

基于多时间尺度复合深度神经网络的股票价格预测

■ 罗鑫 张金林

摘要:本文以2012—2019年沪深300指数为样本,利用深度学习方法对沪深300指数的涨跌方向进行预测。在多时间尺度上分别运用卷积神经网络与长短时记忆模型进行特征提取后,通过将不同时间尺度上的特征矩阵进行拼接而得到最终的预测结果。在使用不同网络结构超参数调整模型结构后,将预测效果与其他模型进行比较,发现本文提出的多时间尺度CNN-LSTM模型能够有效改善对沪深300指数涨跌预测的效果,并在交易回测中获得盈利。本文的研究丰富了金融时间序列数据分析的方法,既能为投资者提供决策参考,也有助于提升对金融市场规律的认知。

关键词:深度学习;神经网络;股票价格;长短时记忆模型;卷积神经网络

中图分类号:F830.91 文献标识码:A 文章编号:1009-3540(2020)09-0032-0009

基金项目:国家社会科学基金项目“金融科技背景下普惠金融机制与路径研究”(19BJY250)。

一、引言

股票价格受到众多因素的共同影响,是一个极为复杂的动力学系统,具有非线性、非平稳性、低信噪比以及长记忆性等特点。经典的金融时间序列预测模型需要假定数据满足某些特定要求,但含有参数的解析方程难以对如此复杂的动力学系统进行准确的刻画,因此股票价格预测一直难以取得较好的效果。自从2016年人工智能程序AlphaGo战胜世界围棋冠军李世石以来,其应用的深度学习方法在图像识别、语音识别、无人驾驶等多个领域取得了极为成功的应用。对于金融领域而言,人工智能无论在学术界还是业界都越来越多地得到研究与应用,与大数据、区块链等其他技术手段共同作为当前金融科技战略的重要技术驱动。

深度学习是一种基于多层神经网络的机器学习方法,其由神经网络输出层和输入层以及两者之间的一系列堆叠的隐含层所构成(Schmidhuber, 2015),通过逐层的传递提炼出学习对象高度抽象、

复杂的特征,并以此特征矩阵作为数据的表现形式,获得输入数据与输出数据之间高度复杂的非线性函数,从而最终提升分类或预测的准确性(Najafabadi et al., 2015)。基于深度学习在处理高维度、非线性、非平稳的数据结构上具有优势,本文要研究的问题是通过深度学习的方法是否能够对金融时间序列数据进行有效的建模以及是否能够提高股票价格的预测准确率。深入研究这些问题,既能为投资者提供决策参考,也可以丰富金融时间序列数据处理与研究方法,有助于提高对金融市场风险的监管和预警能力。

针对以上问题,本文以沪深300指数为例,将目前深度学习领域中应用较多的长短时记忆模型(Long-Short Term Memory, LSTM)与卷积神经网络(Convolutional Neural Networks, CNN)进行融合,以同时利用其特征提取优势,并分别对日K线时间尺度与15分钟K线时间尺度数据应用以上复合深度神经网络进行特征提取,将两种时间尺度上的输入数据分别映射到高维特征空间后,再对其特征矩阵

作者简介:罗鑫(1988—),男,中南财经政法大学金融学院博士研究生,讲师;张金林(1964—),男,博士,中南财经政法大学金融学院教授。

进行拼接,生成涨跌预测的最终结果。在比较不同网络结构超参数后,实证研究结果表明CNN-LSTM复合模型对于沪深300指数的预测效果好于传统浅层神经网络与单模型深度神经网络,利用多时间尺度数据融合预测,可以进一步提升复合深度神经网络的预测效果。

本文的研究贡献主要在以下方面:第一,本文丰富了有关利用机器学习对金融时间序列数据进行建模的研究。已有文献主要利用支持向量机、随机森林、K近邻算法以及深度学习等方法,但对于深度学习在金融时间序列数据的应用研究仅仅局限于使用单一模型。本文将研究模型扩展到复合深度神经网络,不同类型的深度学习模型对于输入数据会通过不同的方式和算法提取特征,因此复合深度神经网络可以结合不同模型的优点。本文分别建立了CNN-LSTM与LSTM-CNN两种复合模型进行比较,发现CNN-LSTM模型可以更有效地分析处理金融时间序列数据。第二,已有文献忽视了结合不同时间尺度数据的信息和特征,往往是用同一频率的数据进行预测,而在实际的交易决策中,投资者一般会参考多个时间尺度上的股价信息。本文分别提取不同时间尺度上的特征后,再将特征矩阵拼接以更充分利用信息,并挖掘出金融时间序列数据在不同时间尺度上的特征,从而更好地对其进行预测。

二、文献综述

(一)金融市场的可预测性

对于资产价格预测的问题,传统的观点以Fama(1965,1970)提出的“有效市场假说(Efficient Markets Hypothesis, EMH)”为代表,认为在有效市场中任何新生信息都将很快被市场吸收并反映在股价上,股票价格因服从随机游走特征而无法准确预测。根据市场参与者可获得信息集的不同,EMH将市场效率分为弱式有效、半强有效以及强式有效三个层次,早期的研究大多发现股市满足弱式有效。Grossman(1976,1981)、Allen(1981)以及Jordan(1992)为有效市场假说建立了一般均衡模型,在不同的约束下论证了竞争性的理性预期均衡价格能够完全反映私人信息。Fama(1991)将弱式有效检验扩展到收益可预测性的检验,并将可预测性描述为未来收益率可由历史收益率或者其他变量预测的能力。早期对于弱式有效的检验方法主要包括方差比检验、单位根检验、鞅过程检验以及随机游走检验等,目前EMH已经摒弃了价格行为应满足随机游走的观点,而是将其作为一个特例,非线性序列相关检

