



一种基于堆叠式自动编码器和长时记忆的金融时间序列深度学习框架

魏宝， 君岳， 饶玉蕾

发布日期：2017年7月14日 <https://doi.org/10.1371/journal.pone.0180944>

摘要

深度学习方法在金融领域的应用受到了投资者和研究者的广泛关注。本文提出了一种新的深度学习框架，将小波变换(WT)、叠加自动编码器(SAES)和长期记忆(LSTM)相结合，用于股票价格预测。首次将层次化提取深度特征的SAE引入到股票价格预测中。深度学习框架由三个阶段组成。首先，对股票价格时间序列进行小波分解，以消除噪声。第二，应用SAE生成深层次特征来预测股票价格。第三，将高水平的去噪特性输入LSTM，以预测第二天的收盘价。选取6个市场指数及其对应的指数期货来检验该模型的性能。结果表明，该模型在预测精度和盈利能力方面均优于其它同类模型。

引用： Bao W, Yue J, Rao Y (2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 12(7): e0180944. https://doi.org/10.1371/journal.pone.0180944
编者： BorisPodobnik， 克罗地亚Rijeka大学
收到： 2016年12月20日；接受： 2017年6月10日；出版： 2017年7月14日
版权： 2017年BaO等人。这是一篇以Creative Commons Attribution许可证，允许在任何介质中不受限制地使用、分发和复制，只要原作者和源被记入帐户。
数据可得性： 所有数据都可以从Figshare数据库获得(DOI: 10.6084/m9.figshare.5028110).
供资： 本课题由国家自然科学基金资助(批准号：71372063和71673306)， http://www.nsf.gov.cn/ 资助者在研究设计、数据收集和分析、决定出版或编写手稿方面没有任何作用。
相互竞争的利益： 提交人宣布，不存在任何相互竞争的利益。

导言

股票市场预测通常被认为是时间序列预测中最具挑战性的问题之一。¹由于它的噪音和不稳定的特性。如何准确地预测股票的流动，对于现代社会的经济和社会组织来说，仍然是一个悬而未决的问题。在过去的几十年里，机器学习模型，如人工神经网络²和支持向量回归(SVR)³，被广泛应用于金融时间序列的预测，并获得了较高的预测精度。^{4–8}然而，在文献中，机器学习和模式识别领域的最新发展趋势认为，时间序列预测应采用深度非线性拓扑。与传统的机器学习模型相比，新的机器学习模型可以通过提取获取相关信息的健壮特征，成功地对复杂的真实世界数据进行建模。⁹取得比以前更好的成绩¹⁰考虑到金融时间序列的复杂性，将深度学习与金融市场预测相结合被认为是具有魅力的课题之一。¹¹然而，这一领域仍相对缺乏探索。

一般说来，有三种主要的深度学习方法在研究中得到了广泛的应用：卷积神经网络¹²，深层信仰网络¹³和堆叠的自动编码器¹⁴在金融深层次学习的相关工作中，引入了前两种研究方法。例如，丁等人¹⁵将神经张量网络和深层卷积神经网络相结合，预测事件对股票价格波动的短期和长期影响。此外，某些著作在金融市场预测中使用了深刻的信念网络，例如Yoshihara等人。¹⁶，沈等人¹⁷和Kuremoto等人¹⁸然而，对于堆叠式自动编码方法能否应用于金融市场预测的问题，很少有人对此进行研究。因此，本文在此基础上提出了一种基于堆叠自动编码方法的股票市场预测模型。

本文提出的模型由三部分组成：小波变换(WT)、叠置自动编码器(SAES)和长时记忆(LSTM)。SAES是该模型的主要组成部分，用于无监督地学习金融时间序列的深层特征。具体来说，它是一个由多个单层自动编码器组成的神经网络，其中每个层的输出特征被连接到连续层的输入端。通过减小输出数据和输入数据之间的误差，一次进行一次声发射的无监督训练。因此，SAES模型可以成功地学习不变量和抽象特征。¹⁹。

另外两种方法结合起来帮助提高预测的准确性。LSTM是一种递归神经网络(RNN)，它将反馈链路附加到网络的某些层。与传统的RNN不同，当存在任意大小的时间步长时，它非常适合于从经验中学习预测时间序列。此外，它还可以通过内存单元在任意时间内保留与时间相关的信息来解决消失梯度的问题。²⁰有证据证明，它比传统的RNN更有效^{21, 22}因此，我们决定使用这个模型来预测股票趋势。将小波变换用于修正金融时间序列的噪声特征。它是一种广泛应用于一维信号滤波和挖掘的技术。^{23–25}我们利用它对输入的金融时间序列进行去噪，然后将其输入到深度学习框架中。总之，本文所介绍的模型是这三种方法的结合，我们将这种新的模型称为WSAES-LSTM(以下简称WSAES-LSTM)。

