

基于卷积神经网络的股指期货套利策略研究

岳鹏飞, 李秀军*

河北工业大学, 理学院, 天津
Email: yue_pf@126.com, *lixijun@hebut.edu.cn

收稿日期: 2021年1月23日; 录用日期: 2021年2月17日; 发布日期: 2021年2月25日

摘要

在人工智能技术迅猛发展的背景下, 神经网络与金融领域的结合越来越密切, 利用神经网络等机器学习方法构建投资策略, 已是当前投资界和学术界的一个热点问题。提出了一种基于卷积神经网络的股指期货套利策略。通过建立卷积神经网络模型对股指期货价差进行预测, 采用动态阈值方法确定套利区间。并对中证500股指期货和上证50股指期货进行了实证分析。对比研究基于SVM模型和XGboost模型的交易策略。实证结果显示, 在最大回撤率处于同一水平下基于卷积神经网络的股指期货套利策略收益率比基于SVM模型和XGboost模型的套利策略收益率表现更好。证实了将卷积神经网络应用在股指期货套利策略当中具有一定的可行性与有效性。

关键词

卷积核, 神经网络, 股指期货, 套利, 动态阈值

Research on Stock Index Futures Arbitrage Strategy Based on Convolutional Neural Network

Pengfei Yue, Xiujun Li*

School of Science, Hebei University of Technology, Tianjin
Email: yue_pf@126.com, *lixijun@hebut.edu.cn

Received: Jan. 23rd, 2021; accepted: Feb. 17th, 2021; published: Feb. 25th, 2021

Abstract

With the rapid development of artificial intelligence technology, the combination of neural network and financial field is closer. It is a hot issue in the field of investment and academia to construct investment strategy by using machine learning methods such as neural network. This paper

*通讯作者。

proposes a stock index futures arbitrage strategy based on convolutional neural network. The convolution neural network model is established to predict the spread of stock index futures, and the dynamic threshold method is used to determine the arbitrage interval. This paper makes an empirical analysis on the CSI 500 stock index futures and the SSE 50 stock index futures. In this paper, we make a comparative study of the trading strategy based on SVM model and XGboost model. The empirical results show that when the maximum withdrawal rate is at the same level, the return rate of stock index futures arbitrage strategy based on convolutional neural network is better than that based on SVM model and XGboost model. It is proved that the convolution neural network is feasible and effective in the arbitrage strategy of stock index futures.

Keywords

Convolution Kernel, Neural Network, Stock Index Futures, Arbitrage, Dynamic Threshold

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

统计套利是一种市场中性策略,主要是将具有一定相关性的两种或两种以上的金融产品进行配对,利用数学工具发现其错误的定价关系,根据错误的定价关系建立头寸,待其价格关系恢复正常后平掉头寸以获取收益。随着计算机技术的快速发展,深度学习技术在金融市场的应用越来越广泛[1],基于神经网络等机器学习方法构建投资策略已是当前学术界的热点研究问题之一[2][3]。

卷积神经网络对于网格型数据的特征有着很好的捕获功能。而时间序列可以看作是一维的网格型数据,根据卷积神经网络的这种特性,卷积神经网络在时间序列特征的抓取方面有着一定的优势。近年来已有许多学者将卷积神经网络用于处理股票数据上[4]。因此,卷积神经网络在金融时间序列上有着很大的应用前景[5]。Cao (2018) [6]对卷积神经网络进行了提升,将卷积神经网络和支持向量机相结合,对1996~2014年的HIS指数进行了预测分析。Alonso (2020) [7]利用卷积神经网络对比特币等加密货币的价格进行了预测,证实了比特币价格的可预测性。其他机器学习方法与金融数据领域的结合也较为密切。Huck [8] [9]利用深度学习方法对美国股市的股票进行分析,通过排序和配对的方法建立了投资组合模型。Krauss (2017) [10]等利用了BP神经网络、梯度提升树、随机森林对标普500指数中包含的成分股进行了研究分析,首先对股票进行排序,然后对排序最高的股票和排序最低的股票之间进行套利交易,最后通过排序的变化从而实现模型的动态性。侯世英[11]等通过将BP-神经网络和GARCH模型相结合建立了统计套利策略,对资产的选取、配比和交易信号的确定都进行了修正。

本文在前人的研究基础上提出了一种基于卷积神经网络的股指期货套利策略。使用卷积神经网络对股指的价差进行预测,根据实际价差的增量与预测价差的增量的相对位置建立套利策略。由于我国的股指期货尚处于发展阶段,存在着一定的套利机会,所以本文对中证500股指期货和上证50股指期货进行了实证研究。