验和长记忆性检验开始得到更多的应用。

但是自从EMH被提出以来,关于市场有效性程度的争论不断。20世纪80年代兴起的行为金融学提出投资者存在“过度自信”“过度交易”以及“处置效应”等非理性行为(Barber和Odean,2001;Shefrin和Statman,1985;Odean,1998),投资者实际上不满足“理性经济人”假设要求的期望效用偏好,而应以前景理论来描述投资者的偏好特征(Kahneman和Tversky,1979)。由于套利限制的存在,理性套利者在许多情况下无法克服非理性行为的影响,非理性投资者对于市场的影响可能是长期和深远的(Shleifer和Vishny,1997),这也就意味着金融市场并不总是具备市场有效性。事实上,从信息经济学的角度来讲,如果市场不具有一定的无效性,反而不可能真正实现有效性(Grossman和Stiglitz,1980)。

随着计量检验方法的进步与相关研究的逐渐深入,研究者发现早期对于股票价格满足随机游走的检验没有考虑到条件异方差、非线性关系以及波动聚集等因素,实证结果的可信度存在疑问(陈灯塔和洪永淼,2003)。这一时期大量的实证检验结果并不支持股市为弱式有效市场。French(1980)和Gultekin et al.(1983)分别发现美国股票市场上股票价格存在“周末效应”和“季节效应”,表明美国股市尚未达到弱式有效,此后这些现象又在其他国家的股票市场被陆续发现。相比于短期,股票市场在中长期会存在“长期反转”与“中期惯性”的现象(De Bondt和Thaler,1985;Jegadeesh和Titman,1993),因而具有更高的可预测性。Keim和Stambaugh(1986)、Campbell和Shiller(1988)通过实证检验发现,一些相关变量如股息价格比(D/P)、盈余价格比(E/P)对未来收益率具有显著的预测能力。基于以上大量研究结论,越来越多的学者相信股票市场存在一定的可预测性。Campbell和Cochrane(1999)将股票收益率存在可预测性称为金融学中“新的事实”(New Facts)。可预测性与市场有效性也不再被认为一定是互相冲突的。Campbell et al.(1996)提出了相对有效的概念,认为完美的效率在实际中不可能存在。Lo(2004,2005)提出的适应性市场假说(Adaptive Markets Hypothesis, AMH)认为在不断变化的市场环境下,可预测性来源于投资者要求的风险溢价的时变性。金融市场一定程度的可预测性,并不代表市场无效率,而是市场维持有效性必要的动力。

(二)基于深度学习的金融市场预测

近些年以来,国内外学者对于利用机器学习方法预测金融市场进行了很多有价值的研究。在方法

上,从过去主要利用支持向量机(张贵生和张信东,2016;苏治和傅晓媛,2013)和随机森林(王淑燕等,2016;曹正凤等,2014)等传统机器学习方法逐步向利用多层BP神经网络、循环神经网络和卷积神经网络等深度学习方法转变。这一转变主要是由金融时间序列的高维度、非线性等特点决定的。一方面,传统的机器学习方法对于复杂的高维度数据的学习效果较差,而且存在维度灾难与特征表示无效的缺陷(Bengio和LeCun,2007);另一方面,其对于复杂非线性数据的建模能力有限,难以对股票价格数据进行有效的拟合(Langkvist et al.,2014)。而深度学习所利用神经网络在数学上已经证明在拥有足够多神经元的情况下,其能够以任意精度逼近任何连续非线性函数(Hornik et al.,1989)。

许多研究结果已经表明深度学习对于金融时间序列数据能够取得更好的预测效果,并将预测结果与传统方法进行了对比。Fischer和Krauss(2018)通过标准普尔500指数对比发现,循环神经网络中的长短时记忆模型对于金融时间序列数据具备更好的适应性,比随机森林、前馈神经网络等模型的预测准确率更高,涨跌准确率可以达到54.3%。杨青和王晨蔚(2019)将长短时记忆模型应用于全球30种股票指数,发现其在不同的预测期限上均较支持向量机、多层感知机以及ARIMA模型有更高的平均预测精度,且预测结果更为稳定。欧阳红兵等(2020)将小波分析与长短时记忆模型相结合,对道琼斯工业指数日收盘价进行预测,发现其预测精度比多层感知机、支持向量机、K近邻以及GARCH四种模型更高。深度学习在金融时间序列数据中的应用优势可以归纳为以下三点:一是不受维度限制,可以将相关数据都作为输入数据纳入模型之中;二是具备更好的非线性拟合能力,更适应于金融时间序列数据的自身特点;三是能够有效减少过拟合以及陷入局部最优解的问题(Heaton et al.,2016)。

在深度学习常用的方法中,不同模型对于金融时间序列数据的适应性也会存在差异。乔若羽(2019)比较了递归神经网络、长短时记忆模型、门控循环单元模型以及卷积神经网络等多种深度学习方法对股票价格的预测效果,发现长短时记忆模型和门控循环单元模型的预测精度更高。Di Persio和Honchar(2017)分别应用递归神经网络、长短时记忆模型、门控循环单元模型对谷歌公司股价趋势进行预测,同样发现长短时记忆模型预测效果更好。潘水洋等(2019)则从神经网络深度的角度,考察了不同层数深度前馈神经网络对股票价格的预测能力,