我们选取了6个股票指数来检验该模型的预测能力。这些指数包括中国内地A股市场的沪深300指数、代表印度股市的Nifty 50指数、香港市场的恒生指数(HangSengindex)、东京的日经225指数(Nikkei 225 Index)、纽约证交所的标准普尔500指数(S&P 500)和DJIA指数。在技术上，我们应用WSAES-LSTM来预测每一种股票指数的走势，并检验我们的模型在预测股票移动趋势方面有多好。

值得注意的是，我们测试了WSAES-LSTM在几个金融市场上的表现，而不仅仅是一个市场。这是因为人们担心取得强有力的结果。根据有效市场假说(EMH)，市场的效率影响其资产的可预测性。换句话说，即使一个市场的预测性能令人满意，也很难将其归因于所提出的模型的作用。在不同的市场条件下测试该模型给我们提供了解决问题的机会，并向我们展示了我们的模型的可预测性有多强。所选择的市场可以达到上述目标。它们代表着金融市场的三个发展阶段。例如，中国内地和印度的股票市场通常被视为发展中市场。尽管这两家公司在过去几十年中都经历了快速增长，但它们的许多监管都是不成熟的。相比之下，纽约股市被认为是最发达的市场。它也是世界上最大的证券交易所和效率最高的市场。除上述情况外，香港和东京的股票市场处于最发达和发展中国家之间的一种中间地带。因此，我们的样本设置可以帮助我们检验我们提出的模型在不同市场状态下的有效性。

对于每个股票指数，使用三种类型的变量作为模型输入。第一套是历史股票交易数据，例如公开、高、低和收盘价格(OHLC)[26–28]第二个是股票交易的技术指标。这些都是以前研究中常用的投入[29]除此之外，我们还引入了宏观经济变量作为第三类投入。由于宏观经济对股市的影响很大，我们的深度学习模型的优势在于能够从输入变量中提取抽象和不变的特征。30, 31]，我们认为，增加宏观经济变量可以改善模型的绩效。

关于预测方法，Chan等人介绍了一种分段预测方法。[32]用于获取每个股票指数的预测结果。然后，从预测精度和盈利能力两个维度对模型的性能进行了评价。采用平均绝对百分比误差(MAPE)、相关系数(R)和泰尔不等式系数(Teil U)三种方法对预测精度进行了评价。所有这些指标都被广泛用于衡量预测值是否与实际值相似[2, 23, 33, 34]为了检查盈利能力，我们制定了买卖策略。35]将该策略应用于基于模型预测结果的交易回报。作为基准，我们还计算了每个股票指数的买入和持有策略的回报[32, 36]其基本思想是，基于WSAES-LSTM的交易回报能否超过这种简单交易策略的收益，这为模型的盈利能力提供了进一步的证据。

为了更好地捕捉WSAES-LSTM的性能，我们还引入了其他三种模型，并评价了它们在预测每一种股票指数时的预测精度和盈利能力，并与我们提出的模型进行了比较。这三个模型包括WLSTM(即WT和LSTM的结合)、LSTM和传统的RNN。前两种模型用于检验SAES方法在提高预测性能方面的有效性。最后一个模型RNN作为性能基准。因为在以往的文献中，它已经成功地应用于金融时间序列的预测。23, 37, 38]，与传统的神经网络相比，它有助于我们更好地了解所提出的模型能提高性能。

本研究的所有样本数据均来自上海风电信息技术有限公司提供的风电数据库、深圳市GTA教育技术公司提供的CSMAR数据库。全球金融门户网站：Investing.com。它由2008年7月至2016年9月大约8年的数据组成。结果表明，WSAES-LSTM模型在可预测性和盈利能力方面均优于其它三种模型。

我们的工作植根于一个日益增长的研究领域，即应用深度学习方法来提高效率。例如，基于深度学习的方法极大地提高了图像识别的先进水平。12, 39–41]，语音识别[42–44]，语文翻译[45, 46和许多其他领域，如药物发现[47和基因组学[48, 49]这项工作的主要贡献是首次尝试应用堆叠式自动编码器产生OHLC的深层次特征、技术指标和宏观经济条件，作为一个多变量信号，以便向LSTM提供信息，以预测未来的股票价格。与传统的长时记忆和递归神经网络(RNN)方法相比，本文提出的深度学习框架WSAES-LSTM能够提取更多的抽象和不变特征。

本文的其余部分分为五个部分。第二节介绍了基于小波、SAES和LSTM的多变量去噪方法。第三节是对输入和数据资源的描述。第四节介绍了实验设计的细节。第五节总结了观测结果，最后一节总结了我们的研究成果。

方法论

为了产生一步的股票价格预测的深度和不变特征，本文提出了一种基于深度学习的预测方法，它集成了堆叠式自动编码器和长期记忆的体系结构，为金融时间序列提供了一个深度学习的框架。图1显示此框架的流程图。该框架包括三个阶段：(1)利用小波变换对数据进行预处理，用于对股票价格时间序列进行分解以消除噪声；(2)应用具有深度结构的堆叠式自动编码器，对数据进行无监督训练；(3)利用具有延时的长时记忆产生单步输出。每个块的详细方法进一步详细如下。

