

深度学习框架下 LSTM 网络的期货价格预测模型研究

Research on Futures Price Prediction Model of LSTM Network Under the Framework of
Deep Learning

王毓谦

(南京财经大学 管理科学与工程学院, 南京 210046)

WANG Yu-qian

(School of Management Science and Engineering, Nanjing University of Finance and Economics, Nanjing 210046, China)

【摘要】期货市场是一个复杂系统,其价格数据具有非线性、高噪声等特性,导致其预测难度较高。论文结合深度学习框架下的神经网络模型对选取的沪深 300 指数期货开盘、收盘价数据进行分析并预测趋势,最终得到较高的预测精度。论文表明,深度学习循环神经网络在期货价格预测方面有着不错的表现并且发展潜力较大。

【Abstract】Futures market is a complex system, and its price data is nonlinear and noisy, which makes it difficult to predict. Combined with the neural network model under the framework of deep learning, this paper analyzes and predicts the trend of the selected opening and closing prices of CSI 300 index futures, and finally obtains a higher prediction accuracy. The paper shows that deep learning recurrent neural network has a good performance in futures price prediction and has great development potential.

【关键词】深度学习;神经网络;长短期记忆;期货市场;价格预测

【Keywords】deep learning; neural network; long and short-term memory; futures market; price prediction

【中图分类号】F724.5;TP183

【文献标志码】A

【文章编号】1673-1069(2021)12-0137-03

1 引言

随着经济的不断发展、理财类产品的不断宣传,各种类型的理财产品逐渐在人们的视野中出现,对资产进行合理配置,让钱变得值钱,使得自己拥有的财富能够在价值上得到最大的提升,引起了我国国民的广泛关注。在一系列的投资交易方式之中,量化交易脱颖而出,成为新时代的宠儿。量化交易也被称为算法交易,顾名思义就是依靠预先编写好的代码由计算机进行拟合预测,通过预先设计的算法,获得交易策略的过程。量化交易采用量化模型代替人为的主观判断,减少了人为因素带来的损失。机器学习作为人工智能中的重要发展领域,现如今发展很快,得到各界人士的认可。深度学习的概念源于人工神经网络的研究,其希望通过创立神经元来模拟人脑的运作机制,从而解释数据并且处理数据,其特点就是模型结构的深度,通常深度学习构建的神经网络能有较高深度。本文根据上述背景,进行深度学习神经网络应用在量化交易策略上的研究。

目前学术界已有相关研究,郭朋(2012)第一次在国内提出了量化交易这一概念,并且对国外的研究进行了综述,从

那之后,国内开始了对量化交易的研究。王新华(2017)在其论文之中提出将人工智能运用到金融领域之中,其在论文中提到,金融领域最多的便是数据,而人工智能则拥有很多处理数据的方式,所以人工智能的相关方法可以在这个领域有很多的运用。胡叶帅(2021)在其论文中提到将深度学习运用到量化交易策略之中,深度学习的特点就是将大量的数据集带入设计好的模型之中,在图像识别领域发展得很好,但是在处理时序金融数据这方面的研究还不是很多。王鑫(2018)将深度学习神经网络中的 LSTM 循环神经网络运用到系统故障预测之中,证明了 LSTM 在时序数据方面的可行性。

文章基于研究现状,以沪深 300 指数的主力合约(合约代码 IF9999)为研究对象,通过深度学习神经网络进行数据预测,预测出未来合约交易的趋势,并进行策略研究,结果证明深度学习神经网络在期货交易价格预测问题方面具有较高的实用性。

2 期货交易价格数据描述

目前资本市场的股票、期货数据获取渠道相对较多,数据的准确度较高,为量化交易行业研究奠定一个较好的基础,文章的期货数据主要是通过聚宽(www.joinquant.com)专业量化投资平台获取,获取的数据主要是开盘价(open)、收盘价(close)、最高价(high)、最低价(low)、成交量(volume)以及成交金额(money)6个常见指标。具体数据形式如表 1 所示。

