1, N])$ ，并令 \mathbf{X}_{id} 为 \mathbf{X}_i 的第 d 个维度，则初始化过程为：

$$\mathbf{X}_{id} = rand * (UB_d - LB_d) + LB_d \quad (5.1)$$

其中 $rand$ 是一个 $[0, 1]$ 之间的随机数， UB_d 和 LB_d 分别是第 d 个维度数值的上界和下界。

2. 迭代更新

在 WOA 中，每个独立个体都会根据当前最优解来更新自己的位置，从而逐渐收敛到全局最优解。而对于每个个体而言，每一轮迭代中都可以有三种方式来更新自己的位置。

(1) 气泡网螺旋更新

这种更新方式模仿了座头鲸捕食时用螺旋逼近的方式聚集猎物的行为。在这种更新策略中，每个个体会根据当前最优解的位置 \mathbf{X}^* ，以螺线的轨迹对其进行逼近。具体而言，对于第 i 个个体，其位置更新方式为：

$$\mathbf{X}_i = \mathbf{D}_i' \cdot e^{bl} \cdot \cos(2\pi l) + \mathbf{X}^* \quad (5.2)$$

其中, \mathbf{D}_i' 是 \mathbf{X}_i 到 \mathbf{X}^* 的距离向量, 可以通过以下公式计算:

$$\mathbf{D}_i' = |\mathbf{X}^* - \mathbf{X}_i| \quad (5.3)$$

l 是一个 $[-1, 1]$ 之间的随机数, b 是算法参数。

通过这样一种更新, 每个个体都会用随机的步长和方向来逼近当前最优解, 这种随机化策略使得每个个体都不急于靠拢最优解, 能够搜索更多的可能。

(2) 收网逼近

在自然界中, 当座头鲸感觉到自己的距离与猎物已经很近时, 它们会直接沿着直线路径逼近猎物, 这种逼近方式被称为收网逼近。为了度量“近”这个概念, WOA 引入了一个参数 a , 它会随着迭代线性的从 2 减小到 0。随后需要计算两个系数:

$$A = 2 \cdot a \cdot rand - a \quad (5.4)$$

$$C = 2 \cdot rand \quad (5.5)$$

可以看出, A 是一个范围在 $[-a, a]$ 之间的随机数, 而 C 是范围在 $[0, 2]$ 之间的随机数。

在 WOA 中, 每个个体都会根据当前最优解的位置 \mathbf{X}^* , 以直线路径对其进行逼近。具体而言, 对于第 i 个个体, 它首先会计算距离最优解的近似距离向量 \mathbf{D}_i :

$$\mathbf{D}_i = |C\mathbf{X}^* - \mathbf{X}_i| \quad (5.6)$$

之后, 它会根据这个距离向量来更新自己的位置:

$$\mathbf{X}_i = \mathbf{X}^* - A \cdot \mathbf{D}_i \quad (5.7)$$

(3) 随机搜索

当座头鲸距离猎物比较远时, 它们会根据其中一头鲸鱼的位置来进行随机搜索, 如此一来, 它们就能够更好地探索周围的环境。在 WOA 中, 每个个体都

会根据一个随机个体的位置 \mathbf{X}_{rand} ，以随机的方式来搜索周围的环境。具体而言，对于第 i 个个体，它的位置更新方式为：

$$\mathbf{D}_i = |C * \mathbf{X}_{\text{rand}} - \mathbf{X}_i| \quad (5.8)$$

$$\mathbf{X}_i = \mathbf{X}_{\text{rand}} - A \cdot \mathbf{D}_{\text{rand}} \quad (5.9)$$

第三节 基于鲸鱼优化算法的改进 LSTM 预测

一、鲸鱼优化算法的优化对象

在之前章节中，我们分析了 LSTM 预测对于极端值的预测效果不佳的问题。在本章中，我们将尝试结合 WOA 来改进 LSTM 预测的效果。具体而言，我们将使用 WOA 来调整 LSTM 模型的超参数，以期提高模型的预测准确性。我们调整的超参数包括：

- LSTM 的隐藏层大小
- 特征转换全连接层神经元个数
- 训练批大小
- LSTM 的 Dropout 层丢弃率

二、算法优化流程

在加入了 WOA 后，数据在模型中的流动方式如图5.2所示。在每一轮迭代中，WOA 会根据当前最优解来更新种群中每个个体的位置，从而逐渐收敛到全局最优解。在每个个体的位置更新中，WOA 会根据当前最优解的位置，以不同的方式来更新自己的位置，从而实现对超参数的调整。