2. 卷积神经网络

2.1. 卷积神经网络模型构建

卷积神经网络(CNN)是通过卷积核对数据进行扫描从而“记忆”数据的特征,所以卷积神经网络不

仅可以对数据的单个特征进行描绘, 而且可以更好的捕捉特征之间的关系。传统的机器学习方法, 比如 SVM 模型和 XGboost 模型并没有对样本特征之间的关联信息做更多的挖掘。此外本文处理的数据为高频时间序列数据, 具有对长期数据相关性不高, 变化频率快、随机性强、难以预测的特点。所以本文通过对卷积核进行改进, 建立了基于一维卷积核的卷积神经网络模型, 一维卷积核可以高效的跟随高频数据的变化, 并且可以更好的挖掘序列数据之间的关系。

因为卷积神经网络的主要组成部分为卷积层和池化层, 所以影响卷积神经网络结构的因素主要有卷积层层数, 每层卷积核的个数, 卷积核的大小和样本的长度。期货价格具有当前价格对长期数据的依赖较弱的特点, 所以输入数据的长度本文考虑 20, 30, 40, 70 四种情况。一个完整的卷积神经网络应该有输入层, 卷积层, 池化层, 输出层, 四个层结构。因为时间序列数据相对于图片数据的复杂度较小, 在对卷积层的考虑过程中, 本文分别考虑一层卷积层和两层卷积层两种情况。关于卷积核的设置, 卷积核的个数分别考虑 5, 8, 16, 32 四种情况, 而卷积核的大小和样本的长度相关, 其中关于输入数据的维度, 输出数据的维度和卷积核的大小设置满足以下公式。

$$\lfloor (L-C+1)/S \rfloor = N, \quad (1)$$

式中: L 表示输入样本的长度, C 表示卷积核的大小, S 为卷积核移动的步长, N 表示输出数据的维度, $\lfloor \rfloor$ 表示向下取整。

Table 1. Relationship between sample length and convolution kernel size

表 1. 样本长度和卷积核大小关系

样本长度	卷积核大小
20	3
30	3 5
40	3 5 9
70	3 5 9 13

表 1 表示样本长度和卷积核大小的相关性, 当样本维度为 20 时, 卷积核大小只设置为 3, 当样本长度扩充到 70 时, 卷积核的大小则可以取 3, 5, 9, 13 四种情况。

2.2. 评价指标

对于预测结果, 本文使用三种评价标准: 均方误差(MSE), 皮尔逊相关系数(r), 和决定系数(R^2)。均方误差反应了预测序列与实际序列的偏差; 相关系数反应了预测序列的趋势与实际的趋势的一致性, 相关系数越接近 1, 预测序列与真实预测的趋势越一致; 而决定系数反应了模型对真实值的拟合效果。通过这三种指标的综合评价, 调整得出最合适的网络参数, 然后构建基于一维卷积核的卷积神经网络。

3. 套利策略

传统的统计套利多是基于标准距离法进行无套利区间的确定, 使用样本均值和样本方差, 来确定套利区间。但是标准距离法只是根据价差的总体均值和总体方差确定套利区间, 不能很好的把握价差特点, 而且当价差出现单边趋势时标准距离法将不再有效[12]。所以, 本文套利策略基于动态阈值的方法, 不再使用固定的标准距离作为套利信号, 而是根据价差的走势对套利区间进行不断地调整实现动态套利的过程。中证 500 指数和上证 50 指数都在一定程度上反映了 A 股的行情走势, 在走势上具有一定的联动性。本文选择中证 500 股指期货合约和上证 50 股指期货合约进行套利分析。股指期货和商品期货不同, 股指

期货的价格是由股指期货的点数乘以合约的乘数构成, 中证 500 股指期货合约和上证 50 股指期货合约的乘数分别为 200 和 300, 本文按照两个期货合约分别为一手的基准仓位建立套利头寸时, 则构建价差时, 中证 500 股指期货价格和上证 50 股指期货价格应分别乘以乘子 2 和 3。

$$spread = 3IH - 2IC, \quad (2)$$

式中: IC 和 IH 分别表示中证 500 和上证 50 股指期货的价格。 $spread$ 表示两种合约的价差, 则该式表示在进行套利交易时, 做多一手价差, 相当于同时买多一手的上证 50 股指期货合约和买空一手中证 500 股指期货合约。买空一手价差, 相当于同时买空一手的上证 50 股指期货合约和买多一手中证 500 股指期货合约。