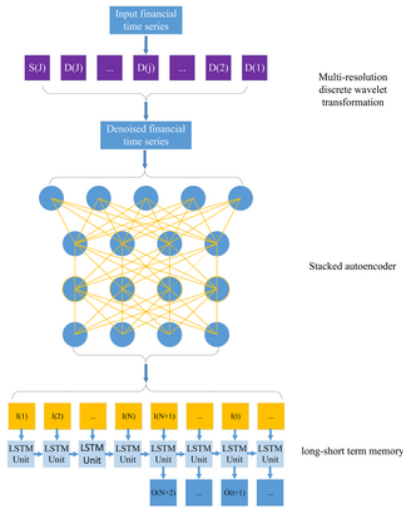


图1.提出的金融时间序列深度学习框架的流程图。
d(J)是j级的详细信号.s(J)是J.i(T)级上最粗的信号，O(T)分别表示去噪特征和时间步长t处的单步输出。N是LSTM的延迟数。
<https://doi.org/10.1371/journal.pone.0180944.g001>

小波变换

由于小波变换具有处理非平稳金融时间序列数据的能力，因此本文将小波变换应用于数据去噪。50]小波变换的关键在于它能同时分析金融时间序列的频率分量，并与傅里叶变换相比较。因此，小波对于处理高度不规则的金融时间序列是有用的。51]。

本研究采用Haar函数作为小波基函数，不仅可以金融时间序列分解为时域和频域，而且大大缩短了处理时间。23]以Haar函数为基的小波变换的时间复杂度为O(n)与n表示时间序列的大小[52]。

对于连续小波变换(CWT)，小波函数可以定义为：

$$\phi_{a,\tau}(t) = \frac{1}{\sqrt{a}}\phi\left(\frac{t-\tau}{a}\right) \tag{1}$$

哪里a和τ分别是尺度因素和翻译因素。φ(t)是基小波，它遵循一个叫做小波可容许条件的规则[53]:

$$C_\phi = \int_0^\infty \frac{|\Phi(\omega)|}{\omega} d\omega < \infty \quad (2)$$

哪里 $\Phi(\omega)$ 是频率的函数 ω 的傅里叶变换 $\Phi(t)$ 放 $x(t)$ 表示一个平方可积函数 $x(t) \in L^2(\mathbb{R})$ ；然后用小波变换进行CWT。 ϕ 可定义为：

$$CWT_x(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \overline{\phi\left(\frac{t-\tau}{a}\right)} dt \quad (3)$$

哪里 $\overline{\phi(t)}$ 表示它的复共轭函数。连续小波变换的逆变换可表示为：

$$x(t) = \frac{1}{C_\phi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} CWT_x(a, \tau) \phi_{a,\tau}(t) d\tau \quad (4)$$

连续小波变换的系数具有大量的冗余信息。因此，为了减少冗余，对系数进行采样是合理的。将时间序列分解为正交分量集，得到离散小波变换(DWT)。马拉特[54]提出用一对高通和低通滤波器对时间序列进行滤波，作为离散小波变换的一种实现。有两种类型的小波，父亲小波 $\varphi(t)$ 和母小波 $\psi(t)$ ，在Mallat算法中。父小波 $\varphi(t)$ 和母小波 $\psi(t)$ 分别集成为1和0，可将其表述为：

$$\int \varphi(t) dt = 1, \int \psi(t) dt = 0 \quad (5)$$

母小波描述高频部分，父亲小波描述时间序列的低频分量。中的母小波和父小波 j -级别可表述为[55]：

$$\varphi_{j,k}(t) = 2^{-\frac{j}{2}} \varphi(2^{-j}t - k) \quad (6)$$

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad (7)$$

财务时间序列可以用父母小波上的一系列预测来重建，多层分析的指标是 $k \in \{0, 1, 2, \dots\}$ 和 $j \in \{0, 1, 2, \dots, J\}$ ，在哪里 J 表示多分辨率尺度的数目。正交小波级数对时间序列的逼近 $x(t)$ 由下列机构拟订：

$$x(t) = \sum_k s_{J,k} \varphi_{J,k}(t) + \sum_k d_{J,k} \psi_{J,k}(t) + \sum_k d_{J-1,k} \psi_{J-1,k}(t) + \dots + \sum_k d_{1,k} \psi_{1,k}(t) \quad (8)$$