三、实验结果分析

1. 训练损失收敛情况

我们首先分析了 WOA 优化后的 LSTM 模型在训练集上的损失收敛情况。如图5.3所示，其最终的 MSE 如表5.1。模型在前期有较多震荡，但其收敛速度明显快于未优化的模型，最终收敛到一个较低的值。

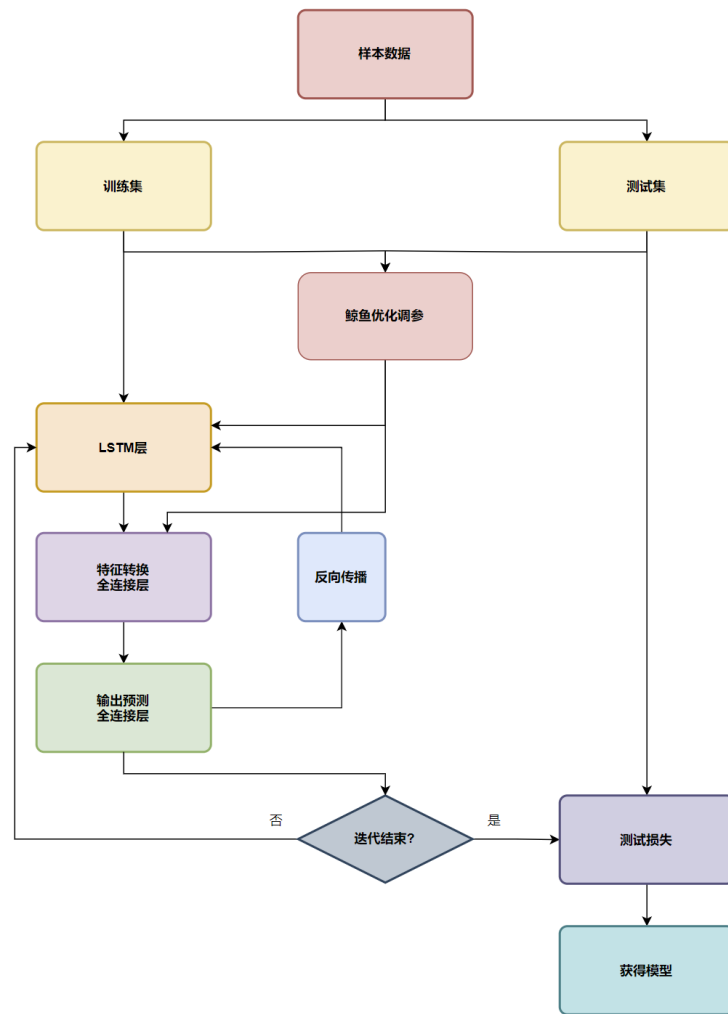


图 5.2 WOA 与 LSTM 模型的数据流动

表 5.1 LSTM 模型的 MSE 损失值

训练集 MSE	测试集 MSE
0.003688221098855138	0.0039420419968390725

2. 预测效果分析

我们还分析了 WOA 优化后的 LSTM 模型在测试集上的预测效果。如图5.4所示，在极端值部分的预测效果有了明显的提升，模型的预测结果与真实值更加接近，预测的波动也与真实值更为吻合。

