

基于注意力的卷积神经网络金融时序数据预测

孙启森¹, 张建新², 程海阳¹, 张强^{1,3*}, 魏小鹏^{1,3}

(1. 先进设计与智能计算省部共建教育部重点实验室(大连大学), 辽宁 大连 116622;

2. 大连民族大学 计算机科学与工程学院, 辽宁 大连 116600; 3. 大连理工大学 计算机科学与技术学院, 辽宁 大连 116024)

(* 通信作者电子邮箱 zhangq@dlu.edu.cn)

摘要:针对现有金融时序数据预测方法在构造金融特征图像的过程中因忽视市场环境变化导致的数据密度分布差异问题,提出一种基于滑动窗口标准化的金融数据预处理方法。所提方法将滑动窗口截取的数据使用独立的标准化转换为金融特征图像,使得依赖价格特征进行训练的卷积神经网络(CNN)模型能够学习到正确的映射关系;同时,针对金融特征图像的特征表达问题,为更好地捕捉其动态变化特征,将注意力机制引入CNN中,进而构建出一种注意力CNN金融时序数据预测模型。对标普500指数未来1天涨跌进行预测的准确率和F1分数分别为61%和0.7397,模拟交易实验投资回报率为23.04%,优于买入并持有策略。此外,消融实验结果也证明了预处理方法、注意力模块引入的有效性。

关键词:卷积神经网络;核密度估计;注意力机制;时间序列;股票指数预测

中图分类号: TP391 文献标志码: A

Financial time series data prediction by attention-based convolutional neural network

SUN Qisen¹, ZHANG Jianxin², CHENG Haiyang¹, ZHANG Qiang^{1,3*}, WEI Xiaopeng^{1,3}

(1. Key laboratory of Advanced Design and Intelligent Computing, Ministry of Education (Dalian University),

Dalian Liaoning 116622, China;

2. School of Computer Science and Engineering, Dalian Minzu University, Dalian Liaoning 116600, China;

3. School of Computer Science and Technology, Dalian University of Technology, Dalian Liaoning 116024, China)

Abstract: To address the problem that existing financial time series data prediction methods ignore the differences in data density distribution caused by market environment changes in the process of constructing financial feature images, a financial data preprocessing method based on sliding window standardization was proposed. The sliding window intercepted data was converted into financial feature images by independent standardization, so that the Convolutional Neural Network (CNN) model trained on price features can learn the correct mapping relationship. Meanwhile, for the feature representation problem of financial feature images, the attention mechanism was introduced into CNN to better capture the dynamic changing features, and an attentional CNN model for financial time series data prediction was constructed. The accuracy and F1-score of predicting the rise and fall of the S&P 500 index in the next day were 61% and 0.7397, respectively, and the return of investment in the simulated trading experiment was 23.04%, indicating that the proposed method is better than the buy-and-hold strategy. In addition, the ablation experiment results demonstrate the effectiveness of the preprocessing method and the attention module introduction.

Key words: Convolutional Neural Network (CNN); Kernel Density Estimation (KDE); attention mechanism; time series; stock index prediction

0 引言

随着我国金融市场的发展,越来越多的人进入到金融投资领域,金融时间序列分析在对冲市场风险和优化投资决策方面发挥着重要作用。在这种背景下,机器学习技术的快速发展为金融数据的预测提供了新的方式,预测的准确性和实

用性得到了迅速提高。文献[1]中使用了4种常见的机器学习方法来预测收盘价的涨跌,分别是逻辑回归(Logistic Regression, LR)、K最近邻(K-Nearest Neighbor, KNN)、随机森林(Random Forest, RF)和受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)。但是,金融资产的价格运动受到多种不同因素的影响,是复杂的市场状态在不同扩散速度下

收稿日期: 2022-04-21; 修回日期: 2022-06-28; 录用日期: 2022-07-01。

基金项目: 国家自然科学基金辽宁省联合基金资助项目(U1908214); 国家自然科学基金资助项目(61972062)。

作者简介: 孙启森(1997—),男,山东淄博人,硕士研究生,CCF会员,主要研究方向:数据挖掘、时序预测; 张建新(1981—),男,江苏徐州人,教授,博士,CCF会员,主要研究方向:医疗影像分析、计算机视觉; 程海阳(1998—),男,江西抚州人,硕士研究生,CCF会员,主要研究方向:数据挖掘、时序预测; 张强(1971—),男,陕西西安人,教授,博士,主要研究方向:生物计算与人工智能、智能大数据处理; 魏小鹏(1959—),男,辽宁大连人,教授,博士,主要研究方向:生物计算与人工智能、智能大数据处理。

