

学校编码: 10173

分类号\_\_密级\_\_

学 号: 2019120290

UDC\_\_

# 东北财经大学

专业硕士学位论文

## LSTM 模型优化及其在股指预测中的 应用研究

Optimization of LSTM Model and Its Application in Stock  
Index Prediction

曹彦彦

指导教师姓名: 刘颖 副教授

专业名称: 应用统计

论文答辩时间: 2022 年 5 月

## 摘要

随着科技的进步,我国股票市场日渐完善,大量的金融数据得以保留,为进行股票市场分析奠定了坚实基础。在股票市场中,高收益也伴随着高风险,然而在高收益的同时降低投资风险是众多投资者一直追求的目标,因此如何对股价进行更精准的预测成为了研究的热点问题。伴随着机器学习和深度学习的不断壮大,研究学者开始将其应用于股票市场中。循环神经网络(RNN)作为深度学习的一种,引入了时序的概念,为分析时间序列数据提供了新的方法。然而RNN存在一定的缺陷,LSTM神经网络作为RNN的一种变种,克服了RNN中存在的梯度消失和爆炸问题,应用领域更加广泛。同时,与其它时间序列预测模型相比,LSTM神经网络在处理非线性数据上具有一定的优势。根据已有研究成果,股票市场常常展现出非线性特征。因此,本文尝试利用LSTM神经网络对股指进行预测,以期得到更精准的预测结果。

本文选取了上证指数近5年的日交易数据作为样本,采用LSTM模型,利用前1天的开盘价、最高价、最低价、收盘价、涨跌额、成交量来预测当天的收盘价。由于LSTM预测结果容易受到网络参数影响,本文通过鸽群算法(PIO)对其网络参数进行优化。但是鸽群算法(PIO)容易陷入局部最优,本文先探讨和研究如何改进鸽群算法。然后利用改进的鸽群算法优化LSTM神经网络模型,从而得到鸽群算法与LSTM模型的组合模型,通过组合模型对上证指数进行预测。最后,文章对预测结果进行了比较,得出结论,和进一步研究展望。

本文的工作重点在于鸽群算法的改进和LPIO\_LSTM预测模型的构建。PIO算法虽然提出时间短,但实际应用有优势,它具有收敛速度快、操作简单、易于实现等特点。因此,该算法逐渐发展成了较热门的群智能优化算法。但它在解决复杂优化问题时,仍然存在着不足,在迭代后期容易陷入局部最优,导致算法的寻优精度有所下降。针对基本PIO算法的不足,本文尝试对鸽群算法进行了改进,在改进过程中主要是考虑将鸽群算法与其他机制相结合,最终形成了LPIO算法。这也是本文的创新之处,本文提出的改进思路主要包括:

- (1) 在地磁导航阶段,引入Levy飞行,利用Levy飞行的跳跃性,帮助种群中的个体跳出局部最优。

- (2) 利用混沌理论的遍历性特点,对每次迭代后的最优个体进行混沌搜索,

可以有效的提高算法的求解精度，同时，也在一定程度提高了收敛性能。

针对 LPIO 算法，本文选取了 5 个测试函数进行测试，测试结果表明 LPIO 算法在复杂的优化问题中不仅能够找到全局最优解，具有良好的全局性能，而且还保留了原有 PIO 算法良好的收敛性能。

在解决了鸽群算法所存在的问题后，本文运用改进的鸽群算法对 LSTM 神经网络的网络参数进行优化，并利用组合模型进行实例分析，得到以下结论：

(1) 在对上证指数收盘价进行预测时，与 ARIMA、BP 神经网络预测结果相比，LSTM，LPIO\_LSTM 的预测效果更好，预测误差更小。证明与其他模型相比，LSTM 模型更适合应用于上证指数收盘价的预测中。

(2) 加入优化算法的 LSTM 神经网络要优于传统 LSTM 神经网络模型。基于 LPIO 算法优化的 LSTM 神经网络模型预测效果最好，表明利用 LPIO 算法优化 LSTM 神经网络的网络结构可以很好的改善传统 LSTM 神经网络的预测性能，具有一定的实用价值。

**关键词：**鸽群算法，LSTM 神经网络，股指预测，参数优化，组合模型

## ABSTRACT

With the progress of science and technology, China's stock market has been improved day by day, and a large number of financial data have been retained, which has laid a solid foundation for stock market analysis. In the stock market, high returns are accompanied by high risks. However, reducing investment risks while high returns is the goal that many investors have been pursuing. Therefore, how to predict the stock price more accurately has become a hot issue. With the continuous growth of machine learning and deep learning, researchers began to apply it to the stock market. As a kind of deep learning, recurrent neural network (RNN) introduces the concept of time series, which provides a new method for analyzing time series data. However, RNN has some defects. As a variant of RNN, LSTM neural network overcomes the problems of gradient disappearance and explosion in RNN, and has a wider application field. At the same time, compared with other time series prediction models, LSTM neural network has certain advantages in dealing with nonlinear data. According to the existing research results, the stock market often shows nonlinear characteristics. Therefore, this paper attempts to use LSTM neural network to predict the stock index in order to get more accurate prediction results.

This paper selects the daily trading data of Shanghai stock index in recent five years as a sample, and uses LSTM model to predict the closing price of the day by using the opening price, highest price, lowest price, closing price, rise and fall volume and trading volume of the previous day. Because LSTM prediction results are easily affected by network parameters, this paper optimizes the network parameters by PIO Algorithm. However, the PIO Algorithm is easy to fall into local optimization. This paper first discusses and studies how to improve the pigeon swarm algorithm. Then the improved PIO Algorithm is used to optimize the LSTM neural network model, so as to obtain the combined model PIO Algorithm and LSTM model, and predict the Shanghai stock index through the combined model. Finally, the paper compares the prediction results, draws a conclusion and prospects for further research.

This paper focuses on the improvement of PIO Algorithm and LPIO Algorithm of LSTM prediction model. Although PIO Algorithm is proposed in a short time, it has advantages in practical application. It has the characteristics of fast convergence, simple operation and easy implementation. Therefore, the algorithm has gradually developed into a popular swarm intelligence optimization algorithm. However, it still has some shortcomings in solving complex optimization problems. It is easy to fall into local optimization in the later stage of iteration, resulting in the decline of optimization accuracy of the algorithm. In view of the shortcomings of the basic PIO Algorithm, this paper attempts to improve the pigeon swarm algorithm. In the process of improvement, we mainly consider the combination of PIO Algorithm and other mechanisms, and finally form the LPIO Algorithm. This is also the innovation of this paper. The improvement ideas proposed in this paper mainly include:

(1) In the geomagnetic navigation stage, Levy flight is introduced to help individuals in the population jump out of the local optimum by using the jumping nature of Levy flight.

(2) Using the ergodic characteristics of chaos theory to search the optimal individual after each iteration can effectively improve the solution accuracy of the algorithm, and improve the convergence performance to a certain extent.

For LPIO Algorithm, this paper selects five test functions to test. The test results show that LPIO Algorithm can not only find the global optimal solution in complex optimization problems, but also retain the good convergence performance of the original LPIO Algorithm.

After solving the problems existing in the pigeon swarm algorithm, this paper uses the improved PIO Algorithm to optimize the network parameters of LSTM neural network, and makes an example analysis by using the combination model. The following conclusions are obtained:

(1) When predicting the closing price of Shanghai stock index, compared with the prediction results of ARIMA and BP neural network, LSTM and LPIO Algorithm. LSTM has better prediction effect and less prediction error. It is proved that compared with other models, LSTM model is more suitable for predicting the closing price of Shanghai stock index.

(2) The LSTM neural network with optimization algorithm is better than the traditional LSTM neural network model. The prediction effect of LSTM neural network model optimized based on LPIO Algorithm is the best, which shows that using LPIO Algorithm to optimize the network structure of LSTM neural network can improve the prediction performance of traditional LSTM neural network, and has certain practical value.

**Key words:** PIO, LSTM, Stock Index Prediction, Parameter Optimization, Combined Model

# 目 录

1 绪论 .....	1
1.1 研究背景 .....	1
1.2 研究意义 .....	2
1.3 国内外研究现状 .....	3
1.3.1 统计模型在股票预测应用中的研究文献 .....	3
1.3.2 神经网络在股票预测应用中的研究文献 .....	4
1.3.3 鸽群算法相关研究文献 .....	6
1.3.4 文献评述 .....	7
1.4 研究思路和研究内容 .....	7
1.4.1 研究思路 .....	7
1.4.2 研究内容 .....	8
1.5 本文创新点与不足 .....	8
2 股票价格预测理论与模型概述 .....	10
2.1 股票价格预测 .....	10
2.1.1 基本面分析 .....	10
2.1.2 技术面分析 .....	11
2.2 常用的股票预测统计模型 .....	11
2.3 神经网络理论 .....	12
2.3.1 BP 神经网络 .....	13
2.3.2 循环神经网络 .....	14
2.3.3 长短时记忆神经网络 .....	15
2.4 本章小结 .....	17
3 改进鸽群算法 .....	18
3.1 鸽群算法 .....	18
3.1.1 鸽群算法基础理论 .....	18
3.1.2 鸽群算法的步骤 .....	20

3.1.3 鸽群算法的优缺点 .....	20
3.2 改进策略 .....	21
3.2.1 Levy 飞行 .....	21
3.2.2 混沌理论 .....	22
3.2.3 改进鸽群算法的具体步骤 .....	23
3.3 仿真测试及结果分析 .....	24
3.3.1 参数设置 .....	24
3.3.2 测试函数 .....	24
3.3.3 改进结果分析 .....	25
3.4 本章小结 .....	29
4 运用改进鸽群算法优化后的 LSTM 进行股指预测 .....	30
4.1 数据的选取及预处理 .....	30
4.1.1 数据的选取 .....	30
4.1.2 数据的预处理 .....	30
4.2 模型的评价标准 .....	31
4.3 LPIO-LSTM 模型的建立过程 .....	32
4.4 实证分析及结果对比 .....	33
4.4.1 实证分析过程及结果 .....	33
4.4.2 结果对比 .....	38
4.5 本章小结 .....	39
5 结论与展望 .....	40
5.1 结论 .....	40
5.2 不足与展望 .....	41
参考文献 .....	42
后 记 .....	46

# 1 绪论

## 1.1 研究背景

在我国随着股票市场的发展,越来越多的人将资金投入其中,以期获得更多的收益。然而,在投资股票的过程中,高收益往往伴随着高风险。在这种情况下,若是盲目的投入资金,很可能造成损失,然而若能对股票价格进行较为精准的预测,就可以更好的帮助投资者进行合理投资,使得他们在获得高收益的同时降低投资风险。因此对股票走势进行研究具有一定的意义。

我国股票市场结构复杂,容易受到众多因素的影响,且股票数据在一定程度上表现出了一定的非线性和不稳定性,使得股指预测问题一直是该行业亟待解决的难题。在股指预测早期,主要采用的预测方法有基本分析法和技术分析,但这些方法只适用于早期比较简单的股票市场。后来,学者们尝试运用统计学的知识建立相应的预测模型来进行预测,比如时间序列分析法、回归分析法等,但随着股票交易数量的增加,股票交易数据表现出了复杂的非线性特征,使得预测模型的预测结果与实际值相差较大。

近年来,随着机器学习和深度学习理论的出现,一系列新型的预测方法应运而生。众人开始尝试将神经网络应用于股票市场中,对未来股票价格进行预测。神经网络的引入,确实一定程度上提高了股票价格的预测精度。LSTM神经网络作为深度学习的一种,同样得到了众多学者的青睐。LSTM神经网络改善了循环神经网络(RNN)中的长期依赖问题,同时,LSTM可以很好的处理非线性数据,也就表明LSTM网络模型在一定程度上可以处理股价数据中存在的非线性及不稳定性等问题。根据这一点,可以初步猜想LSTM对股价进行预测应该是可行的。通过学者们的研究结果可以看出,LSTM神经网络在预测股票价格的过程中确实取得了很好的效果。与其他神经网络一样,LSTM神经网络模型本身容易受到网络结构的影响,因此,神经网络的网络结构成了制约预测结果的一个因素。在现有的研究中,常常通过研究者的经验或多次实验来找到合适的网络模型,但这在消耗时间资源的同时,找到的并不一定是真正意义上的那个最优的网络模型,所以如何确定神经网络的结构,是我们提高预测

能力过程中需要关注和解决的问题。

鸽群算法是在模拟鸽群归巢行为的基础上，提出的一种群智能优化算法。该算法主要通过地磁导航算子和地标导航算子来进行目标的搜索，具有需要调整的参数少、原理简单、收敛速度快等特点。由于其自身的优越性，自提出以来，已被投入到众多应用中。鸽群算法可以根据设定找寻最优的解，倘若我们将神经网络的结构参数看作待优化的解，那理论上我们是可以得到最优的网络结构参数，这在一定程度上可以避免大量的重复实验操作，节省时间和人力资源。但鸽群算法并不是完美的，在解决多峰值函数的优化问题时易陷入局部极值，从而降低了寻优精度。简单地讲就是原有的鸽群算法在处理一些相对复杂的函数时并不能找到真正的最优解，这可能使得得到的结果并不是我们想要的结果。同时，对于该算法的研究还不够成熟，理论成果也不是很多。因此，有必要参考过往群智能优化算法的改进思想对鸽群算法进行相应的改进，从而将鸽群算法运用到更多的应用领域。

鉴于上述研究，本次论文将长短时记忆网络应用于股指预测中，同时，对鸽群算法进行一定的探讨，使其寻优性能得到一定的改进，并利用改进的鸽群算法对长短时记忆网络中的结构参数进行优化，从而构建出预测结果更加准确的股指预测模型。最后，将构造的模型应用于上证股指预测，并与其它预测模型进行对比，以此展示组合模型的优势。

## 1.2 研究意义

首先，受股票市场复杂结构的影响，基本股指预测方法在预测时已经无法得到良好的效果，这就需要去探讨和寻找新的方法来解决这一问题。股指数据在一定程度上表现出了非线性且具有一定复杂性，而深度学习能够处理复杂的非线性问题，所以本文选用了深度学习的一种进行研究。本文将长短时记忆网络应用于股指预测中，为股指预测提供了一种方法，也为更多投资者在选购股票时提供了方向，保护了更多投资者的利益，避免盲目投资。

其次，鸽群算法作为新产生的群智能优化算法，应用广泛，但也暴露出了很多问题。本文尝试对鸽群算法进行了改进，对原有算法的寻优性能进行了再优化，进而在寻优精度方面得到了提高。鸽群算法的优化和改进使得该算法可以应用到更多的场景中，拓宽了算法的应用领域。同时，鸽群算法的改进思路