第四节 本章小结

在本章中，我们介绍了鲸鱼优化算法（WOA）的基本原理和优化流程，并将其应用于 LSTM 模型的超参数调优中。通过 WOA 优化后的 LSTM 模型，我们在

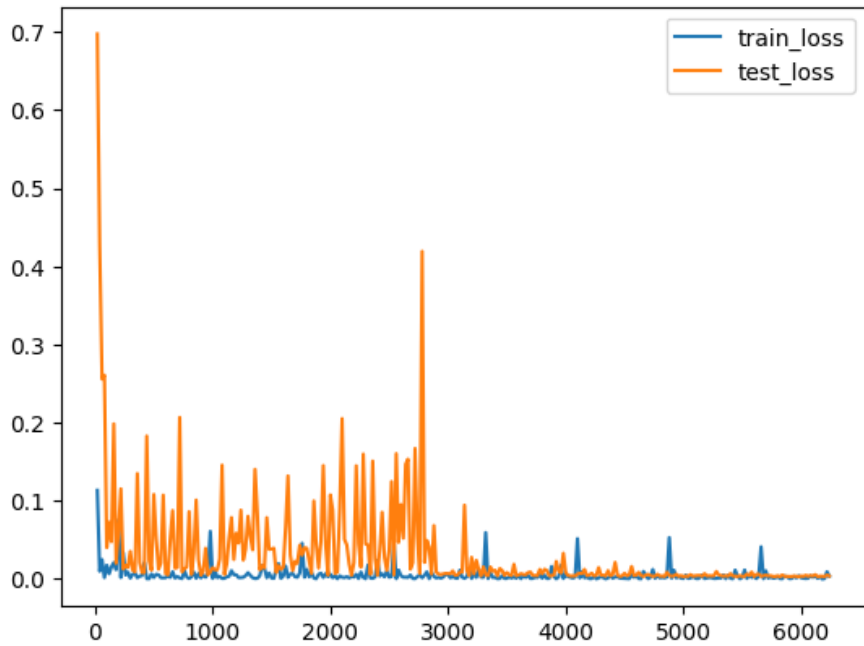


图 5.3 WOA 优化后的 LSTM 模型训练损失

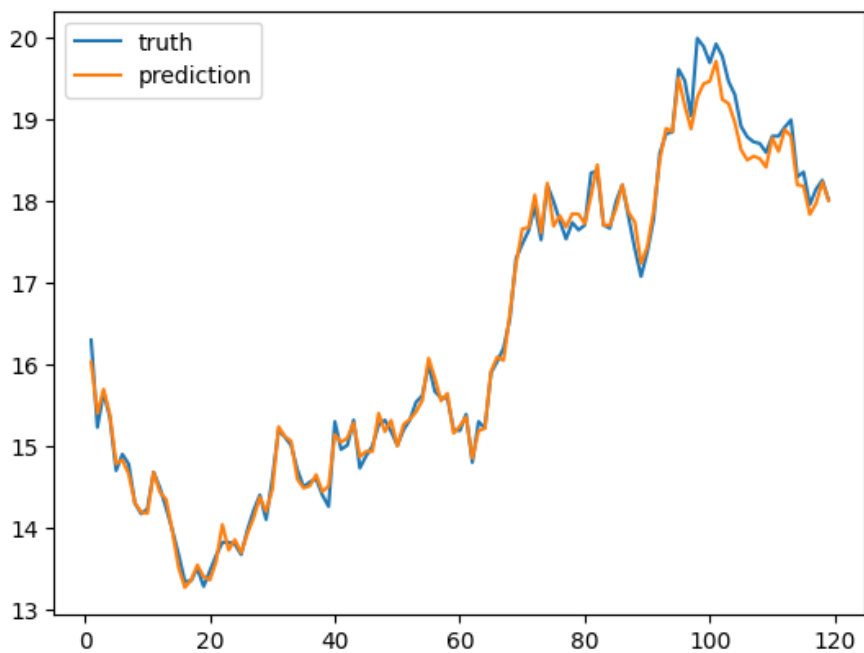


图 5.4 WOA 优化后的 LSTM 模型预测结果

训练集上取得了更快的收敛速度和更低的损失值，同时在测试集上取得了更好的预测效果。WOA 优化后的 LSTM 模型在极端值的预测上有了明显的提升，预测结果更加接近真实值，波动也更为吻合。这表明 WOA 优化算法能够有效提高 LSTM 模型的预测效果，为股票预测问题提供了一种新的解决方案。

第六章 鲸鱼优化算法的改进策略

第一节 鲸鱼算法的问题分析

在上一章节中，我们将鲸鱼优化算法引入了 LSTM 神经网络的超参数优化中，并且通过实验验证了鲸鱼优化算法的有效性。然而，通过我们的实验发现，朴素的鲸鱼优化算法依然存在很多问题，这些问题主要体现在以下几个方面^[19]：

