

## 基于深度学习算法的高频交易策略及其盈利能力

孙达昌, 毕秀春

(中国科学技术大学管理学院统计与金融系, 安徽合肥 230026)

**摘要:** 深度学习算法作为机器学习中的一种重要算法, 在图像处理、语音识别、机器翻译等领域已成功应用. 将深度学习算法应用于高频交易中, 选取卷积神经网络和 LSTM 神经网络分别构建涨跌分类模型, 在此基础上提出高频交易策略, 并以沥青期货主力合约为例进行回测检验, 实证分析策略优良性. 通过与人工神经网络高频交易策略的比较, 回测检验结果表明基于卷积神经网络和 LSTM 神经网络的高频交易策略的盈利能力较强, 泛化能力较好, 两种策略的胜率和期望收益虽有所差异, 但均比人工神经网络高频交易策略高.

**关键词:** 深度学习; 卷积神经网络; LSTM 神经网络; 量化投资; 高频交易

**中图分类号:** F830      **文献标识码:** A      doi: 10.3969/j.issn.0253-2778.2018.11.008

**引用格式:** 孙达昌, 毕秀春. 基于深度学习算法的高频交易策略及其盈利能力研究[J]. 中国科学技术大学学报, 2018, 48(11): 923-932.

SUN Dachang, BI Xiuchung. High-frequency trading strategies based on deep learning algorithms and their profitability[J]. Journal of University of Science and Technology of China, 2018, 48(11): 923-932.

## High-frequency trading strategies based on deep learning algorithms and their profitability

SUN Dachang, BI Xiuchung

(Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei 230026, China)

**Abstract:** As an important algorithm, deep learning has been applied successfully to image processing, speech recognition, machine translation and other fields. Here, deep learning algorithms were applied to high-frequency trading. Convolutional neural network(CNN) and long short-term memory(LSTM) neural network were selected to build up and down classification models, respectively. Based on the models, high-frequency trading strategies were proposed. Then the data of bitumen futures contract was used for back-testing and empirically analyzing the superiority of the strategies. In back-testing, deep learning algorithms were compared with artificial neural network(ANN). The results show that both strategies based on CNN and LSTM neural network exhibit better profitability and generalization ability. In addition, the winning rates and expected returns of the two strategies are also better.

**Key words:** deep learning; convolutional neural network; LSTM neural network; quantitative investment; high-frequency trading

收稿日期: 2018-03-15; 修回日期: 2018-05-30

基金项目: 国家自然科学基金(14401556, 14471304)资助.

作者简介: 孙达昌, 男, 1992年生, 硕士. 研究方向: 金融工程. E-mail: sdc0124@mail.ustc.edu.cn

通讯作者: 毕秀春, 博士/副教授. E-mail: xcbi@ustc.edu.cn

## 0 引言

随着计算机技术的发展,量化投资在全球逐渐兴起.量化投资是通过分析历史数据,运用数学和统计等知识建立模型,经过严密的回测后,形成交易策略,用于预测证券未来的价格和证券交易.量化投资由计算机发出交易决策,这样极大地减少了投资者情绪波动的影响,有效地避免了投资者的非理性交易决策.交易频率指开仓和平仓的频率,即持有仓位的时间.交易按交易频率分类可以分为高频交易和低频交易.高频交易指通过高速计算能力和高速通信网络,利用分笔数据(tick data)识别捕捉市场中极其细微的价格偏离,寻找交易机会,频繁开仓和平仓的交易,其具有持仓时间短、交易次数多的特点.由于高频交易的持仓时间很短,所以每笔收益通常很低.普通投资者在极短的时间内很难发现盈利机会并进行交易,故高频交易大多通过计算机读取市场行情数据,自动实时地完成交易指令.高频交易在给市场带来波动的同时也会提高市场的效率,增加市场的流动性,提升市场价格的发现能力,更好地处理风险,对市场产生诸多良性作用<sup>[1-5]</sup>.

深度学习的概念来源于学者对人工神经网络(artificial neural network, ANN)的研究,由 Hinton 等在 2006 年提出<sup>[6]</sup>,随后引起了学者们的广泛关注,开启了深度学习研究的热潮.深度学习是机器学习研究中的一个新领域,是模拟人脑分析学习问题的神经网络.从数学模型角度来说,深度学习是一种复杂的特征提取方法,将原始数据经过非线性模型变换成为更高层次的抽象表达,再组合多层变换,学习提取出复杂的特征<sup>[7]</sup>.传统的机器学习算法通常需要进行复杂的特征工程来获取有效特征,提高模型准确率.与传统的机器学习相比,深度学习的优点是使用非监督式或半监督式的特征学习和分层特征提取算法来替代手工获取特征.深度学习算法主要包括卷积神经网络(convolutional neural network, CNN),深度信念网络(deep belief network, DBN),循环神经网络(recurrent neural network, RNN),LSTM(long short-term memory)神经网络等<sup>[8]</sup>.其中 CNN 和 LSTM 神经网络应用广泛且效果显著. CNN 主要应用于图像和自然语言领域中,例如交通标志识别<sup>[9]</sup>,人脸识别<sup>[10]</sup>,行为识别<sup>[11]</sup>,视频分类<sup>[12]</sup>,自动驾驶<sup>[13]</sup>,自然语言理解<sup>[14]</sup>等. LSTM 神经网络适用于处理时序问题,主要应用于机器翻

译<sup>[15]</sup>和语音识别<sup>[16]</sup>等领域.

