

应用循环神经网络预测股价

220040009

方琛源

目 录

摘 要	1
Abstract	1
一、 绪论	2
(一) 研究背景与意义	2
1.历史背景	2
2.现实意义	2
(二) 国内外研究综述	2
1. 深度神经网络理论发展	3
2.深度神经网络研究现状	4
3.国内外文献综述	4
(三) 本文的内容与结构	5
(四) 本文的创新点	6
二、 研究方法的设计	7
(一) 循环神经网络介绍	7
1. 全连接神经网络简介	7
2. RNN 神经网络	7
3. LSTM 神经网络	8
(二) 常用激活函数，损失函数与优化器	9

1. 激活函数	9
2. 损失函数	10
3. 优化器	11
(三) LSTM 长短时记忆网络预测股价模型	11
1. 模型简述	11
2. 模型评估	12
(四) LSTM 模型的参数设计	12
1. 超参数的选取	12
2. 神经网络结构的选择	12
三、 LSTM 神经网络预测短期股价实证分析	13
(一) 实证流程概述	13
(二) 数据处理	13
1. 数据获取	13
2. 数据预处理	13
3. 训练集与测试集划分	14
(三) 实证过程详解	15
1. 模型的初步训练	15
2. 优化 DROPOUT 参数	16
3. 优化衰减率	17

4. 优化神经网络结构.....	17
四、 基于 LSTM 模型预测股价实证结论.....	18
（一） 实证结果展示.....	18
1. 初始模型结果展示.....	18
2. Dropout 优化展示	19
3. 衰减率优化展示.....	20
4. 网络结构（各隐藏层节点数）优化展示.....	20
5. 条件最优测试.....	21
（二） 实证结语	22
五、 总结与展望	23
参考文献.....	24

图表目录

图表 1	浅层学习示意图	3
图表 2	深度学习示意图	3
图表 3	深度学习重大事件历程	4
图表 4	RNN 的展开表示	7
图表 5	LSTM 隐藏状态网络模型结构	9
图表 6	实证分析中流程示意图	13
图表 7	获取行情数据表头展示	13
图表 8	标准化后数据表头展示	14
图表 9	纳斯达克调整后收盘价数据（标准化后）	14
图表 10	时间窗口示意图	15
图表 11	整体模型网络结构展示	16
图表 12	初次训练损失函数展示	18
图表 13	初次训练评估指标概览	18
图表 14	初次训练结果可视化	19
图表 15	寻找最佳 dropout 值	19
图表 16	寻找最佳 decay 值	20
图表 17	寻找最优网络结构	20
图表 18	最终实验损失函数	21
图表 19	最终训练结果可视化	22

摘 要

本文浅析了循环神经网络应用于股票价格预测的方法与思路。文章介绍部分介绍了神经网络的发展以及循环神经网络的结构与算法。实证部分基于 LSTM 长短时记忆网络设计了进行实证分析的模型，模型对纳斯达克指数预测结果与真实值偏差较小，对模型进行网格搜索法调优后，预测效果显著提高，验证了循环神经网络预测股价的可行性与潜力。

关 键 词： 股价预测；循环神经网络；LSTM 长短时记忆网络；纳斯达克指数

Abstract

This paper analyzes the method and thinking of the application of recurrent neural network to stock price prediction. In the introduction part, the development of neural network and the structure and algorithm of recurrent neural network are introduced. In the empirical part, a model for empirical analysis is designed based on Long-Short Term Memory network. The model has a small deviation between the predicted results of NASDAQ index and the real values. After the optimization of the model with grid search method, the prediction effect is significantly improved, which verifies that the cyclic neural network can predict the stock price Feasibility and potential.

Key words: Stock price prediction; Recurrent Neural Network; Long-Short Term Memory network; NASDAQ index

一、 绪论

（一） 研究背景与意义

1. 历史背景：随着股票市场与数学理论不断发展，许多有远见卓识的人意识到可以通过数量化研究来辅助股票的交易，赚取更高的收益，追溯历史，最早是一名美国的电子工程师通过一家美国银行建立了一个定量投资系统，在经过了十几年的发展，当今美国金融市场许多的交易都是由机器完成的，而当时在这个领域最成功的两个人就是詹姆斯西蒙斯和大卫肖^[1]。

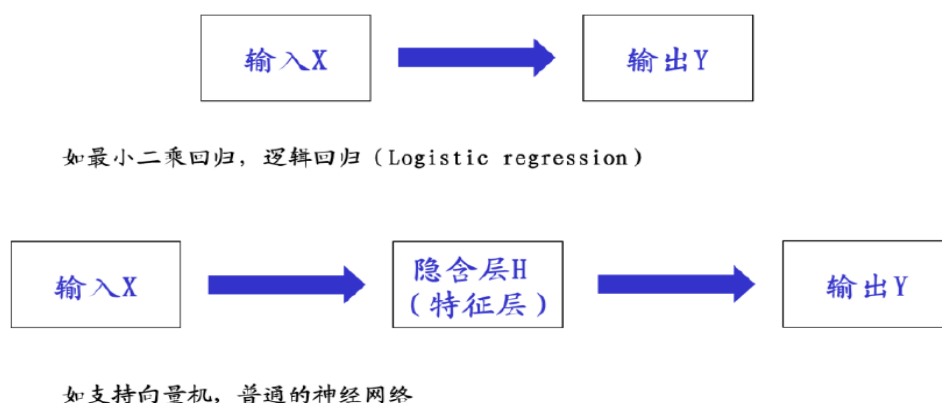
随着时间的流逝与历史的发展，这两人投身金融市场已经三十余年，随着越来越多新的统计学习方法的发展，在 80 年代红遍大江南北的人工神经网络到了 90 年代渐渐的被各种其它机器学习模型所超越，可以说是深度学习在历史上的一个低谷期，人们都将投资策略转而依赖于类似随机森林，XGboost 等机器学习方法，很大的原因是当时的计算机硬件发展水平较低，计算量无法满足神经网络所需要的大量的运算。

到了今天，计算机科学指数爆炸式的发展，运算的成本越来越低，CPU 和 GPU 的性能越来越好，与此同时，金融市场每天产生的数据量越来越大，人们这时候开始意识到数十来年前的神经网络的方法的可行性越来越高，因此深度学习这一概念迎来了新的高潮，人们都坚信，通过深度学习这一领域，金融市场中许许多多的不可知或者是不可预测的东西，都将被慢慢的揭晓。

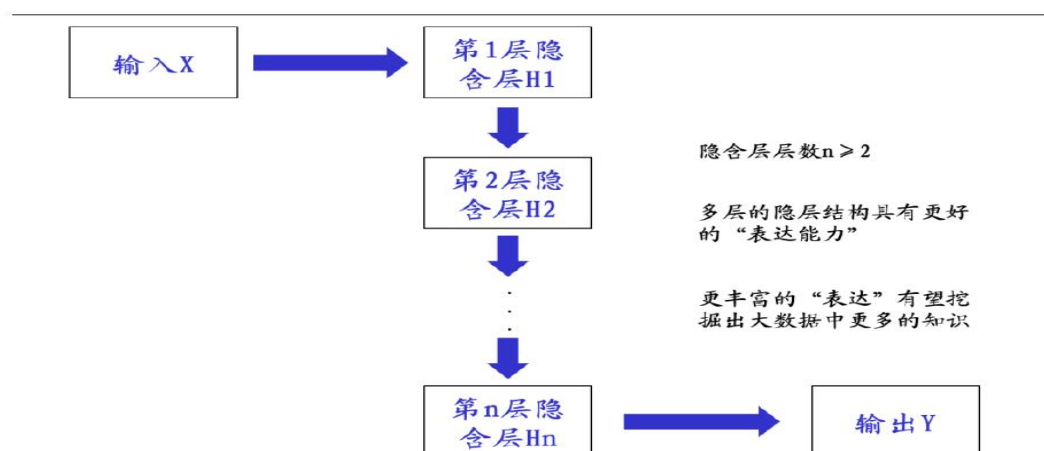
2. 现实意义：在这个大数据时代，深度学习这个概念可以说是掀起了一大股技术领域的波澜，许许多多的互联网巨头们如今都花费大量的人力财力进入到这个领域，就是为了能够引领这个潮流，实现技术的开创性更新迭代，金融领域也不例外，诸如桥水，黑水等全球顶尖对冲基金，都意识到了神经网络将对今后的交易模式产生变革型的影响，也都成立了专门的小组进行相关领域的研究。从微观角度看，市场上的所有的交易行为都是由买卖双方的博弈产生，并通过海量的交易数据展现出来，从理论上来说，只要我们能够学习到交易数据中蕴含的所有信息，我们必然会具备预测短期与长期股价以及市场波动的能力，我们可以通过深度学习模型，从大量的交易数据乃至市场经济数据中学习得到股票短期甚至长期价格变化与过去的经济数据和交易数据的关系，并将这种模型套用到当前数据中对未来的股票价格做出判断，如果模型表现得好，我们就可以从中获取到可观的超额收益，这也是进行股票交易的初衷，不仅如此，我们的模型还可以从历史的金融危机中学习得到金融危机发生之前的信号，这对个人来说，可以及时将多头头寸退出躲避风险，甚至可以加大空头头寸获取危机带来的巨大收益。对于企业来说，可以及时进行经营战略的调整，供应链布局的调整，平稳度过危机，避免在危机之中破产，08 年金融危机就有许许多多的企业因为没有预料到大危机的到来而从历史的长河中永久的消失，对于政府与监管机构，意义更为重大，维持金融市场的平稳运行以及整个国家宏观经济平稳发展是政府与监管局的首要任务，如果能有神经网络预测的数据做支持，辅助政策的制定与经济的调控，将能从很大程度上杜绝系统性风险，保证国家机器的平稳运行和人民福祉。总结而言，运用深度神经网络进行股票预测是量化投资的一个大趋势，也是人工智能应用金融场景的一个重大尝试，具有非常深远的，跨时代的意义。