【基金项目】江苏省研究生科研创新项目(KYCX20_1351):深度学习框架下 LSTM 网络的期货价格预测模型研究。

【作者简介】王毓谦(1997-)男,江苏南通人,研究生在读,从事管理科学、量化交易研究。

表 1 期货交易数据指标

datetime	open	close	high	low	volume	money
2019/9/2 9:31	3793.6	3790.6	3793.6	3788.6	1036	1178341560
2019/9/2 9:32	3791.2	3796.2	3798.2	3790.6	921	1048761060
2019/9/2 9:33	3796.6	3802	3803.4	3795.2	798	909753060
2019/9/2 9:34	3802.8	3806.2	3808	3801.8	779	889472820
2019/9/2 9:35	3806.4	3809.4	3810	3805.8	649	741350160
2019/9/2 9:36	3809.4	3817.8	3818.8	3808.8	1142	1306705200

期货交易市场上午 9:00 开盘,10:15~10:30, 休息 15min,11:30 收盘。下午 13:30 开盘,15:00 收盘。晚上 21:00 开盘,23:00 收盘。文章考虑到数据集的大小问题,选取 2019~2021 年沪深 300 指数主力合约(合约代码 IF9999)的每分钟数据进行研究,原因是沪深 300 指数主力合约属于热门期货交易合约,每分钟的成交量较多,并且成交金额较大,这样有利于学习模型学习到数据趋势,并且沪深 300 指数是国内期货交易市场的主要合约,具有一定的代表性,从某种方面讲具有一定的普适性,在这个合约里能够预测出的结果在其他的合约上也能够实现。

文章获取的数据存在大量缺失值以及噪声数据,文章使用基本的数据处理方式对数据进行处理,对于缺失的数据,我们使用就近原则,取其前一组数据以及后一组数据的平均值将其补上,使得数据的偏差幅度不会过大影响到后面的数据预测。对于偏离整体趋势的数据进行删除并且修正。最后对数据进行归一化处理,并将其做成 64 个数据为 1 组的 batch 型数据方便导入模型之中使用。

3 神经网络选取以及模型搭建

深度学习之中基础的线性回归模型以及卷积神经网络模型一般用于处理数值信息以及将图像信息转化为矩阵模式进行模型迭代计算,而要适配文章所描述的金融时序数据,需要使用新的模型。数据会根据时间进程而进行相应的变化,所以使用新的神经网络模型——循环神经网络。

在神经网络模型之中,有一类含有隐藏层的多层感知机。在对于小批量数据样本 $X \in R^{n \times d}$, 隐藏层的激活函数为 ϕ , 偏差函数为 b , 权重参数为 W , 那么隐藏层输出值 $H \in R^{n \times d}$ 的计算方法为 $H = \phi(XW_{xh} + b_h)$ 。

最终输出的结果为 $O = HW_{ho} + b_o$ 。

一般的多层感知机用于线性回归模型以及卷积块模型,是一个静态模型,并没有加入过时间变量,所以需要用到循环神经网络,它是通过隐藏状态来储存之前的时间段 t 内的信息,并且参与到下一次的迭代之中,对下一次迭代的结果造成影响。

现在考虑输入的数据存在时间相关性的情况,输入的数据样本需要添加一个下标 t 使之变为 $X_t \in R^{n \times d}$, 其代表在序列之中时间为 t 的时候小批量输入的数据值,另外有 H_t 表示时间为 t 的时候的隐藏变量,并且引入新的权重参数 $W_{hh} \in R^{d \times d}$, 其大小表示 $t-1$ 时间段的数据影响 t 时间数据的程度。在这种情况下,隐藏层的隐藏状态输出变化为:

$$H_t = \phi(X_t W_{xh} + b_h + H_{t-1} W_{hh})$$

为解决梯度问题、模型效率不高的问题,使用门控循环神经网络的思路,这种神经网络可以更好地捕捉时间序列之中时间步之间的依赖关系。近年来对于处理时间序列问题,有一类较为热门的神经网络模型——长短期记忆循环神经网络模型。

输入门、遗忘门、输出门的输入是当前时间段的输入 X_t 以及上一个时间段的隐藏状态 H_{t-1} , 通过 sigmoid 函数作为激活函数计算得到。假设时间步 t 的小批量输入 $X_t \in R^{n \times d}$ 以及上一个时间步输出的隐藏状态 $H_{t-1} \in R^{n \times d}$; W_{xx} 与 b_x 为各个门对应的权重参数以及偏差参数,输入门 I_t 、遗忘门 F_t 以及输出门 O_t 的计算方式如下:

$$I_t = \sigma(X_t W_{xi} + b_i + H_{t-1} W_{hi})$$

$$F_t = \sigma(X_t W_{xf} + b_f + H_{t-1} W_{hf})$$

$$O_t = \sigma(X_t W_{xo} + b_o + H_{t-1} W_{ho})$$

计算完输出门、遗忘门以及输出门之后,需要再计算候选记忆细胞 \tilde{C}_t , \tilde{C}_t 使用 tanh 函数作为激活函数进行计算:

$$\tilde{C}_t = \tanh(X_t W_{xc} + b_c + H_{t-1} W_{hc})$$

最终每个神经元内结构如图 1 所示。

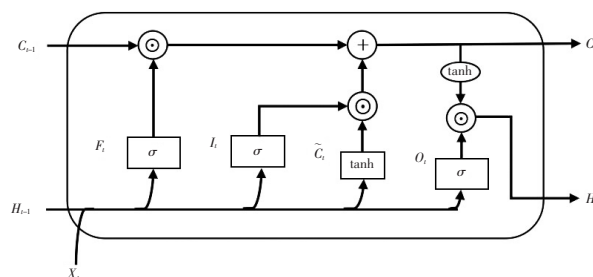


图 1 网络结构图

以上为文章选取的 LSTM 模型的原理, 本文将以此搭建 LSTM 神经网络模型, 并对上文截取的数据进行分析以及研究。

文章在 Windows10 操作系统下搭建 GPU 版本的 LSTM 神经网络, 使用 python 语言作为编译语言, pytorch 包作为搭载。pytorch 包具有灵活性高、扩展性强等优点, 通过继承父类模型来搭建自定义模型, 编写 evaluate 函数对数据进行迭代, 最终得出权重参数。文章采用 model 类, 搭建几种不同的神经网络模型对一支期货进行收盘价的分析预测, 对模型预测准确性使用平方根误差、平均绝对误差进行结果对比。

图 2 展示了搭建的部分神经网络模型: 卷积神经网络、全连接神经网络、长短期记忆神经网络。

```

MyLstm(
  (layer1): LSTM(2, 4, num_layers=2)
  (layer2): Linear(in_features=4, out_features=1, bias=True)
)

MyDense(
  (model): Sequential(
    (0): Linear(in_features=30, out_features=128, bias=True)
    (1): ReLU()
    (2): Linear(in_features=128, out_features=256, bias=True)
    (3): ReLU()
    (4): Linear(in_features=256, out_features=1, bias=True)
  )
)

MyCnnModel(
  (conv1): Sequential(
    (0): Conv2d(1, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU()
    (2): AvgPool2d(kernel_size=2, stride=2, padding=0)
  )
  (conv2): Sequential(
    (0): Conv2d(16, 64, kernel_size=(2, 2), stride=(1, 1))
    (1): ReLU()
    (2): AvgPool2d(kernel_size=2, stride=2, padding=0)
  )
)

```

图 2 搭建的部分网络结构代码

4 实验结果以及分析

根据上文所构建的模型以及获取到的数据,文章基于python 语言对长短期记忆神经网络模型进行代码实现,并且将数据集以 8:2 的比例分为训练集以及测试集,以方便进行模型的训练以及数据预测。

文章使用沪深 300 指数主力合约从 2019 年 9 月~2021 年 9 月的每分钟数据进行数据预测,共有 84864*9 个数据,文章选取期货数据开盘价进行数据预测,并且已经进行数据的预处理(见图 3)。

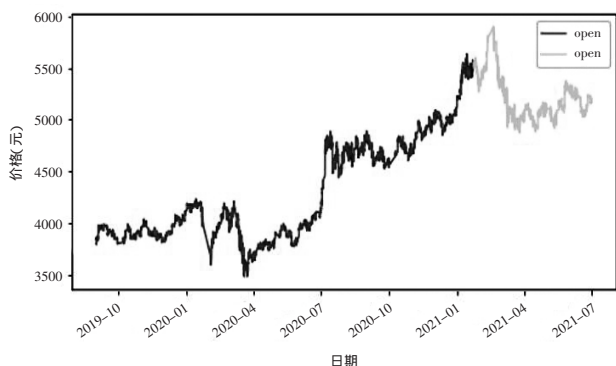


图 3 期货开盘价数据

文章遵循神经网络训练的基本规则,按照 8:2 的比例进行数据集的切分,对 2021 年 2 月之前的数据归为一类作为训练集,将 2021 年 2 月之后的数据归为一类作为测试集,如图 3 所示,训练集的数据由黑色线表示,接着之后的测试集数据由灰色线表示。