1. 局部最优解问题：鲸鱼优化算法容易陷入局部最优解，导致算法无法收敛到全局最优解，这主要和鲸鱼算法过于依赖其余鲸鱼位置、局部搜索策略较为单一有关。
2. 参数选择问题：鲸鱼优化算法的许多参数都为固定或线性的，这使得算法的适应性较差，面对复杂的非线性问题时，算法的性能会大打折扣。
3. 适应性差：股票市场的变化是非常复杂的，不同的情况下，不同的参数设置可能会产生不同的效果，而鲸鱼优化算法的适应性较差，无法很好的适应不同的情况。

以上的这些问题，事实上并不是鲸鱼算法所独有的，许多启发式优化算法都存在这些问题，这也是启发式优化算法的一个共性问题。对于纷繁复杂的股票市场，鲸鱼算法能否在巨大的超参数搜索空间内找到一个较为优秀的解就成为了预测的关键。因此，我们需要对鲸鱼优化算法进行改进，以提高算法的性能。

第二节 混沌策略初始化种群选择

一、初始化种群选择问题

在我们对鲸鱼算法的实践中，我们发现，鲸鱼算法最终给出的超参数结果的预计损失会随着初始种群的不同而有较大的波动，表6.1很好的印证了这一点。通过分析我们可以看出，鲸鱼算法的初始种群是随机生成的，缺乏了物种多样性，而股票市场中经常会出现局部最优解，这就导致了鲸鱼算法容易陷入局部最优解。

二、Tent 混沌序列生成初始化种群

由于我们对于股票市场的预测没有先验知识，故所有的初始种群个体都应该尽可能地平均分配在搜索空间内，并尽可能在这个阶段使用简单的初步训练

表 6.1 不同初始种群下的预计损失

初始种群	预计损失
(7, 0.19, 139, 100)	0.008
(4, 0.5, 209, 96)	0.013
(6, 0.37, 64, 87)	0.024
(4, 0.45, 176, 108)	0.009

找到较为合适的初始种群。因此，我们提出了混沌策略初始化种群选择的方法，使用混沌序列生成初始种群，以增加种群的多样性，提高算法的全局搜索能力。

我们选用了较为简单的 Tent 混沌序列^[20]进行生成。若种群数量为 N ，每个个体的维度为 D ，则我们可以通过如下的公式生成混沌序列 \mathbf{z} ：

$$\mathbf{z}_{i+1,d} = \begin{cases} 2\mathbf{z}_{id} & \text{if } \mathbf{z}_{id} < 0.5 \\ 2(1 - \mathbf{z}_{id}) & \text{if } \mathbf{z}_{id} \geq 0.5 \end{cases} \quad (6.1)$$

生成混沌序列后，我们可以通过如下的公式将混沌序列映射到搜索空间内：

$$\mathbf{Y}_{i,d} = \mathbf{z}_{i,d} \times (UB_d - LB_d) + LB_d \quad (6.2)$$

如此我们就可以得到一个较为均匀分布的初始种群。在此基础上，我们希望能获得这个种群的反向种群：

$$\mathbf{Y}'_{i,d} = UB_d + LB_d - \mathbf{Y}_{i,d} \quad (6.3)$$

反向种群通常能进一步提升初始种群的分布均匀性，增加种群的多样性，提高算法的全局搜索能力。获取反向种群后，我们将种群 \mathbf{Y} 与 \mathbf{Y}' 合并，通过简单的初步训练，找到其中 MSE 最小的 4 个个体，作为鲸鱼算法的初始种群。

第三节 非线性因子的引入

非线性因子在深度学习框架中扮演着至关重要的角色。深度学习模型的强大之处在于其能够学习和理解数据中的复杂关系，而非线性因子则允许模型捕捉和建模这些复杂的非线性关联。在鲸鱼算法中存在两个非常重要的参数，即 a 和 b ，它们分别控制了鲸鱼算法中的收敛速度和螺旋形状。然而，这两个参数在鲸鱼算法中是线性的甚至不变的，这使得鲸鱼算法在适应性和收敛速度上存在一定的问题。我们希望能对这两个因子进行非线性的调整，以提高算法的性能。