其中膨胀系数 $s_{J,k}$ 和 $d_{J,k}$ 是由预测给出的

$$s_{J,k} = \int \varphi_{J,k} x(t) dt \quad (9)$$

$$d_{j,k} = \int \psi_{j,k} x(t) dt \quad (10)$$

时间序列的多尺度逼近 $x(t)$ 如下：

$$S_J(t) = \sum_k s_{J,k} \varphi_{J,k}(t) \quad (11)$$

$$D_j(t) = \sum_k d_{j,k} \psi_{j,k}(t) \quad (12)$$

然后，正交小波级数逼近的简略形式可以表示为：

$$x(t) = S_J(t) + D_J(t) + D_{J-1}(t) + \dots + D_1(t) \quad (13)$$

哪里 $S_J(t)$ 是输入时间序列的最粗近似。 $x(t)$ 多分辨率分解 $x(t)$ 是{的序列。 $S_J(t), D_J(t), D_{J-1}(t), \dots, D_1(t)$ 。当金融时间序列非常粗糙时，可以重复应用离散小波变换，从而降低过拟合的风险。因此，本文将两层小波应用于数据预处理，如本文所建议的。23]。

堆叠式自动编码器

深度学习是利用深度神经网络体系结构从输入数据中提取深层特征的一系列模型。深度学习模式通常有三层以上。深层网络通常通过无监督的分层训练来初始化，然后通过有监督的训练进行调整，标签可以逐层地生成更抽象和更高级别的特征。56]根据最近的研究[57, 58]，与浅层结构的模型相比，深度学习模型能更好地逼近非线性函数。

最近的研究提出了几种深度神经网络结构，包括深度Boltzmann机器(DBMs)[59]、深度信念网络(DBN)[13]和堆叠式自动编码器(SAE)[14]限制性Boltzmann机器(RBM)[60]、卷积神经网络(CNNs)[61]和自动编码器[14]是常用的分层训练模式。本文将自动编码器应用于OHLC变量和技术指标的分层训练，而SAES作为相应的深度神经网络体系结构。

单层AE是一种三层神经网络，如图2。第一层和第三层是输入层和重建层。 k 各单位。第二层是隐藏层 n 单元，这是为了产生这个单层AE的深层特征。单层声发射训练的目的是减小输入矢量与重构矢量之间的误差。单层AE前向传播的第一步是将输入矢量映射到隐藏层，如图2第二步是通过将隐向量映射到重构层来重构输入向量。这两个步骤可以表述为：

$$a(x) = f(W_1x + b_1) \tag{14}$$

$$x' = f(W_2a(x) + b_2) \tag{15}$$

哪里 $x \in R^k$ 和 $x' \in R^k$ 分别是输入向量和重构向量。 $a(x)$ 是由单层AE生成的隐藏向量。 W_1 和 W_2 分别是隐藏层和重构层的权重。 b_1 和 b_2 分别是隐藏层和重构层的偏差。 f 激活函数具有多种形式，如乙状结肠函数、校正线性单元(relu)和双曲正切。在本文中， f 设为乙状结肠函数，如Chen等人所述。[19]将输入向量与重构向量之间的误差最小化的优化函数可以表述为

$$\operatorname{argmin}_{W_1, b_1, W_2, b_2} [J] = \operatorname{argmin}_{W_1, b_1, W_2, b_2} [(1/2) \sum_{i=1}^m \|x_i - x'_i\|^2 + J_{wd} + J_{sp}] \tag{16}$$

哪里 J 是单层声发射的平方重建误差。 x_i 和 x'_i 是 i 输入向量及其相应的重构向量的TH值。 m 是训练数据集的大小，即本文训练阶段的交易日数。 J_{wd} 和 J_{sp} 重量衰减项和稀疏惩罚项，可表述为：

$$J_{wd} = (1/2) \lambda (\|W_1\|_F^2 + \|W_2\|_F^2) \tag{17}$$

$$J_{sp} = \beta \sum_{t=1}^m KL(\rho \parallel \hat{\rho}_t) \tag{18}$$

其中 $\|\cdot\|_F$ 是Frobenius规范。 λ 和 β 控制重量衰减项和稀疏惩罚项。 $KL(\cdot)$ 表示Kullback-Leiber散度。 ρ 是稀疏参数，只有少数隐单元可以大于稀疏参数。 $\hat{\rho}_t$ 的平均激活量。 t 训练数据集中的TH隐藏层，可表述为：

$$\hat{\rho}_t = (1/m) \sum_{i=1}^m a_t(x_i) \tag{19}$$

哪里 $a_t(x_i)$ 表示 k TH单位 t 整个训练数据集中隐藏层。梯度下降算法在求解SAES优化问题中得到了广泛的应用。19, 31]因此，应用梯度下降算法完成参数优化，就像尹等人所提出的。62].