通过比较发现过浅或过深的神经网络预测均会降低其准确度。

综合以上文献,深度学习由于对复杂的非线性问题具备强大的拟合能力而得到了金融行业的高度关注,并已经在金融市场的预测效果上体现出了一定的优势。但总体而言,当前金融学术界对于深度学习的研究还较为不足,尤其是国内针对金融市场的研究更为缺乏(苏治等,2017)。一是绝大多数研究仅使用单一的深度学习模型,而复合模型在组成的单模型均较为精确的情况下可以获得更好的预测效果(Krogh和Vedelsby,1995)。二是多数研究预测的对象是股票价格的数值而非股票价格的涨跌,在这样的预测目标下,模型学习出的预测结果往往是在前一交易日收盘价附近,由于大多数时候股票价格的波动较小,这使得模型预测的结果看上去误差较小,但是转化为涨跌准确率却仅在50%左右,预测结果不具有实用性。三是目前的研究忽视了不同时间尺度下股票市场的相关数据,没有充分提取可用的有效信息。由此,本文使用长短时记忆模型与卷积神经网络组成的融合模型,同时将日K线数据与15分钟K线数据输入模型,对后一交易日沪深300指数的涨跌方向进行预测。

三、研究方法与模型

(一)卷积神经网络

卷积神经网络借鉴自生物的视觉系统,是一种非完全连接的带有卷积结构的深度神经网络,各层神经元之间进行局部连接以实现对输入的分层特征的提取和转换,将拥有相同连接权重的神经元连接到上一层神经网络的不同区域,进而简化网络参数并使得网络具有一定程度的位移、尺度、缩放以及非线性形变的稳定性。卷积神经网络的结构如图1所示,除了输入层和输出层之外,还包含卷积层、池化层和全连接层。其中卷积层通过权值共享的卷积核对输入空间进行特征提取,具体过程可以表示为:

$$h_{j,k} = \sigma(b + \sum_{l=0}^L \sum_{m=0}^M w_{l,m} \alpha_{j+l,k+m}) \quad (1)$$

其中, σ 表示激活函数, b 表示共享的偏置参数, L 和 M 分别表示局部感受野的长和宽, $w_{l,m}$ 表示共享的权重参数, $\alpha_{j+l,k+m}$ 表示卷积层接受的输入矩阵对应的数据。共享权重值意味着该卷积层每一个神经元都在对输入数据进行同样的特征提取,不同神经元只是提取不同位置的数据。因此,一个过滤器只能对一种特征进行提取,考虑到输入的数据实际上有多种特征,一般在每一次卷积操作中使用

多种不同的过滤器提取不同的特征。

经过卷积层提取特征后,卷积神经网络的下一个步骤是池化。池化层通常与卷积层成对出现,作用是对卷积层得到的输出进行降采样,简化其中的信息,增强模型的泛化能力。池化的常见方法主要有三种,分别是最大池化、平均池化和L2-norm池化。在最大池化中一个神经元将对应区域数据中的最大值取作该神经元的输出;在平均池化中一个神经元将对应区域数据中的平均值取作该神经元的输出;在L2-norm池化中一个神经元将对应区域数据的平方和的平方根取作该神经元的输出。经过几组卷积层与池化层的特征提取与抽象后,再进一步通过传统的全连接层对之前得到的输出进行加权和分类。

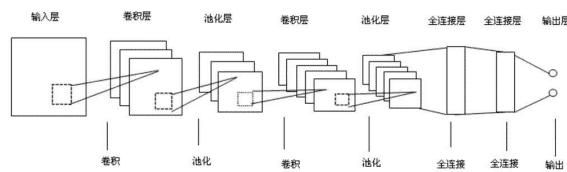


图1 卷积神经网络结构示意图

(二)长短时记忆模型

在传统循环神经网络中,采用基于时间的反向传播算法(Back Propagation Trough Time, BPTT)进行训练时,随着递归层数的增加会出现梯度消失的问题,这会严重影响模型效果,而长短时记忆模型(LSTM)则从神经单元的结构上对此进行了优化。与传统循环神经网络类似,LSTM也具有链式模块结构,但重复模块的组成结构不同。如图2所示,LSTM通过在记忆单元内部设计了遗忘门(Forget Gate)、输入门(Input Gate)以及输出门(Output Gate)三种门限来实现选择性遗忘和记忆信息的功能,并且在记忆单元中增加了细胞状态(C_t)来保存序列中的长期信息,实现选择性遗忘和记忆信息的功能就是体现在对细胞状态的修改中。

LSTM在处理时序数据时,首先通过遗忘门决定丢掉细胞状态的哪些信息。根据 $h(t-1)$ 和 $x(t)$,遗忘门向细胞状态 C_{t-1} 输出一个元素介于0到1之间的系数矩阵,0表示“完全丢弃”,1表示“完全接受”。数学表达式为:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (2)$$

其中, W_f 和 b_f 分别表示遗忘门的权重和偏置参数。下一步就是要确定需要在细胞状态中保存哪些新生信息。输入门分成两部分:第一部分通过将 h_{t-1} 与 x_t 的信息经过一个sigmoid激活函数得到输

出 i_t 来确定哪些信息需要更新进细胞状态;第二部分通过将 h_{t-1} 与 x_t 的信息经过一个tanh激活函数来创建一个新的备选值向量 \tilde{C}_t ,可以用来添加到细胞状态。 i_t 与 \tilde{C}_t 相乘后得到的结果用于更新细胞状态,其数学表达式为:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (4)$$

其中, W_i 和 W_c 分别代表sigmoid激活函数层和tanh激活函数层的权重参数, b_i 和 b_c 分别代表sigmoid激活函数层和tanh激活函数层的偏置参数。原细胞状态 C_{t-1} 通过遗忘门和输入门得到更新,具体过程为将原细胞状态 C_{t-1} 乘以遗忘门输出 f_t ,对原信息进行筛选和过滤,然后加上输入门输出 i_t 乘以备选值向量 \tilde{C}_t 的乘积,可以理解为将部分新生信息加入到遗忘后的原细胞状态中,经过以上过程得到新的细胞状态 C_t ,其数学表达式为:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