具有一定的参考价值，为其他群智能算法的优化和改进提供了一定的方向和思路。

最后，本文将改进的鸽群算法应用于优化长短时神经网络的结构参数，这在一定程度上可以提高股指预测的准确性，从而为投资者提供更准确的预测结果，进而降低投资风险，为投资者带来更高的收益，因此本文的研究具有一定的现实意义。

### 1.3 国内外研究现状

#### 1.3.1 统计模型在股票预测应用中的研究文献

股票价格的变动一直以来都是众多投资者所关注的一个热点话题。早期，学者尝试利用线性时间序列模型对股票价格进行预测，例如移动平均（MA）模型、自回归（AR）模型、自回归移动平均（ARMA）模型等等。李民（2000）在文章中曾利用自回归移动平均（ARMA）模型对深沪市股票指数进行短期预测，但该模型并不适用于长期预测。万建强、文洲（2001）分别利用 ARCH 模型和 ARIMA 模型对香港股指进行预测，结果表明 ARCH 模型要优于 ARIMA 模型。王立凤（2006）利用 ARCH 模型对股价进行了预测，同样得到了较好的预测效果。付燕等（2012）利用 ARMA 模型对中体产业股票进行了预测。经过多次实验研究和验证，结果表明对于金融数据，仅仅运用线性模型并不能得到很好的预测结果，于是学者们开始利用非线性模型进行预测。随着经济的发展和进步，学者们开始对 ARCH 模型进行改进和完善。于志军、杨善林（2013）采用误差修正的 GARCH 模型进行股价预测，经误差修正后，预测精度有所提高。沈传茂（2014）提出一种 RESN 补偿 ARMA 预测误差的体育股票价格预测模型（ARMA-RESN）。首先采用 ARMA 模型对体育股票价格线性特性进行模拟，使得残差序列仅含非线性特性，然后采用 RESN 模型对残差序列进行建模，模拟体育股票价格的非线性特性，最后将残差序列预测值补偿到 ARMA 预测值中，得到体育股票价格最终预测结果。相对于单一 RESN 模型和 ARMA 模型，ARMA-RESN 不仅可以精确刻画体育股票价格的变化特性，可以提高体育股票价格的预测精度。

一直以来，在金融市场上，股票一直扮演者非常重要的角色，经济形势与

股票市场密切相关。股票价格的变动在国外同样受到关注,这也引导更多国外学者开始对股票市场进行研究和探索。Sattayatham 等(2012)通过 GARCH 和 MRS-GARCH 对泰国证券交易所指数的波动性进行预测,结果表明 MRS-GARCH 模型要优于 GARCH 模型,且 MRS-GARCH 模型在长期预测中表现出色。Helmut(2017)使用股票经验收益率的滚动窗口来检验连续 GARCH 信息的独立性,为股价变动提供了有用的信息,使得研究更近一步。Laily 等(2018)分别运用 ARCH、GARCH 和 Elman 递归神经网络模型对股票日收盘价进行预测,结果表明 GARCH 模型预测效果最好。

### 1.3.2 神经网络在股票预测应用中的研究文献

随着社会的进步和发展,学者在金融领域进行统计预测时,不再仅仅局限于应用一些统计模型进行预测研究,开始逐步尝试利用神经网络等机器学习语言对金融市场进行相关的实验和探讨。例如,禹建丽等在(2003)文章中提出了一种基于 BP 神经网络的建模模型,对股票市场进行预测和决策分析。张玉川,张作泉将(2007)支持向量机的分类方法引用到了股票市场中,对股票的涨跌进行了预测,正确率在 60%以上。庞英智等将(2009)RBF 人工神经网络应用到了经济领域,对股票的价格进行了预测,并与 BP 神经网络进行了比较。郝知远将(2017)股市中常用的技术指标作为特征,利用支持向量机对股票价格的变化进行建模,旨在通过机器学习和数据挖掘方法对股票的涨跌进行预测。该方法对股票交易具有预测指导作用,可以帮助投资者简化分析步骤并增加获利。基于机器学习中神经网络存在的问题,学者们开始利用一些改进策略对现有的问题进行改善,以此提高原有神经网络的性能。譬如,杨小平在(2004)对股票数据进行主成分分析的基础上,利用 BP 神经网络对股票价格进行了拟合预测。孙延风(2005)针对 RBF 网络的优点,将最优分割法应用于径向基网络参数的训练过程中,提高了预测的正确率。蔡红(2011)为了提高股票价格预测精度,提出一个基于主成份分析(PCA)的 BP 神经网络(BPNN)股票预测模型(PCA-BPNN),模型预测精度显著提高,是一种高效和准确的股票预测模型。同样,魏文轩(2013)在利用主成分分析对数据进行降维处理的基础上,与 RBF 神经网络进行结合。李坤等采(2014)用小波支持向量机对股票价格进行回归,构造的模型在 3 个大盘指数以及 13 类不同行业的股票中取得了不错的效果。

深度学习的到来为神经网络掀起了一股新的热潮。徐浩然等(2020)经过

研究对股票预测中的机器学习算法进行了总结分析,指出随着机器学习中主要问题的显露,预测方法开始向深度学习研究领域深入发展。随着深度学习热潮的到来,学者们开始尝试将深度学习方法引进到股票市场中。LSTM神经网络作为深度学习的一种,在对股票市场进行预测时显得备受青睐。首先,邓凤欣等(2018)对LSTM神经网络应用于美港股票市场的可能性进行了研究,结果显示是可行的。其次,陈佳等(2019)人将LSTM神经网络应用于特征选取的股指预测中,在对纳斯达克股票指数及标普500指数进行预测时,准确度得到显著提升。梁宇佳等(2021)在对投资者情感进行分析的同时,将LSTM融合进来,对股价未来的走势进行了分析预测。除此之外,有学者对各神经网络的预测性能进行了对比分析。例如,乔若羽(2019)将多层感知机(MLP)、递归神经网络(RNN)、长短期记忆(LSTM)神经网络进行了对比实验。孙伯原学者(2019)将BP、RNN、LSTM三个神经网络应用到了红日药业的股票进行了预测,得到了较高的准确率,同时,LSTM在三个神经网络中预测的准确率最高。黄丽明(2019)针对股票的涨跌预测问题,提出一种基于多路循环神经网络与深度学习的股票预测方法,使用分布式向量表示方法提取出股票相关的新闻文本特征,同时考虑到股票相关信息的时序性以及新闻影响的持续性特质,使用多路循环神经网络模型对所提取的特征与交易信息进行协同训练,从而获得历史信息的低维向量表示。最后将多个循环神经网络的输出进行拼接,利用深度神经网络共同对股票的涨跌进行分类预测。

随着机器学习和深度学习的不断发展,越来越多的学者开始将其中的算法应用到股票市场,对股价进行更精准的预测。Huseyin等(2008)运用支持向量机对短期股票价格进行预测,取得了良好的预测效果。Ritika(2017)将深度学习应用到了股价预测中,发现命中率的准确率提高了15.6%。Sushma(2017)使用了反向传播前馈神经网络对孟买证券交易所上市的SBI进行股票预测,根据结果,反向传播前馈算法是稳健的。Gnanendra(2018)采用的研究方法是神经网络中的多层感知器,它是人工神经网络模型中的一种。人工神经网络模型是利用机器的认知学习进化出的高度灵活的函数算法。股票市场被认为是混沌、复杂、易变和动态的。毫无疑问,它的预测是时间序列预测中最具挑战性的任务之一。Rezaei(2020)采用经验模态分解(EMD)和完全集成经验模态分解(CEEMD)算法将时间序列分解为不同的频谱,提出了新的混合算法,即

CEEMD- CNN- LSTM 和 EMD- CNN- LSTM, 能够提取深度特征和时间序列, 最终应用于预测。Akshit (2021) 介绍了两个在股票预测中具有显著作用的重要算法: 人工神经网络(ANN)和支持向量机(SVM)。此外, 此文章还融合了 ANN 作为 ANN-MLP、GARCH-MLP 等混合模型、BP 算法与多层前馈网络相结合的新方法和新技术, 取得了较好的效果。Marwa (2021) 在文章中提出了一种基于长短期记忆(LSTM)、卷积神经网络(CNN)、支持向量机(SVM)、线性回归、Logistic 回归、K-邻域、决策树、随机森林、Stacked - LSTM、双向 LSTM 等学习模型的股价预测框架。Emioma (2021) 旨在利用机器学习算法对数据集的收盘股票价格进行估计, 以辅助股票价格的预测, 提高预测的准确性。通过研究实验, Kumar (2021) 提出了一种新的机器学习技术, 利用支持向量回归的微调版本对时间序列数据进行股票预测。通过对比发现, 该文章提出的方法对所选数据集的股票预测更加准确。此外, 所提方法所需时间远少于其对应方法。Budiharto (2021) 基于 R 语言和 LSTM 的统计计算, 提出了对印尼证券交易所股票价格预测的数据科学模型。在实验中, 我们利用数据科学对数据进行可视化, 预测并模拟了开放、高、低和关闭(OHLC)等多种参数情况。神经网络的引入, 为金融市场中存在的问题带来了更多的解决方案。

### 1.3.3 鸽群算法相关研究文献

鸽群算法具有一定的优点, 自提出后应用到了很多领域中。段海滨、叶飞 (2017) 阐述了鸽群算法的基本原理和优化模型, 并介绍了鸽群算法的几种典型应用场景, 例如, 控制参数优化、图像处理和无人机编队等。同时, 还指出可以尝试将鸽群算法与其它智能优化算法相结合, 从而增强鸽群算法的优化能力。胡春鹤等 (2019) 将鸽群算法应用到了图像分割领域, 并与 PSO、KWS 算法进行了对比, 经过实验, 结果表明混合后的算法运行结果最好。韩银、张赫 (2018) 将鸽群算法应用于优化粒子滤波算法, 实验结果表明, 与原算法相比, 精度有了很大的提升。鸽群算法属于一种新的群智能优化算法, 研究时间相对较短, 有待进行更深入的研究及探讨。

在国外研究中, 鸽群算法的研究相对较少。Duan 和 Qiao (2014) 在文章中提出了鸽群算法 (PIO), 该算法模拟了鸽群的归巢行为。鸽群在飞行的不同阶段会选取不同的导航工具, 根据鸽群的这一特性提出了地磁算子模型和地标算子模型。Chen 和 Duan (2017) 提出一种改进的图像匹配优化方法, 在鸽群算

法中引入了高斯变异，高斯变异（MGM）机制可以通过变异操作调节鸽子的位置。实验结果表明，提出的方法在收敛速度，鲁棒性和稳定性方面都可以达到较好的效果。

### 1.3.4 文献评述

随着股票交易数据的增加，股票数据表现出了复杂、非线性的特征，利用传统的线性预测方法已经不能得到较好的预测结果。随着社会的发展，深度学习方法用于非线性预测已经逐渐发展成为了学者关注和研究的热点。从国内外已有文献来看，长短时神经网络在股票预测中能够表现出良好的预测性能，可以应用长短时神经网络对股指价格进行相应的预测分析。但在神经网络参数调整方面，为了避免人为因素的影响，需要将其与优化算法搭配组合，对参数进行优化调整。鸽群算法作为新型的群智能优化算法，应用领域广泛，具有一定的研究价值。与此同时，鸽群算法可以进行参数调优，得到参数组合的最优解。因此，本文尝试将鸽群算法应用于优化长短时神经网络，并针对其自身不足进行优化和改进，以提高原有模型的预测性能。

## 1.4 研究思路和研究内容

### 1.4.1 研究思路

股指预测是指通过前一段时间的股指信息来对下一时间的股指价格进行预测的行为。有效的股指预测，不仅可以为投资者降低投资风险，实现收益最大化，而且可以刻画出金融市场的波动情况，从而对金融风险有更好的度量。本文主要研究思路如下：

由于影响股指的因素众多，且具有一定的非线性，所以股指预测是一个非常复杂的问题。而长短时记忆网络具有一定的非线性处理能力，所以本文采用长短时记忆网络对股指进行预测。在长短时记忆网络中，参数选择的不同，会影响模型的预测结果。为此，本文考虑利用鸽群算法较好的寻优性能，寻找长短时记忆网络的参数。

鸽群算法是近些年提出的一种群智能优化算法，应用领域广泛但其本身也具有局限性：在求解复杂问题时，容易出现“早熟”现象，且求解精度低。针对鸽群算法的缺点和不足，本文引入了两种机制，对鸽群算法进行改进。一方

面，在地磁导航阶段引入 Levy 飞行，利用 Levy 飞行步长的“跳跃性”跳出局部极值，从而向真实最优解逼近；另一方面，在地磁导航和地标导航阶段引入混沌搜索策略，每次迭代后，利用混沌理论的遍历性特点，在最优个体附近进行混沌搜索，该过程不仅可以加快收敛速度，也可以提高算法的精度。

利用改进的鸽群算法对长短时记忆网络进行寻优，可以得到更加准确或精确的参数。参数确定后，代入到长短时记忆网络，也就得到了更好的股指预测模型，从而可以更准确的预测出股指的走向。

### 1.4.2 研究内容

本文利用长短时记忆网络对股指进行预测，并运用改进的鸽群算法对记忆网络的参数进行优化，构建相应的股指预测模型，最后，本文利用构建好的混合模型进行预测，并对混合模型的预测性能进行判断和比较。全文一共分为五章：

第一章，首先探讨了文章的理论意义和研究价值。然后是文献综述，对国内外的相关文献进行研读，并对其主要观点进行提炼整合。另外，本章对本文的研究思路、研究内容以及论文中的创新点做了总结。

第二章，介绍了与股票价格预测相关的理论以及该过程中曾运用到的一些统计模型和神经网络的相关概念和原理。

第三章，对鸽群算法进行改进。针对鸽群算法的缺点和不足，引入 Levy 飞行及混沌搜索两种机制对基本鸽群算法进行改进，并选用经典测试函数分别在 5 维和 20 维下进行仿真测试，验证改进算法的寻优性能。

第四章，将改进的鸽群算法用于优化 LSTM 神经网络并应用优化后的组合模型对上证指数收盘价进行预测。将改进后的鸽群算法应用于 LSTM 中，对其结构参数进行优化，并选用上证指数数据进行实验分析。通过与其他模型对比来验证改进鸽群算法与 LSTM 模型的混合模型在进行股票预测时是有效的。