本文使用 pre_In_t 来表示预测的价差的增量, In_t 表示真实值的价差增量,

$$In_t = spread_t - spread_{t-1}, \quad (3)$$

$$pre_In_t = pre_spread_t - spread_{t-1}, \quad (4)$$

式中: pre_spread_t 表示通过卷积神经网络对 $spread_t$ 的预测值, 将 $spread_t$ 的前 T (不包括 t) 个值作为输入, 预测得到的 pre_spread_t 。 In_t 表示实际的价差增量, pre_In_t 表示预测的价差增量, 价差增量表示预测值与前一时间节点的变化值, 通过预测值的价差增量可以提前预判价差的变化浮动, 从而对套利信号做出指导。

交易信号是根据 pre_In_t 与 In_t 的相对位置进行设置, 在本文中交易的阈值设置为移动阈值 $X(t)$ 和 $M(t)$, 当 pre_spread_t 绝对值大于浮动阈值 $X(t)$ 时, 进行开仓交易, 当 $spread$ 反向变化回归移动均值 $M(t)$ 时, 进行平仓交易。具体的交易策略如下:

当 $pre_In_t > X(t)$ 时, 买入一手上证 50 股指期货多头, 买入一手中证 500 股指期货的空头;

当 $pre_In_t > X(t)$ 后, 再次 $pre_In_t < M(t)$ 时平仓;

当 $pre_In_t < X(t)$ 时, 买入一手上证 50 股指期货空头, 买入一手中证 500 股指期货的多头;

当 $pre_In_t < X(t)$ 后, 再次 $pre_In_t > M(t)$ 时平仓。

4. 实证分析

4.1. 数据处理

一般情况当月的期货合约具有成交量大, 市场流动性好, 交易活跃的特点, 在市场中把这样的合约称之为主力合约。本文选取的数据为中证 500 股指期货和上证 50 股指期货每分钟主力合约的行情数据。数据长度为一个自然年: 2019 年 6 月 3 日到 2020 年 6 月 1 日, 共 58,807 条记录。其中 70% 的数据用于训练集, 30% 的数据用于测试集, 即训练数据为 2019 年 6 月 3 日到 2020 年 2 月 14 日的价差数据, 测试数据为 2020 年 2 月 14 日到 2020 年 6 月 1 日的价差数据。数据来源于万得数据库。

在数据进行训练之间本文需要对数据进行预处理, 因为股指价格数据的量级比较大, 不利于神经网络参数的更新。所以, 首先对数据进行归一化处理, 使得数据范围限制在 (0,1) 之间。

$$spread = \frac{spread - MIN(spread)}{MAX(spread) - MIN(spread)}, \quad (5)$$

式中: $MAX(spread)$, $MIN(spread)$ 分别表示 $spread$ 序列的最大值和最小值。

由于金融数据具有强噪声性, 所以本文在训练模型之前对价差数据进行 5 分钟移动平均, 以便去除噪声的干扰, 使得本文的模型可以更好的把握数据的本质趋势。

$$ma5_spread_t = \frac{spread_{t-4} + spread_{t-3} + spread_{t-2} + spread_{t-1} + spread_t}{5}, \quad (6)$$

式中: $spread_t$ 表示第 t 时刻的股指价差。

4.2. 预测过程

关于建立 CNN 卷积神经网络预测模型的过程, 本文通过控制变量法, 分别讨论模型的卷积层数, 卷积核个数、卷积核大小, 样本长度 4 个超参数。基于本文中卷积神经网络是对时间序列进行处理, 时间序列数据相对于图片数据复杂度较小, 所以, 关于卷积层层数本文分别讨论 1 层和 2 层, 为了避免模型的随机性使得对模型的结果的分析产生误判, 所以讨论卷积层层数的过程中, 分别设置样本的长度为 30 和 40 作为对比。因为卷积核的大小和样本长度相关, 所以根据样本长度的不同, 设置卷积核的大小分别为 3 和 5, 当模型的卷积层设置为两层时, 第二层的卷积核个数设置为第一层的二倍, 分别为 5 和 10。