经过遗忘门和输入门对细胞状态进行更新以后,下一步即为通过输出门决定该时刻记忆单元要输出的信息。首先将 h_{t-1} 与 x_t 的信息经过sigmoid激活函数层,以此决定输出细胞状态中的哪些部分,然后将细胞状态的值通过tanh变换调整至 $(-1, 1)$ 之间,将调整后的细胞状态乘以以前一个sigmoid激活函数层的输出即得到该时刻记忆单元最终的输出。其数学表达式为:

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = o_t * \tanh(C_t) \quad (7)$$

其中, W_o 代表sigmoid激活函数层的权重参数; b_o 代表sigmoid激活函数层的偏置参数; o_t 代表sigmoid激活函数层的输出; h_t 代表该时刻记忆单元的输出,即隐含状态。

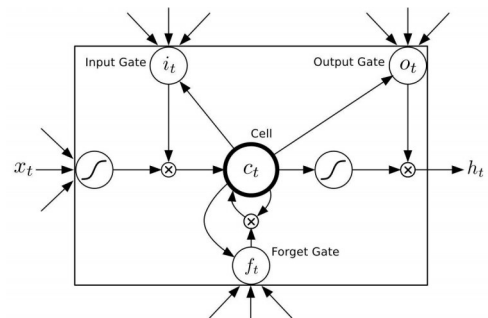


图2 LSTM结构示意图

四、实证分析

(一)数据来源与预处理

本文针对沪深300指数进行下一交易日涨跌方向的预测,选取了2012年1月1日至2019年9月30日的日K线与15分钟K线的开盘价、最高价、最低价、收盘价以及成交量相比前一交易日或者15分钟的涨跌幅数据。数据来源为wind数据库与股票交易通达信软件。其中,日K线时间尺度上共有1883个交易日,15分钟K线时间尺度上共有30129个分钟K线。在日K线时间尺度上使用的时间步长为20,即利用前20个交易日的价格信息预测后一天股价的涨跌方向;在15分钟K线时间尺度上使用的时间步长为16,即利用前一交易日的16个15分钟K线的价格信息预测后一天股价的涨跌方向。因此,本实验中共有1862个可用样本。将其中的80%作为训练集,20%作为测试集,则训练集样本数为1489个,测试集样本数为373个。对于获取的股票数据,本文通过Z-score方法对数据进行预处理。Z-score方法是一种中心化方法,这种方法基于原始数据的均值和标准差对原始数据进行标准化处理,数学表达式为:

$$\bar{x}_i = \frac{x_i - \text{mean}(x)}{\text{std}(x)} \quad (8)$$

其中, x_i 为原始股票价格数据, $\text{mean}(x)$ 为股票数据的均值, $\text{std}(x)$ 为股票数据的标准差, \bar{x}_i 为标准化处理后的股票价格数据。

(二)评价指标

鉴于本文是对股票指数的涨跌进行预测,因此选取在分类任务中经常使用的准确率(Accuracy)、精确率(Precision)、召回率(Recall)、F1值(F1 score)以及曲线下面积(Area Under Curve, AUC)作为预测效果的评价指标。其中,准确率反映的是总样本中预测正确的样本比例;精确率反映的是预测上涨的样本中实际上涨的样本比例;召回率反映的是实际上涨的样本中预测上涨的样本比例;F1值是对精确率和召回率的调和平均值;AUC是接收者操作特征曲线(receiver operating characteristic, ROC)下与坐标轴围成的面积,表示的是随机抽取一个正样本与一个负样本,正样本预测值大于负样本的概率,该指标能够避免因样本不均衡导致的评价失真。

(三)基准模型

为了更好地评价本文提出的多时间尺度CNN-LSTM神经网络的预测效果,本文分别利用BP神经网络、LSTM以及CNN三种模型进行预测,以作为基

准模型。

1. BP神经网络

在应用神经网络进行学习任务时,一个很重要的工作就是寻找合适的网络结构,即隐含层层数以及每一层的神经元个数等超参数,不同超参数的选择对于机器学习任务结果的好坏有非常大的影响,然而目前学术界对于超参数如何选择尚未得到理论性的依据,在实际应用中通常是反复尝试不同的网络结构,并从中选出结果最好的一组超参数。本节在BP神经网络隐含层使用ReLU激活函数,在输出层使用sigmoid激活函数,在迭代优化过程中使用Adam方法,分别尝试隐含层层数为两层、三层、四层与五层,隐含层神经元个数为50、100、150以及200的超参数,从而寻找到一个较为合适的BP神经网络结构。表1所示为应用不同结构的BP神经网络对沪深300指数进行预测的准确率结果,最优组合为三层隐含层且每一层隐含层的神经元个数为100个,测试集最高预测准确率为54.67%。

表1 不同网络结构BP神经网络的预测准确率

	2个隐含层	3个隐含层	4个隐含层	5个隐含层
50个神经元	53.33%	53.87%	53.33%	53.33%
100个神经元	53.87%	54.67%	53.33%	53.07%
150个神经元	53.60%	53.07%	54.13%	53.87%
200个神经元	53.60%	52.80%	54.40%	54.40%

2. LSTM神经网络

在LSTM神经网络中,影响预测结果的主要网络结构包括时间步长、循环层深度以及每个循环层的神经元个数等。本节在LSTM模型中使用Adam优化方法,学习速率为0.001,隐含层激活函数为ReLU函数,输出层激活函数使用sigmoid函数。对20天时间步长下不同的循环层数与神经元个数进行比较,得到的测试集预测结果如表2所示,在单循环层100个神经元的网络结构下可以达到57.64%的预测准确率。

表2 不同网络结构LSTM模型的预测准确率

	单循环层	2个循环层	3个循环层
200个神经元	54.42%	56.30%	55.23%
150个神经元	56.57%	55.77%	55.23%
100个神经元	57.11%	56.84%	54.16%
50个神经元	56.03%	55.50%	54.16%