第五章，结论与展望。总结了本文的主要研究成果，并对未来的探讨方向进行说明。

## 1.5 本文创新点与不足

结合本文内容，主要的创新点主要体现在以下方面：

(1) 本文对鸽群算法进行了改进。针对鸽群算法在求解过程中易陷入“早熟”，且寻优精度低的缺点，在鸽群算法中引入了两种改进机制：第一，在地磁导航阶段引入 Levy 飞行，使得算法能够跳出局部极值；第二，在地磁导航和地标导航阶段引入混沌搜索策略，进行更细化的搜索，提高原有算法的求解精度。

(2) 本文将改进的鸽群算法应用于优化 LSTM 神经网络参数。LSTM 神经网络的预测性能会受到参数的影响，因此，本文利用改进的鸽群算法对参数进行优化，构建相应的混合预测模型，从而提高模型的预测性能。

(3) 本文将改进鸽群算法优化的 LSTM 神经网络应用于股指预测中，提高了股指预测的精度，为股指预测构造了一种新的方法。

虽然本文构造出了相应的预测模型，但仍存在一定的不足：

(1) 在对 LSTM 进行改进时，仅对隐含层的神经元数以及学习率进行了优化，其它参数的优化及改进仍有待探讨。

(2) 在对股指进行预测时，考虑了开盘价、收盘价、最高价、最低价、涨跌额和成交量等因素，但影响股指的因素众多，还需要继续进行研究。

## 2 股票价格预测理论与模型概述

股票市场自建立以来就获得了众多研究学者的关注。通过学者的不断探索，研究出了多种方法对股票价格进行预测。在对股票市场进行探索的早期，主要基于统计学理论和模型对股票价格进行预测。之后随着神经网络的出现，使得神经网络成为众多学者研究的重点，神经网络在时间序列中的预测也受到了广泛的关注。本章主要是对股票价格预测的相关理论方法进行介绍，为后续章节模型的建立奠定基础。

### 2.1 股票价格预测

随着股票市场的壮大，股票交易数量日益上升，股票价格的波动也变得更加复杂。作为投资者，首先最想了解的就是股票市场的变化规律，从而能够理性投资，获取收益。在学术界，股票价格的变动也引起了学者们的关注。随着研究的深入，逐渐形成了两种分析方法：基本面分析和技术面分析。

#### 2.1.1 基本面分析

基本面分析又称为基本分析法，该方法主要从影响股票价格变动的宏观经济政策、行业状况、公司业绩情况及股票的内在价值等因素出发，研究股票市场的现有规律并对股票价格的波动进行一定程度的评估。基本面分析可以概括为三个方面：宏观经济分析、行业分析、公司分析。

宏观经济分析探讨了现有经济指标及政策对价格变动的影响。涉及的经济指标有国内生产总值、通货膨胀率、利率、汇率等等。除此之外，涉及到的相关政策有货币政策、财政政策、税收政策、产业政策等等。公司分析通过对公司财务状况、市场前景、行业地位的研究来评估股票未来的走势及价值。公司盈利以及管理层的正确决策会给投资者带来正的收益，因此财务分析和管理层分析是公司分析的重点。行业分析则主要探讨了行业前景对股价的影响。

综上所述，可以看出基本面分析可以比较全面的对未来股价的变动趋势进行判断，但所需的时间跨度较长，预测精度较低。

### 2.1.2 技术面分析

技术面分析又称技术分析，通过对历史价格及交易数据进行研究，对未来股价进行预测。也就是说，技术分析将过往发生的历史交易信息作为研究对象，并利用图表和公式绘制出股价的变动趋势，进而从中获得股票市场交易的变化规律，对股价未来的变动进行评估和判断。

常用的技术面分析方法有 K 线组合分析，记录了一天、一周、一月中股价的涨跌变化，具有直观、信息量大等特点。

技术面分析对复杂的数理逻辑进行量化，能够清晰的反映出股票价格的变动情况，定量分析出何时是最佳的卖出和买入时机，但是该方法长线投资过程中是不适用的，对于一些突发的状况也不能及时做出响应。

## 2.2 常用的股票预测统计模型

早期的股价预测通常依托于统计学模型，这类预测方法通常有较强的解释性且预测效率高，在一些预测问题中能够表现出良好的预测性能。常用于股票预测的统计模型为 ARIMA 模型。

ARIMA (Autoregressive Integrated Moving Average model) 模型又称为差分整合移动平均自回归模型，属于时间序列数据预测方法的一种，ARIMA 模型在股票预测中的应用实例相对来说比较广泛。

ARIMA(p,d,q)模型的构造方法主要包括以下几点：

(1) 获取时间序列数据。获取的时间序列数据可以是经过实验分析后得到的，也可以是从统计部门获取的。对于取得的数据，要进行初步的检查和判断，查看数据中是否存在突兀点并分析产生这些数据的原因。在进行分析前要保证数据的准确性，从而保障了模型构造的正确性。

(2) 时间序列数据的预处理。在做预处理时，主要关注的是平稳性检验。对于 ARMA 模型，要求数据必须满足具有平稳性且非白噪声的条件。检验数据平稳性在模型构造的过程中非常重要，常常通过时序图和相关图进行检验。时序图直观上看比较简单，但是误差较大。相反，自相关图直观上相对复杂，但结果误差小，更加准确。本文在对时序图进行直观判断的基础上利用自相关图进行更进一步的检验，从而确定数据是否是平稳数据。对于非平稳数据，需要

进一步进行差分处理并进行平稳性检验，直到平稳为止。在此过程中，差分的次数就是模型 ARIMA 参数  $d$  的取值。一般来说，差分的次数越多，提取的信息越充分，但并不是差分次数越多越好。因为在每次差分的过程中都会造成一部分信息的损失，过多的差分可能造成不良的效果。因此，在一般的应用过程中，常常设定差分的次数不超过 2。

(3) 模型定阶。在确定了差分的次数后，还要进一步确定模型的参数  $p$  和  $q$ 。这两个参数的确定，可以采用 BIC 准则来确定。

(4) 模型拟合。模型定阶后，要通过模型进行预测，并进行显著性检验。如果检验过程中参数不显著，需要对参数进行调整。

(5) 残差检验。模型训练好后，需要进一步对模型中的残差进行白噪声检验。若检验显示残差为白噪声，则说明该模型对数据信息的提取比较充分。若检验出残差不为白噪声，则需要进一步调整。

若时间序列  $Y_t$  经过  $d$  次差分后，生成的平稳序列为  $y_t$ ，则可以将 ARIMA( $p,d,q$ ) 模型表示为：

$$y_t = c + \mu_1 y_{t-1} + \mu_2 y_{t-2} + \dots + \mu_p y_{t-p} + \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \varphi_q \varepsilon_{t-q} \quad (2-1)$$

其中， $c$  为常数， $\mu$  为自回归模型系数， $\varphi$  为移动平均回归系数， $\varepsilon_t$  为白噪声序列， $p$  为自回归模型阶数， $q$  为移动平均模型阶数。

根据 ARIMA( $p,d,q$ ) 的建模过程，可以看出该模型的预测结果容易受到参数  $p$ 、 $d$  和  $q$  的影响。参数  $d$  主要依据平稳性检验来确定，而  $p$ 、 $q$  的确定相对来说比较复杂。虽然现在已有自动选择  $p$ 、 $q$  的方法，但通常情况下计算效率低，无法得到很好的效果。同时，ARIMA 模型不支持增量更新操作，每次数据发生更新需要对模型参数进行重新判定，降低了该模型的应用效率，不能很好的刻画序列的变化趋势。

## 2.3 神经网络理论

随着人工智能的进步，神经网络得到了迅速发展，神经网络模型被应用到诸多领域。对于时间序列数据的预测问题，学者也开始尝试使用神经网络模型来解决。股票数据作为时间序列数据的同时还具有一定的非线性，而神经网络在一定程度上恰好能够解决非线性问题，这为神经网络应用于股指预测提供了可能。常用于股指预测的神经网络模型有 BP 模型、RNN 模型及 LSTM 模型，

下面对这 3 种模型的原理进行了相关介绍。

### 2.3.1 BP 神经网络

BP 神经网络是由 Rumelhart 和 McClelland 为首的科学家在 1986 年首次提出的，按误差逆向传播算法训练的多层前馈网络。BP 神经网络在神经网络模型中十分重要，可以很好的处理非线性数据问题，已被应用到众多研究领域。BP 神经网络是包含多个隐含层的网络，层与层之间进行全连接，而同一层之间的神经元不连接，其基本结构如图 2-1 所示，第一层为输入层，最后一层为输出层，中间部分为隐藏层，也可以称为隐含层。隐藏层不直接接受外界的信号，也不直接向外界发送信号，而是将输入层输入的信息转换为输出层可以利用的信息。

在 BP 神经网络的训练过程中，既包含从输入层到隐藏层，再从隐藏层到输出层的正向传播过程，也包含权值系数的反向调整过程。其原理如下：输入的信息按照正向传播的顺序从输入层传递到输出层，并计算实际输出与期望输出的误差大小。若误差值较大，则进入到反向调整阶段，即误差按照输出到输入方向，分配给各层的神经元，神经元再根据得到的误差值对连接权重和阈值进行调整。这是一个动态过程，通过调整权值和阈值，使得输出误差不断地向期望误差逼近。在达到设定的期望误差或迭代次数后，就得到了一个 BP 网络模型，可以利用该模型进行相应的预测。简单理解就是神经网络从训练数据中挖掘出潜在信息，然后利用这些重要信息进行预测。

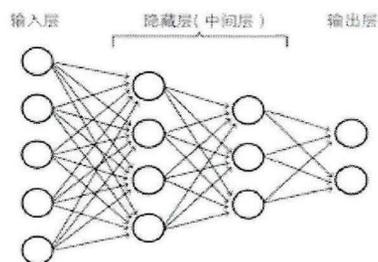


图 2-1 BP 神经网络基本结构

BP 神经网络的算法步骤主要包括以下三点：首先，根据研究问题，搜集原始数据并对数据进行相应的处理，如归一化、拆分训练集和测试集。其次，建立 BP 神经网络并对各个参数赋值，主要包括输入层、隐藏层及输出层的神经元个数，学习速率，训练精度及最大迭代次数等。最后，运用训练集数据训练 BP

神经网络。在得到网络模型后，用测试集数据进行测试，检验并评估该模型的性能。

BP 神经网络在使用过程中表现出很多优点，具体可以表述为以下几点：第一，非线性映射能力强，有理论证明三层的神经网络就可以以任意精度逼近任何非线性连续函数，这就表明一个简单的 BP 神经网络就可以实现复杂的求解问题；第二，BP 神经网络有较强的自学习能力，可以自动学习并提取输入和输出之间的关联信息，并将这些信息记录在权值中；第三，BP 神经网络可以单独处理每个神经元的信息，有较强的信息并行处理能力。除了上面几点外，BP 神经网络还具有信息分布式存储、泛化性强、高鲁棒性等优点。

BP 神经网络在处理非线性数据的预测问题时有着重要的意义，得到了广泛的应用。但随着研究的深入，BP 神经网络也暴露出了一些缺点和不足。主要有以下几点：

(1) 收敛速度慢。BP 神经网络处理的大都是非线性复杂问题，在输出接近 0 或 1 时，存在一些平坦区域，在这些地方，权值的变动很小，造成输出误差无法很好的向期望误差逼近，训练过程几乎处于停滞状态。

(2) 易陷入局部最优。BP 神经网络采用的是梯度下降法，在解决复杂非线性问题时，通过局部搜索不断地调整权值，往往使权值陷入局部极值点，最终导致训练失败。BP 网络对初始的权值和阈值比较敏感，采用不同的初始权重和阈值训练网络，常常会收敛到不同的局部极值点，这也是 BP 神经网络训练不稳定的主要原因。

(3) BP 神经网络的结构选择不统一。BP 神经网络的结构选择至今没有统一的理论指导，一般根据经验确定。结构庞大会使运行速度减慢，甚至可能出现过拟合现象，导致模型预测性能较差。结构设置太小则可能导致模型不收敛。神经网络结构的选择直接影响到模型的性能，因此，在建立模型时要慎重选择模型的结构。

### 2.3.2 循环神经网络

在 1990 年，Jeffrey Elman 提出了第一个全连接的 RNN 循环神经网络，即 Elman 网络。与 BP 神经网络结构类似，RNN 神经网络同样包含输入层、隐藏层及输出层三部分，如图 2-2 所示。 $x$  是一个向量，表示的是输入层所输入的值。 $u$  表示的是输入层到隐藏层的权重矩阵。 $s$  代表的是隐藏层的值，也是一个向量。

$v$  同样是一个权重矩阵，表示隐藏层到输出层权重的大小。 $o$  是向量，代表输出值。这里  $w$  也是一个权重矩阵，代表着权值的大小，但不同的是此权重矩阵并不是某一层到另一层的权值，而是上一隐藏层到此次隐藏层的权重。从 RNN 的网络结构可以看出，RNN 神经网络会对之前的输入进行记忆并应用到最终的输出运算中，也就是说隐藏层之间的节点不再是无连接的，而是有连接的，即隐含层的输入不仅包含输入层的输出还包含着上一次隐藏层的输出。

循环神经网络的训练算法为 BPTT 算法，与 BP 神经网络类似，BPTT 算法的原理也同样包含三个步骤：一是前向计算每个神经元的输出值；二是反向计算每个神经元的误差项值；三是计算每个神经元的梯度。

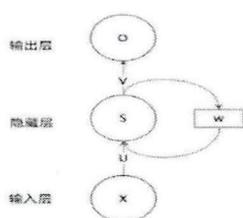


图 2-2 RNN 神经网络基本结构

RNN 神经网络虽然可以处理整个时间序列的信息，但是其记忆最深的还是近期时间段内输入的一些信息。所接收到信号的强度会随着时间的推移变的越来越弱，这就造成早期所产生的信息在模型中只能起到一点辅助作用，使得输出结果仅依赖于近期产生的数据信息。RNN 神经网络对于时间间隔较远的的数据信息难以记忆，因此长时依赖是 RNN 的一个致命缺点。同时，在多层训练过程中，RNN 可能会产生梯度消失和梯度爆炸等问题。RNN 的这些缺点致使该神经网络在早期的应用中产生的作用并不明显，也就慢慢的淡出了人们的视野。

### 2.3.3 长短时记忆神经网络

LSTM (Long Short-Term Memory) 神经网络又称为长短时记忆神经网络，是由 Hochreiter 和 Schmidhuber 在 1997 年首次提出并发表的，在二十一世纪初已经发展成为了深度学习算法的一种。同时，LSTM 也是循环神经网络的一种，但 LSTM 解决了一般循环神经网络 (RNN) 的长期依赖问题，能够对时间序列中间隔和延迟时间非常长的事件做出更好的处理和预测，因此，长短时记