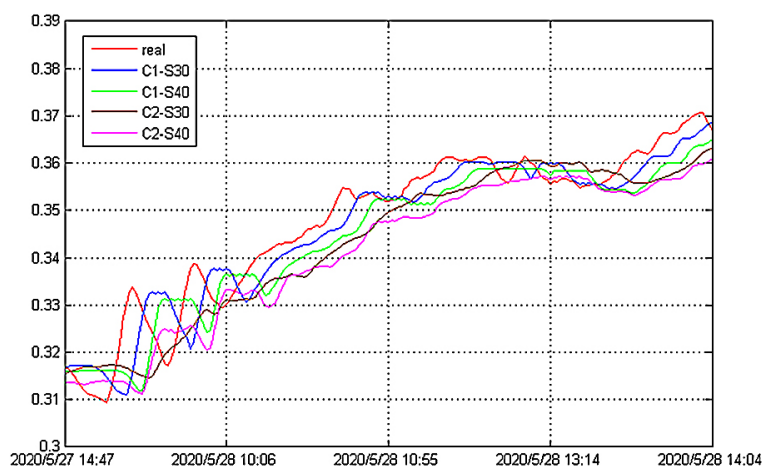


Figure 1. Predicted result based on model with different convolution layers

图 1. 基于不同卷积层层数的模型预测结果

图 1 中 real 表示 $ma5_spread_t$, C1-S30、C1-S40 和 C2-S30、C2-S40 分别为样本长度为 30、40 时卷积层层数分别为一、二时的价差预测值曲线。

Table 2. Predicted result of different convolution layers

表 2. 不同卷积层的预测结果

Conv_num	Kernel_num	Sample_size	Kernel_size	MSE	R^2	r
1	5	30	3	0.00007	0.9908	0.9959
		40	5	0.00014	0.98203	0.9920
2	(5, 10)	30	3	0.00019	0.9748	0.9880
		40	5	0.00016	0.9796	0.9900

通过图 1 和上表 2 中的 MSE、 R^2 和 r 可以看出, 卷积层为 1 层的结果普遍比卷积层为 2 层的结果好, 其中当超参数设置为 (1, 5, 30, 3) 时, 均方误差为 0.00007, 决定系数 R^2 为 0.9908, 皮尔逊相关系数为 0.9959, 最接近于 1, 结果表明此种情况下, 预测数据与真实数据偏离程度最低, 趋势最相似, 对数据变异的解释能力也最强。所以在接下来的讨论过程中模型的卷积层数设置为 1 层。

确定了模型的卷积层数后, 然后分析卷积核的大小, 在对卷积核的研究过程中, 为了产生更明显的对比性, 固定样本长度为 70, 分别设置卷积核的大小为 3, 5, 9, 13。

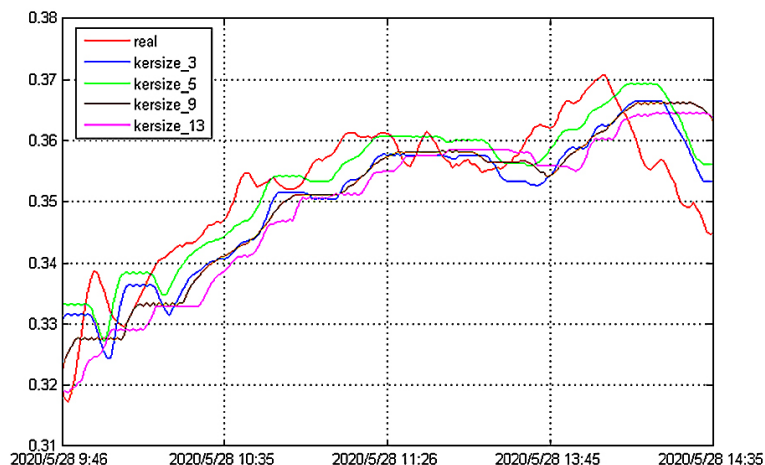


Figure 2. Predicted results based on model with different kernel size
图 2. 基于不同卷积核大小的模型预测结果

图 2 中 real 表示 $ma5_spread_t$ 。kersize_3、kersize_5、kersize_9、kersize_13 分别为卷积核大小为 3、5、9、13 的价差预测值曲线。

表 3 表示当卷积层数为 1，卷积核个数为 5，样本长度为 70，卷积核大小分别为 3、5、9、13 时，卷积神经网络对价差的预测结果。

Table 3. Predicted results of different convolution kernel sizes
表 3. 不同卷积核大小的预测结果

Kernel_size	MSE	R^2	r
3	0.00016	0.9792	0.9918
5	0.00017	0.9784	0.9929
9	0.00024	0.9691	0.9876
13	0.00030	0.9608	0.9837