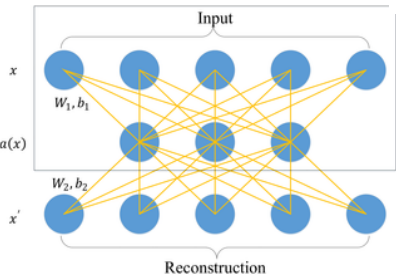


图2.单层自动编码器的流程图。
模型学习隐藏的特征。 $a(x)$ 从输入 x 通过在 x' 。这里 W_1 和 W_2 分别是隐层和重构层的权重。 b_1 和 b_2 分别是隐藏层和重构层的偏差。
<https://doi.org/10.1371/journal.pone.0180944.g002>

堆叠式自动编码器是通过层层堆叠一系列单层aes来构造的。14]. 图3举例说明由4个单层自动编码器组成的5层SAE的实例.单层自动编码器将输入的每日变量映射到第一个隐藏向量中。在训练第一单层自动编码器后，删除第一单层自动编码器的重构层，并保留隐藏层作为第二单层自动编码器的输入层。一般来说，后续声发射的输入层是前一声发射的隐层。每一层都使用与单层AE相同的梯度下降算法，通过求解情商(16)并将隐藏的矢量输入到随后的AE中。值得注意的是，在完成训练后，重建层的权重和偏倚被抛弃。在本工作中，每个数据集的每日输入变量数从18到25不等；然后，通过反复尝试将隐藏层的大小设置为10。深度在SAE中起着重要的作用，因为它决定了提取特征的不变性和抽象性。在这项工作中，SAE的深度被设置为5，正如Chen等人所建议的。[19].

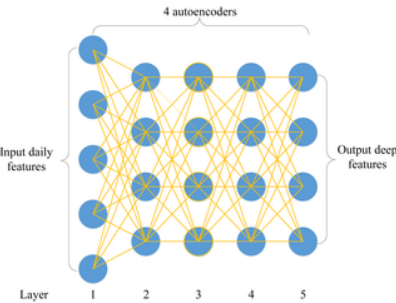


图3.由4个自动编码器训练的具有5个层的堆叠自动编码器的实例。
<https://doi.org/10.1371/journal.pone.0180944.g003>

长期短期记忆是递归神经网络(Rnn)结构的众多变化之一。20]介绍了RNN模型及其预测收盘价的LSTM体系结构。首先从基本的递归神经网络模型入手,然后讨论LSTM模型。

RNN是一种深度神经网络结构^{43, 63}在时间维度上有一个深层次的结构。它已广泛应用于时间序列建模^[21, 22, 64–69]传统神经网络的假设是输入向量的所有单元相互独立。因此,传统的神经网络不能充分利用序列信息。相反,RNN模型增加了由时间序列的顺序信息生成的隐藏状态,其输出依赖于隐藏状态。图4显示了一个RNN模型被展开成一个完整的网络。数学符号图4如下:

1. x_t 是时间上的输入向量。 t
2. s_t 是时间上隐藏的状态吗? t 它是根据输入向量和先前的隐藏状态来计算的。 s_t 计算方法如下:

$$s_t = f(Ux_t + Ws_{t-1}) \quad (20)$$

哪里 f 是激活函数,它有多种选择,如Sigmoid函数和relu函数。初始隐藏状态 s_0 用于计算第一个隐藏状态 s_1 通常初始化为零。

3. o_t 是时间上的输出吗? t ,可表述为:

$$o_t = f(Vs_t) \quad (21)$$

4. U 和 V 分别是隐藏层和输出层的权重。 W 隐藏状态的过渡权重。

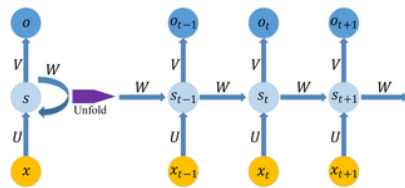


图4.递归神经网络和展开结构。

U , V 和 W 分别是隐藏层、输出层和隐藏状态的权重。 x_t 和 o_t 输入向量和时间输出结果 t 分别。

<https://doi.org/10.1371/journal.pone.0180944.g004>

虽然mn能很好地模拟时间序列,但由于梯度问题的消失,长期依赖关系很难学习。22]LSTM是利用存储单元对抗消失梯度的有效解决方案。70]记忆单元由四个单元组成:输入门、输出门、遗忘门和自回归神经元。图5。门控制相邻存储单元与存储单元本身之间的相互作用。输入信号是否可以改变存储器单元的状态由输入门控制。另一方面,输出门可以根据是否可以改变其它存储器单元的状态来控制存储器单元的状态。此外,遗忘门可以选择记忆或忘记其先前的状态。

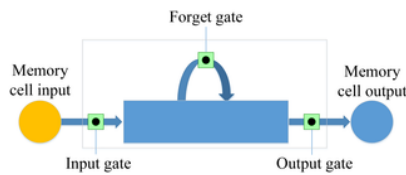


图5.LSTM存储单元的体系结构。

<https://doi.org/10.1371/journal.pone.0180944.g005>

图6显示正在展开为完整网络的LSTM模型,该模型描述如何更新每个门的值。数学符号图6如下:

1. x_t 是内存单元格在时间上的输入向量。 t
2. W_i , W_f , W_c , W_o , U_i , U_f , U_c , U_o 和 V_o 是重量矩阵。
3. b_i , b_f , b_c 和 b_o 是偏差向量。
4. h_t 是内存单元格在时间上的值。 t
5. i_t 和 \tilde{C}_t 是输入门的值和内存单元的候选状态在时间上的值。 t 分别可表述为:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (22)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (23)$$