忆神经网络在自然语言处理、计算机视觉、识别分类及拟合预测领域都有所涉略。

作为 RNN 神经网络的一种变体，LSTM 同样可以处理序列数据。相对于 RNN 网络，LSTM 的最大不同是设置了 3 种门结构，分别是遗忘门、输入门及输出门。LSTM 能够通过这 3 种门的结构决定某些信息的输出，实现了信息的有选择通过，进而使得该模型具有长期记忆性。同时，LSTM 增加了一个信息存储记忆单元，避免了梯度消失和爆炸问题。LSTM 神经网络的结构图如图 2-3 所示。

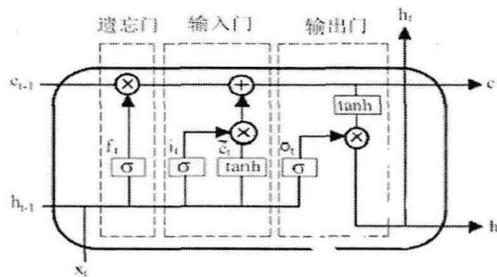


图 2-3 LSTM 神经网络基本结构

根据 LSTM 神经网络的结构图，下面详细的了解一下各部分的工作。

### (1) 遗忘门

LSTM 网络通过遗忘门决定要从细胞状态中丢弃掉什么信息，它是由一个 sigmoid 网络层实现的，用来忘记那些导致错误预测的信息。 $h_{t-1}$ 、 $x_t$ 为遗忘门的输入，通过 sigmoid 函数，返回的是一个介于 0 和 1 之间的数。若靠近 0，则表示应该舍弃该信息。若靠近 1，则表示应该保留该信息。遗忘门的作用就是帮助 LSTM 网络判断哪些信息需要保留，哪些信息需要舍弃。

### (2) 输入门

输入门用来决定要在细胞状态中存储什么新信息。此过程包含两个部分：首先，Tanh 网络层会创建一个可以添加到细胞状态的向量  $\tilde{c}_t$ 。其次，sigmoid 网络层的作用与遗忘门类似，为  $\tilde{c}_t$  中的每个值输出一个介于 0 和 1 之间的数值，以此来决定要更新哪些状态值，以及所要更新的程度。在这一过程中，需要结合这两个网络层来构造状态的更新值。

### (3) 输出门

基于上面所得到的新的细胞状态，需要决定要输出什么信息。同样，通过

sigmoid 网络层判断要输出哪些部分的细胞状态，然后将细胞状态通过 tanh 网络层，并将结果与 sigmoid 层的输出相乘，这就使得输出门只输出所决定输出的部分。

### (4) 单元细胞状态更新

在前面我们了解到遗忘门用于决定是否遗忘之前的信息，输入门用来决定是否更新当前的状态，在这两者的基础，就需要将旧的细胞状态更新为新的细胞状态了，即将  $C_{t-1}$  更新为  $C_t$ 。用旧的状态乘以  $f_t$ ，遗忘门决定忘记的信息，在此基础上再加上更新值  $i_t * \tilde{C}_t$ ，便获取到了新的信息状态。

经过上述过程，可以得到新的隐含状态和新的单元状态，并将此过程产生的结果传递到下一时刻。虽然三个门在任务分工上有所不同，但在操作过程中使用的性质是一致的，都是通过 sigmoid 网络层进行数据信息的选择，通过 tanh 网络层进行数据信息的更新，将两者结合起来，协同实现最终的效果。

LSTM 作为神经网络的一种，可以很好的解决复杂的非线性问题。同时，相比于 RNN，该模型在每一层神经元中都加入了“门”的结构，这使得误差在传播过程中，无论传播多远，都不会发生梯度完全消失的现象。同时，在 LSTM 神经网络中，记忆信息包含了神经元运算结果中未被保存的数据，解决了数据的长期依赖问题。因此，该模型在处理时间序列数据时具有显著优势。

## 2.4 本章小结

本章为股价预测的相关理论和模型，对常用的股价预测的一些常用方法进行了详细的阐述。ARIMA 模型作为常用的时间序列预测模型，原理简单且易于实现，但参数确定比较复杂，只适用于小规模时间序列数据的预测。在神经网络预测模型中，BP 网络模型没有时序概念，在处理时序数据时有一定的限制。RNN 神经网络虽然引入了时序的概念，但该模型在长期记忆方面具有缺陷。LSTM 神经网络在 RNN 的基础上加入了“门”的结构，解决了数据的长期依赖问题，在处理复杂时间序列数据时具有一定的优越性。

### 3 改进鸽群算法

在预测领域，预测精度一直都是学者和业界最为关心的话题。预测精度很大程度上受预测模型的参数选择影响。为了提高 LSTM 模型的预测精度，本文先对其模型参数进行优化。而群智能优化算法作为一种新兴的仿生优化算法，因具有简单易操作且高效等特点，常常被用于参数优化领域。本文选取群智能优化算法中新发展起来的鸽群算法来优化 LSTM 模型参数。并且，本章拟将鸽群算法进行改进，以便更好地对 LSTM 模型进行参数优化，提高其在应用中的预测精度。

#### 3.1 鸽群算法

生物界中一些动物的群体行为总是让人惊叹，科学家们通过观察这些动物的行为，总结出一定的规则和随机过程，并构建出相应的模型，这便是大家所熟知的群智能算法。通过这些算法可以解决实际生活中一些问题，具有一定的研究意义。鸽子在很久以前常常被人们当作通信工具传递信息，这正是利用了鸽子的自主归巢能力。2014 年段海滨等人对鸽群的这种行为进行了归纳，并提出了鸽群算法，亦称为 PIO 算法。

##### 3.1.1 鸽群算法基础理论

PIO 算法是仿生群智能优化算法的一种，根据鸽子的自主归巢行为提出。在鸽子飞行的过程中，主要有 3 个因素会影响鸽群的飞行路径，分别是太阳、地球磁场和地标。在不同飞行阶段，鸽群会在不同导航工具的引领下飞行。有学者发现，在鸽子开始飞行的过程中，主要依靠一种类似指南针的导航工具。在飞行一段时间后，会改变导航工具，即以地标为导航工具，并对当前的路径进行判别和修正。根据鸽子的这种特殊飞行行为，PIO 算法归纳出了两种算子模型：地磁算子模型和地标算子模型。

在地磁导航阶段，鸽群可以根据自身对地磁的感应，将简要的地图绘制在脑中，然后再根据太阳的高度辨别该向那个方向前进。越接近目的地，对导航

工具的依赖性越弱。在地标导航阶段，即在鸽子将要达到目的地时，会将导航工具切换成附近的地标。若鸽子很清楚当前地标的位置，将直接飞向地标所在地。若鸽子对地标位置不够熟悉，就会尾随那些熟悉地标的鸽子飞行。

下面用虚拟的鸽子对鸽群算法的整个过程进行模拟。首先，对初始过程进行模拟。每只鸽子对应问题的一个解，且每只鸽子可以用位置和速度两个因素来定义。假设在  $D$  维空间，有  $N$  只鸽子构成一个鸽群，即  $X=(X_1, X_2, \dots, X_N)$ ,  $V=(V_1, V_2, \dots, V_N)$ ,  $X, V$  分别代表  $N$  只鸽子的位置和速度。其中，第  $i$  只鸽子的位置和速度分别由式 (3-1) 和 (3-2) 表示 ( $i=1, 2, \dots, N$ )。

$$X_i = [x_{i1}, x_{i2}, \dots, x_{iD}] \quad (3-1)$$

$$V_i = [v_{i1}, v_{i2}, \dots, v_{iD}] \quad (3-2)$$

在开始过程中，利用下面的公式初始化鸽群的位置和速度 ( $i=1, 2, \dots, N$ ;  $j=1, 2, \dots, D$ )。

$$x_{ij} = lb + rand(ub - lb) \quad (3-3)$$

$$v_{ij} = rand \quad (3-4)$$

其中， $lb$  为  $x_{ij}$  的下限， $ub$  为  $x_{ij}$  的上限， $rand$  是  $0 \sim 1$  范围内的一个随机数。

然后，对地磁导航过程进行模拟。在地磁导航阶段，鸽子当前的速度不仅由上一次迭代过程中的速度所决定，还要受最优位置和当前所在位置的影响。每只鸽子的速度及位置更新公式如下：

$$V_i' = V_i * e^{-Rt} + rand(X_{gbest} - X_i) \quad (3-5)$$

$$X_i' = X_i + V_i' \quad (3-6)$$

其中， $R$  代表地磁因子， $t$  为当前所达到的迭代次数， $X_{gbest}$  是前一次迭代过程中全局最优的鸽子所在的位置。在地磁导航阶段达到设定的迭代次数时，开始转向地标导航阶段。

最后，模拟了地标导航过程。在地标导航阶段，如式 (3-7) 所示，每迭代一次，鸽群的数量就会相应的减半。这是因为在迭代过程中质量好的鸽子对地标比较熟悉，能够很好的分辨飞行路径，所以保留了下来。而那些质量差的鸽子不熟悉地标，没有辨别路径的能力，因此被淘汰掉。留下来的鸽群，会以中

心位置为地标，即向中心位置前进，如式（3-8）和（3-9）所示。在迭代到最大迭代次数后，地标导航也就停止了。

$$N' = N/2 \quad (3-7)$$

$$X_c = \frac{\sum X_i * fitness(X_i)}{N' * \sum fitness(X_i)} \quad (3-8)$$

$$X_i' = X_i + rand(X_c - X_i) \quad (3-9)$$

其中， $X_c$ 是鸽群数量减半后的中心位置。 $fitness(X_i)$ 是第*i*只鸽子的适应度值。

### 3.1.2 鸽群算法的步骤

假设要求解的目标函数为  $f(x)$ ，记迭代次数为  $NC$ ，鸽群算法的主要步骤如下：

步骤 1：参数初始化，这里包括种群数  $N$ ，地磁因子  $R$ ，地磁导航次数  $NC1$  和地标导航次数  $NC2$ ，且  $NC1 > NC2$ 。

步骤 2：利用公式（3-3）和（3-4）初始化种群，并比较每只鸽子的适应度值，将适应度值最好的鸽子记为  $X_{gbest}$ 。

步骤 3：进行地磁导航。首先，对每只鸽子的位置和速度进行更新。然后，计算适应度值并与  $X_{gbest}$  处的适应度值比较，若  $X_i$  处的适应度值优于  $X_{gbest}$  处适应度值，则令  $X_{gbest} = X_i$ ，否则，保持  $X_{gbest}$  不变。

步骤 4：若迭代次数  $NC > NC1$ ，停止地磁导航并转向步骤 5。否则，转向步骤 3。

步骤 5：进行地标导航。首先，对所有的鸽子按照适应度值大小进行排序。然后，按照公式（3-7）减半，并计算出当前鸽子的中心位置  $X_c$ 。最后，根据公式（3-9）更新鸽子的位置并计算当前位置的适应度值。与  $X_{gbest}$  处的适应度值进行比较，并判断是否更新  $X_{gbest}$ 。

步骤 6：若  $NC > NC1 + NC2$ ，则停止地标导航并输出结果。否则，转至步骤 5。

### 3.1.3 鸽群算法的优缺点

与其它群智能算法相比，PIO 算法有以下特点：第一，需要调整的参数少、容易实现。针对不同的优化问题，除了迭代次数、种群规模等必要信息外，只需要对适应度函数和取值范围做出调整，便可以得到相应目标函数的最优解。第二，鲁棒性强。第三，收敛速度快。相比之下，鸽群算法有较快的收敛速度，

能够在迭代初期就逼近于最优解附近。

鸽群算法的提出为解决优化问题提供了新的思路，得到了广泛的应用。但随着研究的深入，鸽群算法的缺点和不足也逐渐开始显露出来。在解决复杂优化问题时，鸽群算法容易陷入到局部最优，使得鸽群算法不能向真正的最优值靠近，进而降低了算法的精度。

## 3.2 改进策略

PIO 算法收敛速度快、容易实现、需要调整的参数少，为众多研究学者提供了新的方法和思路，但是其缺点和不足也在研究过程中暴露出来。在前面的章节中已经明确的指出了鸽群算法的缺点，这为 PIO 算法的改进提供了空间。针对鸽群算法的不足，本文提出了两种机制，下面主要讲述了这两种机制的原理、作用以及引入到原有鸽群算法过程中的具体操作步骤。

### 3.2.1 Levy 飞行

Levy 飞行是随机游走的一种，有研究表明，很多生物活动的轨迹都是满足 Levy 飞行的。Levy 飞行过程中的步长满足 Levy 分布，其密度函数如式 (3-10) 所示：

$$L(s) \sim |s|^{1-\beta}, \quad 0 < \beta \leq 2 \quad (3-10)$$

式中， $s$  为 Levy 步长，即  $Levy(\beta)$ ， $s$  的具体计算公式为：

$$s = \frac{\mu}{|v|^{1/\beta}} \quad (3-11)$$

公式 (3-11) 中，参数  $\beta$  一般取值为 1.5，参数  $\mu$ 、 $v$  均服从正态分布，即参数  $\mu$  的均值为 0，方差为  $\sigma_\mu^2$ ，参数  $v$  的均值为 0，方差为  $\sigma_v^2$ ，同时， $\sigma_\mu$ 、 $\sigma_v$  满足如下条件：

$$\begin{cases} \sigma_\mu = \left[ \frac{\sin(\pi\beta/2) + \Gamma(1+\beta)}{\Gamma(1+\beta/2) * \beta * 2^{(\beta-1)/2}} \right]^{1/\beta} \\ \sigma_v = 1 \end{cases} \quad (3-12)$$

Levy 飞行展现的是一种随机状态，图 3-1 展示了进行 500 次随机 Levy 飞行时的每次的飞行步长，从图中可以看出，Levy 飞行是长时间短距离搜索与偶尔长距离搜索的混合，这种“跳跃”性增加了算法跳出局部极值的可能性，进而提高了算法的精度。