随着机器学习和深度学习理论与技术的不断发展,越来越多的学者开始使用机器学习和深度学习算法建立量化交易模型,预测证券价格.王宣承<sup>[17]</sup>基于 LASSO 和人工神经网络建立量化交易智能系统,在沪深 300 股指期货上进行回测检验,结果表明该系统可以获取超额收益.张贵勇<sup>[18]</sup>建立卷积神经网络和支持向量机融合的混合模型,预测股票指数和汇率,结果验证了该模型的可行性和有效性. Maknickienė 和 Maknickas<sup>[19]</sup>利用 LSTM 神经网络对欧元兑美元(EURUSD)的汇率进行预测并建立交易策略,结果证明这种预测和交易策略是可行有效的. Persio 和 Honchar<sup>[20]</sup>将 MLP(multi-layer perceptron), CNN 和 RNN 应用于标普 500 指数价格变动的预测中,并提出基于小波和 CNN 的一种新算法,结果发现这种新算法在价格预测中要优于其他神经网络算法. Lu<sup>[21]</sup>运用强化学习和 LSTM 神经网络构建交易模型,并应用在英镑兑美元(GBPUSD)的汇率预测中,结果显示该模型有效地控制了交易次数,保证了盈利的稳定性.龙奥明等<sup>[22]</sup>利用 LSTM 神经网络构建黑色金属期货套利模型,与神经网络套利策略进行对比,结果表明基于 LSTM 神经网络的黑色金属期货套利策略模型可行有效,并比神经网络套利策略表现更好.

目前国内外学者对深度学习算法在高频交易中的应用研究较少,主要是研究神经网络的应用<sup>[23-24]</sup>.范翔<sup>[23]</sup>选取中国证券市场的股票案例对配对模型和协整模型的统计套利进行实例性的论证分析和研究,并进一步研究协整模型的统计套利在高频交易下的实例分析,最后使用神经网络模型进行案例检验,结果发现该模型的预测精度较高,可以获得较高收益. Silva 等<sup>[24]</sup>使用神经网络研究高频交易中的做市过程,通过实证检验神经网络在短时间段(5, 10 和 15min)内的涨幅预测,发现与长周期(10 和 15 min)获得的结果相比,神经网络在短周期(5 min)涨幅预测中效果更好.考虑到深度学习算法的优良性质<sup>[8]</sup>,本文采用深度学习中两种主要的神经网络算法(CNN 和 LSTM 神经网络),使用 Python 语言构建基于 CNN 和 LSTM 神经网络的两种高频交易策略,以上海期货交易所的沥青期货主力合约为例进行回测检验,并与构建的人工神经网络高频交易策略进行对比与分析.

# 1 方法与分类模型构建

本节主要分别介绍三种神经网络的概念与特点,并提出基于这些神经网络的分类模型,模型均用于涨跌预测。

## 1.1 人工神经网络

人工神经网络是模仿生物神经元的原理设计出的一种模型,由大量的节点(神经元)相互连接构成。神经网络具有非线性、自适应、自组织、自学习、非局限性等特点,适用于对非线性复杂函数的映射和对模糊信息的处理<sup>[25]</sup>。在人工神经网络中,神经元是其基本的组成元素。神经元是一个包含输入输出与计算功能的模型,如图 1 所示。

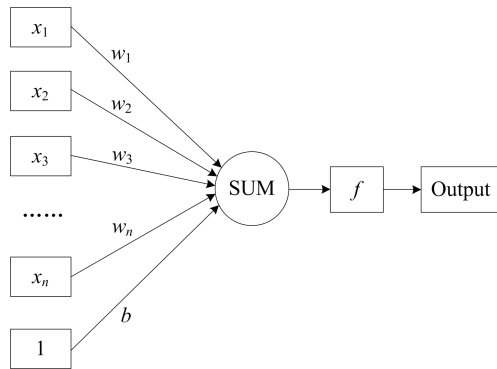


图 1 人工神经元模型

Fig. 1 Artificial neuron model

图 1 中,设  $x_1, x_2, x_3, \dots, x_n, 1$  为神经元模型的输入,  $w_1, w_2, w_3, \dots, w_n, b$  分别为  $x_1, x_2, x_3, \dots, x_n, 1$  的权重,神经元的输出用  $y$  表示,则输入输出的映射关系如下:

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right) \quad (1)$$

式中,  $f(\cdot)$  为激活函数,也称为激励函数。在生物神经元中,当膜电位超过一定的阈值时,神经元处于“兴奋”状态,当膜电位没有超过一定的阈值,神经元处于“抑制”状态。模仿生物神经元的特点,在人工神经网络中,激活函数我们选择阈值函数,也就是大于某个值输出 1,小于等于则输出 0。常用的激活函数有 sigmoid 函数, tanh 函数, softplus 函数, ReLU 函数等。

人工神经网络是由多个神经元连接构成的,一个神经元的输出可以是另一个神经元的输入。人工神经网络模型包括输入层,隐藏层和输出层,如图 2 所示。

图 2 中,我们使用圆圈来表示神经网络的神经

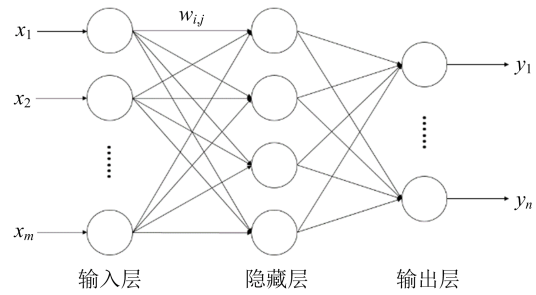


图 2 人工神经网络模型

Fig. 2 Artificial neural network model

元。神经网络最左边的一层称为输入层,最右边的一层称为输出层,中间所有神经元组成的一层称为隐藏层,隐藏的含义是我们不能在训练样本中观测到这些神经元的值。设神经网络一共有  $m$  个输入神经元,  $n$  个输出神经元,  $k$  个隐藏神经元,则隐藏层第  $i$  个神经元的输出为

$$x_i = f\left(\sum_{j=1}^m w_{i,j} x_j + b_i\right) \quad (2)$$

输出层第  $i$  个神经元的输出为

$$y_i = \sum_{j=1}^k w_{i,j} x_j + b_i \quad (3)$$

人工神经网络可以使用误差反向传播(back propagation, BP)<sup>[26]</sup>算法训练模型,计算权重值。BP 算法的主要特点是信号前向传播,误差后向传播,通过不断调整数据的权重,使神经网络的输出与期望输出尽可能接近,达到模型训练的目的。