（二） 国内外研究综述

1. 深度神经网络理论发展：深度神经网络最早的发现与生理医学上的发现有关，1981 年的诺贝尔医学奖得主 David Hubel 和 Torsten Wiesel 发现“方向选择性细胞”的神经元细胞。当人的眼镜注意到某些东西的边缘部分，然后这些边缘部分指向一个固定的方向，这个所谓的“方向性选择细胞”，也可以称作神经元，就变得非常的活跃起来，这个实验让人们的大脑的工作原理进行了重新的思考，人的大脑学习事物可能是一个不断迭代更新的过程，从最开始信息的获取，然后是预处理，然后一层一层的不断抽象，换言之，就是通过一层的特征抽象提取，构成一个神经网络，实现人脑特有的各种复杂功能。基于这个理论，人工神经网络面世，这是一种模仿大脑神经突触连接结构，通过结构之间的联系以及连接信息的不断更新而获得新的学习能力的数学模型，统称为神经网络，这种网络由非常多的神经元相互连接组成，每个神经元背后都定义与其它神经元之间联系的数据，成为权重数据，也定义着一种特殊的函数，统称为激活函数。每个节点之间都有对应的权重关系。网络的输出结果与网络的权重值，激活函数与连接方式都有关系。神经网络的发展大致经历了两个阶段，第一个是浅层学习阶段，第二个是深度学习阶段，浅层学习与深度学习如图表 1 与图表 2 所示，以 2006 年为发展分界线，在 2006 年以前，深度学习主要以神经网络和连接主义名义发展，经历了二起二落；2006 年，Geoffrey Hinton 将深度神经网络正式的称为深度学习，也就是从 2006 年开始，深度学习这一概念开始了自己的第三次复兴之路。

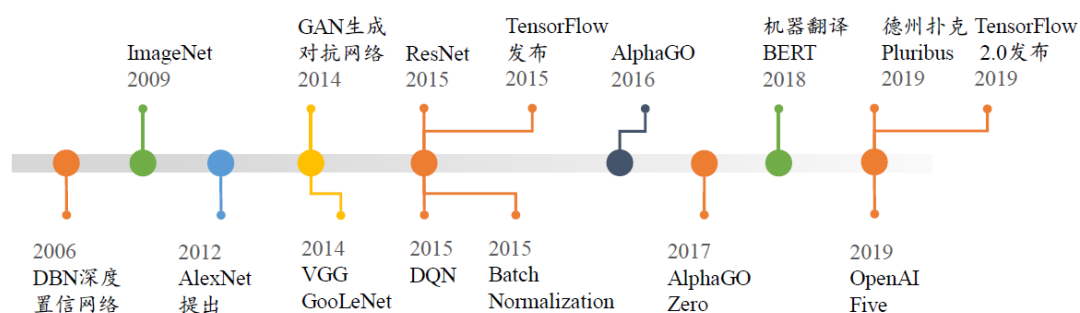


图表 1 浅层学习示意图



图表 2 深度学习示意图

2. 深度神经网络研究现状：2006 年，Geoffrey Hinton 等人发现通过一层层的预处理可以非常好的进行神经网络的训练，在公开的 MNIST 手写数字训练集上，他们的网络取得了优异的成绩，正确率远超 SVM。这使得第三种人工智能的复兴。2012 年，亚历克斯·克里热夫斯基提出了一种 8 层深层神经网络——亚历克斯网，它采用了 ReLU 激活函数，同时采用 Dropout 手段防止过拟合的发生，直接在英伟达支持深度学习的 GPU 上进行训练。2014 年，Ian Goodfellow 发明了生成对抗网络，运用对抗网络训练方法来生成逼近真实图片程度非常高的计算机生成图片，已经无法又肉眼辨别，2016 年 DeepMind 公司将神经网络应用到了强化学习的领域，在 Atari 游戏平台的许多个游戏中击败人类，知名的击败全世界顶尖围棋高手的 AlphaGo 就是由这家公司发明的。而最具推广性意义的还是谷歌公司发布并开源的 TensorFlow 深度学习系统，首先该系统的计算速度非常快，能有效的进行大规模训练且兼容性强方便代码迁移，其次从计算平台上说，Tensorflow 不仅支持单机计算，也能支持 GPU 大规模分布式系统。在一般支持方面，TensorFlow 支持对现有深度学习模型的扩展^[2]。深度学习并不是指一个特定的数学模型，因此对于特定的机器学习场景和数据，模型调整甚至模型结构变化都是经常发生的。TensorFlow 提供自动梯度计算，允许用户自行选择特定的模型函数。即使它不是一个深度学习模型，它也可以在 TensorFlow 上进行训练。随着 TensorFlow 等开源软件的引入，深度学习将逐步从少数机器学习专家手中的“黑技术”转变为普通定量研究者可以控制的工具。期望研究者从繁琐的参数调整和优化计算中解脱出来，在市场数据中寻找隐藏的投资机会。



图表 3 深度学习重大事件历程

3. 国内外文献综述：近些年来，伴随着人工智能的飞速发展，越来越多的投资人群开始尝试利用机器学习的方法对股票价格进行建模分析与预测，都取得了较为可观的预测效果，而对于存在时间序列的股票数据预测来说，相对较为传统的自回归滑动平均模型和多元线性回归模型大多数情况下依靠许许多多的假设条件，仅仅能体现出股票价格与一些相关因素之间的线性关系，然而众所周知股票价格波动具有极强的非线性特征，因此这些方法的局限性就被凸显出来，因而许多研究者都会机器学习的方法来预测，最为广泛的就是采用 RNN（循环神经网络）方法，RNN 是一种专门研究出来用于处理与时间序列有关的数据类型，当然也局限于广义的时间序列数据，很多图像处理问题也用到了 RNN 网络，理论创立初期，许多学者喜欢应用 RNN 来进行对股票等金融资产价格的预测，但传统 RNN 也有其局限性，因为网络结构的原因，RNN 不得不面对因为超长的时间序列而导致的梯度消失以及梯度爆炸的问题^[3]。从而随着网络的不断学习更新，早期的数

据特点无法通过梯度得到体现，导致预测效果的不准确。随着深度学习的不断发展，学者们意识到 RNN 梯度消失的本质原因是隐藏层状态的计算方式的问题，该问题容易导致梯度被表示为连乘积的形式^[4]。1997 年，Hochreiter 和 Schmidhuber 提出了长短期记忆网络 LSTM，该网络巧妙的重新构造了隐藏层结构，相当大程度上缓解了传统 RNN 的梯度消失问题，后文将详细阐述 LSTM 和 RNN 相比是如何缓解了梯度消失的问题。随着近些年硬件水平的发展已经深度学习框架的完善，比如谷歌的 Tensorflow 深度学习框架，让许多个人与机构能够较低门槛的进入到这一领域进行研究探索，许多金融机构如证券研究所都成立了相应的小组来探索深度学习在金融领域的应用，其中很重要的一个方向就是用循环神经网络如 LSTM 来预测股票价格走势，这也是本文的主要研究方向，随着技术与算法的逐渐成熟，现在国内外探索的主要方向是如何调出最优的超参数与选择最优化的激活函数与优化函数甚至调整隐藏层网络结构来适应不同的股票或行业指数，如在 2014 年 Cho 提出的 GRU 门控循环单元，将 LSTM 中的遗忘门和输出门合并，成为单一的更新门，减少了很多的矩阵乘法运算，在大数据条件下能很大程度的节省模型训练时间，即使很多情况下准确率和 LSTM 模型并没有较大的差异。