的反映,传统机器学习方法的性能不足以满足人们的需要,于是深度学习模型开始受到更多的关注。深度学习技术在过去几年中取得了巨大的成功,在包括股市预测在内的任务上表现出比传统线性模型和机器学习模型更好的性能^[2-3]。

考虑到卷积神经网络(Convolutional Neural Network, CNN)模型在计算机视觉领域的出色表现,近年来,越来越多的研究使用CNN模型来预测金融资产的价格。文献[4]中为了验证CNN在股票市场预测中的适用性,与人工神经网络(Artificial Neural Network, ANN)和支持向量机(Support Vector Machine, SVM)的预测精度进行了比较。研究结果表明,CNN模型取得了更好的预测结果。

在最近的一些研究中,不同的独立研究小组创新性地将一维金融时序数据转化为不同类型的图像,通过自适应滤波器和隐式降维来利用CNN的能力并且取得了更高的预测精度。文献[5]中将每日价格序列绘制成连续柱状图形式,并采用滑动窗口方式,每次滑动1 d,滑动窗口大小为30 d截取图像,CNN模型直接使用这些图像来预测交易点。文献[6]中定义了不同的股票图像特征,使用CNN分析了这些图像特征对股票价格预测的影响。

上述这些转换特征图像的方法包含的信息有限,模型能利用的信息很少,因此,研究者开始尝试将更多的金融市场信息融入到模型训练中。文献[7]中将15种反映市场变化的技术指标数据转换成类图像数据,使用CNN预测交易点,并且与其他股票投资策略或模型预测结果进行比较。文献[8]中为了验证使用外部因素(如黄金价格和利率)进行预测的适用性,使用了8种不同的输入特征组合来预测未来股市趋势。此外,还有研究者通过将长短期记忆(Long Short-Term Memory, LSTM)网络与CNN相结合,构造混合模型来综合利用时序数据与图像数据^[9-12]。

然而,目前已有的绝大部分研究都是基于价格数据或者与价格相关的技术指标进行预测,CNN模型试图找到预测目标和价格之间的映射关系,但这些关系可能会受到市场变化的影响^[13]。这是因为从长期来看,价格的变化可能较大,而模型的单次输入只能是一小段时间的价格信息,因此既有研究在应对不断变化的市场环境方面存在问题。为此,在以往研究的基础上,本文提出了一种新的数据预处理方式,从数据特征的密度分布角度出发,对金融特征图像转换的方法进行了改进,使得模型可以在一定程度上适应不断变化的市场环境;同时,为了提高CNN模型的特征提取能力,将注意力机制融入到模型中,并讨论其在金融时序预测领域的适用性。实验结果表明,注意力机制与这种新的预处理方式可以提高以金融特征图像作为输入的CNN模型的预测精度。在模型性能评估上,所提方法在测试集上的表现优于最新的模型,并且在模拟交易实验中收益率超过了买入并持有策略。

1 研究方法

1.1 金融图像数据预处理

在金融时序数据预测领域中,使用金融图像作为输入,利用CNN的能力,可以提高预测精度,但如何将一维金融时序数据转化成二维金融特征图像是一个必须解决的问题。在以往的研究^[4,7-8]中,研究者将指标维度作为横轴,时间维度作为纵轴,构造二维金融特征图像数据。同时,为了消除指标间的量纲差异,使用最大最小标准化或者Z-score标准化对数据整

体进行处理,然后划分训练集、验证集和测试集。然而在时序预测任务中,测试集的时间段要在训练集之后,但是由于金融时间序列数据中某些指标随着时间的推移变化很大,如每日收盘价(Close)、指数移动平均值(Exponential Moving Average, EMA)、异同移动平均线(Moving Average Convergence and Divergence, MACD)等,容易导致训练集和测试集的数据分布有明显差异,这在一定程度上会影响到模型的预测性能

为了解决这个问题,需要保证训练集和测试集数据的密度分布要尽可能地一致,模型的输入既要保留局部变化信息,同时又不会因为输入时间段的不同导致数值产生较大变化。基于以上考虑,本文提出了一种新的预处理方式,即滑动窗口标准化(Sliding Window Standardization, SWS):将滑动窗口截取的二维矩阵数据使用独立的Z-score标准化,使得每一个输入 $X \in \mathbf{R}^{w \times n}$ 转换成一个二维图像 X' ,其中 w 是滑动窗口大小, n 是特征数。每个元素 $x'_{ij} \in X'$ 的计算方法如式(1)所示:

$$x'_{ij} = (x_{ij} - \mu_j) / s_j \quad (1)$$

其中: $x_{ij} \in X$; μ_j 和 s_j 分别是 X 第 j 列的平均值和标准差; $i \in [1, w]$, $j \in [1, n]$ 。在传统的方法中, μ_j 和 s_j 分别是整个数据集的第 j 个特征的平均值和标准差,它们的值是不变的,但金融资产在短期的波动往往相对较小,因此进行标准化之后,局部的变化并不明显。然而,将滑动窗口截取的二维矩阵数据使用独立的标准化后, μ_j 和 s_j 分别是一个滑动窗口内部值的平均值和标准差,它们的值反映金融资产短期内的信息,经过标准化后数值的变化也会更加明显,实际上是增强了金融特征图像局部的动态变化特征,而CNN模型对于局部的变化更加敏感,可以更容易地提取这些变化的特征。同时,每一个输入与所处的市场环境无关,每张图像之间相互独立,它既保留了滑动窗口内部的变化信息,也可以保证数值分布在相似的区间内,CNN通过自适应滤波器来综合利用这些变化的特征,可以更好地对输入进行区分鉴别。这样,对于金融时序数据的预测问题就转换成了对图像的分类问题,而CNN模型在解决该类问题上具有明显的优势。

为比较两种预处理方法,采用核密度估计(Kernel Density Estimation, KDE)对各个指标的分布进行了可视化。KDE是一种用于估计概率密度函数的非参数方法,设其概率密度函数为 f ,独立同分布的 m 个样本点,计算公式如下:

$$f_h(x) = \frac{1}{m} \sum_{i=1}^m K_h(x - x_i) = \frac{1}{mh} \sum_{i=1}^m K\left(\frac{x - x_i}{h}\right) \quad (2)$$

其中: $K(\cdot)$ 为核函数; h 为一个平滑参数,称作带宽。图1给出以文献[8]中所采用的数据预处理方式在整个数据集上进行标准化处理后,训练集和测试集部分指标的数据密度分布情况。采用改进后的预处理方式后,训练集和测试集的数据密度分布如图2所示。图1~2中横坐标表示输入指标的取值 x ,纵坐标表示该指标值对应的核密度值 $f_h(x)$ 。

由图1可以看到,与价格有关的指标如Close、EMA、MACD,在测试集与训练集上的分布差异很大,这是由于测试集所在期间市场环境相较于训练集已经发生了巨大变化,资产的单位价格上涨到了之前从未出现过的高值,而表征市场变化的变动率(Rate of Change, ROC)指标受到的影响相对较小。以上结果说明以这种方式进行处理不能保证输入的一致性,这样即便模型在训练集上效果很好,但在测试集上可能无

法取得同样的效果,因为模型从来没有“见过”这么高的价格,换句话说模型没有正确学习到输入与输出之间的映射关系。

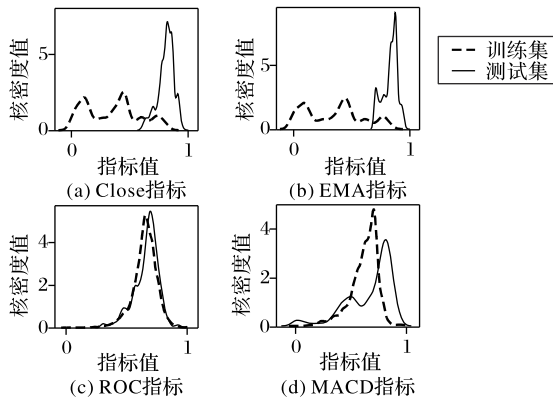


图1 传统预处理方法数据密度分布情况

在图2中可以看到,当采用所提出改进预处理方法后,训练集和测试集的数据密度分布几乎一致,模型可以正确学习到测试集上的映射关系,使得结果更加准确可靠。

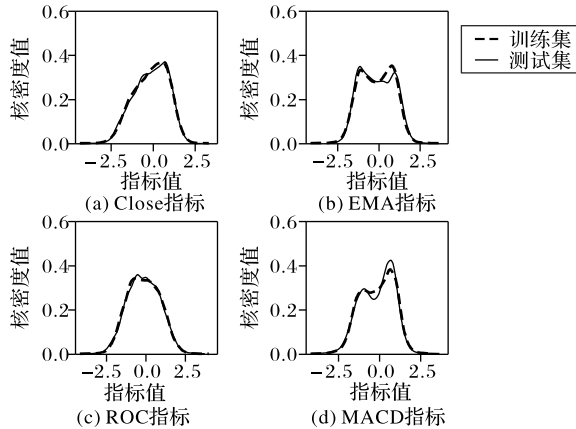


图2 改进后预处理方法数据密度分布情况

1.2 基于注意力的卷积神经网络模型

最近几年基于注意力的模型在深度学习的各个领域被广泛使用,注意力机制已被证明是在深层神经网络中嵌入的一种强大机制^[14]。它的主要作用是在计算高级特征表示时,通过为每个输入分配权重或重要性来筛选对当前任务目标更关键的信息。因此,为了提高模型对于金融图像动态变化特征的提取能力^[15],尝试引入注意力机制来构建神经网络模型。