6. f_t 和 C_t 是遗忘门的值和内存单元在时间上的状态。 t 分别可通过以下方法计算:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (24)$$

$$C_t = i_t * \tilde{C}_t + f_t * C_{t-1} \quad (25)$$

7. o_t 和 h_t 是时间上输出门的值和内存单元的值。 t 分别可表述为:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o C_t + b_o)$$

(26)

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

(27)

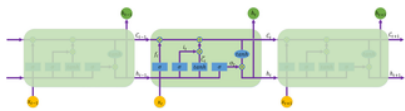


图6.LSTM中的重复模块。
这里 x_t 和 h_t 是将输入向量和输出结果在时间上输出到内存单元。 t 分别。 h_t 内存单元的值。 i_t , f_t 和 o_t 输入门、遗忘门和输出门的时间值。 t 分别。 C_t 是内存单元格在时间上的候选状态的值。 t 。
<https://doi.org/10.1371/journal.pone.0180944.g006>

LSTM网络的体系结构包括隐藏层数和延迟数，即过去用于培训和测试的数据数。目前，没有经验法则可以选择延迟和隐藏层的数量[21, 22]在这项工作中，隐藏层的数目和延迟被设置为5和4通过试验和错误。财务时间序列分为三个子集：训练集、验证集和测试集，其中80%的训练、10%的验证和10%的测试。采用反向传播算法训练实验控制组的WSAES-LSTM模型，包括WLSTM模型、LSTM模型和RNN模型。学习率为0.05，批次大小为60，历元数为5000。收敛速度受学习速度的控制，是时间的递减函数。将训练次数和学习速率设为5000和0.05，可以达到训练的收敛性。当参数组合发生变化时，一旦收敛，实验结果将趋于稳定。30]

数据描述

在这一部分中，我们详细介绍了我们的样本选择和我们选择的模型预测的输入变量。另外，本节还提供了数据资源。

样本选择和输入变量

我们选择的六个指数是：沪深300指数、Nifty 50指数、恒生指数、日经225指数、标准普尔500指数和DJIA指数。正如我们前面提到的，市场状况可能会影响神经网络的有效性。不同市场条件下的样本有助于解决这一问题。标准普尔500指数和DJIA指数在纽约证券交易所交易，纽约证券交易所被普遍认为是世界上最先进的金融市场。因此，它们是指发展水平最高的这类市场。相反，中国大陆和印度的金融市场往往被归类为新市场。事实上，他们的大部分市场机构仍远未完全完成。因此，我们选择CSI 300和Nifty 50来代表发展中市场。除上述市场外，香港恒生指数和东京日经225指数代表了介于发达市场和发展中市场之间的市场状况。老实说，在大多数情况下，香港和东京的金融市场通常被视为发达市场。然而，在本文中，我们可以说，与美国股市相比，这两个市场还不如美国市场成熟。因此，这六种股指都为我们在不同的市场条件下检验模型的稳健性提供了一个自然的环境。

我们选择三组变量作为输入。表1描述详细信息。面板A中的第一组变量是每个指数的历史交易数据。根据之前的文献，数据包括开放，高，低，收盘价(OHLC变量)以及交易量。这些变量表示每个指数的基本交易信息。另一组投入是12项广泛使用的每个指标的技术指标。小组B给出了细节。

Name	Definition/Implication
Panel A. Daily Trading Data	
Open/Close Price	nominal daily open/close price
High/Low Price	nominal daily highest/lowest price
Trading volume	Daily trading volume
Panel B. Technical Indicator	
MACD	Moving average convergence divergence: displays trend following characteristics and momentum characteristics.
CCI	Commodity channel index: helps to find the start and the end of a trend.
ATR	Average true range: measures the volatility of price.
BOLL	Bollinger Band: provides a relative definition of high and low, which aids in rigorous pattern recognition
EMA20	20 day Exponential Moving Average
MA5/MA10	5/10 day Moving Average
MTM5/MTM10	6/12 month Momentum: helps pinpoint the end of a decline or advance
ROC	Price rate of change: shows the speed at which a stock's price is changing
SMI	Stochastic Momentum Index: shows where the close price is relative to the midpoint of the same range.
WVAD	Williams's Variable Accumulation/Distribution: measures the buying and selling pressure.
Panel C. Macroeconomic Variable	
Exchange rate	US dollar Index
Interest rate	Interbank Offered Rate