（三）本文的内容与结构

本文将分为五部分来进行叙述，内容概括与具体分布如下

第一部分：绪论部分，介绍本文要研究的深度学习神经网络的历史发展，是如何被提出并且被一代一代人不断完善，并且介绍了发展深度神经网络的现实意义，介绍了神经网络有哪些的潜在优势可以促进人类社会的发展。最后是文献综述部分叙述的本文要重点研究的利用循环神经网络进行股价预测的发展与现状，并简单介绍了 RNN，LSTM，GRU 循环神经网络的。最后简单阐述了本文在研究利用循环神经网络预测股票价格波动的一些创新之处。

第二部分：研究方法设计部分，该部分将简单介绍最基础的全连接神经网络，并详细介绍 RNN 循环神经网络与 LSTM 长短时记忆网络的算法原理与具体实现方法，以及介绍 LSTM 网络很大程度上缓解 RNN 网络梯度消失问题的原理，最后将介绍本文设计实现的将 LSTM 神经网络应用于股票价格的预测分析模型框架。

第三部分：实证分析部分，该部分将对于第二部分设计的模型与算法的落地实现，本文中运用 Tensorflow 框架进行 python 编程，构建 LSTM 长短时记忆网络的训练模型，并对纳斯达克指数进行了短时预测，并通过参数的调整，类似网格搜索的方法，找到最优参数，使模型的损失函数达到最小，完成对 LSTM 神经网络的实证分析。

第四部分：结论部分，对第三部分实证分析产生的结果进行总结，对各循环神经网络应用与股价预测的可行性与准确性进行总结并综合对比各个网络用于股票预测时的优劣，通过数据可视化的方法将模型的预测成绩直观的展现出来。

第五部分：总结与展望部分，对本文的各类循环神经网络模型与实证分析结果进行概括总结，阐述模型在股票预测方面的优势，并提出目前模型与算法的局限性与不足的地方，最后做出展望，如何更好的优化模型并且结合各个地区股票市场的交易规则是循环神经网络模型能够落地，真正作为实盘的策略来进行机器

交易，获取超额收益。

（四）本文的创新点

第一：应用神经网络预测股价以及金融市场波动，与传统的时间序列模型或者逻辑回归模型相比，能够更好的表达出金融市场波动的非线性性质，当我们尝试用线性模型去解释非线性波动时，显然是不够合理的，而深度神经网络能通过多个隐藏层的参数更新，以及非线性激活函数的选取，能更好的体现出股票价格波动的非线性性质，达到更好的预测效果。

第二：循环神经网络参数很多，并且模型的训练速度比较慢，对于参数的调整也缺乏方向性，十分困难，需要人为的更改参数，通过网格搜索的方法以及分析 LSTM 网络训练过程的权重更新，我们成功的揭开了神经网络这一个“黑箱”，让我们的模型具备可读性与可解析性。

第三：相比于相对较为传统的 RNN 网络，本文与时俱进创新性的引入了 LSTM 与 GRU 两种新型循环神经网络，能够高效的缓解 RNN 网络在训练过程中的梯度消失问题，并且对长时间序列数据很够通过记忆门长久的保留住早期数据的信息，使得预测更加精准可靠，而 GRU 网络通过进一步隐藏层的调整优化，大大降低了运算量的同时还保证准确性没有发生较大改变。

第四：在进行模型网络建立的过程中，我们考虑到了不同市场下有着不同交易机制，如 T+0 与 T+1 等等，并针对特定的市场设计出可行性较高的方案，使得本文中设计的循环神经网络模型具有较高的现实可行性，且在模型评估的时候，我们增加了对模型是否具备识别短期黑天鹅事件的能力的评估力度，不仅仅拘泥于预测值与真实值的偏差。在金融市场，具备预判黑天鹅事件的能力，远比准确预测平稳序列中的股价波动要更加的有意义。

二、 研究方法的设计

（一） 循环神经网络介绍

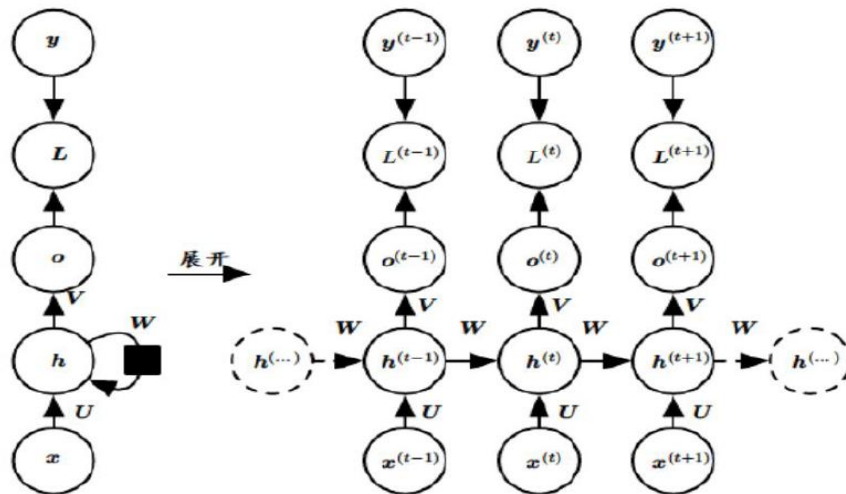
1. 全连接神经网络简介

神经网络已经逐渐发展成熟，其中网络类型也繁杂多样，诸如 BP 神经网络卷积神经网络，循环神经网络等等，而其中最底层也是最核心的，就是全连接神经网络，毫不夸张的说，这是所有神经网络的雏形，也是所有神经网络的基本结构，理解与领会全连接神经网络对学习循环神经网络至关重要。

全连接神经网络模型一般由输入层，若干个隐藏层以及输出层组成，每一层包含着数目不定的节点，节点个数属于超参数范畴，由网络构建者自由指定。该网络的精髓在于两点，前向传播与反向传播，前向传播指的对于给定数据和模型参数情况下，由输入层输入数据后经过一层层网络和激活函数的处理后输出预测值，尔后进入反向传播过程，由损失函数计算出模型预测值与真实值的差距，后反向传播来更新参数值，此处的核心是梯度下降法，通过更新参数从而更新模型，周而复始，直到损失函数的值越来越小，神经网络则越来越精准。无论是哪种神经网络，都将用到全连接网络的基本结构以及反向传播机制。其中提到的激活函数与损失函数将在本节后面详细介绍。

2. RNN 神经网络

概述：如英文名称所示以，RNN 称循环神经网络，如图 4 所示，RNN 网络读取输入数据 x ，然后经过 h 神经元部分的处理，输出到神经元 o 部分，一层一层往下运行直至输出最后的结果，这样看 RNN 网络看似是一个黑箱，于是我们在途中将它仔细的展开如图所示， x_t 代表了 t 刻的训练样本的输入，同理我们也能得知 x_{t-1} 与 x_{t+1} 的含义，此处 h_t 代表了 t 时刻模型隐藏层的状态，而此处的关键在于 h_t 不仅受到反向传播的影响，也会受到 $t-1$ 时刻，也就是 h_{t-1} 的影响， o_t 作为模型的输出值， L_t 为损失函数的计算， y_t 为 t 时刻真值，通过输入值与真值来计算损失函数，完成模型的迭代更新，更新的数据便是 U, V, W 三个存储权重数据的矩阵，这三个矩阵也是整个网络能够实现数据处理的核心。可以看出，经过这样设计的网络将具有处理时间序列数据的能力，能够将过去的的数据所产生的影响考虑进来，不断迭代的过程中，产生的数据都对未来的数据处理产生影响^[4]。



图表 4 RNN 的展开表示

算法详解：在任意的 t 时刻，隐藏层中 h_t 由 x_t 与 h_{t-1} 经过以下函数共同决定

$$h_t = \tanh(Ux_t + Wh_{t-1} + b)$$

其中 \tanh 为双曲正切函数， U 和 W 为存储权重参数的矩阵， b 为偏置项，然后经由输出函数输出 o_t

$$o_t = Vh_t + c$$

最后经过 RNN 网络计算得出的预测值为

$$y_t = \sigma(o_t)$$

此处的激活函数 σ 可根据实际情况自行选择，一般情况下若是探讨分类问题采用 softmax 函数，在股票价格预测问题上根据先前学者们的经验一般采用 tanh 或者 relu 函数。