下面我们构建基于人工神经网络的分类模型,模型包括输入层,隐藏层和输出层,结构如图 3 所示。人工神经网络分类模型包括 1 层输入层, 3 层隐藏层和 1 层输出层。添加 3 层隐藏层的目的是为了模型更好地发现数据中的特征。本文使用 ReLU 函数作为该模型的激活函数,使用随机梯度下降作为该模型的优化器。使用 ReLU 函数可以加快收敛速度,缩短模型训练时间<sup>[27]</sup>。随机梯度下降是指在计算梯度时,随机抽取一部分样本计算该样本的梯度并更新参数,而不是遍历所有样本计算梯度并更新参数,这样可以减少模型的训练时间。

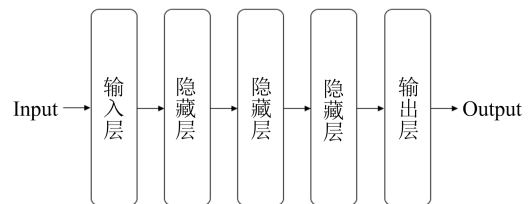


图 3 人工神经网络分类模型结构

Fig. 3 Classification model architecture based on ANN

### 1.2 卷积神经网络

卷积神经网络的概念最早由 Hubel 和 Wiesel 提出,他们通过研究猫的视觉皮层,发现每一个神经元只会处理一小块区域的图像,于是提出感受野(receptive)的概念<sup>[28]</sup>. 1984 年, Fukushima 在感受野的基础之上提出神经认知机(neocognitron)<sup>[29]</sup>. 1998 年, LeCun 等在前人的基础上提出了著名的 LeNet-5 卷积神经网络,并将其成功应用在图像识别中,获得了良好的效果<sup>[30]</sup>.

CNN 是一种前馈神经网络,网络中的神经元可以响应一部分范围内的单元. CNN 中主要包括卷积层和池化层. 卷积层中包含若干个卷积核,每一个卷积核都会提取数据中最有效的特征,得到一类特征的映射,即一个 feature map. 池化层可以对卷积层输出的结果进行池化操作,即降采样,常用的有最大池化和平均值池化. 经过池化层,可以保留最显著的特征,同时降低模型的参数量.

CNN 具有局部连接,权值共享的特点. 传统的人工神经网络中神经元之间为全连接,即每个神经元与前一层全部连接,而 CNN 中的神经元则是与前一层的局部进行连接. 通过局部连接,可以提取局部特征,获取局部信息,然后在更高层将局部信息融合获取全局信息. 权值共享是指卷积层中的卷积核在进行卷积运算时,在原始数据各个位置的权值相等. 权值共享使卷积神经网络含有较少的参数,这样便降低了网络的复杂性.

下面我们构建基于卷积神经网络的分类模型. 如图 4 所示,卷积神经网络分类模型包含 3 层卷积层,3 层池化层,3 层 Dropout 层,1 层 Flatten 层和 1 层 Dense 层. 传统的神经网络具有容易过拟合的缺点,加入 Dropout 层可以有效地防止神经网络过拟合<sup>[31]</sup>. 在模型每一次训练过程中,对于每一个神经网络中的神经元,Dropout 层按照一定的概率  $p$  将其从神经网络中暂时丢弃,当一个神经元被丢弃时,无论输入或者相关的参数是什么,它的输出值都为 0. Flatten 层用于卷积层到 Dense 层的过渡,它可以把多维数据转换为一维数据. Dense 层为全连接层,用于模型的最后一层,经过 Dense 层,模型输出一个  $n \times 2$  矩阵,  $n$  为输入数据的数量,2 代表分类类别数. 输出矩阵中第  $i$  行  $j$  列的元素表示输入数据中第  $i$  个样本预测为类别  $j$  的概率,每一行的元素之和为 1,模型选取概率最高的类别作为预测类别. 该模型中,激活函数使用 ReLU 函数,优化器使用随

机梯度下降.

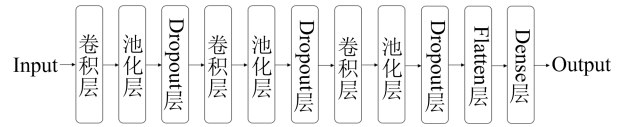


图 4 卷积神经网络分类模型结构

Fig. 4 Classification Model architecture based on CNN

### 1.3 LSTM 神经网络

LSTM 神经网络是循环神经网络<sup>[32]</sup>中的一种,由 Hochreiter 和 Schmidhuber<sup>[33]</sup>在 1997 年提出. Gers 等<sup>[34]</sup>在 2000 年通过引入遗忘门改进了 LSTM 神经网络. LSTM 神经网络可以解决普通循环神经网络的梯度消失和梯度爆炸问题,适用于处理序列长期依赖问题,对序列数据进行预测. 该网络中包含多个判断信息有用与否的 cell,每一个 cell 中有 3 扇门,分别为输入门,遗忘门和输出门. 输入门控制输入信息中有多少可以流入到 cell 中,遗忘门控制上一时刻 cell 中的信息有多少可以累积到当前时刻的 cell 中,输出门控制当前时刻 cell 中的信息有多少可以流入到当前隐藏状态中. LSTM 神经网络的结构如图 5 所示,其中的数学公式如下:

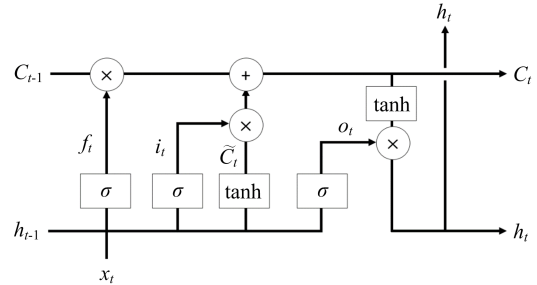


图 5 LSTM 神经网络结构

Fig. 5 Structure of LSTM neural network

输入门:

$$i_t = \sigma(W_{hi}h_{t-1} + W_{xi}x_t + b_c) \quad (4)$$

遗忘门:

$$f_t = \sigma(W_{hf}h_{t-1} + W_{xf}x_t + b_f) \quad (5)$$

输出门:

$$o_t = \sigma(W_{ho}h_{t-1} + W_{x0}x_t + b_o) \quad (6)$$

更新 cell 状态:

$$\tilde{C}_t = \tanh(W_{hc}h_{t-1} + W_{xc}x_t + b_c) \quad (7)$$

最终 cell 状态:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (8)$$

隐藏层:

$$h_t = o_t \times \tanh(C_t) \quad (9)$$

式中,  $\sigma(x) = \frac{1}{1+e^{-x}}$ ,  $\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ ,  $x_t$  表示  $t$  时刻的输入值,  $h_t$  表示  $t$  时刻隐藏状态值,  $\times$  表示元素积,  $W_{ij}$  为连接神经元  $i, j$  的权重值,  $b$  为常数,  $W_{ij}$  和  $b$  可以通过模型训练得到。

下面我们构建基于 LSTM 神经网络的分类模型。模型包括 LSTM 层, Flatten 层和 Dense 层, 结构如图 6 所示。LSTM 神经网络分类模型中的 LSTM 层中包括多个 cell, Flatten 层和 Dense 层的作用与卷积神经网络分类模型中的相同。该模型中, 激活函数使用 ReLU 函数, 优化器使用随机梯度下降。

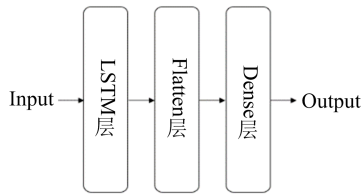


图 6 LSTM 神经网络分类模型结构

Fig. 6 Classification model architecture based on LSTM neural network

## 2 实证检验

本节以期货交易为例来检验我们构建的高频交易策略。首先选取上海期货交易所沥青期货主力合约 tick 级别的行情数据, 并在数据中添加特征, 然后选取训练样本并训练三种分类模型, 随后构建三种交易策略并分别对三者进行回测检验, 最后对这些

回测检验结果进行对比与分析。

构建交易模型的流程图如图 7 所示。构建涨跌分类模型已在上一章中详细介绍, 下面详细介绍图 7 中的其余各个部分并进行回测检验, 对回测检验结果进行对比与分析。

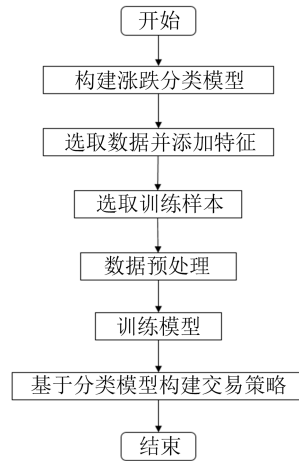


图 7 交易模型流程图

Fig. 7 Trading model flow chart

### 2.1 选取数据与添加特征

tick 级别数据是指每秒两条的数据快照, 本文选取上海期货交易所沥青(BU)期货主力合约 2016-10-10~2017-9-14 期间 tick 级别的行情数据作为实证检验数据, 共 8013191 条, 每一条行情数据都是一维向量, 其中包括价格, 成交量, 持仓量变化, 委买, 委卖, 委比, 委差。其样例如表 1 所示。

表 1 行情数据样例表

Tab. 1 Samples of market data

日期时间	价格	成交量	持仓量变化	委买	委卖	委比	委差
2016-10-10T09:02:10	1 750	38	22	841	147	5.72	694
2016-10-10T09:02:10	1 750	40	34	865	431	2.01	434
2016-10-10T09:02:11	1 752	42	20	847	441	1.92	406
2016-10-10T09:02:11	1 750	262	252	719	561	1.28	158
2016-10-10T09:02:12	1 750	1 144	956	152	558	0.27	-406
2016-10-10T09:02:12	1 750	324	-100	546	469	1.16	77
2016-10-10T09:02:13	1 750	1 100	738	49	475	0.10	-426
2016-10-10T09:02:13	1 748	120	-84	466	217	2.15	249
2016-10-10T09:02:14	1 748	384	-306	524	417	1.26	107
2016-10-10T09:02:14	1 750	80	-56	493	380	1.30	113
2016-10-10T09:02:15	1 750	18	0	494	376	1.31	118
2016-10-10T09:02:15	1 748	40	-40	575	376	1.53	200

选取期货主力合约作为研究对象可以解决期货价格的非连贯性. 本文实证检验数据来源于天软数据库. 在所选取的时间范围内, 沥青期货主力合约行情走势如图 8 所示. 通过图 8 我们可以发现, 沥青期货主力合约的价格波动十分明显, 存在大量获取价差收益的机会.