3. 卷积神经网络

目前常用的卷积神经网络主要是二维卷积神经网络,其通过二维卷积核可以有效提取数据中的空间关系和特征,但是一般的金融时间序列数据的存储结构中并不存在有意义的二维空间关系。一维卷积神经网络与二维卷积神经网络的主要区别在于卷积核的尺寸与窗口滑动方式,在二维卷积神经网络中,卷积窗口会按照横向以及纵向两个方向移动提取输入数据记性特征。在一维卷积神经网络中,通过将卷积核的尺寸设为与输入数据的尺寸同宽,从而卷积窗口仅做纵向的滑动。这种网络结构主要用于提取输入数据在单个空间方向上的平移特征,因而对于金融时间序列数据可能具有更好的适用性。因此,本文分别使用二维卷积神经网络与一维卷积神经网络两种模型在多种网络结构下进行预测对比。结果如表3所示,括号内为卷积核的尺寸。

表3 不同网络结构CNN模型的预测准确率

二维CNN (2*2)	二维CNN (3*3)	二维CNN (4*4)	二维CNN (5*5)	一维CNN (5*1)
54.42%	53.89%	55.76%	53.62%	56.57%

(四) 多时间尺度复合深度神经网络

1. 模型建立

为了更好地应用深度学习方法对股票价格进行预测,本文利用同时提取长短时记忆模型与卷积神经网络的特征优势,将这两种深度学习模型融合进行预测,以更好地处理和分析金融时间序列数据,并进一步通过对日K线与15分钟K线数据分别提取特征进行特征矩阵拼接,从而同时提取股票数据中的中长期信息与短期信息,利用不同时间尺度的特征取得更好的预测效果。

本文使用的多时间尺度复合深度神经网络包括两种模型,即多时间尺度CNN-LSTM神经网络与多时间尺度LSTM-CNN神经网络。在前文基准模型的对比分析中,一维卷积神经网络和单循环层100个神经元的LSTM模型预测表现较好。此处以这两种模型结构组成的多时间尺度CNN-LSTM神经网络为例,详细介绍其具体的数据处理和特征学习过程。如图3所示,该模型首先构造一个接收日K线时间尺度数据的模型Model1与一个接收15分钟K线时间尺度数据的模型Model2。在Model1与Model2中,输入的数据会先经过一个卷积层,利用一维卷积核提取输入数据提取出一级特征,再将卷积层的输出数据输入到循环层中提取出二级特征。然后将

两个时间尺度下得到的二级特征合并后输入到全连接层,经过Dropout层处理后输入到输出层,得到最终的预测结果。具体数据处理和特征学习过程为:

(1)输入层。日K线时间尺度的输入为经过z-score处理过的开盘价、最高价、最低价、收盘价以及成交量较前一交易日的变动幅度。每一个样本包含20个交易日的上述数据,因此输入数据的单个样本为20行5列的矩阵。15分钟K线时间尺度的输入为经过z-score处理过的开盘价、最高价、最低价、收盘价以及成交量较前一个15分钟的变动幅度。每一个样本包含16个15分钟K线的上述数据,因此输入数据的单个样本为16行5列的矩阵。

(2)卷积层。以尺寸为5的一维卷积核分别对两个时间尺度的输入数据进行卷积处理,卷积核个数为16,即每一个卷积核输入空间按照1*5的窗口大小向下滑动提取特征,共提取16种不同的特征。该层单个样本在两个时间尺度上的输出分别是维度为(20,16)和(16,16)的矩阵。

(3)循环层。将卷积层的输出作为该层的输入,即输入数据在两个时间尺度上分别为时间步长度为20、特征维度为16的数据结构,以及时间步长度为16、特征维度为16的数据结构,均使用一个包含100个神经元的LSTM隐含层对输入数据分别进行处理,该层单个样本在两个时间尺度上的输出均是一个维度为(1,100)的矩阵。

(4)全连接层。将循环层在两个时间尺度上的输出合并后,该全连接层单个样本的输入数据结构为(1,200)的矩阵,再使用含有20个神经元的全连接层对上一层合并后输出的数据进行处理,该层单个样本的输出是一个维度为(1,20)的矩阵。

(5)Dropout层。Dropout层的作用为在每一次

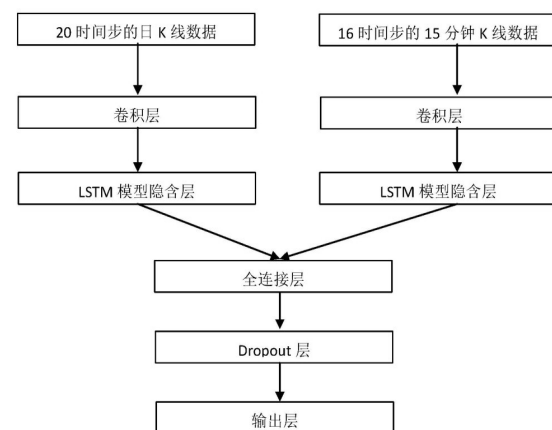


图3 多时间尺度CNN-LSTM神经网络结构

迭代训练中随机地从神经网络中冻结一部分隐含层的神经元,并通过这样的方式降低神经元之间的关联性以及模型的复杂度,从而达到正则化的效果。本模型中使用0.2的Dropout比例。

(6)输出层。使用仅含有1个神经元的全连接层对上一层的输出进行处理,得到最终的输出,即模型的预测结果。该输出为沪深300指数下一交易日上涨的概率。当上涨概率超过0.5时即认为预测的结果为上涨,反之则视为下跌。