与全连接网络不同，由于每个时刻都会有一个损失函数值的输出，并且由于权重参数矩阵是整个网络共享的，因此在进行损失函数梯度计算从而更新参数的时候，需要对每个时刻的损失函数值进行求和

$$L = \sum_{i=1}^n L_i$$

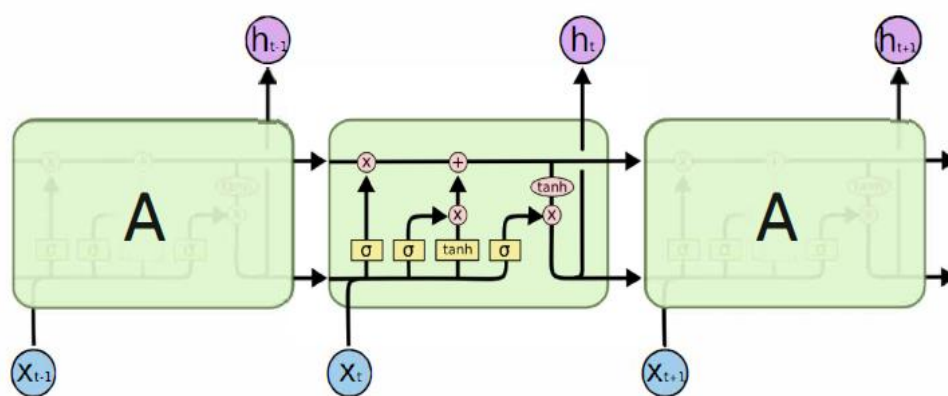
其中， L 为关于 U, W, V, b, c 的复合函数，由 L 关于这几个矩阵求偏导，得到每个权重矩阵的关于损失函数梯度，再乘上设定的学习率，即为网络运行一次对权重参数的更新值，以 W 为例，参数更新过程如公式所示，其中 η 为学习率，是一个超参数

$$w_{new} = w - \eta \frac{\partial L}{\partial w}$$

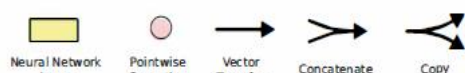
需要注意的是，此处的参数更新方法优化器选择了最简单的 vanilla SGD 优化器，而在实际训练过程中，将会选择更复杂的优化器，如 Adam 等等，后文介绍优化器的部分将详细叙述。经过不断的迭代更新，由 RNN 网络输出的数据会与真实值越来越接近，损失函数值也会越来越小。

3. LSTM 神经网络

概述：前面的文章我们可以看到 RNN 模型已经能够进行与时间序列有关的数据处理预测，已经能较好的用于股票价格等金融时间序列指标的预测与分析，但 RNN 网络有一个非常致命的缺陷，那便是在处理长序列数据的情况下，RNN 几乎不可避免的会面临梯度消失的问题，因为随着序列的大量增加，最开始的数据特点渐渐被网络“遗忘”，权重越来越小，导致很久以前的数据几乎对未来的数据没有任何影响，而梯度的消失更加导致了模型参数无法经过梯度计算来更新。LSTM 神经网络在 RNN 中 H 隐藏层的基础上增加了 C 隐藏层，正是由于新的隐藏层的引入，使得即使很长时间之前的数据的影响也能保留下来，极大程度上解决了梯度消失的问题，具体的 LSTM 隐藏状态网络结构如下图所示



The repeating module in an LSTM contains four interacting layers.



图表 5 LSTM 隐藏状态网络模型结构

算法详解：在 LSTM 模型当中，除了与 RNN 相同的 h_t 受到 h_{t-1} 的影响，还多了一个被称之为 **Cell State** 的隐藏状态，整个细胞状态与 LSTM 特有的门控结构以及 H 状态有关，门控结构由遗忘部分，输入部分，输出部分组成。将所有部分的公式陈列如下：

$$\begin{aligned}
 f_t &= \sigma(W_f h_{t-1} + U_f x_t + b_f) \\
 i_t &= \sigma(W_i h_{t-1} + U_i x_t + b_i) \\
 N_t &= \tanh(W_N h_{t-1} + U_N x_t + b_N) \\
 C_t &= C_{t-1} \odot f_t + i_t \odot N_t \\
 o_t &= \sigma(W_o h_{t-1} + U_o x_t + b_o) \\
 h_t &= o_t \odot \tanh(N_t)
 \end{aligned}$$

其中， \odot 表示矩阵的哈达玛积 (HADAMARD PRODUCT)。由上面详细的公式我们可以很清晰的认识到 h_t 不仅仅与 h_{t-1} 有关，也受到 **Cell State** 的影响， f 和 i 分别为两个窗口，一出一进，易理解为 **forget** 和 **in**， f 代表着遗忘部分，而 i 代表着输入部分，其中权重参数矩阵 W_N 为解决了 RNN 中梯度消失问题的关键所在，经过了这个 **Cell State** 的巧妙处理，LSTM 便成功的解决了梯度消失问题，使得我们有机会通过 LSTM 来进行长时间序列的学习以及预测，这对于以长时间序列著称的股票价格变化来说，无疑是至关重要的。

（二）常用激活函数，损失函数与优化器

本小节将介绍组成神经网络和深度学习算法的三个非常关键的部分，只有运用好这三个部分才能搭建出稳定高效的神经网络模型。

1. 激活函数

激活函数神经网络中用来引入非线性因素的函数，由于线性模型的表达能力不够，因而引入非线性的激活函数，可以使得深度神经网络能表达出更多非线性的性质。

一个好的激活函数应该满足一下的几个特点：非线性、可微性、单调性、近似恒等性，当激活函数输出为有限值时，基于梯度的优化方法十分可靠，反之当激活函数的输出是无限的值的时候，调小学习率是一个非常好的解决方案。

常见的激活函数有：sigmoid, tanh, ReLU, PReLU, ELU, softsign, softmax 等，下面将介绍几个常用的激活函数^[5]

Sigmoid 函数：

$$f(x) = \frac{1}{1 + e^{-x}}$$

该函数的优势在于输出映射在 0 和 1 的区间内，输出范围固定，优化与求导也稳定，劣势在于容易造成梯度的消失以及收敛速度慢，同时输出并不是以 0 为均值，且因为幂运算相对来说较为复杂会导致模型运算效率不高，耗费大量的算力。

Tanh 函数：

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$$

该函数的优势在于收敛速度较快并且以 0 为均值输出，劣势在于已造成梯度消失且运算量大。

ReLU 函数：

$$f(x) = \max(0, x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

该函数的优势在于在正区间内能解决梯度消失的问题，收敛速度也非常快，同时还能提供神经网络的稀疏表达能力，劣势在于非 0 均值输出，并且存在某些神经元永远不会被激活的可能。

Softmax 函数：

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K$$

该函数对神经网络全连接层输出进行变换，使其服从概率分布，即每个值都位于[0,1]区间且和为 1。

2. 损失函数

常用的损失函数有两种，一种称为均方误差损失函数，另一种称为交叉熵损失函数，在股票价格的预测模型中，如果最后输出的结果是类似一个分类问题，则选择交叉熵作为损失函数较为合理，若最后输出的是准确的股价，损失函数设为均方误差损失函数较为合理，当然这也不是绝对的，在一些特定的场景中，我们时常会自定义损失函数来满足我们特定的需求，但在大多数情况下，这两种损失函数被使用的最多。

均方误差损失函数：

$$MSE(y, y') = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n}$$

其中是 y_i 真实值， y'_i 是经由神经网络处理的预测值。

交叉熵损失函数：交叉熵（Cross Entropy）表征两个概率分布之间的距离，

交叉熵越小，说明二者分布越接近，是分类问题中使用较广泛的损失函数。计算公式如下：

$$H(y, y') = - \sum y * \ln y'$$

其中， y 代表数据的真实值， y' 代表数据的预测值，对于多分类问题，神经网络的输出通常情况下并非概率分布，因而需要使用softmax函数，这样一来输出的数值并符合加和为1的概率分布。

3. 优化器

文章前面提到过简单的参数的更新公式为新参数等于旧参数减去学习率与梯度的乘积，而大多数情况下，由于神经网络模型较为复杂，过于简单的梯度更新算法容易发生梯度爆炸或梯度消失的问题，因此需要引入优化器，优化器直接作用在参数更新的过程中。

优化算法可以分成一阶优化和二阶优化算法，其中一阶优化就是指的梯度算法及其变种，而二阶优化一般是用二阶导数（Hessian 矩阵）来计算，如牛顿法，由于需要计算 Hessian 阵和其逆矩阵，计算量较大，因此没有流行开来。这里主要总结一阶优化的各种梯度下降方法。总的来说，深度学习优化算法经历了 SGD -> SGDM -> NAG -> AdaGrad -> AdaDelta -> Adam -> Nadam 这样的发展历程。下面我们只介绍在神经网络应用金融市场是最为广泛使用的 Adam 优化器，也是本文实证分析中使用的优化器。

Adam: 名字来源于 adaptive moment estimation，可以说 Adam 融合了 Adagrad 和 RMSprop，事实上，SGDM 是在 SGD 基础上增加了一阶动量，AdaGrad、RMSProp 和 AdaDelta 在 SGD 基础上增加了二阶动量^[6]。而我们自然而然的想到，可以将一二阶的动量相结合，然后修正偏差，就是 Adam 了，公式如下：

$$\begin{aligned} g_t &= \frac{\partial L}{\partial W_t} \\ m_t &= \beta_1 * m_{t-1} + (1 - \beta_1) * g_t \\ V_t &= \beta_2 * V_{t-1} + (1 - \beta_2) * g_t^2 \\ \eta_t &= a * m_t / \sqrt{V_t} \\ W_{t+1} &= W_t - \eta_t \end{aligned}$$