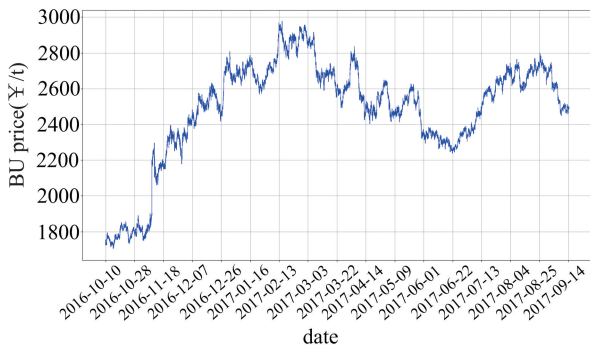


图 8 沥青行情走势  
Fig. 8 BU price chart

为了提高模型的精度, 在选取数据后, 需要向原始 tick 级别行情数据中添加特征. 本文通过计算技术指标, 在每一条行情数据中添加特征. 添加的特征如表 2 所示.

表 2 技术指标与特征个数

Tab. 2 Technical indicators and number of features

技术指标	特征个数
价格 MA5	1
价格 MA10	1
价格 MA20	1
成交量 MA5	1
成交量 MA10	1
成交量 MA20	1
价格涨速	1
成交量涨速	1
KDJ	3
MACD	3
BOLL	2

表 2 中, MA (moving average) 指移动平均, 在本文中是指一定时期内价格或成交量的平均. 价格涨速指当前价格与上一 tick 时价格的比值, 成交量涨速与之同理. KDJ 又称随机指标, 包括 K 值, D 值和 J 值, 计算方法如下:

$$RSV_t = (Close - Low_n) / (High_n - Low_n) \times 100 \quad (10)$$

$$K_t = SMA(RSV_t, m_1, 1) \quad (11)$$

$$D_t = SMA(K_t, m_2, 1) \quad (12)$$

$$J_t = 3 \times K_t - 2 \times D_t \quad (13)$$

式中, Close 指收盘价,  $Low_n$  指  $n$  个周期的最低价,  $High_n$  指  $n$  个周期的最高价, SMA (simple moving average) 是简单移动平均,  $n, m_1, m_2$  为参数, 本文设定  $n=9, m_1=3, m_2=3$ . MACD 称为指数平滑移动平均线, 包括 DIF 值, DEA 值和 MACD 值, 计算方法如下:

$$DIF = EMA(Close, Short) - EMA(Close, Long) \quad (14)$$

$$DEA = EMA(DIF, Mid) \quad (15)$$

$$MACD = (DIF - DEA) \times 2 \quad (16)$$

式中, Close 指收盘价, EMA (exponential moving average) 是指数平均, Short, Long, Mid 为参数, 本文设定 Short=12, Long=26, Mid=9. BOLL 指布林线指标, 包括 BOLL 值, UB 值和 LB 值, 计算方法如下:

$$BOLL = MA(Close, M) \quad (17)$$

$$UB = BOLL + 2 \times STD(Close, M) \quad (18)$$

$$LB = BOLL - 2 \times STD(Close, M) \quad (19)$$

式中, Close 指收盘价, STD 是标准差函数,  $M$  为参数, 本文设定  $M=20$ . 添加特征后, 每一条行情数据中都含有 23 维特征.

## 2.2 选取训练样本并进行数据预处理

因为只有上涨或者下跌才可以通过期货合约价差获取收益, 所以我们只选择未来上涨或者下跌的行情数据作为训练样本. 本文定义 10 s 后价格涨幅大于 1% 为上涨 (类别 1); 10 s 后价格跌幅大于 1% 为下跌 (类别 2), 在前 1000000 条行情数据中, 选择满足上涨, 下跌定义的 20 条 tick 行情数据构成的矩阵作为训练样本. 选择涨幅 (跌幅) 大于 1% 为上涨 (下跌) 是因为这样可以获得更多的训练样本, 对于机器学习和深度学习来说, 大量的训练样本可以提高模型的泛化能力. 同时, 在扣除手续费后, 如果预测正确, 价格变动 1% 可以保证交易盈利. 本文共选取 186270 个上涨训练样本个体, 184673 个下跌训练样本个体.

在选取训练样本后, 需要对数据进行预处理. 为了消除计量单位对训练和预测结果的影响, 提高模型的精度, 缩短模型的训练时间, 需要对训练样本数据中的每一列数据进行标准化和归一化处理. 本文标准化处理公式如下:

$$y_{ij} = (x_{ij} - \text{mean}_j) / \text{std}_j \quad (20)$$

式中,  $x_{ij}$  为原始数据中第  $i$  行第  $j$  列的元素,  $\text{mean}_j$  为原始数据中第  $j$  列的均值,  $\text{std}_j$  为原始数据中第  $j$  列的标准差,  $y_{ij}$  为标准化后第  $i$  行第  $j$  列的元素. 进行标准化处理之后, 需要对数据的每一列进行归一化处理. 归一化处理公式如下:

$$z_{ij} = (y_{ij} - \min_j) / (\max_j - \min_j) \quad (21)$$

式中,  $y_{ij}$  为标准化后的数据中第  $i$  行第  $j$  列的元素,  $\min_j$  为标准化后数据中第  $j$  列的最小值,  $\max_j$  为标准化后数据中第  $j$  列的最大值,  $z_{ij}$  为归一化后第  $i$  行第  $j$  列的元素. 进行归一化后, 所有的数据元素都在  $[0, 1]$  内. 使用预处理过后的训练样本分别对三种分类模型进行训练.