2.神经网络模型与结构的比较

多时间尺度复合神经网络在不同网络结构下的预测准确率如表4所示,其中N1、N2、N3分别表示神经元个数为50、100以及200。可以看出,多时间尺度CNN-LSTM的预测准确率在多种模型结构下均高于多时间尺度LSTM-CNN,这表明多时间尺度CNN-LSTM更适应于金融时间序列数据的分析预测任务。当使用含有100个神经元的单层循环层时,多时间尺度CNN-LSTM的预测准确率达到60.86%,预测效果优于其他网络结构。其ROC曲线如图4所示,横坐标轴为假正类率(False Positive Rate, FPR),纵坐标轴为真正类率(True Positive Rate, TPR),曲线下面积即AUC值为0.612。以上结果表明模型在涨跌两种情况下预测效果均表现良好且较为平衡。本文的预测对象是沪深300指数后一交易日的涨跌方向,为了更全面地反映多时间尺度CNN-LSTM对股票价格数据的预测效果,本文同时利用该模型对沪深300指数的数值进行了预测。预测结果与实际值的对比如图5所示。

表4 多时间尺度复合神经网络预测准确率

	单层循环层			两层循环层		
	N1	N2	N3	N1	N2	N3
多尺度CNN-LSTM	60.32%	60.86%	58.71%	58.45%	59.51%	58.18%
多尺度LSTM-CNN	57.11%	56.57%	57.91%	57.91%	58.45%	57.91%

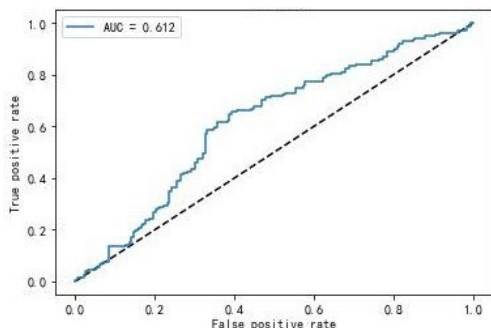


图4 多时间尺度CNN-LSTM的ROC曲线

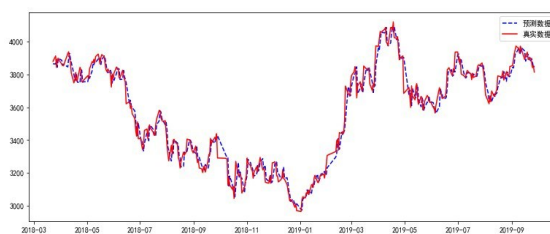


图5 多时间尺度CNN-LSTM的预测效果

为了更好地衡量多时间尺度CNN-LSTM神经网络的预测效果,本文将其预测结果的评价指标与BP神经网络、LSTM神经网络、CNN神经网络以及复合深度神经网络进行比较分析。如表5所示,深度神经网络的预测效果好于BP神经网络,复合深度神经网络的预测效果好于单模型的深度神经网络,而通过将多时间尺度的数据在复合深度神经网络上进行融合又可以进一步提高预测的效果。多时间尺度CNN-LSTM模型相比浅层BP神经网络的预测准确率提高了6.19%,相比目前在金融时间序列数据中应用最广的LSTM模型,其预测准确率提高了3.75%。其他评价指标同样有较为明显的提升。

表5 各模型的预测效果比较

	Accuracy	Precision	Recall	F1-score	AUC
BP	54.67%	55.65%	54.67%	54.46%	0.552
二维CNN	55.76%	55.62%	55.76%	55.60%	0.543
一维CNN	56.57%	56.59%	56.57%	56.58%	0.538
LSTM	57.11%	57.65%	57.10%	57.10%	0.577
LSTM-CNN	56.84%	57.45%	56.84%	56.69%	0.554
CNN-LSTM	57.91%	57.92%	57.91%	57.91%	0.561
多尺度LSTM-CNN	59.25%	59.60%	59.25%	59.29%	0.587
多尺度CNN-LSTM	60.86%	61.50%	60.86%	60.83%	0.612

接下来通过多时间尺度CNN-LSTM神经网络以及其他模型进行交易回测,进一步分析其预测效果和盈利能力。由于本文主要针对股票市场进行预测,因此在回测中不进行做空交易。出于模型计算耗时和实际交易可行性的考虑,买卖指令均在后一天以开盘价执行。因此使用的交易策略为当第 t 日预测后一交易日标的股票将上涨,若此时持有现金,则在第 $t+1$ 日以开盘价买入标的股票,若此时持有股票,则继续持有。当预测后一交易日标的股票将下跌,若此时持有股票,则在第 $t+1$ 日以开盘价卖出标的股票,若此时不持有股票,则继续持有现金。

当不考虑交易费用时,各模型的交易回测的结果如图6所示,在回测期间内沪深300指数从

4020.35 下降到 3852.65, 累计跌幅为 4.17%; 而基于多时间尺度 CNN-LSTM 神经网络预测结果进行交易则可以获得 22.47% 的盈利, 超出沪深 300 同期收益率的 27.80%, 其年化收益率为 14.47%。回测期间内资金净值最大回撤比例为 10.58%, 夏普比率达到 0.79。当计算交易成本时, 假设交易佣金为万分之三, 印花税为千分之一, 则回测期间收益比例为 14.75%, 年化收益率为 9.61%, 夏普比率为 0.48。以上回测结果表明利用多时间尺度 CNN-LSTM 神经网络进行交易的投资收益高于其他模型结构。