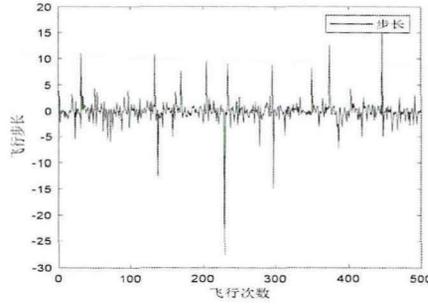


图 3-1 500 次 Levy 飞行步长

本文将 Levy 飞行过程产生的随机步长引入到鸽群算法的地磁导航阶段，对鸽群中的个体进行更新，利用步长的这种“跳跃”性，改善鸽群算法易陷入局部极值的缺陷。个体更新方式为：

$$X_i^l = X_i + \alpha \otimes \text{Levy}(\beta) \quad (3-13)$$

$$\alpha = \alpha_0 * (X_i - X_{gbest}) \quad (3-14)$$

其中， $i=1, 2, \dots, N$ ； $N$  为种群规模； $X_i^l = (x_{i1}^l, x_{i2}^l, \dots, x_{iD}^l)$  为引入 Levy 飞行后更新得到的个体的位置； $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$  是基本鸽群算法进行地磁导航更新后得到的个体位置； $X_{gbest} = (x_{gbest,1}, x_{gbest,2}, \dots, x_{gbest,D})$  为当前适应度值最优的个体所在位置； $\alpha$  为步长信息，用来对 Levy 步长进行动态调节， $\alpha_0$  为调节参数； $\text{Levy}(\beta)$  是通过 Levy 飞行得到的飞行步长； $\otimes$  为点乘符号。

Levy 飞行操作的具体步骤如下：在 PIO 算法的地磁算子中，利用式 (3-13) 对地磁导航产生的新个体  $X_i$  进行再更新得到  $X_i^l$ ， $i=1, 2, \dots, N$ 。若  $X_i^l$  处的适应度值优于  $X_i$  处的适应度值，则更新当前  $X_i$ ，即令  $X_i = X_i^l$ ；否则，保持  $X_i$  不变。Levy 飞行的引入增强了算法的全局搜索能力，可以有效的避免陷入局部最优。

### 3.2.2 混沌理论

混沌是一种伪随机现象，即看似随机，却有着一定的规律。除了上述特点外，混沌还具有遍历性，即混沌变量可以经历混沌区间内的所有状态，且在有限时间内，随着时间的推移所经历的状态是不同的。本文采用 Logistic 映射公式产生混沌序列：

$$K_{t+1} = \mu K_t [1 - K_t], K_t \in [0,1] \quad (3-15)$$

式中,  $\mu$  为参数, 取值范围在 3.56 到 4 之间; 当  $\mu = 4$  且  $0 \leq K_t(0) \leq 1$  时, 为完全混沌状态。

混沌的遍历性和不可预测性使其可以作为一种优化机制引入到算法中。混沌的引入, 进一步增强了算法的局部搜索能力, 加快了算法的收敛速度。本文将混沌搜索引入到地磁导航和地标导航阶段, 针对两个算子模型迭代过程中的最优个体, 添加混沌变量进行更细化的搜索:

$$X'_{gbest}(t) = X_{gbest} + b * H_j(t) \quad (3-16)$$

式中,  $b$  为调节参数;  $H_j(t)$  是在  $[-1, 1]$  区间按照混沌公式产生的随机数, 是  $K_t$  经过放大平移后得到的。

引入混沌搜索的具体步骤如下: 针对地磁算子和地标算子迭代过程中得到的最优个体, 运用式 (3-7) 进行混沌搜索, 构造出多个  $X'_{gbest}$ , 并与  $X_{gbest}$  处对应的适应度值进行比较, 找到其中适应度最优的个体, 并更新  $X_{gbest}$ 。利用混沌变量进行细化的局部搜索, 可以有效改善算法的精度, 得到更优质的解。

### 3.2.3 改进鸽群算法的具体步骤

针对 PIO 算法的缺陷, 本文引入了两种机制来改进基本的 PIO 算法。首先, 在地磁算子模型中, 引入了 Levy 飞行, 对鸽群中的每个体进行更新, 促使鸽群跳出局部极值点, 具有全局性; 其次, 在地磁算子和地标算子模型中, 对迭代过程中的最优个体添加混沌扰动, 进行更细化的搜索, 提高了算法的局部搜索能力, 保证了算法的寻优精度。在后续章节中, 将改进的鸽群算法简记为 LPIO 算法。

LPIO 算法的具体步骤如下:

步骤 1: 设定相关参数, 并进行种群初始化。

步骤 2: 计算适应度值, 并保留适应度值最优的鸽子;

步骤 3: 进入地磁导航阶段;

步骤 4: 判断是否要引入 Levy 飞行。若最优值在连续多次迭代后变化很小, 则开始利用 Levy 飞行对每个鸽子的位置进行更新, 否则, 直接跳至步骤 5;

步骤 5: 利用地磁导航公式更新鸽子的位置和速度;

步骤 6: 计算各个鸽子的适应度值, 找到适应度值最优的鸽子所在位置, 并在其附近进行混沌搜索, 产生新的鸽子, 并与最优鸽子进行比较。若适应度值

优于最优鸽子的适应度值，则更新最优鸽子的位置；

步骤 7：判断是否达到地磁导航的最大迭代次数，若已达到，则停止地磁导航并转向步骤 8；否则，转向步骤 3；

步骤 8：启用地标导航；

步骤 9：利用地标导航公式对鸽群位置和速度进行更新；

步骤 10：同步骤 5。

步骤 11：判断是否达到地标导航的最大迭代次数，若已达到，则停止地标导航并输出结果；否则，转至步骤 8。

### 3.3 仿真测试及结果分析

本节主要对改进的鸽群算法进行仿真测试，从收敛速度和收敛精度两个方面来展示改进算法的性能。本次实验使用了 5 个测试函数，包含 3 个单峰函数和 2 个多峰函数。实验时，分别对改进前和改进后的 PIO 算法独立运行 50 次，测试结果取 50 次实验的平均值。同时，本文将本次改进的鸽群算法与朱震宇等学者（2019）提出的高斯鸽群算法（GPIO）、段海滨、杨之元（2018）提出的柯西变异鸽群算法（CMPIO）进行了对比，以彰显本次改进的优越性。

#### 3.3.1 参数设置

本次测试在 MATLAB 2020a 中运行，具体的参数设置如表 3-1 所示。

表 3-1 鸽群算法运行时参数

参数	参数含义	参数值	参数应用于
N	种群规模	50	
T1	地磁导航最大迭代次数	450	PIO 算法、
T2	地标导航最大迭代次数	50	LPIO 算法
R	地磁因子	0.3	
$\alpha$	Levy 飞行调节参数	5	
b1	地磁导航过程混沌搜索的调节参数	1	LPIO 鸽群算法
b2	地标导航过程混沌搜索的调节参数	0.1	

#### 3.3.2 测试函数

本文为了测试改进后的鸽群算法的性能，选择了 5 个典型的测试函数进行仿真实验，其中， $F_1$ 、 $F_2$ 、 $F_3$  为单峰函数， $F_4$ 、 $F_5$  为多峰函数。本次测试分别

在 5 维和 20 维情况下进行仿真,根据算法求解的目标值情况来判断算法的性能,各个测试函数的公式、定义域、峰值信息如表 3-2 所示。

表 3-2 测试函数及其定义域、峰值

测试函数	定义域	峰值
$F_1 = \sum_{i=1}^D x_i^2$	[-100,100]	0
$F_2 = \sum_{i=1}^D ix_i^2$	[-5.12,5.12]	0
$F_3 = -\exp(-0.5 \sum_{i=1}^D x_i^2)$	[-1,-1]	-1
$F_4 = -20 \exp\left(-0.2 \sqrt{\frac{1}{D} \sum_{i=1}^D x_i^2}\right) - \exp\left(\frac{1}{D} \sum_{i=1}^D \cos(2\pi x_i)\right) - 20 + e$	[-32,32]	0
$F_5 = \sum_{i=1}^D  x_i \sin(x_i) + 0.1 x_i $	[-10,10]	0

### 3.3.3 改进结果分析

为了体现改进的鸽群在寻优过程中的表现,将 5 个测试函数分别在 5 维和 20 维情况下进行了仿真,结果主要以仿真得到的数据和迭代曲线图展示。通过仿真得到了 PIO、LPIO 算法在求解各个函数时的最优解,最差解,平均值等数据,这些数据可以直接的展现算法的优化性能,而迭代曲线图则可以看出算法在整个迭代过程中的寻优趋势,也可以根据迭代曲线判断算法是否陷入局部极值。同时,迭代曲线的斜率大小展示了算法的寻优速率,曲线越陡,寻优速率越快,表明算法能够更快的跌落在目标值附近。迭代曲线越接近 x 轴,就越接近目标解。

表 3-3 和 3-4 是 PIO、LPIO 算法为在 5 维和 20 维情况下的仿真结果,包含了最优解、最差解及平均解数据,从结果可以看出,LPIO 算法在 50 次独立实验后的平均适应度值要优于原有的 PIO 算法,尤其在高维函数中表现的更为明显。在 5 维仿真结果中,LPIO 算法在求解  $F_1$ 、 $F_2$ 、 $F_3$ 、 $F_5$  函数时基本都能够找到真实最优解。在求解  $F_4$  时虽然没有达到真实最优解,但也已经十分逼近。在 20 维仿真结果中,原有的 PIO 算法得到的效果不太理想,不能很好的靠近最

优解。而 LPIO 算法仍能够很好的靠近真实最优解，甚至 50 次实验得到的最差解都要远远优于原有 PIO 算法得到的最优解。由此可见，与 PIO 算法相比，LPIO 算法在求解精度上有了很大程度的提升。同时，从结果中可以看出 LPIO 算法的寻优能力要优于 GPIO 和 CMPIO，进一步说明本文对鸽群算法的改进是可行的。

表 3-3 PIO 及其改进算法在 5 维测试函数上的仿真结果

测试函数	算法	最优解	最差解	平均值
$F_1$	PIO	4.5730e-04	16.6553	1.0353
	GPIO	2.1449e-12	5.3674	0.1867
	CMPIO	1.5206e-04	0.1708	0.0576
	LPIO	0	0	0
$F_2$	PIO	0	1.2039	0.1248e-12
	GPIO	3.1564e-13	0.7355	0.0790
	CMPIO	5.7528e-05	0.1897	0.0288
	LPIO	0	0	0
$F_3$	PIO	-1	-0.9488	-0.9923
	GPIO	-1	-0.9842	-0.9989
	CMPIO	-1	-0.9981	-0.9996
	LPIO	-1	-1	-1
$F_4$	PIO	2.3232e-08	4.1682	1.6431
	GPIO	1.8148e-06	3.1777	0.5679
	CMPIO	0.0246	1.9379	0.8963
	LPIO	8.8818e-16	2.2204e-14	7.0699e-15
$F_5$	PIO	1.9963e-19	2.2192	0.4902
	GPIO	2.5888e-09	1.2090	0.2636
	CMPIO	0.0027	0.9621	0.1741
	LPIO	0	1.3117e-13	9.7050e-15

表 3-4 PIO 及其改进算法在 20 维测试函数上的仿真结果

测试函数	算法	最优解	最差解	平均值
$F_1$	PIO	415.3671	2.1757e+03	1.1283e+03
	GPIO	432.2194	2.3782e+03	1.1277e+03
	CMPIO	2.0332	21.6011	8.1246
	LPIO	3.3010e-17	3.7746e-10	1.0196e-11
$F_2$	PIO	43.0898	233.9764	127.9887
	GPIO	0.3273	234.0546	131.6008
	CMPIO	8.0314	84.9189	28.2923
	LPIO	2.1664e-15	1.7163e-05	3.5357e-08

### 3 改进鸽群算法

续表 3-4 PIO 及其改进算法在 20 维测试函数上的仿真结果

测试函数	算法	最优解	最差解	平均值
$F_3$	PIO	-0.6692	-0.3311	-0.5414
	GPIO	-1.0000	-0.5004	-0.7050
	CMPIO	-1.0000	-0.9911	-0.9978
	LPIO	-1.0000	-1.0000	-1.0000
$F_4$	PIO	7.7432	13.4752	10.3512
	GPIO	0.1742	12.7682	8.3670
	CMPIO	3.7503	5.8170	5.0442
	LPIO	1.6114e-09	1.8400	0.1213
$F_5$	PIO	5.0153	17.8932	11.0322
	GPIO	0.0394	19.1913	9.6886
	CMPIO	3.5552	14.8518	8.6591
	LPIO	8.1496e-11	0.0103	3.7491e-04

表 3-5 PIO 和 LPIO 算法的稳定性统计

维数	算法	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$
5	PIO	7.3974	0.0537	1.2975e-04	1.2966	0.2621
	GPIO	0.5866	0.0340	1.0501e-05	0.6240	0.1134
	CMPIO	0.0016	0.0017	1.2989e-07	0.3738	0.0388
	LPIO	0	0	0	1.8448e-29	5.8513e-28
20	PIO	1.3160e+05	2.3442e+03	0.0072	2.1811	6.7134
	GPIO	2.9635e-21	2.4344e+03	0.0185	11.0843	22.0593
	CMPIO	2.3828e+05	299.4628	3.0444e-06	0.2106	6.7303
	LPIO	4.9683e-09	5.8865e-14	3.8055e-20	0.1522	2.2778e-06

表 3-5 为算法独立运行 50 次的稳定性统计，是 50 次所求解的方差。方差越小，证明算法的稳定性越好。从表中可以看出，不管是在 5 维还是 20 维，LPIO 算法的方差均比其它算法的方差小，且 LPIO 的方差可以稳定在一个很小的数值范围内，表明 LPIO 算法有很好的稳定性。

图 3-2 是两种算法的在 20 维情况下的迭代曲线图，是根据 50 次独立实验后的均值绘制的。从迭代曲线图，可以看出 PIO 算法在收敛速度方面表现很好，但该算法在处理复杂的目标函数时寻优性能不够稳定，容易陷入局部最优，导致算法精度下降。LPIO 算法保留了原有算法前期的寻优速率，能够很快跌落在最优解范围附近，同时，改进后的算法可以通过不断地迭代跳出局部极值，向