一、递减收敛因子的非线性调整

在朴素的鲸鱼算法中，参数 a 是一个线性的参数，它的取值范围是 $[0, 2]$ 并不断线性递减。然而我们发现，这种线性收敛很容易导致鲸鱼算法在探索阶段错失最优解，并导致后续算法陷入局部最优解。我们希望在算法的前期能找到一个计算函数，使得算法可以以不同的概率对 a 进行计算，并能够根据实际面向问题的需要，手动调整概率。所以，我们使用了一种类似于指数函数的衰减方式，对参数 a 进行非线性调整：

$$a = \begin{cases} 2 - \frac{3}{2} \sqrt{e^{\frac{k}{max_iter}} - 1} & \text{if } rand \leq threshold \\ 2(\frac{k}{max_iter})^2 - \frac{4k}{max_iter} + 2 & \text{if } rand > threshold \end{cases} \quad (6.4)$$

这两种收敛方式的函数曲线如图6.1所示。其中， $threshold$ 是一个阈值，它表示了收敛因子指数衰减的概率， $rand$ 是一个随机数，通过 $rand$ 和 $threshold$ 的比较，我们可以以不同的概率对 a 进行计算。若 $threshold$ 比 0.5 要大，则算法使用指数收敛，在前期能够更大范围地搜索，而在后期则更加集中于局部搜索；若 $threshold$ 比 0.5 要小，则算法使用多项式更容易收敛，整体上提升了算法的性能。

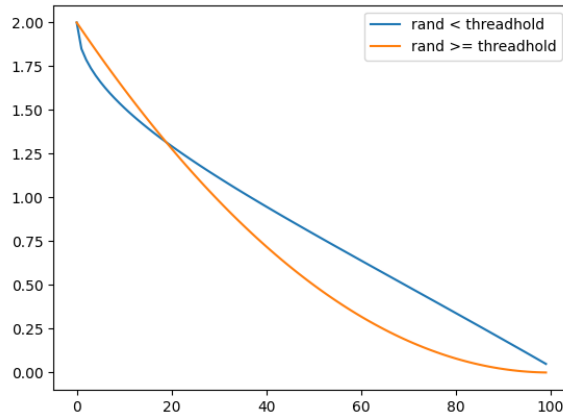


图 6.1 参数 a 的非线性调整

二、螺旋形状因子的非线性调整

在气泡网攻击的过程中，鲸鱼会沿着螺旋形状的路径进行搜索，这个路径的形状由参数 b 和随机数 l 共同控制。在朴素的算法中，参数 b 是一个定值，这会导致螺旋的形状较为固定，不能体现座头鲸在捕食后期快速收敛的特点。因此，我们希望能对参数 b 进行非线性调整，使得螺旋形状能够更好地适应训练的过

程。我们的参数 b 的非线性调整如下：

$$b = 1 - \frac{5}{l} \cos\left(\frac{\pi k}{\max_iter}\right) \quad (6.5)$$

回顾气泡网攻击的位置更新公式我们可以看到，螺旋的半径与 b 和 l 的乘积相关，通过这种非线性处理，训练前期鲸鱼的搜索范围更大，更能找到潜在的最优解；而到了训练后期，螺旋的半径会逐渐缩小，使得鲸鱼更容易收敛到局部最优解。

第四节 自适应权重与最优局部抖动

一、自适应权重

在鲸鱼算法中，如果鲸鱼不处在随机搜索更新的状态，那么鲸鱼会根据当前最优个体的位置进行气泡网攻击更新或收网更新。在这两种更新中，对于最优个体的位置的权重是固定的，这使得算法在不同的情况下表现不稳定。我们希望能够根据当前的情况，自适应地调整最优个体的权重，以提高算法的性能。具体来说，我们需要定义一个自适应权重函数 ω ：

$$\omega = 1 - \frac{e^{\frac{k}{\max_iter}} - 1}{e - 1} \quad (6.6)$$

这个函数的曲线如图6.2所示。我们可以看到，在训练初期，最优解位置的权重较大，这使得算法能够更多地利用最优解的信息，更好地进行搜索；而到了训练后期，最优解位置的权重逐渐减小，这使得个体更新有更多的随机性，更容易跳出局部最优解。

基于这个更新函数，我们可以将个体更新位置的公式改为：

$$\mathbf{X}_i = \omega \cdot \mathbf{X}^* - A \cdot \mathbf{D}_i \quad (6.7)$$

$$\mathbf{X}_i = \omega * \mathbf{X}^* + \mathbf{D}_i' \cdot e^{bl} \cdot \cos(2\pi l) \quad (6.8)$$