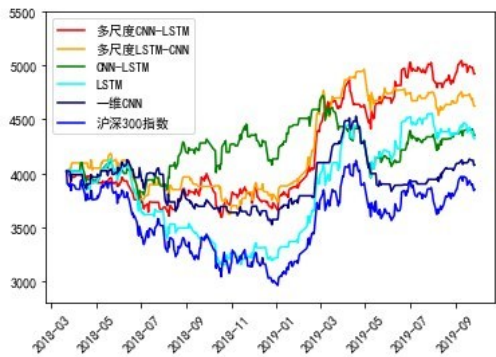


图6 回测结果

3. 数据时间长度对预测结果的影响

在本文多时间尺度 CNN-LSTM 神经网络的数据处理与建模过程中, 输入数据的时间长度是在日 K 线尺度上使用 20 个时间步的数据, 在 15 分钟 K 线上使用 16 个时间步的数据。数据时间长度对于模型的预测结果有着重要影响, 如果时间长度过短, 则输入数据不能包含足够多的有效信息。如果数据长度过长, 则一方面由于较久远的股票信息对于未来股价的影响微小而造成数据冗余, 另一方面因循环神经网络时间步数过长造成梯度消失效应更严重, 从而降低其预测效果。

本小节分别在日 K 线尺度上使用 5、10、20 以及 40 个时间步, 在 15 分钟 K 线尺度上使用 16 与 32 个时间步, 通过多种时间步组合下准确率 (Accuracy)、曲线下面积 (Area Under Curve, AUC) 以及年化收益率 (Annualized Rate Of Return, ARR) 的比较, 来分析数据时间长度对于预测效果的影响。如表 6 所示, 可以看出在日 K 线尺度上使用 20 个时间步的预测效果好于其他时间步数目, 在 15 分钟 K 线尺度上使用 16 个时间步的预测效果好于 32 个时间步。这一结果表明前 20 天的日 K 线数据以及前一天的分钟级数据对于预测后一天股价涨跌有较大的帮助。

表6 不同时间长度下的模型预测效果比较

	16 个时间步 (分钟)			32 个时间步 (分钟)		
	Accuracy	AUC	ARR	Accuracy	AUC	ARR
5 个时间步 (日)	58.51%	0.573	5.27%	56.65%	0.560	2.15%
10 个时间步 (日)	58.13%	0.578	9.38%	56.53%	0.568	2.63%
20 个时间步 (日)	60.86%	0.612	14.47%	58.45%	0.571	8.21%
40 个时间步 (日)	58.00%	0.587	12.65%	58.27%	0.583	9.86%

五、结语

股票市场价格的预测对于公司经营决策、投资者交易策略以及政府监管政策都具有重要影响, 然而股票价格具有非线性、非平稳性、低信噪比以及长记忆性等特点, 长期以来一直难以取得令人满意的预测效果。本文以沪深 300 指数为例, 运用多时间尺度复合深度神经网络进行涨跌预测, 预测准确率达到了 60.86%, 当不考虑交易成本时, 在回测中获得的年化收益率为 14.47%, 夏普比率为 0.79。本文研究得到以下结论: 第一, 常用的深度学习方法 CNN 与 LSTM 对于股价涨跌的预测准确率均高于传统的浅层神经网络。其中, 一维卷积神经网络的预测准确率高于二维卷积神经网络, 而二维卷积神经网络的特征提取方式是基于数据空间位置关系, 提取输入数据中的多种空间特征, 这一结果表明金融时间序列数据中不存在明显的空间特征。以时间序列关系提取特征的 LSTM 模型的预测效果好于卷积神经网络, 这与股价数据的时间序列性质相一致。第二, 复合深度神经网络的预测效果好于单独使用 LSTM 或者 CNN 模型。这一结果表明, 复合深度神经网络可以结合不同深度学习模型提取特征的优势, 从不同的算法逻辑上更全面地获取金融时间序列数据中蕴含的特征和规律。第三, 股价数据在不同时间尺度上包含不同的信息, 通过同时对日 K 线与 15 分钟 K 线尺度上的信息进行特征提取可以进一步改善预测效果。

深度学习在处理高维非线性系统上具有巨大的优势。作为一种非参数分析方法, 深度学习不需要假设分析对象符合特定的方程结构或者分布函数, 因此其对于金融时间序列数据有较好的适用性和研究前景。本文的研究丰富了金融时间序列数据分析的方法, 不同于传统计量方法关注模型参数的显著性, 深度学习方法更关注模型结构、特征提取以及拟合效果。■

[参考文献]