1.2.1 神经网络模型架构

本文研究所构建的模型结构如图3所示,主要由6部分组成,分别为输入层、两个带有注意力模块的卷积层、最大池化层、完全连接层和输出层。

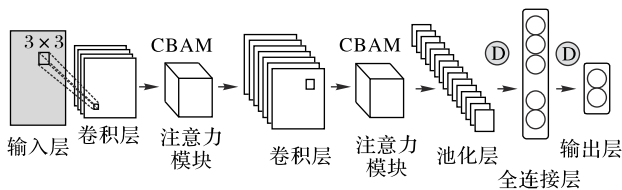


图3 基于注意力的卷积神经网络结构

CNN是一种前馈神经网络,其输入为二维矩阵。CNN的主要特点是能够处理多通道输入数据,因此在本研究中,它非常适合处理具有多维输入和输出的时间序列数据。此外,

CNN还有局部感知和参数共享两大优点。对于局部感知来说,CNN与多层感知机(Multi-Layer Perceptron, MLP)等完全连接的神经网络不同,输入的向量(或矩阵)中数据的局部性非常重要。因此,应仔细选择矩阵中的相邻数据点。对于金融时序数据来讲,经过转化后的特征图像相邻像素点在纵向和横向两个维度上表示的都是连续的金融市场信息,其中所包含的信息更加紧密也更加丰富。因此,相邻像素点在两个方向上彼此相关,直接满足上述要求。而对于参数共享,这可以大幅减少参数的数量,从而提高学习过程的效率。

在卷积层中,卷积核大小的设置尤为重要,选择较小的卷积核大小通常会捕获图像的更多细节。在本研究中,考虑到用作输入的金融特征图像像素点相对较少,故设置卷积核的大小为 3×3 。通过使用 3×3 的卷积核,最近邻信息被用作卷积运算的输入,因此,模型可以注意到图像中的显著强度变化。在卷积层之后使用卷积注意力模块(Convolutional Block Attention Module, CBAM)^[16]强化重要特征,使得模型可以关注到最重要的信息。同时,考虑到数据集的规模较小,模型只构建了两个卷积层,分配的卷积核数量为32和64。在CNN结构中添加更多层会增加模型的复杂性,如果不使用大型数据集训练,这种复杂的网络结构可能会导致过拟合的发生,并降低模型在测试集的准确性^[7]。同时,CBAM的引入虽然会增加一些参数数量和计算量,但由于本文研究所用模型中卷积层较少,对于模型整体来说,参数数量和计算量在本研究中分别增加了0.4%和1.8%,因此该模块的加入对模型整体复杂度的影响较小。

随后,使用最大池化来选择特征图中的最大值。池大小为 2×2 ,压缩图像并保留主要特征的同时减少参数和计算量。为防止过拟合,使用Dropout层去除部分神经元,最后使用全连接层将神经元展平并连接到下一层,输出层对提取的特征进行分类,由于本文解决的是一个二分类问题,故输出层神经元个数为2。输出层使用了Softmax激活函数。

1.2.2 卷积注意力机制模块

CNN在图像和视频识别中的优势在于,不同的卷积层能够提取图像中特征的变化,而注意力机制的加入,可以使得模型更加关注这些变化的特征,进而提高模型的特征提取能力。在本研究中,采用CBAM实现注意力机制,它是一种为CNN设计的并且简单有效的注意力模块。图4给出了CBAM的结构。

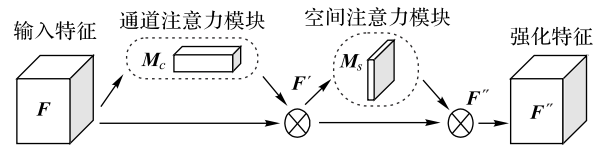


图4 CBAM结构

CBAM有两个子模块:通道注意力模块和空间注意力模块,两个注意力子模块呈现不同的功能。通道注意力模块突出输入特征图像提供的代表性信息,其表达式如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F))) + \text{MLP}(\text{MaxPool}(F)) \quad (3)$$

其中: F 表示输入特征; σ 表示应用的Sigmoid函数;MLP表示多层感知器;AvgPool和MaxPool分别表示平均池化和最大池化。

与通道注意力不同,空间注意力子模块侧重于突出特征图像内容中有代表性贡献的区域,是对通道注意力的补充,其表

达式如下:

$$M_s(F') = \sigma \left(f \left(\left[\text{AvgPool}(F'), \text{MaxPool}(F') \right] \right) \right) \quad (4)$$