如此，我们就可以根据当前的训练情况，自适应地调整最优解位置的权重，提高算法的性能。

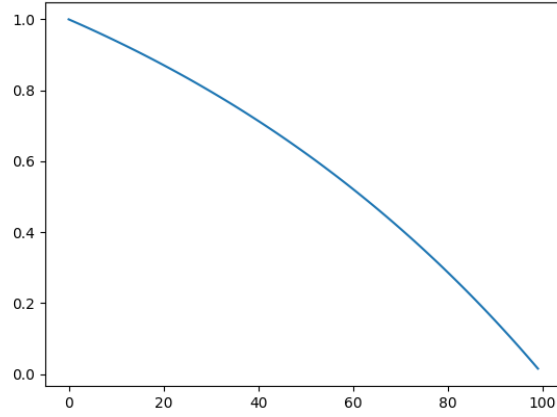


图 6.2 自适应权重函数

二、最优局部抖动

自适应权重主要是为了解决其他个体对最优个体的过度依赖问题，而最优局部抖动是为了让最优个体不要故步自封，而是能够在一个局部更好地探索其它可能性。具体而言，我们希望在每次迭代中，最优个体都能在以它为圆心的一个小的超球体内随机抖动，如果发现了更好的解，则更新最优个体的位置。这样一来，最优个体就能够更好地探索局部空间，提高算法的全局搜索能力。

在每次迭代中，我们都使用随机数，基于当前最优解生成一个可能替代解：

$$\mathbf{X}'^* = \mathbf{X}^* + rand \cdot \mathbf{X}^* \quad (6.9)$$

其中， $rand$ 是一个 $[0, 0.5]$ 之间的随机数。我们将 \mathbf{X}'^* 与当前最优解通过损失函数 $Loss$ 比较，如果 \mathbf{X}'^* 的预计损失更小，则更新最优解的位置，否则保持原来的位置：

$$\mathbf{X}^* = \begin{cases} \mathbf{X}'^* & \text{if } Loss(\mathbf{X}'^*) < Loss(\mathbf{X}^*) \\ \mathbf{X}^* & \text{otherwise} \end{cases} \quad (6.10)$$

通过这种方式，我们就可以让最优个体在局部空间内进行抖动，提高算法的全局搜索能力。

第五节 实验结果

一、训练损失收敛情况

我们将上述策略应用到鲸鱼算法中，通过实验验证了这些策略的有效性。我们使用了同样的 LSTM 神经网络模型，以及同样的股票数据集，对比了朴素的鲸鱼算法和改进后的鲸鱼算法的训练损失收敛情况。如图6.3和表6.2所示，我们可以看到，改进后的鲸鱼算法在训练初期的收敛速度更快，而在训练后期更容易跳出局部最优解，这表明我们的改进策略能够有效提高算法的性能。

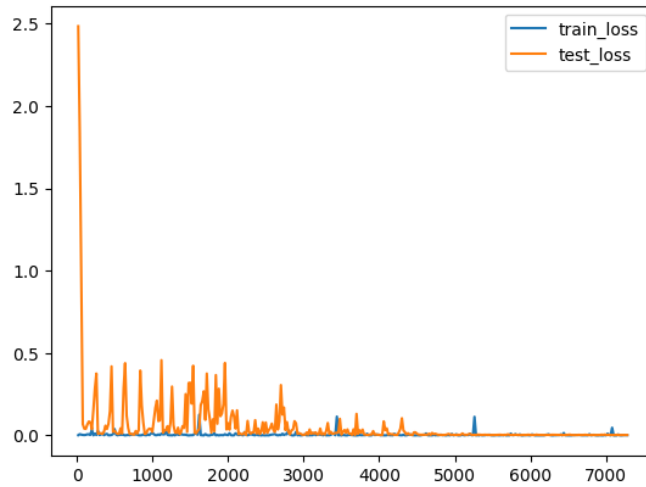


图 6.3 训练损失收敛情况

表 6.2 LSTM 模型的 MSE 损失值

训练集 MSE	测试集 MSE
0.0019056115997955203	0.0029083097339025698

二、预测结果对比

我们将改进后的鲸鱼算法应用到 LSTM 神经网络中，对股票数据集进行预测。我们将预测结果与朴素的鲸鱼算法进行了对比，如图6.4所示。我们可以看到，改进后的鲸鱼算法在预测结果上表现更好，预测的曲线更加接近真实值，这表明我们的改进策略能够有效提高算法的性能。