真正的最优解靠近，有效的提高了算法的寻优性能。通过迭代曲线图，可以看

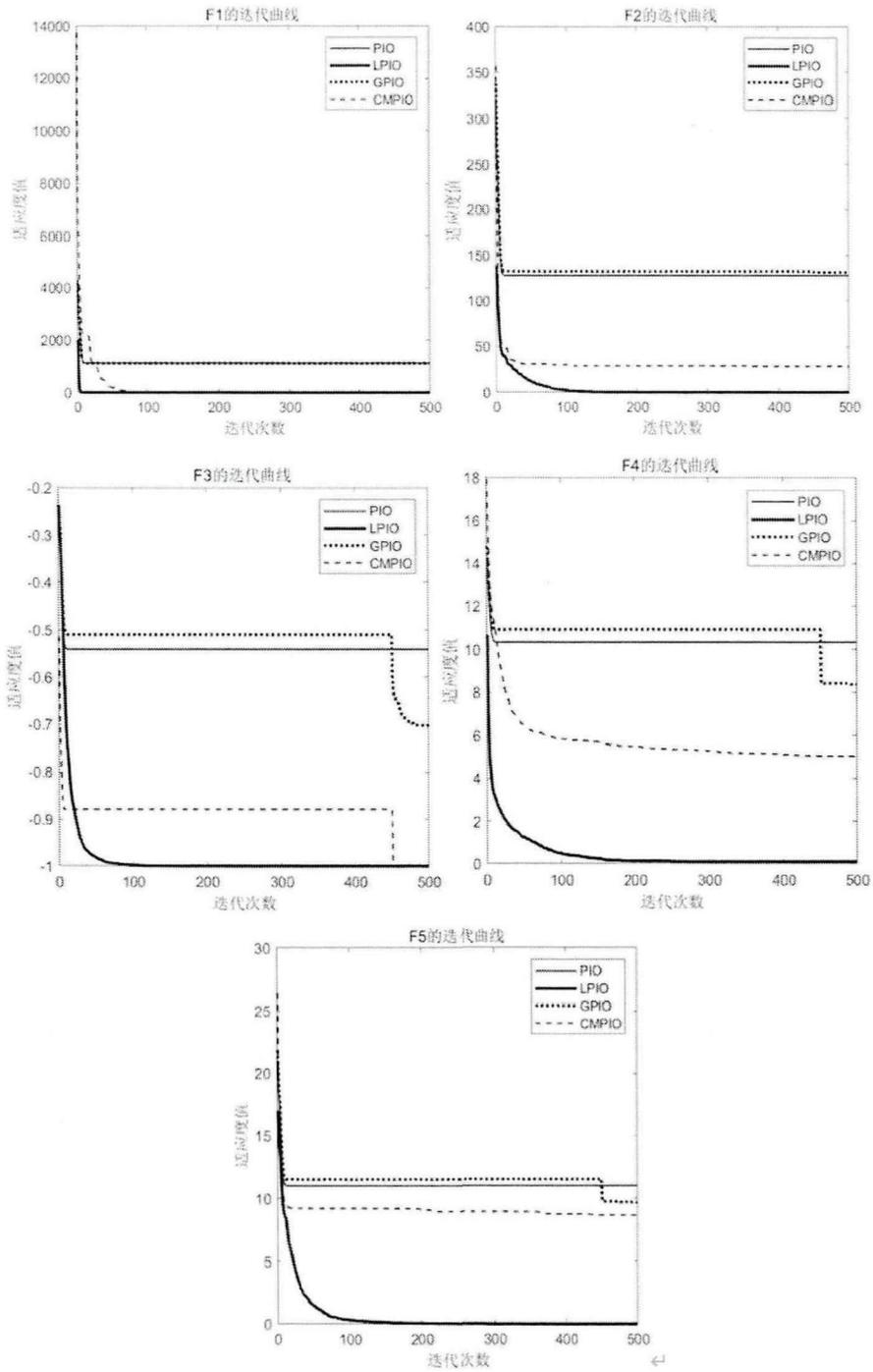


图 3-2 PIO、LPIO 算法迭代曲线

出, 与 GPIO 和 CMPIO 相比, 本文所提出的改进算法 LPIO 算法在寻优速度和稳定性方面都表现良好。同时, 可以看出, 改进的鸽群算法在寻优过程中具有更强的全局搜索性能。

综合以上结果, 可以得到以下结论: 第一, LPIO 算法具有更优的搜索精度, 能够跳出局部最优, 搜索到真实最优解, 可以得到满意的结果, 具有很好的全局搜索性能。第二, LPIO 的算法继承了原有算法的收敛速率, 在早期迭代过程中, 能够快速的逼近最优解。第三, LPIO 算法在可靠性和稳定性上表现的更好, 即使在求解复杂目标函数时, 也能够保证较高的正确率。由此可以看出, 本文所提出的改进策略是有效的, 能够改善原有算法的不足。

#### 3.4 本章小结

本章针对 PIO 算法在运行过程中的不足, 引入了两种改进机制, 并详细的阐述了两种机制在 PIO 算法中是如何起作用的, 即构造出了 LPIO 算法。为了测试 LPIO 算法的寻优性能, 选取了 5 个典型的测试函数进行仿真实验, 并与基本鸽群算法、基于高斯变异的鸽群算法以及基于柯西变异的鸽群算法在寻优结果方面进行了对比分析, 实验结果表明, 与其它改进算法相比, LPIO 算法更具有全局性, 求解精度更高且稳定性更强。

## 4 运用改进鸽群算法优化后的 LSTM 进行股指预测

股票价格数据是一种特殊的时间序列数据，具有高噪声、复杂性以及非线性等特征。这些特征使得现有的统计模型不能很好的对股票价格的走势进行很好的预测，然而 LSTM 神经网络在处理这类时间序列时能够表现出良好的性能，因此，本章将利用 LSTM 的优势对股票价格进行预测。同时，本章在发挥 LSTM 优势的基础上会对其进行进一步的优化及改进，将寻优性能和收敛速度良好的 LPIO 算法与 LSTM 算法进行融合，从而提高股票价格的预测精度，得到更好的预测结果。本章主要包括四个部分，分别是实验数据的选取和预处理、模型的评价标准、LPIO-LSTM 预测模型的构建过程以及真实数据在模型中的实验结果及分析。

### 4.1 数据的选取及预处理

#### 4.1.1 数据的选取

本文选择上证指数进行预测，在本次实验中，从网易财经官方网站选取了 2016 年 11 月 1 日到 2021 年 10 月 29 日共 1216 条数据进行分析，如图 4-1 所示，为所选数据的分布情况。在神经网络模型中，本文采用前 1 天的开盘价、最高价、最低价、收盘价、涨跌额、成交量来预测当天的收盘价格，因此每相邻 2 天的数据可构造一个样本。以此类推，一共可以组建出 1215 条样本数据。在选用神经网络对数据进行预测时，要将得到的样本数据划分为训练集数据和测试集数据。本文中的实验统一将 90% 的数据设定为训练集，将 10% 的数据设定为测试集，即训练集数据 1093 条，测试集数据 122 条。

#### 4.1.2 数据的预处理

在构造模型前，常常需要先对数据进行预处理，数据处理是否得当对训练和测试结果的影响非常大。数据量级的差异可能会使得数量级大的数据占据主导地位，从而与实际情况相悖。同时，量级差异会使得迭代速度减慢。尤其是

在使用梯度下降法寻求最优解时，很可能走“之”字型路线，从而导致迭代多次才能收敛。因此，在进行实验前对数据进行无量纲化处理至关重要。常用的无量纲化处理方法有标准化和归一化，可以根据数据的特点选用合适的处理方法。本次操作采用数据预处理方法中的标准化方法将数据映射到[-1, 1]之间，其公式如下：

$$\tilde{x} = \frac{x-\mu}{s} \quad (4-3)$$

式中， $x$ 分别为变量的原始值， $\mu$ 为变量的均值， $s$ 为变量的标准差。采用该方法对数据进行归一化，可以有效的提高模型预测的精度，得到更好的预测模型。

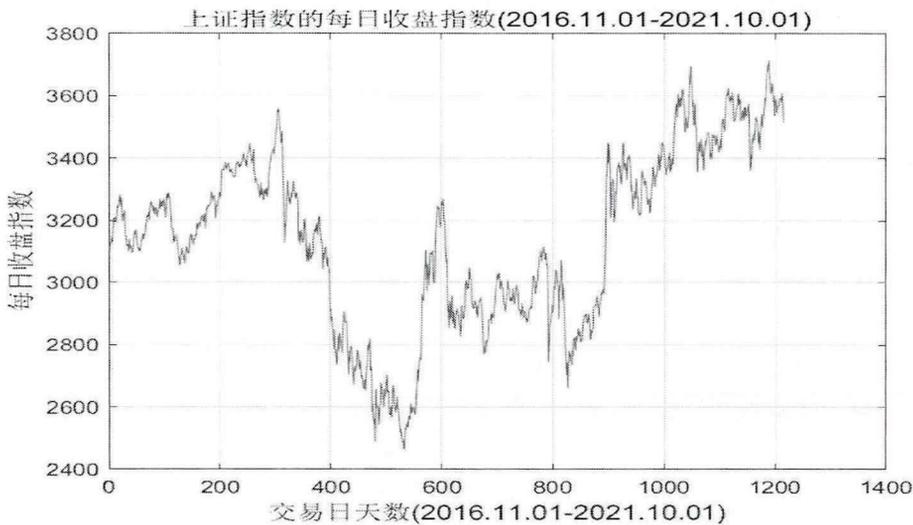


图 4-1 上证指数每日收盘价

## 4.2 模型的评价标准

在预测过程中，得到的预测值总会与实际情况产生偏差，这种偏差在统计中被称为误差。在实际应用中，通常会选择一些统计指标对模型的预测效果和准确性进行评估，常用的统计指标有均方根误差(RMSE)、平均绝对误差(MAE)和平均绝对百分比误差(MAPE)。

表 4-1 预测模型评价指标

指标名称	表达式
均方根误差	$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}}$
平均绝对误差	$\text{MAE} = \frac{\sum_{i=1}^N  y_i - \hat{y}_i }{N}$
平均绝对百分比误差	$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left  \frac{y_i - \hat{y}_i}{y_i} \right  \times 100\%$

表 4-1 为这些常用统计指标的具体展示形式。其中,  $y_i$  为真实值,  $\hat{y}_i$  为预测值,  $N$  为预测值的个数。本文将采用表中的 3 个统计指标来评估优化模型的预测效果。通过比较以上指标大小的, 可以直观的看出每个模型的预测性能, 从而进行优劣判断, 选择出最优的模型。

### 4.3 LPIO-LSTM 模型的建立过程

LSTM 作为神经网络的一种, 与其它神经网络一样, 会受到自身网络结构参数的影响。网络结构参数选择的不同, 将直接影响到神经网络模型的性能, 也就是说, 找到最优的参数可以进一步提高模型的预测效果, 使得投资者能够更好的了解股票市场的变动, 从而进行合理的投资。

在 LSTM 神经网络中, 关键参数的选取在模型构造的过程中及其重要。然而, 在实际的操作过程中, 常常通过个人经验或大量的实验操作来确定参数, 这无疑会带来时间资源和人力资源的浪费。因此, 本节考虑在 LSTM 神经网络中引入优化算法 LPIO, 去寻求最优的参数, 从而得到比较理想的网络模型。经过多次实验, 可以发现隐藏层神经元的数目及学习率的大小设置的不同会引起神经网络预测结果的变化, 所以这里将各隐藏层神经元的数目及学习率作为待优化的参数, 进行优化分析。

利用 LPIO 算法优化 LSTM 神经网络的基本思想是: 将 LSTM 神经网络训练过程中的均方误差作为 LPIO 算法的适应度值, 将学习率及隐藏层神经元个数作为鸽群的位置信息并通过获得的均方误差大小不断的对位置进行迭代和更新, 直到达到预定的误差精度或最大迭代次数为止。迭代结束后, 以最优个体的位置信息作为训练 LSTM 神经网络的初始神经元个数和学习率进行训练, 达到改

善原有神经网络模型的目的。根据 LSTM 神经网络的改进思想,可以知道,通过 LPIO 算法优化 LSTM 神经网络主要包含 4 个部分:第一,确定 LSTM 神经网络待优化的结构参数;第二,确定 LPIO 算法的适应度函数;第三,利用 LPIO 算法的寻优能力找到较优的参数组合;第四,利用得到组合参数训练神经网络。

假设 LSTM 神经网络含有  $n$  个隐含层,则每个隐含层的神经元个数记为  $H_1$ 、 $H_2$ 、 $\dots$ 、 $H_n$ , LSTM 神经网络的学习率记为  $L_r$ 。依据此定义,可以明确神经网络中待优化的参数总数,也就确定了 LPIO 算法求解过程中的维数信息。

利用 LPIO 算法优化 LSTM 神经网络的详细步骤如下:

步骤 1: 确定神经网络的隐含层数  $n$ 。

步骤 2: 根据隐含层数  $n$  确定 LPIO 算法的求解维度  $D$ , 即令  $D=n+1$ 。

步骤 3: 设置相应的参数,并初始化种群。

步骤 4: 将鸽群位置信息作为 LSTM 神经网络的神经元个数和学习率大小,代入到神经网络中,进行训练,并选取均方误差作为 LPIO 算法的适应度值。

步骤 5: 根据 LPIO 算法的寻优规则不断的调整鸽群的位置。

步骤 6: 迭代停止后,得到最优个体的位置信息,并将其转化为 LSTM 神经网络的对应的参数进行网络模型的构建,自此,LPIO-LSTM 网络模型构造完成。

### 4.4 实证分析及结果对比

本小节将会将上证股票指数数据代入到上节所构造的预测模型中,并将结果与其它模型进行对比,以此来展示并验证本文所构造模型的优势。

#### 4.4.1 实证分析过程及结果

为了展示本文所提出的 LPIO-LSTM 模型的预测效果,利用 Matlab 2020a 编写了 LSTM、LPIO\_LSTM 的相关程序。本小节共进行了 4 个模型的构建及分析。

##### (1) ARIMA 实证过程及结果分析

在图 4-1 中,可以看到上证指数收盘价的分布呈现出一定的非线性和非平稳性。经过平稳性检验,发现上证指数收盘价确实是非平稳的。这时需要进行差分来确定模型的阶数。首先进行一阶差分,并进行平稳性检验,结果显示一

阶差分下的数据是平稳的，所以先暂定 ARIMA(p,d,q)模型中的 d 为 1。在确定了差分阶数后，根据 BIC 准则，得到 p、q 的值分别为 1 和 0。

确定的阶数是否合适，要进一步经过残差检验确定。对残差进行检验有两个方法可以实现，一是判断残差是否是随机正态分布的、不自相关的；二是通过 Durbin-Watson 检验对相关性进行检验。从残差检验图可以看出残差是接近于正态分布的，ACF 和 PACF 检验了残差的自相关和偏自相关。同时，得到的 Durbin-Watson 值为 2.0008，接近于 2，说明残差不存在一阶相关性。经过上述检验可以证明，残差接近于正态分布，且相互之间是独立的，符合 ARIMA 模型的建模要求，因此得到了 ARIMA (1, 1, 0) 模型用于收盘价预测。