<https://doi.org/10.1371/journal.pone.0180944.t001>

表1.输入变量的描述。
<https://doi.org/10.1371/journal.pone.0180944.t001>

最后一组投入是宏观经济变量。毫无疑问，各区域的宏观经济状况在影响股票市场的表现方面也发挥着关键作用。赵等人[71]结论]人民币汇率的波动会影响中国内地A股市场的走势。因此，宏观经济变量的加入有助于在神经网络预测中引入更多的信息。我们选择了两种宏观变量：汇率和利率。这两种利率都可能影响股票市场的资金流动，最终影响股票的表现。具体来说，我们选择美元指数作为汇率的代名词。人们认识到，美元在货币市场中扮演着最重要的角色。因此，仅凭这一点就足以反映货币市场对股市的影响。在利率方面，我们选择各个市场的银行间同业拆借利率(SHIBOR)、孟买银行同业拆借利率(MIBOR)、香港银行同业拆借利率(HIBOR)、东京银行同业拆借利率(TIBOR)和美国联邦基金利率作为代理利率。

数据资源

我们所有的样本数据都来自风能数据库(<http://www.wind.com.cn>)上海风电信息有限公司提供的CSMAR数据库(<http://www.gtarsc.com>)由深圳GTA教育技术有限公司提供。全球金融门户网站: Investing.com。本文的样本周期为1。²⁶2008年7月至30日²⁷2016年9月。我们所有的变量都是每日数据，可以在Figshare网站(figshare.com/s/acdfb4918c0695405e33、DOI: 10.6084/m9.figshare.5028110)。

具体来说，我们在本文中没有进行先验时间序列分析。实际上，应用经典的时间序列模型，如自回归积分移动平均(ARIMA)，通常需对时间序列的分布和平稳性进行严格的假设。由于金融时间序列通常具有非常复杂、非平稳、噪声大等特点，在应用经典时间序列模型之前，必须先了解时间序列的性质。72, 73]否则，预测工作将是无效的。然而，利用人工神经网络对时间序列进行先验分析并不

是必不可少的。首先，ans不需要事先知道时间序列的结构，因为它们的黑匣子特性[74]同时，时间序列的平稳性对ANS预测能力的影响也很小。有关证据表明，将ANS应用于预测时，放宽非平稳时间序列的平稳条件是可行的[75]因此，我们简化了先验数据分析的过程，并将数据直接放入模型中。

实验设计

我们详细介绍了如何获得预测值和评估每个模型的性能。

预测方法

预测程序遵循Chan等人所描述的分段预测方法。[32]特别是，这一程序由三个部分组成。第一部分是训练部分，用于模型的训练和模型参数的更新。第二部分是验证部分。我们利用它来调整超参数，并得到一个最优的模型设置。最后是试验部分，利用最优模型对数据进行预测。具体而言，由于数据有限，我们每个部分的时间框架与Chan等人的时间框架不一致。[32]在训练部分，我们利用过去两年的数据来训练模型。下面的三个月期间(一个日历季度)用于验证部分。在测试部分，根据目前流行的投资组合管理实践，对每个模型的季度性能进行了预测。从2010年10月到2016年9月，这一过程每季度持续6年。最后，对于每种股票指数，有24个季度和6个年度的预测结果。预测过程在图7。



图7.连续的数据集安排，在整个样本期间进行培训、验证和测试。
<https://doi.org/10.1371/journal.pone.0180944.g007>

为了简化结果的演示，我们报告了年度业绩，而不是季度业绩。在此基础上，给出了模型预测精度和交易回报在六年内的表现.有关六年间隔的详情，请参阅表2.

Year	Time interval
Year 1	2010-10-01–2011-09-30
Year 2	2011-10-01–2012-09-30
Year 3	2012-10-01–2013-09-30
Year 4	2013-10-01–2014-09-30
Year 5	2014-10-01–2015-09-30
Year 6	2015-10-01–2016-09-30

<https://doi.org/10.1371/journal.pone.0180944.t002>

表2.六个预测年的时间间隔。
<https://doi.org/10.1371/journal.pone.0180944.t002>

性能测量

我们在这一部分中讨论了性能度量。我们首先演示选择的精确性度量来判断预测性能。接下来，我们将讨论如何检验每个模型的盈利绩效。

预测精度性能

以往的论文选取了几个指标来衡量模型对金融市场走势的预测效果。2, 23, 33, 34]本文按照它们的方法，选取三个经典指标(MAPE、R和Teil U)来衡量每个模型的预测精度。这些指标的定义如下：

$$MAPE = \frac{\sum_{t=1}^N \left| \frac{y_t - y_t^*}{y_t} \right|}{N} \tag{28}$$

$$R = \frac{\sum_{t=1}^N (y_t - \bar{y}_t)(y_t^* - \bar{y}_t^*)}{\sqrt{\sum_{t=1}^N (y_t - \bar{y}_t)^2 (y_t^* - \bar{y}_t^*)^2}} \tag{29}$$

$$Theil\ U = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - y_t^*)^2}}{\sqrt{\frac{1}{N} \sum_{t=1}^N (y_t)^2} + \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t^*)^2}} \tag{30}$$

在这些方程中， y_t 是实际价值和 y_t^* 是预测值。 n 表示预测周期。Mape测量错误的大小。它是以误差的相对平均值来计算的。 r 是两个变量之间线性相关的度量。它们的U是两个变量之间差异的相对度量。它使偏差平方，使较大的误差具有更大的权重，并夸大误差。如果R较大，则表示预测值与实际值相似，而如果MAPE和Theil U较小，则表示预测值接近实际值。23, 76].