其中， L 为模型设置的损失函数， W 为要更新的参数矩阵， m_t 与 V_t 分别为一阶动量与二阶动量， β_1 与 β_2 分别为经验参数值，属于超参数由模型作者自行制定，在 Tensorflow 框架中，默认情况下取值分别为 0.9 与 0.999。

（三）LSTM 长短时记忆网络预测股价模型

1. 模型简述

由于 RNN 模型的局限性以及 LSTM 模型相对 GRU 模型更为完善，本文将重点与实证分析放在了 LSTM 长短时记忆网络上，构建 LSTM 模型对选定的股票进行短期的预测，其中模型的输入为多段选定时间区间股票各种价格数据与成交量数据，验证数据为最近等长时间区间的股价数据，在循环神经网络范畴中属于 manyto many 多对多数据类型，并建立标准的 LSTM 网络隐藏层，隐藏层具体分析前文已经给出，模型最后还设计了全连接神经网络，数据经过 LSTM 处理后再

通过全连接网络，这样设计可以增加网络的复杂性，提高预测准确率。

2. 模型评估

在模型的构建与完善过程中，评估一个模型的好坏是十分关键的，合适的方法评估一个预测模型有利于预测模型的不断完善。本文的模型功能是通过循环神经网络预测短期股票价格，顾名思义，预测价格与实际价格的偏差，偏差比例都可以作为评估的标准，以及损失函数的收敛速度与效率，也是评估的标准之一。本文还考虑到一点更具实际意义的评估方向，在股票市场上，我们可能会面临金融危机等黑天鹅事件突变，此时如果模型具备预测短时间市场风格迅速切变的能力，那这个模型将会变得非常的有价值，这也是本文在模型评估标准上更加独特的地方。

（四）LSTM 模型的参数设计

1. 超参数的选取

对于 LSTM 模型参数初始化采取生成截断标准正态分布的随机数，因为本文着眼于循环神经网络对股票价格的预测，为保证模型的广泛应用性，学习率采用 Tensorflow 框架中的默认学习率，Adam 优化器中的经验参数值也采取框架内置默认值，而 dropout 参数，即随机冻结神经元个数比例，将作为一个待探讨超参数，将由网格搜索法找到待建模型的最优 dropout 值，同时 decay 衰减率也将作为待探讨超参数出现。

2. 神经网络结构的选择

在深度神经网络模型的构建过程中，首先是确立模型的种类，本文实证中采用的是 LSTM 神经网络，由卷积神经网络的发展历程受到启发，神经网络的模型的结构设计至关重要，卷积神经网络在图像识别领域的不断更新发展，识别的准确率越来越高，很大程度上都是依赖于模型结构的不断完善，神经网络层数的不断增加与每一层神经元个数的优化调整。因此在本文的 LSTM 模型实证分析中，神经网络结构的设计被设定为待考量参数，将通过不断变换 LSTM 模型的结构，如层数与每层神经元个数来找到在股票短期价格预测模型中效果相对较优的模型结构。

三、 LSTM 神经网络预测短期股价实证分析

（一） 实证流程概述

本文将采用 LSTM 网络模型进行实证研究，以美股市场的纳斯达克指数为研究对象，获取纳斯达克指数的有关的历史数据，进行预处理后送入构建的网络模型进行训练与测试，然后通过不断调整模型参数与结构找到最佳模型，最后进行模型预测效果的分析与模型评价，流程图如下图所示



图表 6 实证分析中流程示意图

（二） 数据处理

1. 数据获取

本模型的数据研究对象是美股市场的纳斯达克指数，纳斯达克指数即为美股市场上在纳斯达克证券交易所交易的股票的一揽子指数，通过对所有成分股的加权平均处理，得到一个交易所股票的综合价格数据，该数据也很大程度上意味着美股的总体行情。本文中获取数据的方法是利用 python 的 pandas_datareader.data 库中从财经网站获取行情数据的功能，从雅虎（YaHoo）财经网站获取纳斯达克指数的历史数据，再通过 pandas 将历史行情数据转化为 DataFrame 格式，数据具有五个维度的标签，以天为单位，分别为开盘价（open price）、最高价（high price）、最低价（low price）与调整后收盘价（adjusted close price）。通过运行代码将获取的数据的表头示例展示如下

Date	High	Low	Open	Adj close
1980-01-02	148.169998	148.169998	148.169998	148.169998
1980-01-03	145.970001	145.970001	145.970001	145.970001
1980-01-04	148.020004	148.020004	148.020004	148.020004
1980-01-07	148.619995	148.619995	148.619995	148.619995
1980-01-08	150.679993	150.679993	150.679993	150.679993

图表 7 获取行情数据表头展示

其中，我们所选取的数据是纳斯达克指数从 1980 的第一个交易日到模型运行当天的数据，选取的时间段非常的长，对模型的挑战更大。

2. 数据预处理

缺失值处理：对于雅虎财经纳斯达克指数数据中若存在缺失的部分，编写识别代码对数据中缺失的地方做出标记，并用前后各三天的相同指标数据均值进行

补全。

异常值处理：设定阈值找到数据中可能存在的异常值进行标记，并与 Wind 数据库中的相同标签数据进行比对，判断数据的真实性，若数据失真，则采用 Wind 数据库中的数据进行替换。

数据标准化：在将数据送入神经网络训练以前，都会采取标准化的处理，特别是对于股价数据来说，其数值很大，且不同标签对应的数值在绝对值上可能有着巨大的差别，本文模型采用的是极值标准化的方法，公式如下

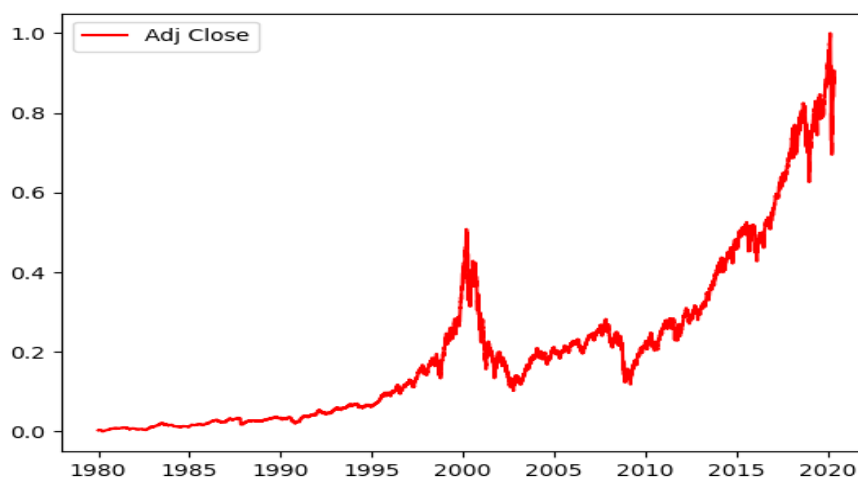
$$X^* = \frac{X - \min}{\max - \min}$$

对所有行情数据进行标准化以后，数据表头展示如下

Date	High	Low	Open	Adj close
1980-01-02	0.002479	0.002495	0.002489	0.002484
1980-01-03	0.002252	0.002267	0.002261	0.002257
1980-01-04	0.002463	0.002479	0.002473	0.002469
1980-01-07	0.002525	0.002541	0.002535	0.002531
1980-01-08	0.002737	0.002755	0.002748	0.002743

图表 8 标准化后数据表头展示

并将所有标准化后的收盘价的数据按时间顺序展示出来，如下图所示



图表 9 纳斯达克调整后收盘价数据（标准化后）

3. 训练集与测试集划分

常用的神经网络训练过程中为了是模型更加具备普遍性和准确率更高，常使用交叉验证的方法，即测试集与训练集的比例的不断变化和身份的互换，从而取得最优的参数组合，而本文只采用了最简单的训练集与测试集的划分方式，及单纯的按比例划分训练集与测试集并且不再改变，因为对于具备时间序列属性的行情数据来说，我们只能通过过去的的数据，来预测未来的数据，选择新数据作为训练集，旧数据作为测试集，显然是不具备现实意义和预测价值的，因此本文的模

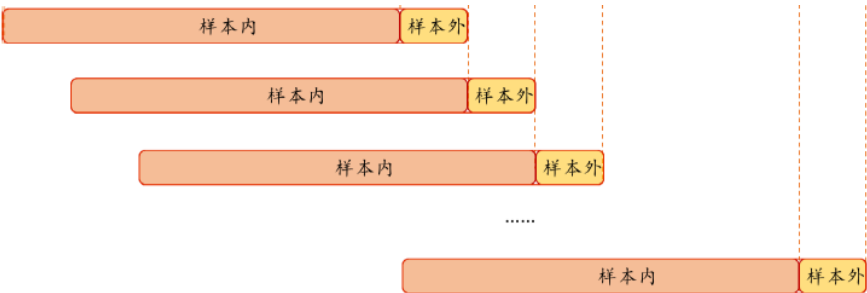
型中将所有输入数据的 90%标记为训练集，而剩下的 10%标记为测试集，并且不再改变，训练集与测试集的数据在网络模型训练测试过程中，永远不会有交集，这也符合神经网络训练原则。