其中: f 表示卷积运算。对于从卷积层提取到的特征图, CBAM 依次从通道和空间维度计算特征图的注意力权重, 然后将注意力权重与特征图相乘, 来实现特征的自适应学习。CBAM 使得模型可以关注重要信息, 同时忽视无用信息。图4显示了在神经网络的卷积块之间处理信息流时两个子模块的顺序, 研究结果表明, 这种顺序相较于空间注意力模块在前, 通道注意力模块在后的顺序产生了更好的结果^[16]。

2 实验与结果分析

本章首先介绍了实验数据集的设置情况以及二分类任务的数据标记方法; 其次给出了实验的一些具体细节, 包括参数设置等; 最后, 对模型的预测结果进行了分析讨论。依据模型的预测结果进行模拟交易并计算投资回报率, 以此来评估预测模型的有效性。

2.1 数据集设置与数据标记

本研究的实验数据使用2010年1月1日至2019年12月31日标准普尔500(S&P500)指数(GSPC)的每日交易数据, 包括开盘价(Open)、收盘价(Close)、最高价(High)、最低价(Low)、交易量(Volume), 即OCHLV。其中2010年至2018年的数据按照8:2划分为训练集和验证集, 2019年的数据作为测试集。技术指标数据是由OCHLV计算得出。参照文献[4], 本文选取的技术指标为简单移动平均线(Simple Moving Average, SMA)、EMA、ROC、MACD、Fast%K(快速随机指标)、Slow%D(慢速随机指标)、Upper band(布林线上轨线)、Lower band(布林线下轨线)、%B(价格在布林线中的位置)。

选择S&P500指数的原因是标准普尔指数在金融资本市场上具有举足轻重的地位, 指数通常比个股波动性小, 因为它们由不同行业的多只股票组成, 更能反映整体经济势头和总体状况; 并且, 其代表着相对发达完备的市场环境, 技术分析认为市场包容消化一切信息, 因此模型可以更容易地利用历史数据来预测市场走势。图5给出了在20 d窗口期内10个输入变量绘制的两类特征图像示例。

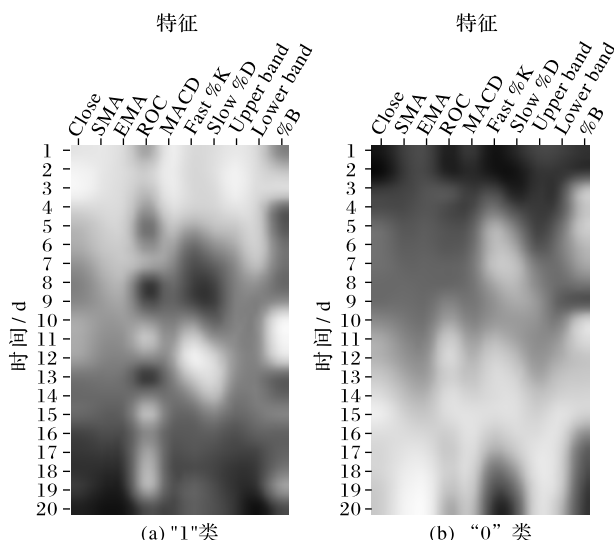


图5 两类金融特征图像示例

如图5所示, 本文选取收盘价以及上述9个技术指标共10个特征作为横轴, 即特征数 $n = 10$ 。纵轴表示时间, 选择20 d为一个周期, 即窗口大小 $w = 20$, 采用SWS的方式截取数据, 这样对于每一个输入 $X \in \mathbf{R}^{w \times n}$, 是一个 20×10 的二维矩阵, 然后将其转换为金融特征图像。每一幅图像包含20 d的市场信息, 将最后一天作为第 T 天, 则每一个输入对应的输出 Y 是根据第 $T + 1$ 天的收盘价涨跌确定的, 若 $T + 1$ 天的收盘价大于第 T 天, 则标记为“1”类(上涨), 否则标记为“0”类(下跌)。

2.2 实验设置

本文使用Adam优化器进行训练, 训练epoch设置为200, 利用earlystop和checkpoint方法避免训练过拟合并保存验证集上最优的模型。实验采用贝叶斯优化(Bayesian Optimization, BO)算法对模型进行超参数优化, 该方法可以有效提高神经网络的预测性能^[8]。BO算法进行优化的参数有批次大小、学习率、两个辍学层的辍学率。实验给出的结果均是使用经过优化后的最优模型进行预测的结果。