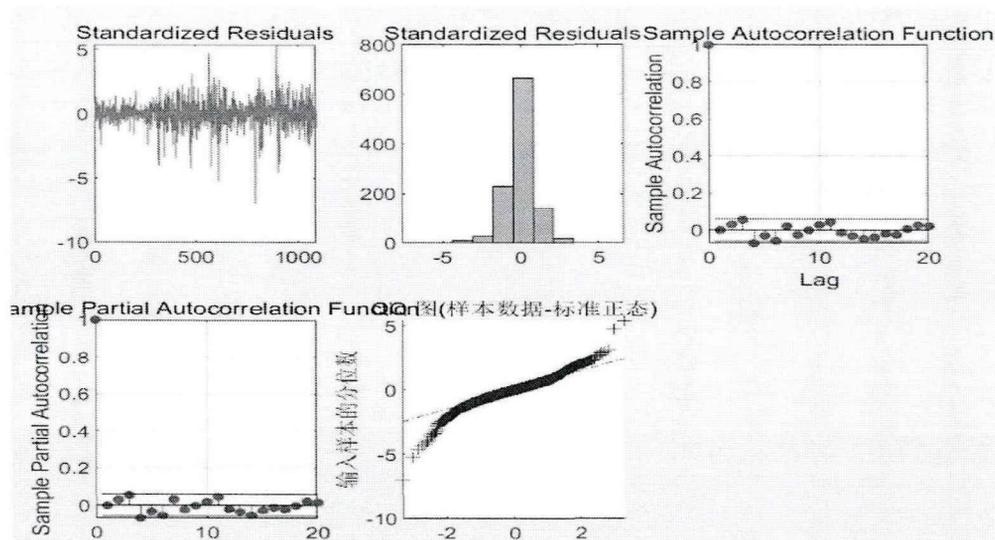


图 4-2 残差检验图

从 ARIMA(1,1,0)模型的预测结果中，可以看出 ARIMA 模型能够捕捉到股价的走势，得到较好的预测结果。将观测值与预测值进行对比，得到均方根误差为 29.7361，平均绝对误差为 23.1352，平均绝对百分比误差为 0.66%。

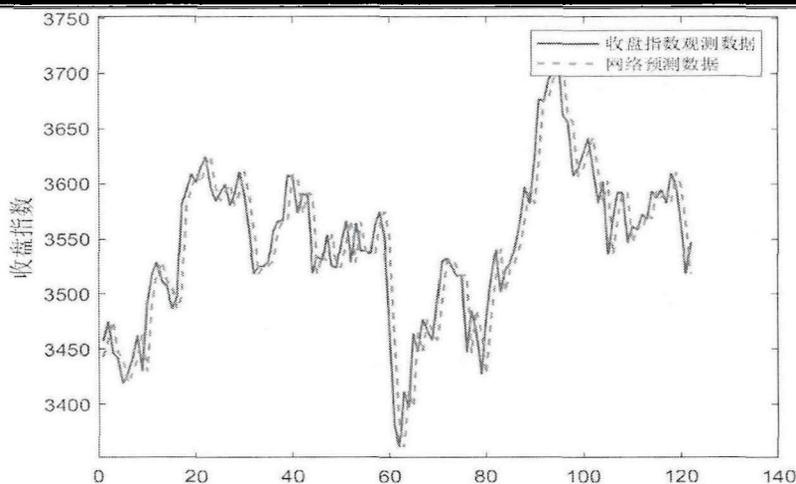


图 4-3 ARIMA 股价预测效果图

## (2) BP 网络实证过程及结果分析

在使用 BP 神经网络建立模型前, 要对模型进行相应的设置。例如, 输入层、隐含层、输出层节点个数、隐含层层数以及其它相关参数的确定。

第一、神经网络的输入层和输出层节点数设定。本文在构造 BP 预测模型时, 同样选取前 1 天上证指数开盘价、最高价、最低价、收盘价、涨跌额、成交量作为特征变量来预测当天的上证指数收盘价, 所以确定输入层有 7 个神经元节点, 输出层有 1 个神经元节点。

第二、隐层层数设定。关于隐含层层数的设定, 至今没有一个确定的方案。一般来说, 隐含层层数越多, 预测误差越小, 但这也提升了网络模型的复杂度, 增加了训练时间, 甚至可能会出现过拟合现象, 进而导致测试集预测效果不好。通过文献阅读发现, 大部分学者在运用 BP 神经网络时优先考虑一个隐含层, 若效果不够理想, 再考虑适当的增加隐含层层数。本文设定隐含层层数为 1 时, 效果不太理想。因此, 增加隐含层层数, 设定隐含层层数为 2。

第三、隐含层节点数的设定。对于隐含层节点数的设定, 同样没有具体的解决方案。通常情况下, 若隐含层中神经元个数太少, 则不能够提取出足够多的特征信息, 进而使得网络模型的性能较差, 达不到我们期望的效果。相反, 若个数太多, 则容易出现过拟合现象, 也不能得到较好的预测结果。因此, 应该合理的选择隐含层节点数。本文通过多次实验, 设定两隐含层节点数分别为 50, 10。

第四、其他参数设定。本文设定 BP 网络的迭代次数为 200 次, 学习率为

0.01, 误差精度为 0.00001, 训练方式选择梯度下降法。在 BP 神经网络中, 通常将 sigmoid 函数和线性函数作为激励函数, 但由于本文网络输出时将数据归一到  $[-1, 1]$  内, 所以选取 S 型对数函数 tansig 作为激励函数。

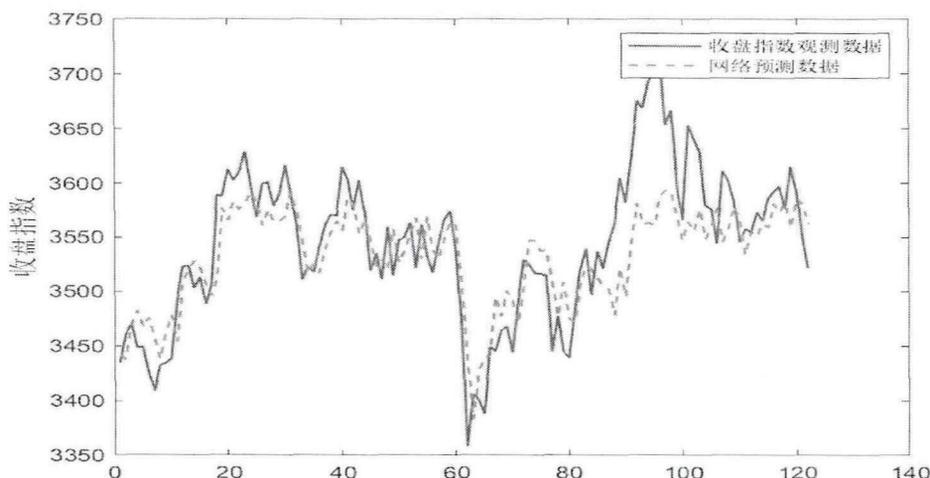


图 4-4 BP 股价预测效果图

通过 BP 神经网络对上证指数收盘价进行预测后的结果如图 4-4 所示, 由此可知, 通过 BP 神经网络并不能很好的预测股票价格, 预测值与真实值相差较大, 根据预测结果看出该网络无法对股价进行很好预测。此次实验过程中, 得到均方根误差为 43.5744, 平均绝对误差为 33.4372, 平均绝对百分比误差为 0.93%。

### (3) LSTM 网络实证过程及结果分析

本模型以上证指数为研究对象, 并将收盘价作为该过程唯一输出变量进行预测。在本次实验中, 选用上证指数前一天的开盘价、最高价、最低价、收盘价、涨跌额、成交量来预测当天的收盘价格, 因此本次 LSTM 模型的输入节点个数为 7, 输出节点个数为 1。该模型选择的优化器是 adam 算法, 损失函数设定为均方误差 MSE, 隐含层层数设定为 2, 最大迭代次数设置为 50。之所以将隐含层层数设定为 2 是因为, 通过多次实验验证发现在 1 层隐含层下的预测效果并不理想, 不能达到预期效果。当把隐含层层数设定为 2 时, 预测效果得到显著提升, 因此本文将隐含层层数设定为 2。

经过多次实验比较和调整, 发现两隐含层神经元个数分别设定为 60 和 20, 学习率设定为 0.01 时, 可以达到比较好的效果, 因此本次设定两隐含层神经元的个数为 60 和 20, 学习率为 0.01。

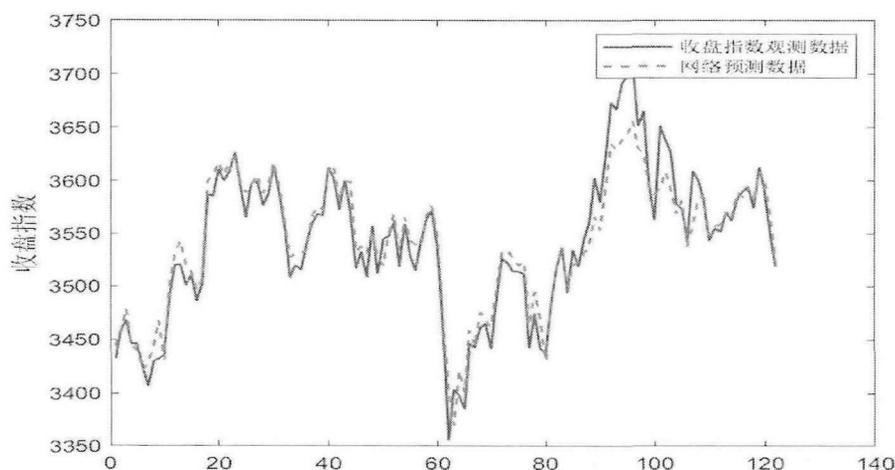


图 4-5 LSTM 股价预测效果图

针对 LSTM 网络模型，按照设定的参数进行训练之后，得到了如图 4-5 所示的结果。图 4-5 直观的描述出了观察值和预测值的变动情况，可以清晰的看到 LSTM 模型的预测效果。此次实验过程中，得到均方根误差为 19.1027，平均绝对误差为 13.9831，平均绝对百分比误差为 0.39%。

由此可以看出，LSTM 模型经过预测可以很好的展现出股票的走势，能够达到良好的预测效果。

#### (4) LPIO-LSTM 实证过程及结果分析

在本次实验中，为了体现出该组合模型的性能优势，除了两隐含层神经元数及学习率这 3 个待优化的参数外，LSTM 神经网络的其余参数与上面保持一致。

对于改进的鸽群算法部分，设置种群规模  $N$  为 10，地磁导航最大迭代次数  $T1$  为 10，地标导航最大迭代次数  $T2$  为 5，地磁因子  $R$  为 0.3，Levy 飞行调节参数  $\alpha$  为 5，地磁导航过程混沌搜索的调节参数  $b1$  为 1，地标导航过程混沌搜索的调节参数  $b2$  为 0.1。同时，将训练集的均方根误差 (RMSE) 最小作为最终寻优目标。

通过 LPIO 算法，得到了相应的隐含层层数分别为 96, 96，学习率为 0.0013。将这 3 个参数代入到 LSTM 中，得到了图 4-6 所示的预测图。

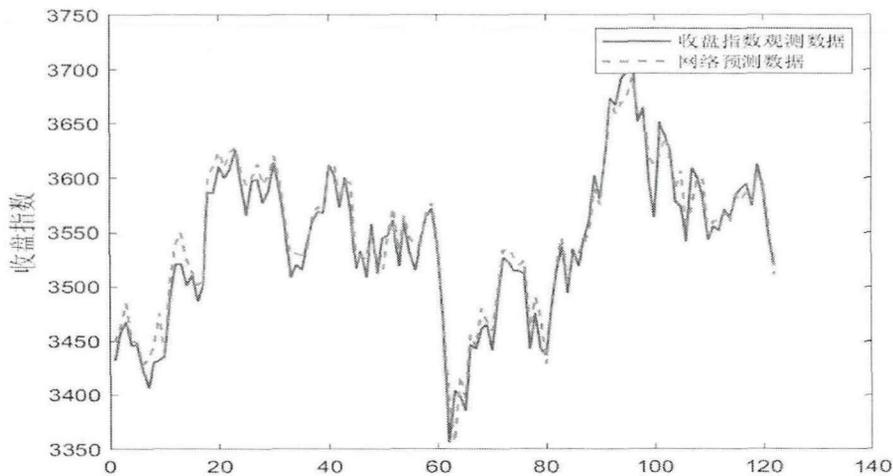


图 4-6 LPIO\_LSTM 股价预测效果图

从组合模型的预测结果来看，LPIO\_LSTM 模型所预测的结果与真实结果比较接近，可以达到很好的预测效果。根据模型评价指标，将测试集中的真实值与预测值进行了比较，得到均方根误差为 15.7928，平均绝对误差为 12.2453，平均绝对百分比误差为 0.34%。

#### 4.4.2 结果对比

根据上面的实证分析，得到了各个模型下的预测结果。通过预测效果图，可以直观的判断出每个模型的预测情况。表 4-2 展示了在利用各模型进行预测的过程中，真实值与预测值相比所得到的均方根误差，平均绝对误差，平均绝对百分比误差。

经过对比可以看出 ARIMA、LSTM、LPIO\_LSTM 三个模型都能够描摹出股价的走势，其中，LSTM 神经网络的预测效果要优于 ARIMA 模型，预测的误差更小，更精确。BP 神经网络不能够很好的处理时序数据，在所有模型中表现最差，预测误差大。相比 LSTM 神经网络，LPIO\_LSTM 神经网络的预测误差更小，更接近于真实值，这就进一步说明本文所提出的模型是有效的，可以利用该模型对上证指数的收盘价进行预测。同时，再次证明了 LPIO 算法的有效性，表明改进的鸽群算法在优化 LSTM 神经网络的过程中发挥了作用，对其参数进行了优化。

#### 4 运用改进鸽群算法优化后的 LSTM 进行股指预测

表 4-2

预测结果汇总

	RMSE	MAE	MAPE
ARIMA	29.74	23.14	0.66%
BP	43.57	33.44	0.93%
LSTM	19.10	13.98	0.39%
LPIO_LSTM	15.7928	12.25	0.34%

#### 4.5 本章小结

本章以上证指数数据为例,主要利用 ARIMA 模型、BP 神经网络模型、LSTM 模型以及 LPIO\_LSTM 模型对上证指数的收盘价进行了预测,结果发现,LSTM 网络模型、LPIO\_LSTM 网络模型的预测误差相比于其它模型来说较小,表明 LSTM 模型在预测上证指数时更具有优势。同时,LSTM 模型与 LPIO\_LSTM 模型对比发现,LPIO\_LSTM 模型的预测误差要小于 LSTM 模型的预测误差,表明 LPIO\_LSTM 模型的预测结果表现更好,证明本文所提出的 LPIO\_LSTM 模型是可行的,进一步说明 LPIO 算法在对 LSTM 模型的网络参数进行优化时,起到了该有的优化作用。