### 2.3 构建交易策略与回测检验

使用训练完后的三种分类模型在回测数据集上分别进行回测检验. 本文选取数据中后 7013191 条行情数据作为回测数据集. 训练后的模型进行涨跌预测时, 会输出预测为某一类别的概率值. 本文构建的交易策略如下: 在回测时间区间内任意 tick 时间节点下, 当模型预测当前行情数据属于类别 1 (上涨) 的概率大于  $p$  ( $p$  为交易策略参数) 时, 进行模拟开多仓交易, 开仓价格为当前价格, 成交量为 1 手, 在开仓 10 s 后进行模拟平仓交易; 当模型预测当前行情数据属于类别 2 (下跌) 的概率大于  $p$  时, 进行模拟开空仓交易, 开仓价格为当前价格, 成交量为 1 手, 在开仓 10 s 后进行模拟平仓交易. 在进行涨跌预测时, 输入的行情数据也需进行标准化和归一化处理. 模拟交易的平仓价格均为平仓时的当前价格. 通过查找上海期货交易所提供的信息, 原油期货的交易手续费为成交金额的 0.1‰, 由于期货公

司会对客户收取交易佣金, 实际的交易手续费要比成交金额的 0.1‰ 高. 本文在进行模拟交易时使用成交金额的 0.15‰ 作为交易手续费, 开仓和平仓交易均需要支付交易手续费. 模拟平仓后计算本次模拟交易的收益率, 做多交易的收益率的计算公式为

$$\frac{P_2 \times \text{Vol} - P_1 \times \text{Vol} - P_1 \times \text{Vol} \times s - P_2 \times \text{Vol} \times s}{P_1 \times \text{Vol}} \quad (22)$$

做空交易的收益率的计算公式为

$$\frac{P_1 \times \text{Vol} - P_2 \times \text{Vol} - P_1 \times \text{Vol} \times s - P_2 \times \text{Vol} \times s}{P_1 \times \text{Vol}} \quad (23)$$

式中,  $P_1$  为开仓价格,  $P_2$  为平仓价格,  $\text{Vol}$  为成交量,  $s$  为手续费比例, 本文设定  $s = 0.00015$ .

回测检验结束后, 需要计算胜率, 平均盈利, 平均亏损, 盈亏比和期望收益. 胜率是指所有交易中盈利的交易次数占总交易次数的比率. 平均盈利指在所有盈利的交易中一笔交易的收益率的平均值. 平均亏损指在所有亏损的交易中一笔交易的收益率的平均值. 通过定义可知, 平均盈利大于 0, 平均亏损小于 0. 盈亏比是指平均盈利与平均亏损的比值的绝对值. 期望收益是指一笔交易的收益的期望值, 其计算公式如下:

$$P_w \times W + (1 - P_w) \times L \quad (24)$$

式中,  $P_w$  为胜率,  $W$  为平均盈利,  $L$  为平均亏损.

### 2.4 回测检验结果与对比分析

由于  $p$  为交易策略参数, 故本文需在  $p$  为不同值的情况下对三种策略进行回测检验, 结果如表 3~5 所示.

表 3 神经网络高频交易策略的回测结果

Tab. 3 Back-testing results of high-frequency trading strategy based on ANN

$p$	交易次数	胜率	平均盈利	平均亏损	盈亏比	期望收益
0.50	7 013 151	39.10%	0.5588‰	-0.4313‰	1.30	-0.0441‰
0.55	6 807 015	39.61%	0.5584‰	-0.4273‰	1.31	-0.0368‰
0.60	6 593 208	40.64%	0.5581‰	-0.4234‰	1.32	-0.0245‰
0.65	6 367 272	41.21%	0.5578‰	-0.4197‰	1.33	-0.0168‰
0.70	6 117 356	41.20%	0.5574‰	-0.4157‰	1.34	-0.0147‰
0.75	5 830 391	41.82%	0.5572‰	-0.4117‰	1.35	-0.0065‰
0.80	5 484 446	42.53%	0.5569‰	-0.4074‰	1.37	0.0027‰
0.85	5 041 836	43.42%	0.5567‰	-0.4025‰	1.38	0.0139‰
0.90	4 415 969	44.64%	0.5567‰	-0.3967‰	1.40	0.0288‰
0.95	3 357 756	46.70%	0.5573‰	-0.3880‰	1.44	0.0534‰

表 4 卷积神经网络高频交易策略的回测结果

Tab. 4 Back-testing results of high-frequency trading strategy based on CNN

$p$	交易次数	胜率	平均盈利	平均亏损	盈亏比	期望收益
0.50	7 013 151	40.79%	0.5592‰	-0.4102‰	1.36	-0.0147‰
0.55	6 748 041	41.45%	0.5579‰	-0.4025‰	1.39	-0.0044‰
0.60	6 633 440	41.78%	0.5577‰	-0.4002‰	1.39	0.00001‰
0.65	6 491 209	42.13%	0.5568‰	-0.3966‰	1.40	0.0050‰
0.70	6 198 399	42.80%	0.5548‰	-0.3899‰	1.42	0.0144‰
0.75	5 758 877	43.96%	0.5546‰	-0.3843‰	1.44	0.0284‰
0.80	5 631 410	44.26%	0.5538‰	-0.3818‰	1.45	0.0322‰
0.85	4 740 119	46.52%	0.5523‰	-0.3708‰	1.49	0.0586‰
0.90	4 393 566	47.39%	0.5524‰	-0.3680‰	1.50	0.0681‰
0.95	3 043 224	51.10%	0.5534‰	-0.3575‰	1.55	0.1079‰

表 5 LSTM 神经网络高频交易策略的回测结果

Tab. 5 Back-testing results of high-frequency trading strategy based on LSTM neural network