(三) 实证过程详解

本文的实证过程由 python 语言编码实现，充分利用谷歌开源的 **Tensorflow2.1** 框架与 **Keras** 框架，在这两个框架的基础山进行 LSTM 神经网络的构建，所有的有关实证处理都在该框架下进行。

1. 模型的初步训练

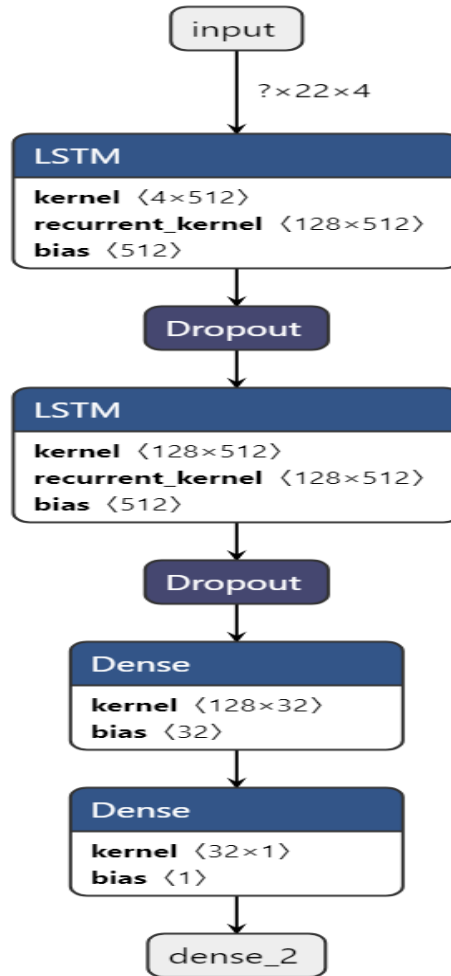
输入数据：首先将数据集进行初步的划分，模型的设计为取一段时间窗口的历史数据作为训练数据，并预测这一段滑窗的后一天的价格，选取一个月作为滑窗，通常情况下一个月有 22 个交易日，**模型则将每 22 天的行情数据作为一个窗口**，通过前 21 天的行情数据来输出预测第 22 天的调整后收盘价数据，为充分利用数据，我们每个新滑窗的起点都是前一个滑窗起点往后推一天，每次输入模型中的数据维度为[21, 4]，输出为 1 个单值，时间窗口可视化理解如下图所示



图表 10 时间窗口示意图

训练轮数与参数设置：在初始参数与模型结构的设置中，保守选择**设置神经元的随机屏蔽率为 20%**，即 **dropout 值为 0.2**，该操作的目的是避免发生过拟合，根据经验初步将**模型的循环轮数 epochs 设置为 200 轮**，可以通过编写 python 代码对权重参数进行本低存储，这也多次运行时，都相当于在原先的基础上继续运行，运行 N 次程序训练时，整体循环轮数变为 N*epochs。

初始模型结构设置：本文中的模型主要部分为 LSTM 长短时记忆网络结构，并且由于最后输出判断的是单个数值，于是最后一层网络只能具有一个神经元对应一个输出。在 LSTM 网络的选择上，为了使得训练效果显著，模型采用双层 LSTM 网络，**两个 LSTM 层的初始的节点数目分别设置为 128, 128**。这个数目设置是依据阅读的相关文献优秀结果得出的经验。由于最后由单层单个节点输出，因此在输出前需要再添加一个隐藏层，这一层采用**全连接网络的连接方式**，节点数设置为 **32**。于是整体模型结构经过可视化工具处理展示如下



图表 11 整体模型网络结构展示

其中，kernel、recurrent_kernel、bias 分别代表各个权重参数矩阵的维度。

激活函数、损失函数与优化器选择：本文模型中，LSTM 层的激活函数与超参数采用的是 Keras 框架中对于 LSTM 模型默认的参数值，在最后两层的全连接层中，倒数第二层采用 **ReLU 激活函数**，最后一层由于只输出一个单值，则采用最简单的**线性激活函数**。由于本实证中对比的是预测股价与实际股价之间的差距，因此采用 **MSE** 作为损失函数，Adam 算法作为参数更新时的优化算法^[8]。

到这里本文的实证部分的模型搭建已经完成，在 Tensorflow2.1 框架和 python3.7 环境下，便可以执行模型的训练然后得到结果

2. 优化 DROPOUT 参数

在本文前面部分提到过，本文将 dropout 值作为变量，将考察 dropout 取不同的值时，模型的表现，本文中由训练集与测试集损失函数值，以及去标准化后的实际值与预测值的偏差比例作为评估模型优劣的标准，这个标准也同样适用于其它参数变量的考量。此处采用网格搜索的方法^[9]，遍历 dropout 值取 **0.2 到 0.8**，间隔为 **0.1**。一共 7 个 dropout 值，找出最优的一个。

3. 优化衰减率

本模型中要调整的衰减率为 Adam 优化器中的衰减率,也属于超参数的范畴,衰减率对优化器的影响非常大,而优化器对于参数的更新起到了决定性的影响,于是选取不同的衰减率(decay)进行模型训练,从而得到该模型的最优衰减率,本文中**衰减率取 0.1 到 0.9, 间隔为 0.1, 一共 9 个值进行测试。**

4. 优化神经网络结构

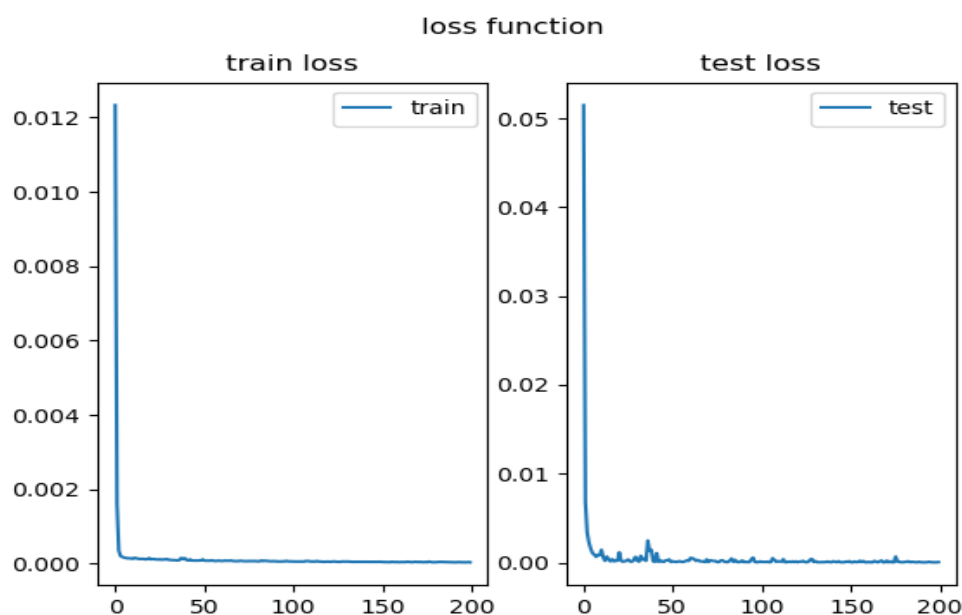
前文部分多次提到了神经网络结构的选择对模型的预测准确率以及普适性的重要意义,由于改变模型层数会使计算量剧增,于是本文的优化着眼于各层节点数的改变,本模型的最后一层固定为单节点不可更改,于是选择更改两层 LSTM 的节点数,此处由 LSTM 模型的特点决定,这两层的节点数需保持一致^[10],因此更改其中一层的节点数另一层也需要保持一致,**分别取 32, 64, 128, 256, 512,**并于第三层全连接层节点数进行网格搜索法下的随机组合,**第三层待选节点数为 16, 32, 64。**运用网格搜索法得出相对最优的模型层数选择。

四、 基于 LSTM 模型预测股价实证结论

（一） 实证结果展示

1. 初始模型结果展示

在第一次训练中，我们如前文所述，根据经验以及其它参考文献提供的思路设定了一系列的超参数与网络结构，接下来对初次训练结果作出展示，其中训练集与测试集中损失函数值变化经可视化如下图所示：