2.3 评估指标

本研究使用准确率(Accuracy, A)、精确率(Precision, P)、召回率(Recall, R)和F1分数(F1-Score, $F1$)来评估模型的表现, $F1$ 分数是精确率和召回率之间的调和平均。准确率、精确率、召回率以及 $F1$ 分数, 这些指标是衡量分类模型的常用评估指标, 其计算方法由式(5)~(8)给出。其中: TP (True Positive) 为正类预测正确的数量; TN (True Negative) 为负类预测正确的数量; FP (False Positive) 为正类预测错误的数量; FN (False Negative) 为负类预测错误的数量。

$$A = TP + TN / (TP + FP + FN + TN) \quad (5)$$

$$P = TP / (TP + FP) \quad (6)$$

$$R = TP / (TP + FN) \quad (7)$$

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (8)$$

此外, 依据模型的预测结果进行模拟交易是检验模型盈利能力的重要方式。本文使用投资回报率(Return of Investment, ROI)来衡量模型的盈利能力。ROI的计算公式如下:

$$V_{ROI} = \text{净投资损益} / \text{总投资花费} \times 100\% \quad (9)$$

2.4 实验结果及分析

2.4.1 实验结果

1) 预处理方法与注意力机制消融实验。

本文首先对提出的预处理方法进行了验证, 使用的模型为不带有注意力模块的CNN, 其余设置与1.2.1节中提出的CNN模型结构相同。采用不同的预处理方法进行训练, 输入待预测日期前20 d的数据, 对S&P500指数2019年 $T + 1$ 日的涨跌情况进行预测。 $T + 1$ 表示预测的是输入时间段未来1天的涨跌。结果如表1所示, “SWS+CNN”表示采用了新的预处理方式进行训练的模型。

表1 $T+1$ 日涨跌情况预测结果

模型	准确率	精确率	召回率	F1分数
CNN	0.518 4	0.589 0	0.653 1	0.619 4
SWS+CNN	0.567 3	0.591 9	0.788 0	0.713 5

由表1可以看出, 采用新的预处理方法的预测准确率提高了4.89个百分点, 召回率也得到了明显提升, 结果证明提

出的新预处理方法确实可以提高模型的预测精度。

此外,对于注意力机制的验证,采用消融实验的方式验证注意力机制的有效性。首先使用传统的数据预处理方法,在不加注意力的CNN模型和增加注意力的CNN模型上进行训练,比较预测结果。然后,再结合本文提出的预处理方法进行实验,结果如表2所示。“ATT”表示注意力(Attention)，“ATT+CNN”表示加入了注意力的CNN。

表2 消融实验结果

模型	准确率	精确率	召回率	F1 分数
CNN	0.518 4	0.589 0	0.653 1	0.619 4
ATT+CNN	0.546 9	0.602 3	0.721 1	0.656 3
SWS+CNN	0.567 3	0.591 9	0.788 0	0.713 5
SWS+ATT+CNN	0.612 2	0.619 3	0.918 4	0.739 7

由表2结果可以看出,注意力机制的加入同样提高了模型的预测精度。值得注意的是,对采用传统的预处理方式进行训练的CNN模型而言,注意力的加入使得模型预测的准确率提高了2.85个百分点。而采用了新的SWS预处理方式进行训练的模型,注意力机制对模型预测的准确率贡献了4.49个百分点的精度提升。可以推测,本文所提出的预处理方法,将金融时序数据转换为相互独立的特征图像数据,在一定程度上更加适合带有注意力机制的CNN模型,注意力机制的加入可以更好地捕捉图像特征,使得模型可以更好地作出判断。

2)方法稳定性验证。

为了进一步验证所提方法的稳定性,参照文献[8]中的研究,对 $T+N(N=5,10,15,20)$ 日后的涨跌情况进行了预测,实验结果如表3所示。结果表明,采用新的预处理方法与使用注意力机制的模型都在一定程度上提高了预测结果,与 $T+1$ 日涨跌预测情况相符。以上实验结果证明,模型在不同的预测任务上均有较好的表现,提出的改进方法也可以稳定提升模型的预测精度。