$p$	交易次数	胜率	平均盈利	平均亏损	盈亏比	期望收益
0.50	7 013 150	38.57%	0.5569‰	-0.4387‰	1.27	-0.0546‰
0.55	6 668 948	39.36%	0.5561‰	-0.4330‰	1.28	-0.0436‰
0.60	6 258 512	40.84%	0.5562‰	-0.4268‰	1.30	-0.0253‰
0.65	5 814 531	42.56%	0.5565‰	-0.4218‰	1.32	-0.0054‰
0.70	5 611 819	44.09%	0.5571‰	-0.4180‰	1.33	0.0119‰
0.75	5 227 912	45.65%	0.5580‰	-0.4135‰	1.35	0.0299‰
0.80	5 031 622	47.40%	0.5587‰	-0.4083‰	1.37	0.0500‰
0.85	4 571 186	49.57%	0.5595‰	-0.4013‰	1.39	0.0749‰
0.90	4 025 918	52.37%	0.5604‰	-0.3916‰	1.43	0.1069‰
0.95	2 972 234	56.45%	0.5605‰	-0.3761‰	1.49	0.1526‰

表 3~5 中的胜率,平均盈利,平均亏损,盈亏比和期望收益均是判别投资策略优良的评价指标.策略的预测能力可以用胜率这一指标来衡量,策略的胜率越高,策略的预测能力就越强.策略的盈利能力可以用期望收益这一指标来衡量,策略的期望收益越大,策略的盈利能力就越强.从表 3~5 中我们可以看出,对于三种高频交易策略,随着  $p$  的增大,策略的胜率和期望收益都随之提高,策略的预测能力和盈利能力也都随之增强.

接下来我们比较三种高频交易策略的预测能力和盈利能力.如图 9 和 10 所示,在相同  $p$  值的情况下,卷积神经网络策略的胜率和期望收益都要比人工神经网络策略的高,即卷积神经网络策略的预测

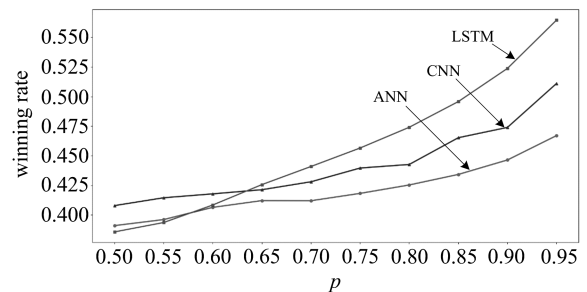


图 9 三种策略的胜率

Fig. 9 Winning rates of three strategies

能力和盈利能力都要比人工神经网络策略的好.当  $p$  较小时,LSTM 神经网络策略的表现不如另外两种策略,但是随着  $p$  增大,LSTM 神经网络策略的表现变好.当  $p > 0.65$  时,LSTM 神经网络策略的



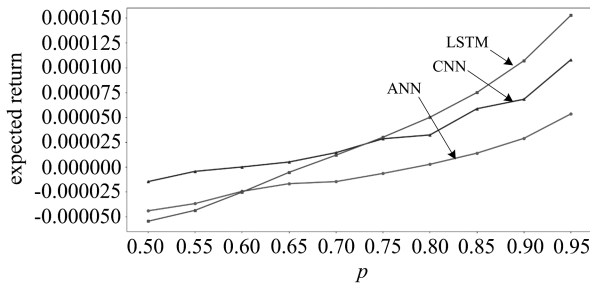


图 10 三种策略的期望收益

Fig. 10 Expected returns of three strategies

胜率最高. 当  $p > 0.75$  时, LSTM 神经网络策略的期望收益最高. 当  $p = 0.95$  时, LSTM 神经网络策略的胜率为 56.45%, 期望收益为 0.1526‰, 卷积神经网络策略的胜率为 51.10%, 期望收益为 0.1079‰, 而人工神经网络策略的胜率为 46.70%, 期望收益仅为 0.0534‰, 预测和盈利效果明显不如前两种策略. 同人工神经网络策略相比, 在  $p = 0.95$  时, 卷积神经网络策略的胜率提高了 4.4%, 期望收益大幅提升, 是人工神经网络的 2.02 倍; LSTM 神经网络策略的胜率提高了 9.7%, 期望收益是人工神经网络的 2.86 倍, 为三者中最高. 在策略的实际应用中, 可以将  $p$  值设为 0.95, 这样可以保证收益的稳定性.

### 3 结论

随着计算机技术与人工智能技术的不断发展和我国资本市场的不断壮大, 量化投资逐渐兴起, 并有取代人为主观投资的趋势. 人工智能技术已在图像处理、语音识别和机器翻译等领域广泛应用, 本文将其在期货高频交易中. 本文基于人工神经网络、卷积神经网络和 LSTM 神经网络提出三种高频交易策略, 并以沥青期货主力合约为例进行回测检验, 结果显示基于卷积神经网络和 LSTM 神经网络的高频交易策略效果较好, 当  $p > 0.65$  时, 两种策略的胜率和期望收益均比基于人工神经网络的策略高; 当  $p = 0.95$  时, LSTM 神经网络策略的胜率和期望收益最高, 卷积神经网络策略次之, 这证明了基于深度学习算法的高频交易策略的有效性和较高的盈利能力. 在实际应用中, 深度学习算法模型可以通过在训练集上训练模型, 然后将训练后的模型应用在其他期货品种或者股票当中. 投资者可以使用训练后的量化交易模型进行计算机程序化交易, 代替人为主观交易, 这样可以减少情绪对投资者的不利影响, 避免投资者的不理智行为, 并为投资者带来高

额的收益. 本文的不足在于只对沥青期货进行了回测检验, 没有对其他期货品种进行回测检验, 不能充分说明深度学习应用在期货高频交易中的有效性.

深度学习算法不仅适用于高频率短周期 (tick 级别) 的量化交易, 在未来的研究中, 也可以将深度学习算法应用在低频率中长周期 (日级别) 的量化交易模型中. 同时, 在模型的训练数据选取中, 不仅可以选取技术指标, 也可以选择宏观经济指标、公司财务指标和基本面指标等; 在交易策略中, 可以添加止盈条件和止损条件, 控制仓位, 丰富交易框架与逻辑.