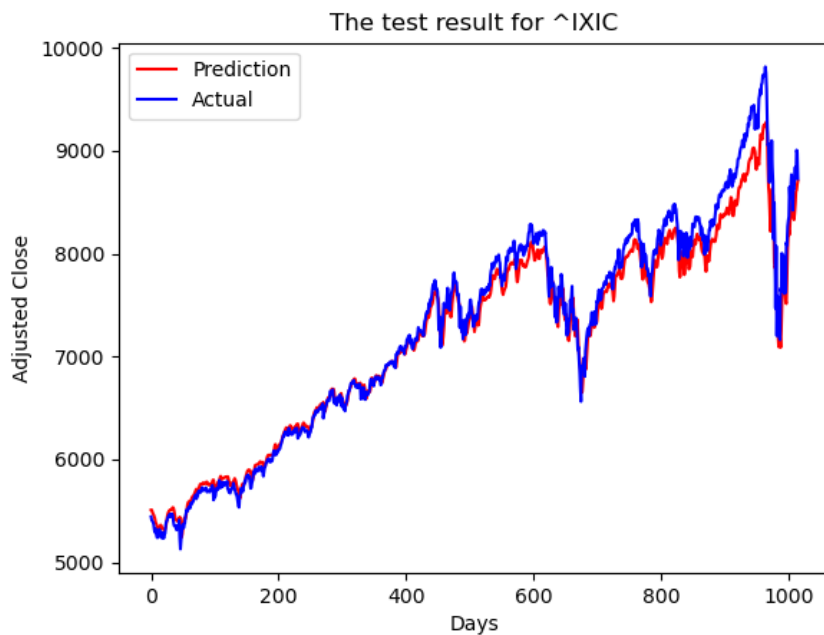
图表 12 初次训练损失函数展示

如图所示训练集与测试集的损失函数都在前 20 次循环中快速收敛，并在后面 180 次运行中逐渐收敛。并计算模型的训练集与测试集中预测数据与真实值的均方误差与均方根误差，同时并计算测试集中，模型预测数据与真实数据的平均偏差比例，计算结果如下

TYPE	MSE	RMSE
Train Score	0.00003	0.01
Test Score	0.00037	0.02
Percentage Different = 31.4112%		

图表 13 初次训练评估指标概览

再通过 matplotlib 画图工具绘制出测试集内，预测纳斯达克指数值与实际纳斯达克指数值的图像叠加在一张图中，如下所示

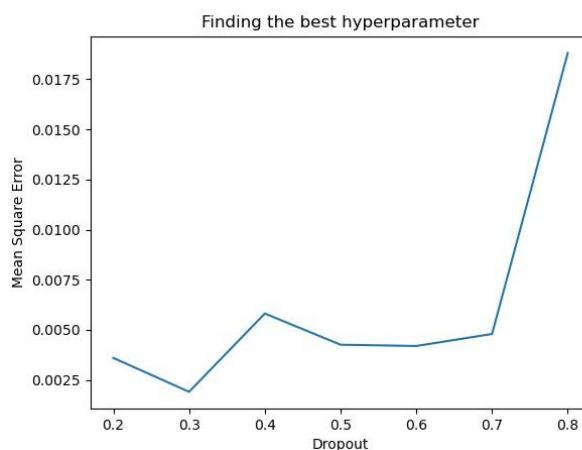


图表 14 初次训练结果可视化

由此图可以看出，借助于简单的 LSTM 网络，我们成功的实现了对短期股价的预测，整个预测值与实际值只有较小的偏离与迟滞，这对于一个模型的初次尝试来说可以说是较为成功的，在整体趋势上也紧跟着真实指数。

2. Dropout 优化展示

本小节根据实证部分所述，对进行 dropout 值优化选择的训练过程与结果进行可视化与结果分析，并给出对于当前选取的股票与预测模型下最优的 dropout 值。为了计算方便以及满足单一变量的原则，epochs 值将修改为 100 次，即总体循环轮数缩减为 100 轮。得出的优化结果如下所示

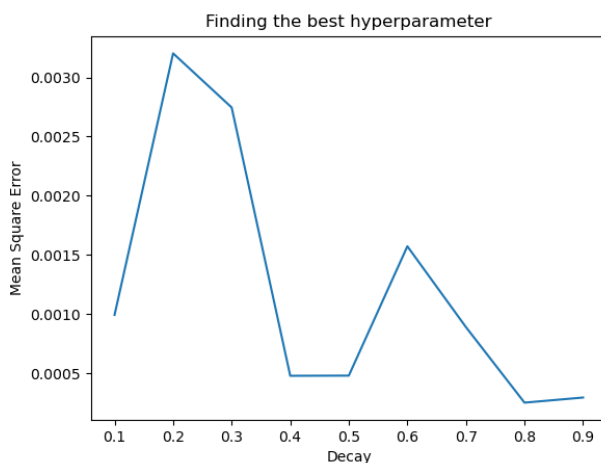


图表 15 寻找最佳 dropout 值

如图可以看出，当 dropout 值取 0.3 的时候，整个系统的均方误差达到最小，模型训练的效果最好，并且随着 dropout 值的增加，系统有渐渐失控的趋势。于是后面的训练将固定 dropout 参数为 0.3。

3. 衰减率优化展示

本小节将根据实证部分所述，对进行衰减率（decay）值优化选择的训练过程与结果进行可视化与结果分析，并给出对于当前选取的股票与预测模型下最优的 decay 值。同样在 epochs = 100 的条件下进行模型训练，结果如下

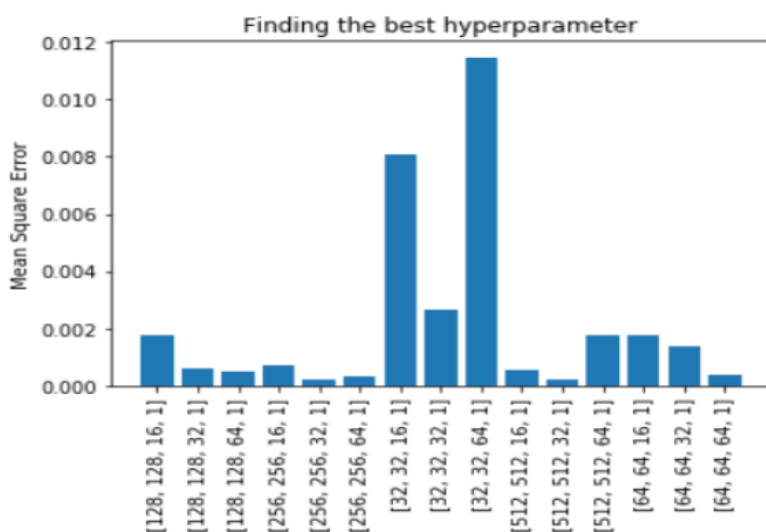


图表 16 寻找最佳 decay 值

如图所示，最适合本模型的 decay 值为 0.8。于是在后面的训练将 decay 值固定为 0.8，将使得模型保持最佳化。

4. 网络结构（各隐藏层节点数）优化展示

本小节同样将根据实证部分所述，进行网络结构的网格搜索优化选择，对训练过程与结果进行可视化与结果分析，并给出对于当前选取的股票与预测模型下最优的网络结构。同样在 epochs = 100 的条件下进行模型训练，结果如下