表3 不同预测目标实验结果

目标	模型	准确率	精确率	召回率	F1 分数
$T+5$	CNN	0.481 3	0.666 7	0.503 0	0.573 4
	ATT+CNN	0.510 4	0.684 2	0.549 9	0.606 7
	SWS+CNN	0.547 7	0.693 3	0.622 8	0.656 2
	SWS+ATT+CNN	0.597 5	0.710 8	0.706 6	0.708 7
$T+10$	CNN	0.597 5	0.789 5	0.610 5	0.688 5
	ATT+CNN	0.610 2	0.798 5	0.622 1	0.699 3
	SWS+CNN	0.605 9	0.733 7	0.720 9	0.727 3
	SWS+ATT+CNN	0.648 3	0.740 5	0.796 5	0.767 5
$T+15$	CNN	0.519 5	0.746 0	0.543 4	0.628 8
	ATT+CNN	0.562 8	0.746 6	0.630 1	0.683 4
	SWS+CNN	0.575 8	0.719 3	0.711 0	0.715 1
	SWS+ATT+CNN	0.614 7	0.747 1	0.734 1	0.740 5
$T+20$	CNN	0.500 0	0.717 6	0.552 9	0.624 6
	ATT+CNN	0.561 9	0.732 0	0.658 8	0.693 5
	SWS+CNN	0.566 4	0.709 3	0.717 6	0.713 5
	SWS+ATT+CNN	0.623 9	0.724 9	0.805 9	0.763 2

3)已发表研究比较。

在金融时序数据预测领域,文献中使用的金融资产的种类和数据非常广泛,难以直接进行比较,表4列出了已发表的研究中所使用的金融资产数据以及时间段,并给出了研究报告的对于金融资产一天后涨跌预测的结果作为参照。值得

注意的是,本文所使用的数据集与文献[8]中是一致的,因而可以直接进行准确率的比较。

表4 与已发表研究的比较结果

来源	数据集	时间段	准确率
文献[6]	KOSPI指数	2015.01—2019.12	0.64
文献[17]	S&P500指数	2011.10—2017.07	0.55
文献[18]	9个行业88家公司股票数据	2014.01—2016.01	0.58
文献[19]	上证A股数据	2006.01—2016.10	0.57
文献[20]	公司股票数据	2017.07—2020.07	0.63
文献[8]	S&P500指数	2010.01—2019.12	0.59
本文	S&P500指数	2010.01—2019.12	0.61

由表4可以看出,金融资产选取的时间段长短对模型预测准确率有显著的影响,时间跨度越长,模型更加难以把握当下的市场环境,导致预测的准确率普遍偏低。而选取较短的时间跨度,市场环境还未发生明显的变化,模型可以做出更好的预测,但这些模型往往无法持续发挥作用。本文所提方法降低了模型对于市场环境的依赖性,从准确率的结果来看,本研究使用的测试集与文献[8]是一致的,均是在上行市场环境下进行的测试,所提方法在文献的基础上有了进一步的提高。

2.4.2 模型盈利能力评估

进行金融数据预测一个非常重要的目的是帮助投资者获得更多的收益,因此有必要对模型的盈利能力进行评估。为此,基于模型在测试集上对 $T+1$ 日的预测结果,进行模拟交易实验,起始资金设为10000\$,每笔交易产生的手续费为交易金额的0.09%^[8]。依据预测结果,一共产生了18次交易,交易的具体情况由表5给出。

表5 交易详情

交易序号	买入日期	买入单价/\$	卖出日期	卖出单价/\$	手续费/\$	收益/\$
1	1/2	2510.03	1/28	2643.85	7.14	394.32
2	1/29	2640.00	2/5	2737.70	7.39	285.71
3	2/7	2706.05	3/1	2803.69	7.57	285.35
4	3/5	2789.65	3/14	2808.48	7.58	48.91
5	3/15	2822.48	4/9	2878.20	7.77	159.39
6	4/11	2888.32	4/30	2945.83	7.95	164.58
7	5/3	2945.64	6/25	2917.38	7.88	-92.66
8	6/26	2913.78	7/5	2990.41	8.07	221.82
9	7/9	2979.63	7/10	2993.07	8.08	32.24
10	7/11	2999.91	7/31	2980.38	8.05	-66.64
11	8/5	2844.74	9/3	2906.27	10.46	235.66
12	9/5	2976.00	9/17	3005.70	8.12	80.98
13	9/18	3006.73	9/25	2984.87	8.06	-73.64
14	9/26	2977.62	10/4	2952.01	7.97	-84.80
15	10/7	2938.79	11/8	3093.08	8.35	454.52
16	11/11	3087.01	12/4	3112.76	8.40	68.85
17	12/6	3145.91	12/10	3132.52	8.46	-48.63
18	12/11	3141.63	12/23	3224.01	8.70	238.44
总计						2304.39

由表5可以看出,模型的盈利能力良好,在18次交易中,有13次收益为正,正收益率超过了70%。此外,依据式(9)计算的ROI达到了23.04%,超过了文献[8]中在相同时间段报告的13.23%以及买入并持有策略的21.33%,说明模型可以在一定程度上帮助投资者规避市场波动带来的损失,提高收益率。