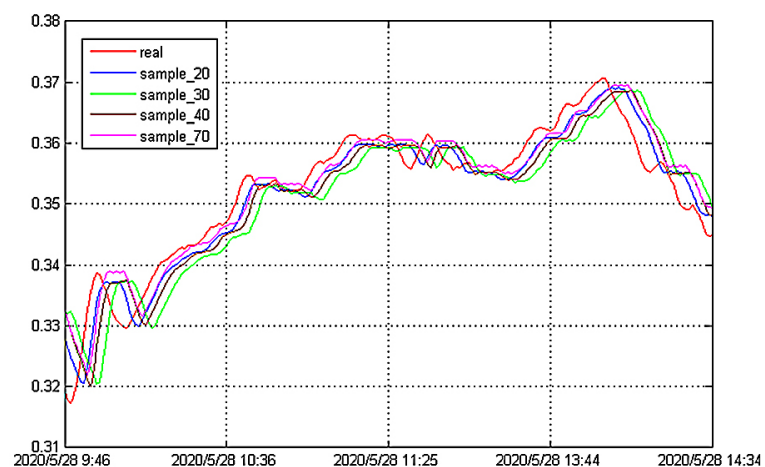


Figure 3. Predicted results based on model with different sample size
图 3. 基于不同样本长度的模型预测结果

通过上图和上表可以看出当 Kernel_size 为 3 时，均方误差 MSE 和 R^2 都是最优的，皮尔逊相关系数

r 也为 0.9918。表示当卷积核大小设置为 3 时, 与真实数据的偏离情况最低和与真实情况的趋势相似性最高。所以, 在接下来的情况中卷积核的大小设置为 3。

确定了卷积层的层数为 1 和卷积核的大小为 3 后, 接下来讨论样本的长度问题, 因为训练数据为股指期货的分钟级价差数据, 根据其特点, 当前的价差受长期时间序列的影响较弱, 所以设置样本长度分别为 20、30、40、70。

图 3 中 $real$ 表示 $ma5_spread_t$, $sample_20$ 、 $sample_30$ 、 $sample_40$ 、 $sample_70$ 分别为样本长度为 20、30、40、70 的价差预测值曲线。

表 4 表示当卷积层数为 1, 卷积核数为 5, 卷积核大小为 3, 样本长度分别为 20、30、40、70 时, 卷积神经网络的预测结果。

Table 4. Predicted results without sample length

表 4. 不同样本长度的预测结果

Sample_size	MSE	R^2	r
20	0.000029	0.9963	0.9983
30	0.000084	0.9893	0.9951
40	0.000060	0.9923	0.9968
70	0.00094	0.8799	0.9581

当样本长度为 20 时, 均方误差 MSE 为 0.000029, 决定系数 R^2 为 0.9963, 皮尔逊相关系数 r 为 0.9983, 当样本长度为 70 时, 均方误差 MSE 为 0.00094, 决定系数 R^2 为 0.8799, 皮尔逊相关系数 r 为 0.9581, 可以观测到模型的表现效果随着样本长度的增加而衰退。这种结果也再次证实了高频股指数据对于短期的时间序列更加具有依赖性。

最后本文来研究模型卷积核个数, 通过前面的分析可以固定卷积层的层数, 卷积核的大小, 样本的长度。然后设置模型的卷积核的个数分别为 5、8、16、32。

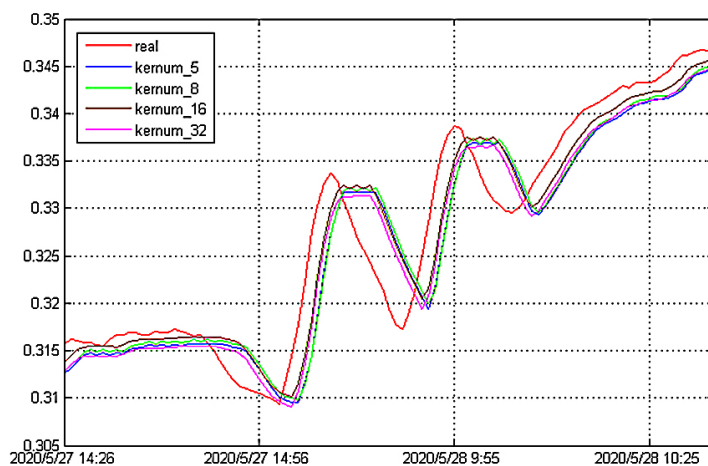


Figure 4. Predicted results based on model with different kernel numbers

图 4. 基于不同卷积核个数的模型预测结果

图 4 中 $real$ 表示真实值。kernum_5、kernum_8、kernum_16、kernum_32、分别表示卷积核个数为 5、8、16、32 的价差预测值曲线。