如图 2 所示模型以及参数,本文搭建 LSTM 神经网络、卷积神经网络、全连接多层感知器几类模型进行数据的分析以及预测。

在每个模型进行 1000 次迭代后,得出的预测精度如表 2 所示。

表 2 模型预测结果

模型名称	精确性(acc)	误差值(loss)
多层感知器	0.623	0.12
卷积神经网络	0.712	0.09
LSTM 神经网络	0.791	0.016

从表 1 可以看出,LSTM 神经网络的预测精度较好并且稳定性较强,于是使用 LSTM 进行数据的预测并使用 matplotlib 包进行绘图(见图 4)。

如图 4 所示,LSTM 神经网络对期货开盘价数据进行了一次较好的拟合,虽然与真实数据仍有差距,但依旧能展示出大概的趋势。

5 结语与展望

文章通过代码构建神经网络模型实现期货数据预测,并且经过多个模型的对比选取较为精确的网络,虽仍有误差,但是能够对整体趋势有一个较好的预测,对于投资者而言有

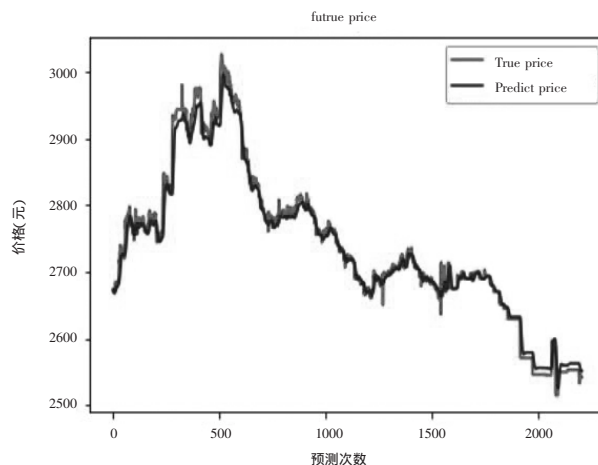


图 4 真实值以及预测值

参考价值,文章仍存在以下不足之处:

文章中采用的模型是最基础的神经网络模型,模型深度并没有达到多深,并且使用的超参数都是默认参数,没有进行修改,在之后的文章中可以使用网格排序法对每种超参数进行枚举,并得出最优的参数,并且对模型的结构进行调整,找到一个最适合的模型结构。本文使用的是滑动窗口式的预测方法,容易造成数据误差的累计,导致长期的模型预测精度较低,固考虑采用 S2S 的搭建模型与读取数据方式,加入注意力机制让模型预测精度有更多的提升。

本文仅通过数据预测得出开盘价未来变动的趋势,未来可以考虑在此数据基础上使用量化交易中的策略进行期货交易,并预估收益,这都是未来文章研究的方向。

【参考文献】

- [1]郭朋.国外高频交易的发展现状及启示[J].证券市场导报,2012(07):56-61.
- [2]王新华,肖波.人工智能及其在金融领域的应用[J].银行家,2017(12):126-128.
- [3]胡叶帅.期货深度量化交易策略的研究[D].沈阳:辽宁大学,2021.
- [4]王鑫,吴际,刘超,等.基于 LSTM 循环神经网络的故障时间序列预测[J].北京航空航天大学学报,2018,44(04):772-784.
- [5]岑跃峰,张晨光,岑岗,等.基于近端强化学习的股价预测方法[J].控制与决策,2021,36(04):967-973.
- [6]余凯,贾磊,陈雨强,等.深度学习的昨天、今天和明天[J].计算机研究与发展,2013,50(09):1799-1804.
- [7]Wen Hu,Yuxue Shi. Prediction of online consumers' buying behavior based on LSTM-RF model [A]. Sichuan University. Proceedings of the 5th International Conference on Communication, Image and Signal Processing (CCISP 2020)[C].Sichuan University:成都夏洛克教育咨询有限公司,2020:5.
- [8]孙志军,薛磊,许阳明,等.深度学习研究综述[J].计算机应用研究,2012,29(08):2806-2810.
- [9]李泽艳,陈银钧.LSTM Deep Learning Stock Prediction System Based on PyTorch Framework[J].运筹与模糊学,2021,11(02):137-146.