### 参考文献 (References)

- [1] BROGAARD J A. High frequency trading and its impact on market quality [R]. Evanston, IL: Northwestern University, 2010.
- [2] ALDRIDGE I. High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems [M]. Hoboken, NJ: Wiley, 2010.
- [3] MARTINEZ V H, ROSU I. High Frequency Traders, News and Volatility [C]// AFA 2013 San Diego Meetings Paper. Aldan, PA: American Finance Association, 2013.
- [4] BROGAARD J, HENDERSHOTT T, RIORDAN R. High frequency trading and price discovery [J]. The Review of Financial Studies, 2014, 27(8): 2267-2306.
- [5] ANGEL J, MCCABE D. Fairness in financial markets: The case of high frequency trading [J]. Journal of Business Ethics, 2013, 112: 585-595.
- [6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18: 1527-1554.
- [7] 毛勇华, 桂小林, 李前, 等. 深度学习应用技术研究 [J]. 计算机应用研究, 2016, 33(11): 3201-3205.  
MAO Yonghua, GUI Xiaoling, LI Qian, et al. Study on application technology of deep learning [J]. Application Research of Computers, 2016, 33(11): 3201-3205.
- [8] LECUN Y, BENGIO Y, HINTON G E. Deep learning [J]. Nature, 2015, 521(7553): 436.
- [9] CIRESAN D, MEIER U, MASCI J, et al. Multi-column deep neural network for traffic sign classification [J]. Neural Networks, 2012, 32: 333-338.
- [10] TAIGMAN Y, YANG M, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern

- Recognition. Washington, DC: IEEE Computer Society, 2014: 1939-1946.
- [11] JI S, XU W, YANG M, et al. 3D convolutional neural networks for human action recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 221-231.
- [12] KARPATY A, TODERICI G, SHETTY S, et al. Pedestrian detection with unsupervised multi-stage feature learning[C]// Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2013: 3626-3633.
- [13] HADSELL R, SERMANET P, BEN J, et al. Learning long-range vision for autonomous off-road driving[J]. Journal of Field Robotics, 2009, 26(2): 120-144.
- [14] COLLOBERT R, WESTON J, BOTTOU L, et al. Natural language processing (almost) from scratch[J]. Journal of Machine Learning Research, 2011, 12: 2493-2537.
- [15] BAHDANAU D, CHO K, BENGIOY. Neural machine translation by jointly learning to align and translate[EB/OL]. [2016-5-19]. <https://arxiv.org/abs/1409.0473>.
- [16] WU Y, SCHUSTER M, CHENZ, et al. Google's neural machine translation system: Bridging the gap between human and machine translation [EB/OL]. [2016-10-08] <https://arxiv.org/abs/1609.08144>.
- [17] 王宣承. 基于 LASSO 和神经网络的量化交易系统构建—以沪深 300 股指期货为例[J]. 投资研究, 2014, 33(9): 23-29.  
WANG Xuancheng. Construct intelligent quantitative trading systems based on LASSO and ANNs: A case study of CSI300 futures [J]. Review of Investment Studies, 2014, 33(9): 23-29.
- [18] 张贵勇. 改进的卷积神经网络在金融预测中的应用研究[D]. 郑州: 郑州大学, 2016.
- [19] MAKNIKIENĖ N, MAKNICKAS A. Application of neural network for forecasting of exchange rates and Forex trading [C]// The 7th International Scientific Conference "Business and Management 2012". Vilnius, Lithuanian; Vilnius Gediminas Technical University, 2012: 122-127.
- [20] PERSIO L D, HONCHAR O. Artificial neural networks architectures for stock price prediction: Comparisons and applications[J]. International Journal of Circuits, Systems and Signal Processing, 2016, 10: 403-413.
- [21] LU D W. Agent inspired trading using recurrent reinforcement learning and LSTM neural networks [EB/OL]. [2017-07-23] <https://arxiv.org/abs/1707.07338>.
- [22] 龙奥明, 毕秀春, 张曙光. 基于 LSTM 神经网络的黑色金属期货套利策略模型[J]. 中国科学技术大学学报, 2018, 48(2): 125-132.  
LONG Aoming, BI Xiuchong, ZHANG Shuguang. An arbitrage strategy model for ferrous metal futures based on LSTM neural network [J]. Journal of University of Science and Technology of China, 2018, 48(2): 125-132.
- [23] 范翔. 基于自动化交易平台的高频交易及统计套利分析和研究[D]. 上海: 复旦大学, 2014.
- [24] SILVA E, CASTILHO D, PEREIRA A, et al. A neural network based approach to support the market making strategies in high-frequency trading[C]// 2014 International Joint Conference on Neural Networks. IEEE, 2014: 845-852.
- [25] 张德丰. MATLAB 神经网络应用设计[M]. 北京: 机械工业出版社, 2012.
- [26] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature, 1986, 323(6088): 533-536.
- [27] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C]// Advances in Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2012: 1097-1105.
- [28] HUBEL D H, WIESEL T N. Receptive fields, binocular and functional architecture in the cat's visual cortex[J]. The Journal of Physiology, 1962, 160(1): 106-154.
- [29] FUKUSHIMA K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. Biological Cybernetics, 1980, 36(4): 193-202.
- [30] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [31] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[EB/OL]. [2012-07-03] <https://arxiv.org/abs/1207.0580>.
- [32] GRAVES A. Supervised Sequence Labelling with Recurrent Neural Networks [M]. Berlin: Springer, 2012.
- [33] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [34] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: Continual prediction with LSTM [J]. Neural Computation, 2000, 12(10): 2451-2471.