表 5 表示当卷积层数为 1, 卷积核大小为 3, 样本长度为 20, 卷积核个数分别为 5、8、16、32 时,

卷积神经网络的预测结果。

Table 5. Prediction results of different convolution kernels

表 5. 不同卷积核个数的预测结果

Kernel_num	MSE	R^2	r
5	0.000033	0.9957	0.9980
8	0.000034	0.9956	0.9979
16	0.000026	0.9965	0.9985
32	0.000023	0.9970	0.9985

通过实验分析可以得出当卷积核个数为 32 时, 均方误差 MSE 为 0.000023, 决定系数 R^2 为 0.9970, 皮尔逊相关系数 r 为 0.9985。分别在 4 种情况中表现最好。

根据上面对模型中超参数的讨论, 最后可以设置 CNN 卷积神经网络模型中卷积层的层数为 1, 卷积核的大小为 3, 样本长度为 20, 卷积核的个数设置 32。为了对 CNN 模型的预测结果进行对比分析, 本文采用国内应用较为广泛的 SVM 模型和 XGboost 模型, 它们的预测结果分别如表 6 所示。

Table 6. Comparison between predicted value and actual value

表 6. 预测值与实际值的对比

	Mean	Std	Min	25%	50%	75%	Max
真实值	0.296950	0.088837	0.000000	0.241792	0.299989	0.343670	0.515797
CNN	0.298125	0.087845	0.016718	0.243841	0.301164	0.344240	0.512961
SVM	0.400073	0.069181	0.173027	0.357542	0.402243	0.436633	0.569122
XGboost	0.542577	0.045901	0.391998	0.514417	0.544066	0.566803	0.654671

表 6 显示真实价差值的均值为 0.296950, CNN 模型预测值的均值为 0.298125, 真实价差的最大值为 0.515797, CNN 模型预测价差的最大值为 0.512961。而 SVM 模型 XGboost 模型预测值的均值分别为 0.400073 和 0.542577, 最大值分别为 0.569122 和 0.654671, 和真值均相差较远。所以从预测序列分布来看, 在股指期货序列的预测方面, CNN 模型要好于 SVM 模型和 XGboost 模型。下面将基于三种不同机器学习模型的预测结果根据套利策略进行套利交易。

4.3. 套利策略

本文对 2020 年 2 月 14 日到 2020 年 6 月 1 日期间的上证 50 股指期货主力合约和中证 500 股指期货主力合约进行回测。根据中金所要求, 股指期货交易的手续费率万分之 0.23。关于止损策略, 本文设置为 4 个点位, 即当策略连续亏损 4 个价差点位的时候强行平仓, 由于股指期货的保证金额为 10%, 按照操作资金占总仓位的 80% 计算, 本文中假设初始保证金为 280,000 元。通过卷积神经网络的预测模型, 可以得到价差的预测序列 pre_spread_t 。通过 pre_spread_t 得到预测值的价差增量 pre_In_t 。

$$In_t = ma5_spread_t - ma5_spread_{t-1}, \quad (7)$$

$$pre_In_t = pre_spread_t - ma5_spread_{t-1}. \quad (8)$$

套利过程中开仓和正常性平仓的阈值由 In_t 的移动平均值和 In_t 的移动标准差共同构成。开多仓和开空仓的阈值分别设置为 $X_l(t)$ 和 $X_s(t)$, 平仓的阈值为 $M(t)$ 。

$$\begin{aligned} X_l(t) &= M_s(t) + 2 \times Std_s(t), \\ X_s(t) &= M_s(t) - 2 \times Std_s(t), \end{aligned} \quad (9)$$

式中: $M(t)$ 表示在 t 时刻的前 s 个 In_t 的移动平均值, $Std(t)$ 表示在 t 时刻前 s 个 In_t 的移动标准差值。

套利过程分为开仓、正常性平仓和止损性平仓, 其中开多仓、开空仓和正常性平仓的交易阈值分别为 $X_l(t)$ 、 $X_S(t)$ 和 $M(t)$ 。则交易策略为:

当 $pre_In_t > X_l(t)$ 时, 买入一手上证 50 股指期货多头, 买入一手中证 500 股指期货的空头;

当 $pre_In_t > X_l(t)$ 后, 再次 $pre_In_t < M(t)$ 时或亏损超过 4 个价差点位时进行平仓;