第六节 本章小结

在本章中，我们对鲸鱼算法进行了改进，提出了混沌策略初始化种群选择、非线性因子的引入、自适应权重与最优局部抖动等策略。这些策略主要是为了解决鲸鱼算法依赖初始点选取、参数选择固定、适应性差等问题，以提高算法的性

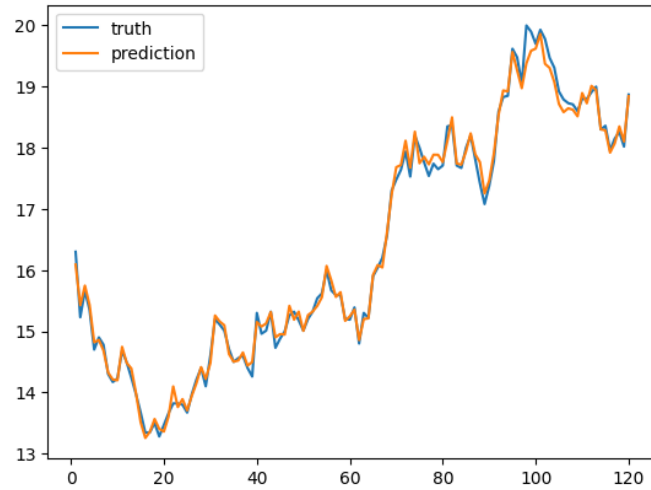


图 6.4 预测结果对比

能。

我们通过实验验证了这些策略的有效性，实验结果表明，这些策略能够有效提高鲸鱼算法的性能，提高算法的全局搜索能力，提高算法的适应性，提高算法的收敛速度。这些改进策略为鲸鱼算法的应用提供了新的思路，为算法的进一步优化提供了参考。

第七章 成果总结与展望

第一节 成果总结

本文通过 LSTM 和改进后的鲸鱼优化算法相结合，提出了一种用于股票市场预测的超参数优化方法和预测方法。通过实验验证，我们发现改进后的鲸鱼优化算法在超参数优化问题上具有较好的性能，能够有效地提高 LSTM 模型的预测准确率。具体来说，我们的工作主要有以下几个方面的贡献：

1. 我们提出了一种基于鲸鱼优化算法的超参数优化方法，通过模拟鲸鱼觅食的过程，来搜索超参数空间，并尝试找到最优的超参数组合。我们将这种方法应用于 LSTM 模型的超参数优化中，通过实验验证，我们发现改进后的鲸鱼优化算法能够有效地提高 LSTM 模型的预测准确率。
2. 我们提出了一种混沌策略初始化种群选择的方法，通过使用混沌序列生成初始种群，以增加种群的多样性，提高算法的全局搜索能力。我们将这种方法应用于鲸鱼优化算法中，通过实验验证，我们发现混沌策略初始化种群选择的方法能够有效地提高算法的性能。
3. 我们提出了一种基于 LSTM 的股票市场预测方法，通过使用 LSTM 模型来预测股票市场的走势。我们将这种方法应用于实际的股票市场数据中，通过实验验证，我们发现 LSTM 模型能够有效地预测股票市场的走势。

第二节 未来展望

本文提出的基于鲸鱼优化算法的超参数优化方法和基于 LSTM 的股票市场预测方法在实验中取得了较好的效果，但仍然存在一些问题和不足之处，有待进一步的研究和改进。未来的工作可以从以下几个方面展开：

1. 在超参数优化方面，我们可以进一步研究鲸鱼优化算法的性能，尝试改进算法的搜索策略，提高算法的全局搜索能力。我们还可以尝试将鲸鱼优化算法应用于其他的深度学习模型中，验证算法的通用性。
2. 在股票市场预测方面，我们可以进一步研究 LSTM 模型的性能，尝试改进模型的结构，提高模型的预测准确率。我们还可以尝试将其他的深度学习模型应用于股票市场预测中，验证模型的有效性。
3. 在实验方面，我们可以进一步扩大实验数据集的规模，验证算法和模型的