- [1] SCHMIDHUBER J. Deep Learning in Neural Networks: An Overview[J]. *Neural Networks*, 2015, 61: 85-117.
- [2] NAJAFABADI M M, VILLANUSTRE F, KHOSH-GOFTAAR T M, et al. Deep Learning Applications and Challenges in Big Data Analytics[J]. *Journal of Big Data*, 2015, 2(1): 1-21.
- [3] FAMA E F. The Behavior of Stock-Market Prices[J]. *The Journal of Business*, 1965, 38(1): 34-105.
- [4] FAMA E F. Efficient Capital Markets: A Review of Theory and Empirical Work[J]. *The Journal of Finance*, 1970, 25(2): 383-417.
- [5] GROSSMAN S J. On the Efficiency of Competitive Stock Markets where Traders Have Diverse Information[J]. *Journal of Finance*, 1976, 31(2): 573-585.
- [6] GROSSMAN S J. An Introduction to the Theory of Rational Expectations under Asymmetric Information[J]. *Review of Economic Studies*, 1981, 48(4): 541-559.
- [7] ALLEN B. Generic Existence of Equilibria for Economics with Uncertainty When Prices Convey Information[J]. *Econometrica*, 1981, 49(5): 1173-1199.
- [8] JORDAN J S. Convergence to Rational Expectations in a Stationary Linear Game[J]. *Review of Economic Studies*, 1992, 59(1): 109-123.
- [9] FAMA E F. Efficient Capital Markets: II [J]. *Journal of Finance*, 1991, 46(5): 1575-1617.
- [10] BARBER B M, ODEAN T. Boys Will be Boys: Gender, Overconfidence and Common Stock Investment[J]. *Quarterly Journal of Economics*, 2001, 116(1): 261-292.
- [11] SHEFRIN H, STATMAN M. The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence[J]. *Journal of Finance*, 1985, 40(3): 777-790.
- [12] ODEAN T. Are Investors Reluctant to Realize Their Losses?[J]. *Journal of Finance*, 1998, 53(5): 1775-1798.
- [13] KAHNEMAN D, TVERSKY A. Prospect Theory: An Analysis of Decision Under Risk[J]. *Econometrica*, 1979, 47(2): 263-292.
- [14] SHLEIFER A, VISHNY R W. The Limits of Arbitrage[J]. *Journal of Finance*, 1997, 52(1): 35-55.
- [15] GROSSMAN S J, STIGLITZ J E. On the Impossibility of Informationally Efficient Markets[J]. *The American Economic Review*, 1980, 70(3): 393-408.
- [16] 陈灯塔, 洪永淼. 中国股市是弱式有效的吗: 基于一种新方法的实证研究[J]. *经济学*, 2003(1): 97-124.
- [17] FRENCH K R. Stock Returns and the Weekend Effect[J]. *Journal of Financial Economics*, 1980, 8(1): 55-69.
- [18] GULTEKIN M N, GULTEKIN N B. Stock Market Seasonality: International Evidence[J]. *Journal of Financial Economics*, 1983, 12(4): 469-481.
- [19] DE BONDT W, THALER R. Does the Stock Market Overreact?[J]. *Journal of Finance*, 1985, 40(3): 793-808.
- [20] JEGADEESH N, TITMAN S. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency[J]. *Journal of Finance*, 1993, 48(1): 65-91.
- [21] KEIM D B, STAMBAUGH R F. Predicting Returns in the Stock and Bond Markets[J]. *Journal of Financial Economics*, 1986, 17(2): 357-390.
- [22] CAMPBELL J Y, SHILLER R J. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors[J]. *The Review of Financial Studies*, 1988, 1(3): 195-228.
- [23] CAMPBELL J Y, COCHRANE J H. By Force of Habit: A Consumption - Based Explanation of Aggregate Stock Market Behavior[J]. *Journal of Political Economy*, 1999, 107(2): 205-251.
- [24] CAMPBELL J Y, LO A W, MACKINLAY A C. *The Econometrics of Financial Markets*[M]. Princeton University Press, 1996.
- [25] LO A W. The Adaptive Markets Hypothesis[J]. *The Journal of Portfolio Management*, 2004, 30(5): 15-29.
- [26] LO A W. Reconciling Efficient Markets with Behavioral Finance: The Adaptive Markets Hypothesis[J]. *Journal of Investment and Consulting*, 2005, 7(2): 21-44.
- [27] 张贵生, 张信东. 基于近邻互信息的SVM-GARCH股票价格预测模型研究[J]. *中国管理科学*, 2016(9): 11-20.
- [28] 苏治, 傅晓媛. 核主成分遗传算法与SVR选股模型改进[J]. *统计研究*, 2013(5): 54-62.
- [29] 王淑燕, 曹正凤, 陈铭芷. 随机森林在量化选股中的应用研究[J]. *运筹与管理*, 2016(3): 163-168+177.
- [30] 曹正凤, 纪宏, 谢邦昌. 使用随机森林算法实现优质股票的选择[J]. *首都经济贸易大学学报*, 2014(2): 21-27.
- [31] BENGIO Y, LECUN Y. Scaling Learning Algorithms Towards AI[J]. *Large-Scale Kernel Machines*, 2007, 34(5): 1-41.
- [32] LANGKVIST M J, KARLSSON L, LOUTFI A. A Review of Unsupervised Feature Learning and Deep Learning for Time-series Modeling[J]. *Pattern Recognition Letters*, 2014, 42(1): 11-24.
- [33] HORNIK K, STINCHCOMBE M, WHITE H. Multilayer Feedforward Networks are Universal Approximators[J]. *Neural Networks*, 1989, 2(5): 359-366.
- [34] FISCHER T, KRAUSS C. Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions[J]. *European Journal of Operational Research*, 2018, 270(2): 654-669.
- [35] 杨青, 王晨蔚. 基于深度学习LSTM神经网络的全球股票指数预测研究[J]. *统计研究*, 2019(3): 65-77.
- [36] 欧阳红兵, 黄亢, 闫洪举. 基于LSTM神经网络的金融时间序列预测[J]. *中国管理科学*, 2020(4): 27-35.
- [37] HEATON J, POLSON N, WITTE J. Deep Learning for Finance: Deep Portfolios[J]. *Applied Stochastic Models in Business and Industry*, 2016, 33(1): 3-12.
- [38] 乔若羽. 基于神经网络的股票预测模型[J]. *运筹与管理*, 2019(10): 132-140.
- [39] DI PERSIO L, HONCHAR O. Recurrent Neural Networks Approach to the Financial Forecast of Google Assets[J]. *International Journal of Mathematics and Computers in Simulation*, 2017(11): 7-13.
- [40] 潘水洋, 刘俊玮, 王一鸣. 基于神经网络的股票收益率预测研究[J]. *浙江大学学报(理学版)*, 2019(5): 550-555.
- [41] 苏治, 卢曼, 李德轩. 深度学习的金融实证应用: 动态、贡献与展望[J]. *金融研究*, 2017(5): 111-126.
- [42] KROGH A, VEDELSBY J. Neural Network Ensembles, Cross Validation and Active Learning[J]. *Neural Computing and Application*, 1995, 25(7): 231-238.

(责任编辑:GW / 校对:XY)