当 $pre_In_t < X_S(t)$ 时, 买入一手上证 50 股指期货空头, 买入一手中证 500 股指期货的多头;

当 $pre_In_t < X_S(t)$ 后, 再次 $pre_In_t > M(t)$ 时或亏损超过 4 个价差点位时进行平仓。

4.4. 结果分析

为了使回测结果更加具有对比性, 防止回测结果具有偶然性。本文将策略的开平仓阈值的参数 s 分别设置为 40、50 和 60。即策略开仓阈值分别设置为 In_t 的前 40、50、60 期的移动均值加减两倍标准差, 策略的平仓阈值分别设置为 In_t 的前 40、50、60 期的移动均值。策略的回测结果如下。

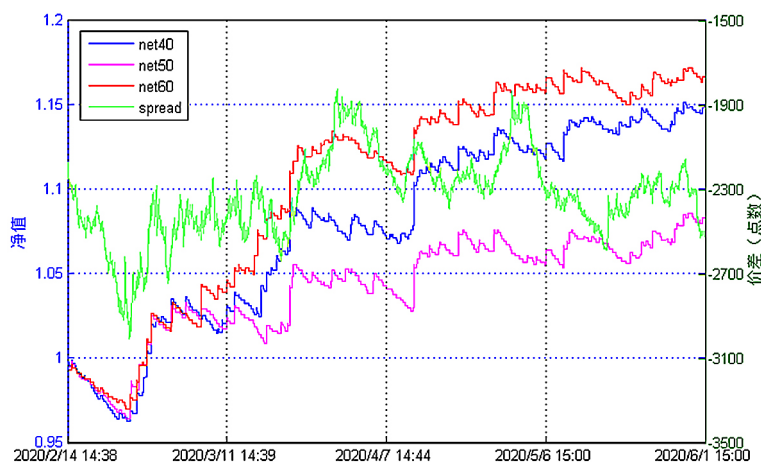


Figure 5. The curve net value and trend of spreads

图 5. 策略净值曲线与价差走势曲线

图 5 中的蓝色、紫色、红色的曲线分别表示策略在阈值参数为 40、50、60 时的净值曲线。绿色曲线表示价差曲线。从图 5 中可以看出三种策略的净值曲线虽然在局部时有波动, 但总体上表现为上升趋势。其中当阈值参数为 40 和 60 时策略的收益表现相对较好。并且通过图 5 可以看出无论价差的走势是在波动型走势 (2020/3/11 左右), 还是趋势型走势 (2020/4/7 左右), 策略的净值曲线总体都是上升趋势。

为了和基于卷积神经网络的套利交易策略进行对比分析, 分别将基于 SVM 模型和 XGboost 模型的预测结果进行相同的交易过程, 设置阈值参数为 40, 得出结果如表 7 所示。

Table 7. Return of arbitrage strategy

表 7. 套利策略收益情况表

模型	CNN			SVM	XGboost
	40	50	60		
回测时间段	2020 年 2 月 14 日到 2020 年 6 月 1 日				
初始保证金(元)	280,000				
交易次数	328	281	267	300	352
止损次数	201	189	170	230	255

Continued

总盈利(元)	41,449.4	23,039.4	46,577.3	11,664.3	13,480
最大回撤率	3.7%	3.5%	2.9%	3.2%	3.4%
收益率	15%	8%	17%	4.1%	4.7%
年化收益率	50%	28%	57%	13%	15%

通过表 7 可以看出在阈值参数同为 40 的情况下, 基于 CNN 模型、SVM 模型、XGboost 模型的收益率分别为 15%、4.1%、4.7%, 交易次数分别为 328、300、352; 最大回撤率分别为 3.7%、3.2%、3.4%。所以可以看出在相同阈值参数的设置下基于 CNN 模型交易策略的收益率要好于基于 SVM 模型交易策略和基于 XGboost 模型交易策略的收益率。并且基于 CNN 模型的交易结果随着 s 值越大交易次数越少, 这是因为 s 越大, 阈值受越多数据的影响, 因此套利区间越宽松, 套利机会就越多。对于 CNN 模型, 基于三种不同阈值参数模型的收益率分别为 15%、8%、17%, 最大回撤率分别为 3.7%、3.5%、2.9%。年化收益率分别为 50%、28%、57%。阈值参数 s 为 60 时, 策略的年化收益率最大, 并且同时最大回撤也最小。但阈值参数 s 为 50 时, 策略的收益率相对最小, 说明了策略的表现与 s 的大小并不呈单调关系。