3 结语

在金融时序数据预测任务中,金融市场环境随时间的变化程度相对较大,使得依赖价格特征进行训练的模型难以学习到正确的映射关系。针对以上问题,本研究通过核密度估计分析数据分布,提出了一种基于滑动窗口标准化的金融图像数据预处理方式,该方法可以解决CNN模型应对不断变化的市场环境能力薄弱的问题。此外,将注意力机制与CNN结合,构建了一个基于注意力的CNN模型结构。注意力机制能有效捕捉数据的动态变化特征,提高了模型的特征提取能力。

实验结果表明,提出的预处理方法显著改善了数据密度分布情况,降低了模型对于市场环境的依赖性,模型可以正确学习到输入与输出的映射关系,注意力加入也使得模型效果得到了提升。结合实际应用场景,该方法可以用于预测S&P500指数未来一天收盘价的涨跌,并且基于预测结果可以获得超额收益。

需要注意的是,本研究也存在一定局限性。首先实验数据采用的S&P500指数在实验选取的时间段内整体上是上行趋势,对于下行市场或者平市情况还需要进一步进行实验。其次,虽然模型预测结果有了一定提高,但是在研究过程中也发现受限于数据量很难再有进一步的提高。未来的工作可以结合更多的与金融市场有关的信息,如新闻媒体数据、宏观经济数据等信息,理论上可以进一步提高预测的精度。

参考文献 (References)

- [1] LIN Y, LIU S, YANG H, et al. Improving stock trading decisions based on pattern recognition using machine learning technology[J]. PLoS One, 2021, 16(8): e0255558.
- [2] JIANG W. Applications of deep learning in stock market prediction: recent progress[J]. Expert Systems with Applications, 2021, 184: 115537.
- [3] 张倩倩,林天华,祁旭阳,等. 基于机器学习的股票预测研究综述[J]. 河北省科学院学报, 2020, 37(4): 15-21.
- [4] SIM H S, KIM H I, AHN J J. Is deep learning for image recognition applicable to stock market prediction?[J]. Complexity, 2019, 2019: 4324878.
- [5] SEZER O B, OZBAYOGLU M. Financial trading model with stock bar chart image time series with deep convolutional neural networks[J]. Intelligent Automation and Soft Computing, 2019, 26(2): 2007-2022.
- [6] JIN G, KWON O. Impact of chart image characteristics on stock price prediction with a convolutional neural network[J]. PLoS One, 2021, 16(6): e0253121.
- [7] SEZER O B, OZBAYOGLU A M. Algorithmic financial trading with deep convolutional neural networks: time series to image conversion approach[J]. Applied Soft Computing, 2018, 70: 525-538.
- [8] CHEN Y-C, HUANG W-C. Constructing a stock-price forecast CNN model with gold and crude oil indicators[J]. Applied Soft Computing, 2021, 112: 107760.
- [9] KIM T, KIM H Y. Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data[J]. PLoS One, 2019, 14(2): e0212320.
- [10] MOGHADDAM H A, MOMTAZI S. Image processing meets time series analysis: Predicting forex profitable technical pattern positions[J]. Applied Soft Computing, 2021, 108: 107460.
- [11] 方义秋,卢壮,葛君伟. 联合RMSE损失LSTM-CNN模型的股价预测[J]. 计算机工程与应用, 2022, 58(9): 294-302.
- [12] 赵红蕊,薛雷. 基于LSTM-CNN-CBAM模型的股票预测研究[J]. 计算机工程与应用, 2021, 57(3): 203-207.
- [13] SHAH D, ISAH H, ZULKERNINE F. Stock market analysis: a review and taxonomy of prediction techniques[J]. International Journal of Financial Studies, 2019, 7(2): 26.
- [14] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.
- [15] 刘翀,杜军平. 一种基于深度LSTM和注意力机制的金融数据预测方法[J]. 计算机科学, 2020, 47(12): 125-130.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Proceedings of the 2018 European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [17] LIU Y, ZENG Q G, YANG H R, et al. Stock price movement prediction from financial news with deep learning and knowledge graph embedding [C]// Proceedings of the 15th Pacific Rim Knowledge Acquisition Workshop and 2018 Conference on Knowledge Management and Acquisition for Intelligent Systems. Cham: Springer, 2018: 102-113.
- [18] XU Y, COHEN S B. Stock movement prediction from tweets and historical prices [C]// Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: Association for Computational Linguistics, 2018: 1970-1979.
- [19] PANG X, ZHOU Y, WANG P, et al. An innovative neural network approach for stock market prediction[J]. The Journal of Supercomputing, 2020, 76(3): 2098-2118.
- [20] ZHANG Q, YANG L J, ZHOU F. Attention enhanced long short-term memory network with multi-source heterogeneous information fusion: An application to BGI Genomics [J]. Information Sciences, 2021, 553: 305-330.