稳定性和鲁棒性。我们还可以尝试将算法和模型应用于其他的金融领域，验证其在其他领域的适用性。

参 考 文 献

- [1] HUANG Y, et al. Machine learning for stock prediction based on fundamental analysis[A/OL]. 2022: 1-10. <https://doi.org/10.48550/arXiv.2202.05702>.
- [2] SHAHI T B, SHRESTHA A, NEUPANE A, et al. Stock price forecasting with deep learning: A comparative study[J/OL]. Mathematics, 2020, 8(9): 1441. <https://doi.org/10.3390/math8091441>.
- [3] FENG Y, WANG B, CHEONG M. Deep learning for stock prediction using numerical and textual information[J]. Proceedings of the International Conference on Web Intelligence (WI), 2015: 11-20.
- [4] YOO J, et al. Hyperparameter optimization for forecasting stock returns[A/OL]. 2022: 1-10. <https://doi.org/10.48550/arXiv.2201.06834>.
- [5] DANIEL K, HIRSHLEIFER, D. Journal of finance[J]. Investor Psychology and Security Market Under- and Overreactions, 1998, 53(6): 12-13.
- [6] LO, W A. Journal of portfolio management[J]. The Adaptive Markets Hypothesis: Market Efficiency from an Evolutionary Perspective, 2003, 30(5): 15-29.
- [7] ANG, A. Regime changes and financial markets[J]. Annual Review of Financial Economics, 2012, 4(1): 15-29.
- [8] 朱倩, 陈浩, 樊翔. 中国股票市场波动性的预测: 基于 GARCH 模型和 SVR 模型比较研究[J]. 系统工程理论与实践, 2016, 36(1): 15-29.
- [9] 孙婧, 王莹. 基于马尔可夫模型的中国股票市场短期预测.[J]. 电子科技大学学报, 2019, 48(4): 15-29.
- [10] 韩思远, 尹洋, 白鹤. 中国 A 股市场交易数据对股票价格波动率的预测[J]. 金融研究, 2020, 45(7): 15-29.
- [11] KAUR J, DHARNI K. Data mining-based stock price prediction using hybridization of technical and fundamental analysis[J/OL]. Data Technologies and Applications, 2023, 57(5): 780-800. <https://doi.org/10.1108/DTA-04-2022-0142>.
- [12] VIDHYA A. Stock market forecasting using arima model[J/OL]. Analytics Vidhya, 2023: 1-6. <https://www.analyticsvidhya.com/blog/2023/01/stock-market-forecasting-using-arima-model/>.
- [13] BEN ABDALLAH M, FEKETE-FARKAS M, NATHAN R J, et al. Garch (1,1)

- models and analysis of stock market turmoil during covid-19 outbreak in an emerging and developed economy[J/OL]. *Journal of Risk and Financial Management*, 2021, 14(12): 576. <https://doi.org/10.3390/jrfm14120576>.
- [14] TITARMARE N, WAZARKAR M, PALKAR P, et al. Stock market predictions using linear regression[J/OL]. *AIP Conference Proceedings*, 2022, 2424(1): 050004. <https://doi.org/10.1063/5.0076964>.
- [15] GUPTA V, HEWETT R. Adaptive normalization in streaming data[A/OL]. 2019: 1-12. <https://arxiv.org/abs/1910.07696>.
- [16] ALOM M Z, TAHA T M, YAKOPCIC C, et al. Activation functions in deep learning: A comprehensive survey and benchmark[A/OL]. 2021: 3-4. <https://arxiv.org/abs/2109.14545>.
- [17] DIETRICH S A, SICK B. Survey of dropout methods for deep neural networks [A/OL]. 2019: 1-21. <https://arxiv.org/abs/1904.13310>.
- [18] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. *Advances in Engineering Software*, 2016, 95: 51-67.
- [19] BISCHL B, BINDER M, LANG M, et al. Hyperparameter optimization: Foundations, algorithms, best practices and open challenges[A]. 2021: 1-50.
- [20] BUNIMOVICH A, OTHERS. Characterizing the effects of randomness in the tent map[A]. 2018: 1-14.

致 谢

在研究学习期间，我有幸得到了我的导师，中国科学技术大学李勇老师的悉心指导。老师学术功底，严谨的工作态度和敏锐的科学洞察力使我受益良多。衷心感谢老师一直以来给予我的悉心教导和热情帮助。

感谢李勇老师在实验方面的指导以及帮助。中国科学技术大学的龙浩泉同学和徐航宇同学为项目设计提出了宝贵意见，在此深表谢意。

2024 年 6 月