5. 结论

通过对卷积神经网络的卷积层层数、卷积核大小、卷积核数量和样本长度进行分析, 建立了基于一维卷积核的卷积神经网络模型。利用卷积神经网络模型对股指期货价差进行预测, 预测结果显示价差的预测值与价差的实际值具有相似分布。根据价差的预测增量与实际增量的相对关系确定动态的套利区间。对 2020 年 2 月 14 日到 2020 年 6 月 1 日的中证 500 股指期货主力合约和上证 50 股指期货主力合约进行了实证研究。实证结果显示在阈值参数为 40 的情况下, 基于卷积神经网络模型交易策略的最大回撤率和年化收益率分别为 3.7% 和 50%。在阈值参数同为 40 的情况下将卷积神经网络替换为 SVM 模型和 XGboost 模型进行相同的交易策略, 策略分别获得了 3.2%、3.4% 的最大回撤率和 13%、15% 的年化收益率。在最大回撤率保持平稳的情况下, 基于卷积神经网络模型的交易策略在策略收益方面获得了巨大的提高。在不同的阈值参数下, 基于卷积神经网络模型的交易策略也分别取得了 50%、28%、57% 的年化收益率。因此, 在股指期货价差预测方面, 卷积神经网络要优于 SVM 模型和 XGboost 模型, 并且当最大回撤率处于同一水平下, 基于卷积神经网络的股指期货套利策略在要优于基于 SVM 模型和基于 XGboost 模型的套利策略。所以, 将卷积神经网络应用在股指期货套利方面有一定的研究价值和应用价值。

参考文献

- [1] Xu, Z.Y., Zhang, J., Wang, J.Y. and Xu, Z.M. (2020) Prediction Research of Financial Time Series Based on Deep Learning. *Soft Computing*, **11**, 8295-8312. <https://doi.org/10.1007/s00500-020-04788-w>
- [2] 王文波, 费浦生, 羿旭明. 基于 EMD 与神经网络的中国股票市场预测[J]. 系统工程理论与实践, 2010, 30(6): 1027-1033.
- [3] 林杰, 龚正. 基于人工神经网络的沪锌期货价格预测研究[J]. 财经理论与实践, 2017, 38(2): 54-57.
- [4] Sezer, O.B. and Ozbayoglu, A.M. (2018) Algorithmic Financial Trading with Deep Convolutional Neural Networks: Time Series to Image Conversion Approach. *Applied Soft Computing*, **70**, 525-538. <https://doi.org/10.1016/j.asoc.2018.04.024>
- [5] Spelta, A. (2017) Financial Market Predictability with Tensor Decomposition and Links Forecast. *Applied Network Science*, **1**, 7. <https://doi.org/10.1007/s41109-017-0028-1>
- [6] Cao, J.S. and Wang, J.G. (2019) Stock Price Forecasting Model Based on Modified Convolution Neural Network and Financial Time Series Analysis. *International Journal of Communication Systems*, **12**, e3987.

<https://doi.org/10.1002/dac.3987>

- [7] Alonsomonsalve, S., Suarezcetrulo, A.L., Cervantes, A., *et al.* (2020) Convolution on Neural Networks for High-Frequency Trend Prediction of Cryptocurrency Exchange Rates Using Technical Indicators. *Expert Systems with Applications*, **149**, 113250. <https://doi.org/10.1016/j.eswa.2020.113250>
- [8] Huck, H. (2019) Large Data Sets and Machine Learning: Applications to Statistical Arbitrage. *European Journal of Operational Research*, **1**, 330-342. <https://doi.org/10.1016/j.ejor.2019.04.013>
- [9] Huck, H. (2009) Pairs Selection and Outranking: An Application to the S & P 100 Index. *European Journal of Operational Research*, **196**, 819-825. <https://doi.org/10.1016/j.ejor.2008.03.025>
- [10] Krauss, C., Do, X.A., Huck, N., *et al.* (2017) Deep Neural Networks, Gradient-Boosted Trees, Random Forests: Statistical Arbitrage on the S & P 500. *European Journal of Operational Research*, **259**, 689-702. <https://doi.org/10.1016/j.ejor.2016.10.031>
- [11] 侯世英, 宋良荣, 王国俊. 基于 BP-GARCH 模型的统计套利策略[J]. 统计与决策, 2020, 36(10): 149-152.
- [12] 周亮. 基于价差预测的商品期货跨期套利研究[J]. 金融理论与实践, 2019(7): 84-90.