## 5 结论与展望

### 5.1 结论

本文以 LSTM 神经网络为研究基础,对上证指数收盘价进行预测,探讨了 LSTM 神经网络在股价预测中的可行性。同时,针对 LSTM 神经网络预测结果容易受到网络结构参数影响这一问题,本文选用了鸽群算法对其网络结构参数进行优化,但鸽群算法在复杂问题中容易陷入局部最优,这又为本文带来了新的研究点。在对 LSTM 神经网络进行优化前,本文首先对鸽群算法的改进进行了探讨,然后将改进的鸽群算法应用于 LSTM 神经网络中,最后将构造的组合模型应用于上证指数收盘价的预测中。本文的研究主要侧重两个方面:一是鸽群算法(PIO)的改进,二是探讨改进鸽群算法(LPIO)在 LSTM 神经网络中的应用。

PIO 算法虽然提出时间短,但在实际应用中,具有收敛速度快、操作简单、易于实现等特点。因此,该算法自提出以来受到了大批学者的关注,也被广泛应用到众多领域,逐渐发展成了较热门的群智能优化算法。然而,该算法在解决复杂优化问题时,仍然存在着不足,在迭代后期容易陷入局部最优,导致算法的寻优精度有所下降。针对基本 PIO 算法的不足,本文对鸽群算法进行了改进,在改进过程中主要是考虑将鸽群算法与其他机制相结合,最终形成了 LPIO 算法。改进的思路主要包括下面两点:

- (1) 在地磁导航阶段,引入 levy 飞行,利用 levy 飞行的跳跃性,帮助种群中的个体跳出局部最优,并向真实最优解靠近。
- (2) 利用混沌理论的遍历性特点,对每次迭代后的最优个体进行混沌搜索,此项改进主要是对个体进行更细化的搜索,可以有效提高算法的求解精度,同时,也在一定程度提高了收敛性能。

针对 LPIO 算法,本文选取了 5 个测试函数进行测试,测试结果表明 LPIO 算法在复杂的优化问题中依然能够找到全局最优解,具有良好的全局性能,同时,也保留了原有 PIO 算法良好的收敛性能。

其次,本文运用 LPIO 算法对 LSTM 神经网络的结构参数进行优化,并运用

实际数据集对组合模型的性能加以验证，得到以下结论：

(1) 在对上证指数收盘价进行预测时，与 ARIMA、BP 神经网络预测结果相比，LSTM，LPIO\_LSTM 的预测效果更好，预测误差更小。证明与其他模型相比，长短时记忆神经网络更适合应用于上证指数收盘价的预测中。

(2) 加入优化算法的 LSTM 神经网络要优于传统 LSTM 神经网络模型。基于 LPIO 算法优化的 LSTM 神经网络模型预测效果最好，表明利用 LPIO 算法优化 LSTM 神经网络的网络结构可以很好的改善传统 LSTM 神经网络的预测性能，具有一定的实用价值。

### 5.2 不足与展望

尽管本文所提出的组合模型在对上证指数进行预测时有着较好的预测性能，但由于时间及自身能力的欠缺，模型仍有改进的空间。展望未来，文章可以在以下几个方面进行拓展研究：

(1) 本文在考虑鸽群算法的改进时，仅仅考虑了鸽群算法自身的改进，也可以考虑与其他具有全局性的优化算法相结合。例如，免疫算法、人工鱼群算法等。如何将这些优化算法与鸽群算法结合起来，构造出具有全局性的组合优化算法，将是之后要继续探索和研究的方

(2) 在利用 LPIO 算法对 LSTM 神经网络进行优化时，本文仅考虑了对 LSTM 神经网络的隐含层神经元个数和学习率进行优化，其它相关的参数在后续的研究中也可加入到所要优化的行列中。

(3) 由于时间因素，在对上证股指进行预测时，本文仅仅选取了开盘价、收盘价、最高价、最低价、涨跌额和成交量等历史交易数据作为样本数据。在未来的研究中，可以考虑将技术指标、情感因素等信息也加入到样本数据中。

(4) 本文仅将提出的组合模型应用到了上证指数的预测中，未来可以考虑将该模型应用到其它领域。

(5) 本文采用归一化对样本数据进行预处理，未来可在归一化的基础上，尝试采用小波分析和重构算法将原始时间序列数据分解成低频子序列和高频子序列，再将每一子序列代入到组合模型 LPIO\_LSTM 中进行训练，相加得到最终预测结果。

## 参考文献

- [1]蔡红,陈荣耀.基于 PCA-BP 神经网络的股票价格预测研究[J].计算机仿真,2011,28(03):365-368.
- [2]陈佳,刘冬雪,武大硕.基于特征选取与 LSTM 模型的股指预测方法研究[J].计算机工程与应用,2019,55(06):108-112.
- [3]邓凤欣,王洪良.LSTM 神经网络在股票价格趋势预测中的应用——基于美港股票市场个股数据的研究[J].金融经济,2018(14):96-98.
- [4]段海滨,叶飞.鸽群优化算法研究进展[J].北京工业大学学报,2017,43(01):1-7.
- [5]段海滨,杨之元.基于柯西变异鸽群优化的大型民用飞机滚动时域控制[J].中国科学:技术科学,2018,48(03):277-288.
- [6]付燕,栗锋.ARMA 模型在我国体育股票价格预测中的应用[J].统计与决策,2012(21):101-103.
- [7]郝知远.基于改进的支持向量机的股票预测方法[J].江苏科技大学学报(自然科学版),2017,31(03):339-343.
- [8]黄丽明,陈维政,闫宏飞,等.基于循环神经网络和深度学习的股票预测方法[J].广西师范大学学报(自然科学版),2019,37(01):13-22.
- [9]胡春鹤,王依帆,朱书豪,等.基于鸽群优化算法的图像分割方法研究[J].郑州大学学报(工学版),2019,40(04):42-47.
- [10]韩锬,张赫.基于鸽群优化改进的粒子滤波算法[J].传感器与微系统,2018,37(11):139-141+144.
- [11]李民,邹捷中,李俊平,等.用 ARMA 模型预测深沪股市[J].长沙铁道学院学报,2000(01):78-84.
- [12]李坤,谭梦羽.基于小波支持向量机回归的股票预测[J].统计与决策,2014(06):32-36.
- [13]梁宇佳,宋东峰.基于 LSTM 和情感分析的股票预测[J].科技与创新,2021(21):126-127.
- [14]庞英智,赵晓丹,郭伟业.一种基于 RBF 人工神经网络的预测方法研究[J].情报科学,2009,27(11):1707-1709.

- [15]乔若羽.基于神经网络的股票预测模型[J].运筹与管理,2019,28(10):132-140.
- [16]孙延风,梁艳春,张文力,等.RBF神经网络最优分割算法及其在股市预测中的应用[J].模式识别与人工智能,2005,18(03):374-379.
- [17]孙伯原.基于神经网络模型的股票预测与研究[J].电子测试,2019(12):69-70.
- [18]沈传茂.回声状态网络补偿 ARMA 预测误差的体育股票价格预测[J].统计与决策,2014(23).2014.23.022.
- [19]万建强,文洲.ARIMA 模型与 ARCH 模型在香港股指预测方面的应用比较[J].数理统计与管理,2001(06):1-4.
- [20]王立凤.ARCH 模型在预测股价变动中的应用[J].统计与决策,2006(07):142.1
- [21]魏文轩.改进型 RBF 神经网络在股票市场预测中的应用[J].统计与决策,2013(15):70-72.
- [22]徐浩然,许波,徐可文.机器学习在股票预测中的应用综述[J].计算机工程与应用,2020,56(12):19-24.
- [23]于志军,杨善林.基于误差校正的 GARCH 股票价格预测模型[J].中国管理科学,2013,21(S1):341-345.
- [24]禹建丽,孙增圻,刘治军,等.基于 BP 神经网络的股市建模与决策[J].系统工程理论与实践,2003(05):15-19+58.
- [25]杨小平.基于主成分与 BP 神经网络的股票价格预测分析[J].统计与决策,2004(12):42-43.
- [26]张玉川,张作泉.支持向量机在股票价格预测中的应用[J].北京交通大学学报,2007(06):73-76.
- [27]朱震宇,王艳,纪志成.基于高斯鸽群优化算法的典型工艺知识发现方法[J].信息与控制,2019,48(01).
- [28]Akshit Kurani,Pavan Doshi,Aarya Vakharia,et al. A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting[J]. Annals of Data Science,2021.
- [29]Budiharto Widodo. Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM)[J]. Journal of Big Data,2021,8(1).
- [30]Chen S.J., Duan H.B. Fast image matching via multi-scale Gaussian mutation

pigeon-inspired optimization for low cost quadrotor[J]. Aircraft Engineering and Aerospace Technology,2017,89(6).

[31]DUAN H.B., QIAO P.X. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning [ J ] . International Journal of Intelligent Computing and Cybernetics, 2014, 7(1): 24-37.

[32]Emioma C. C.,Edeki S. O.. Stock price prediction using machine learning on least-squares linear regression basis[J]. Journal of Physics: Conference Series,2021,1734(1).

[33]Helmut Herwartz. Stock return prediction under GARCH — An empirical assessment[J]. International Journal of Forecasting,2017,33(3).

[34]Laily V.O.N., Warsito B., Maruddani D.A.I. Comparison of ARCH / GARCH model and Elman Recurrent Neural Network on data return of closing price stock[J]. Journal of Physics: Conference Series,2018,1025(1).

[35]Huseyin Ince,Theodore B. Trafalis. Short term forecasting with support vector machines and application to stock price prediction[J]. International Journal of General Systems,2008,37(6).

[36] Lipton Z C,Berkowitz J,Elkan C.A critical review of recurrent neural networks for sequence learning[J].arXiv:1506.00019,2015.

[37] Hochreiter S,Schmidhuber J.Long short-term memory[J]. Neural Computation, 1997,9(8):1735-1780.

[38]M Gnanendra,Jesna Jana. Stock Prediction using Artificial Neural Network; Evidence with Nifty 50 IT Companies[J]. Asian Journal of Research in Business Economics and Management,2018,8(4).

[39]Marwa Sharaf,Ezz El-Din Hemdan,Ayman El-Sayed,etal. El-Bahnasawy. StockPred: a framework for stock Price prediction[J]. Multimedia Tools and Applications,2021.

[40]Ritika Singh,Shashi Srivastava. Stock prediction using deep learning[J]. Multimedia Tools and Applications,2017,76(18).

[41]Rezaei Hadi,Faaljou Hamidreza,Mansourfar Gholamreza. Stock price prediction using deep learning and frequency decomposition[J]. Expert Systems with

Applications,2020.

[42]Ranjan Kumar Dash,Tu N. Nguyen,Korhan Cengiz,Aditi Sharma. Fine-tuned support vector regression model for stock predictions[J]. Neural Computing and Applications,2021.

[43]Rumelhart D E, Hinton G E, Williams R J. Learning Representations by Back-propagating Errors[ J]. Nature,1986, 323(9):533 - 536.

[44]Sattayatham P., Sopipan N., Premanode B.Forecasting the Stock Exchange of Thailand uses Day of the Week Effect and Markov Regime Switching GARCH[J]. American Journal of Economics and Business Administration,2012,4(1).

[45]Sushma Devi Patel,Daisy Quadros,Vidyullata Patil,et al. Stock Prediction using Neural Networks[J]. International Journal of Engineering and Management Research (IJEMR),2017,7(2).

[46]Yongbin Sun,Haibin Duan. Pigeon-inspired optimization and lateral inhibition for image matching of autonomous aerial refueling[J]. Proceedings of the Institution of Mechanical Engineers, 2018,232(8).

[47] Yang Z , Duan H , Fan Y , et al. Automatic Carrier Landing System multilayer parameter design based on Cauchy Mutation Pigeon-Inspired Optimization[J]. Aerospace Science and Technology, 2018, 79.

[48] Zhang D H, Lou S.The application research of neural network and BP algorithm in stock price pattern classification and prediction [ J ] .Future Generation Computer Systems, 2021,115: 872-879.

## 后 记

转眼之间，我的研究生生涯就要结束了。回忆起在东财的点点滴滴，依然历历在目。在毕业论文完成之际，我由衷的向过去给予我关心和帮助的所有老师、同学及亲人朋友们表达我最诚挚的谢意。

在此我要感谢我的导师刘颖老师，本文是在刘老师的悉心指导下完成的。从论文的选题、开题报告到初稿的完成，刘老师给了我极大的鼓励和帮助。在修改过程中，刘老师也是对论文的每一字每一句进行把关，提出了很多宝贵的意见，使我在修改过程中不再迷茫。本论文能够顺利完成，离不开刘老师的关心和帮助。借此机会，我要向刘老师致以深深的谢意。

感谢统计学院的所有老师。在你们的悉心教导下，我掌握了扎实的专业知识，正是这些知识的沉淀，使我有能力完成这篇论文的撰写。感恩之余，恳请各位老师对我的论文加以批评和指正，使我及时对论文中的不足之处进行完善。同时，我要感谢参与我论文评审和答辩的各位老师，你们给了我一个审视自己学习成果的机会，你们的批评和建议将是我进一步学习的动力。

曹彦彦

2022年2月17号

## 东北财经大学研究生学位论文原创性声明

本人郑重声明：此处所提交的硕士学位论文《经鸽群算法优化的 LSTM 模型在股指预测中的应用研究》，是本人在导师指导下，在东北财经大学攻读硕士学位期间独立进行研究所取得的成果。论文中除已注明的部分外，不包含他人已发表或撰写过的研究成果，对本文的研究工作做出重要贡献的个人和集体均已注明。本声明的法律结果将完全由本人承担。

作者签名：曹彦彦

日期：2022 年 5 月 1 日

## 东北财经大学研究生学位论文使用授权书

本论文作者完全了解学校关于保存、使用学位论文的规定，即按照学校要求提交学位论文的印刷本和电子版本；同意学校保留论文的复印件和电子版本，允许论文被查阅和借阅；本人授权东北财经大学可以将本学位论文的全部或部分内 容编入有关数据库进行检索，学校可以采用影印、缩印或其他复制手段保存或汇编论文。

作者签名：曹彦彦

日期：2022 年 5 月 1 日

导师签名：刘颖

日期：2022 年 5 月 1 日