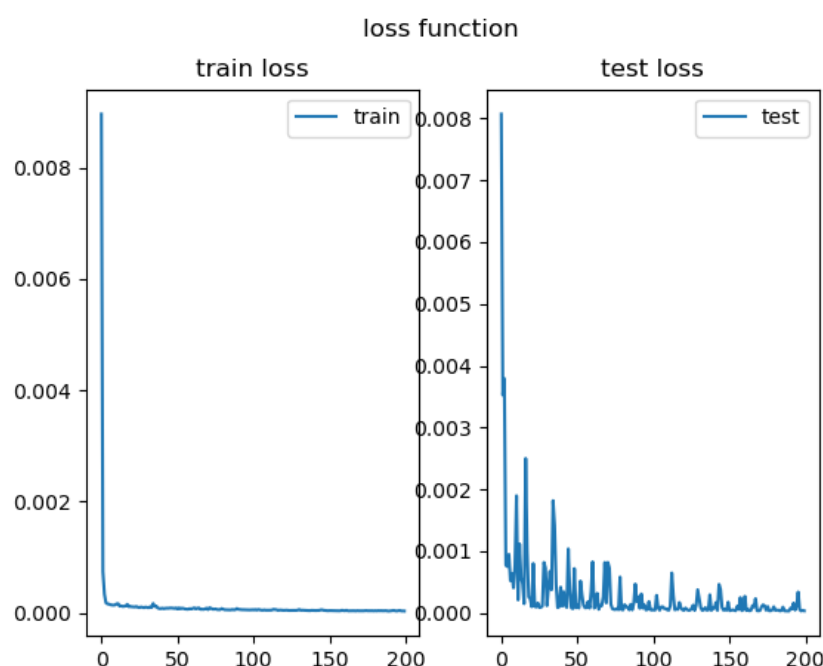


图表 17 寻找最优网络结构

如图所示，最适合本模型的神经网络结构为[256, 256, 32, 1]，其中前两个数字代表两个 LSTM 层的节点数，后两个数字代表全连接层的节点数。

5. 条件最优测试

本小节为最终实验，将所有的超参数设置为经过优化的相对最优超参数，并重新进行模型的训练与误差的计算，下面对结果做出展示，其中，画出损失函数图像如下

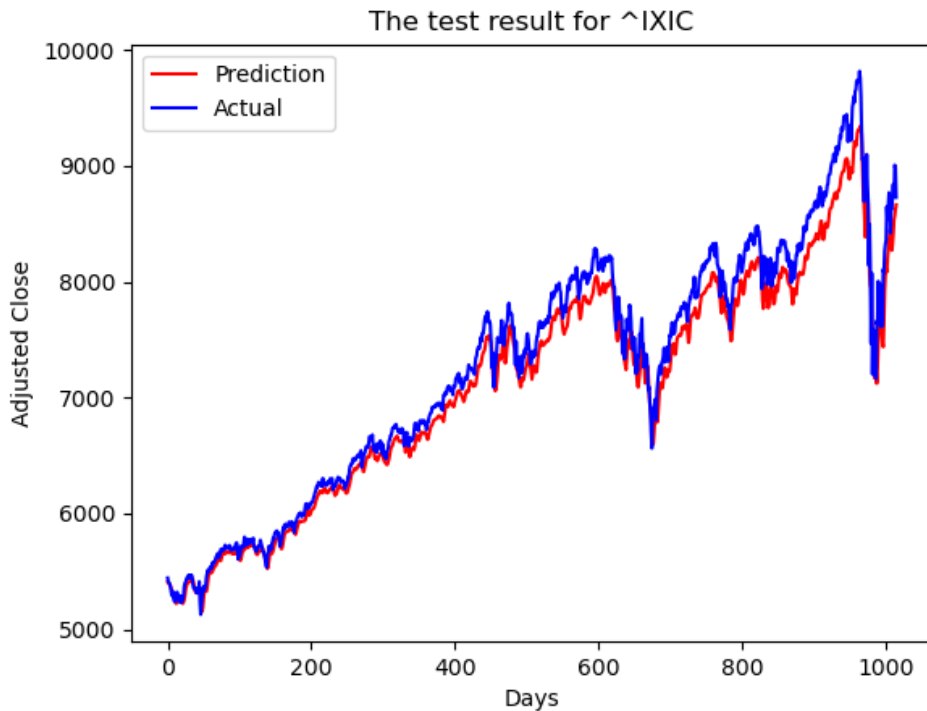


图表 18 最终实验损失函数

损失函数情况如图所示，可以看出，训练集的损失函数图像快速收敛，而测试集的损失函数有规律的呈现波段收敛，整体慢慢的趋于稳定的极小。再计算最终模型的训练集与测试集中预测数据与真实值的均方误差与均方根误差，同时并计算测试集中，模型预测数据与真实数据的平均偏差比例，将结果展示如下

TYPE	MSE	RMSE
Train Score	0.00002	0.00
Test Score	0.00017	0.01
Percentage Different = 26.2352%		

偏差数据与最开始的模型结果进行对比，明显可以看出模型的准确率得到了大幅度的优化，各项偏差值都有所减小，再通过 matplotlib 画图工具绘制出测试集内，与第一次执行训练时一样，展示预测纳斯达克指数值与实际纳斯达克指数值的图像，叠加在一张图中，如下图所示。



图表 19 最终训练结果可视化

对图像进行分析，与初次训练相比，最终训练图像的预测曲线明显的与实际曲线更加的贴合，说明模型的训练效果更好，预测股价变得更加的精准，同时对于市场风格的切换识别也是更加的迅速，因此最终的模型与最初的模型相比，毫无疑问经过优化后的最终模型在实际应用上能获得更多的超额收益，由此也可以看出，LSTM 为架构的神经网络模型在预测短期股票价格波动上还是非常的有潜力与应用价值的。

（二）实证结语

本文的实证分析到这里也就告一段落了，在实证过程中，本文先是建立了基于 LSTM 神经网络框架的网络模型，并进行股票数据的训练与预测，此时已经得到一个预测结果较为可观的模型，然后再借助网格搜索优化的思想，对模型的一些关键参数进行调优，得到一个不断优化过后的模型，经过实证，本文成功的建立并优化了一个能够预测未来一天股价并且具有较高的准确率的模型，这一点说明了 LSTM 长短时记忆网络在股票预测上具备一定的可行性与潜力，值得深入研究在股票市场的应用，这也许是机器交易的下一个风口所在。同时 LSTM 巧妙的构造也使得能够实证中避免梯度消失的问题，较成功的预测了 1980 至今如此长时间段的股价。总而言之，对于循环神经网络应用股票市场的初探来说，本模型的构建与预测及最终成绩，也算是获得了初步的成功。

五、 总结与展望

总结：顺应着科学技术的发展，以及神经网络算法与应用的飞速发展，本文与时俱进的着眼于应用循环神经网络来预测短期的股票价格，文章由浅入深，既有神经网络基础内容的详细介绍和引入，也有算法设计和实证分析的深入应用，可以说是一次对循环神经网络应用于金融市场的非常有意义的实践，最后通过实证成功构建出了一个用于预测股价的较可靠的模型。

展望：本文的也有许多地方有不小的局限性，例如在标的的选取上，仅仅选取纳斯达克指数作为对象，在实际应用中，可能随着标的物的转变，模型的风格都会大相径庭，并且若想将模型应用于中国市场，还有很长的路要走，其原因便是纳斯达克与上海深圳证券交易所的交易规则具有很大的区别，美股市场普遍采取 T+0 而 A 股市场采取 T+1。在数据选择输入方面，本文模型仅仅采纳了几个较为简单的行情指标，而真实的应用场景中，有数不胜数的相关指标值得我们去尝试，这甚至触及到了量化交易的另一个领域，Alpha 因子的探究，只要是与股价波动有关的数据，我们都能用来作为输入，都有可能产生意想不到的效果与准确率，这都是值得我们在未来继续进行探索的方向。与此同时，LSTM 在不断的发展已经衍生出了许许多多的变种，内核虽然还是 LSTM 的结构，但就像都是基于 RNN 的结构一样，网络的结构一直在优化发展，如 GRU 结构就是一个很好的例子，或许循环神经网络随着结构的不断优化，能够衍生出一种对股票预测适应性极强的模型，这也是未来该方向探索的目标。

本文前面曾提出过一个创新点，就是本文将网络模型是否具有判断黑天鹅事件，即金融市场重大风格切边的能力，从实证结果可视化来看，我们的初始模型没有做到这一点，优化后的模型只能说是跟进风格切边更加迅速，但并不具备提前识别风格切边的能力，这其中可能的原因是输入数据并不具备这方面的信息，因此无论如何调节网络，都无法实现这一点的需求。但本文认为这一点是当今甚至未来神经网络发展的一个极其重要的方向，在现实中黑天鹅事件的到来就像一个黑箱，无人可以预判，而神经网络本身也具备许许多多黑箱的特质，因此用黑箱来预测黑箱，这可能是一个非常可行的方向，当神经网络具备了预测金融市场黑天鹅的能力时，对全人类的影响都是变革性的，因此本文也将这一点视为最大的展望。

由于时间以及篇幅等限制，本文对循环神经网络的探讨还停留在一个浅析的阶段，展望未来，我将进行更深入的研究，在这个领域继续深耕。

参考文献

- [1] 安宁宁. 深度学习股指期货日内交易策略[EB/OL]. 广发证券发展研究中心. 2014
- [2] 张超. 谷歌开源 Tensorflow 点评[EB/OL]. 广发证券发展研究中心. 2015. 11. 13
- [3] 王理同, 薛腾腾, 王惠敏, 刘震. 基于循环神经网络的股指价格预测研究[J]. 浙江工业大学学报, 2019, 47(02):186-191.
- [4] 林晓明, 陈烨. 人工智能选股之循环神经网络模型[EB/OL]. 华泰证券研究所. 2017. 11. 24
- [5] 李树阳. 基于循环神经网络及百度指数预测 A 股市场波动率[D]. 山东大学, 2019
- [6] 丁鲁明. 大数据、机器学习、深度学习在投资领域应用的方法论概述[EB/OL]. 中信建投证券研究发展部. 2017. 10. 8
- [7] 文巧钧, 安宁宁. 风险中性的深度学习选股策略[EB/OL]. 广发证券发展研究中心. 2018. 7. 14
- [8] Bing Yang. Stock Market Index Prediction Using Deep Neural Network Ensemble[C]. 中国自动化学会控制理论专业委员会. 第 36 届中国控制会议论文集(C). 中国自动化学会控制理论专业委员会:中国自动化学会控制理论专业委员会, 2017:580-585.
- [9] 陈奥林, 杨能. 基于深度学习理念的高频交易策略[EB/OL]. 国泰君安证券金融工程团队. 2020. 3. 19
- [10] 林晓明, 陈烨. 人工智能选股之卷积神经网络[EB/OL]. 华泰证券研究所. 2019. 02. 13