盈利能力表现。

根据每个模型的预测结果，建立了一个买卖策略.这意味着，在相同的交易策略下，我们希望找到最有价值的模式，为投资者赚取最高的利润。实际上，买卖交易策略被广泛用于盈利能力的表现。35].

该策略建议投资者在下一时期的预测价值高于当前实际价值时买入。相反，它建议投资者在预测值小于当前实际价值时卖出。具体而言，这一战略可以用下列方程式来描述：

$$Buy_signal : y_{t+1}^* > y_t \tag{31}$$

$$Sell_{signal} : y_{t+1}^* < y_t$$

(32)

这，这个，那，那个 y_t 表示当前的实际值，以及 y_{t+1}^* 下一个时间段的预测值。战略收益的定义是：

$$R = 100 \times \left(\sum_{t=1}^b \frac{y_{t+1} - y_t + (y_t * B + y_{t+1} * S)}{y_t} + \sum_{t=1}^s \frac{y_t - y_{t+1} + (y_{t+1} * B + y_t * S)}{y_t} \right)$$

(33)

其中R是策略回报。b和s分别表示购买和卖出的总天数。B和S分别是买卖的交易成本。

由于现货市场进行一篮子股票卖空的难度和产生的巨大交易成本，我们通过交易相应的指数期货合约来执行这一策略，而不是使用股票指数。然而，在执行之前，一个主要的问题是，股指期货是否与其标的股指密切相关。事实上，从理论和实证两方面的证据都证明了股指与其相应的指数期货之间存在着密切的联系。77–81]此外，为了获得更有力的证据，我们进一步检验了六种股票指数与其相应的指数期货之间的长期关系。Spearman相关和协整检验结果表明，我们的所有指数与其相应的指数期货有稳定的长期关系(S1表)因此，我们相信现货市场的预测结果可以成功地应用到相应的指数期货市场中。

根据上述交易规则，当预测价格低于当前价格时，我们卖空指数期货合约；当预测价格高于当前价格时，我们买入该指数期货合约。我们注意到，一些市场有一个以上的未来产品交易市场。例如，恒生指数和标准普尔500指数都有两类期货产品：标准期货合约和迷你期货合约。然而，与前两个市场不同的是，中国只有标准的沪深300指数未来。因此，为了市场之间的一致性，我们选择标准的未来产品来执行交易策略。

To make the results more realistic, we consider the influence of transaction cost on profit. As the cost rates are different among the markets and would be occasionally adjusted for the regulation purpose, we unify the cost rates among our sample markets into one rate within our sample period in order to simplify the calculation procedure. Finally, the chosen cost rate of unilateral trading is 0.01%.

除买卖交易策略外，我们亦采用买卖策略，提供了一个被动的门槛，以测试根据以往文献提出的模型的盈利能力。32, 36]每个模型的交易回报将与买入和持有策略的回报进行比较。具体来说，由于长期持有期货合约在现实中存在很大的风险，我们在现货市场而不是在指数期货市场进行买卖策略。现货股票市场交易成本的计算过程遵循上述规则。最后，现货市场的统一成本为买入的0.25%和销售0.45%。

结果

对于每一种股票指数，我们从这四种模型中得到了年度预测数据，并在图中给出了相应的实际数据。图8说明第一年的结果，其余的第二年至第六年的数字可在沙一–S5无花果。根据图8和沙一–S5图1，我们可以发现，LSTM和RNN比WSAES-LSTM和WLSTM具有更大的变化和距离。此外，将WSAE-LSTM与WLSTM进行比较，前者优于后者：WSAE-LSTM比WLSTM具有更小的波动性，更接近实际的交易数据。具体来说，WSAES-LSTM在预测方面的优势在欠发达市场比在发达市场更加明显。

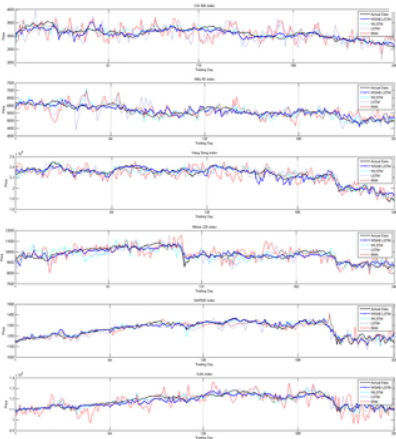


图8. Displays the actual data and the predicted data from the four models for each stock index in Year 1 from 2010.10.01 to 2011.09.30. <https://doi.org/10.1371/journal.pone.0180944.g008>

预测精度试验

每个模型的预测精度测试结果见表。3到5。每个表都包含了两个在类似市场条件下交易的股票指数的测试结果。在每个表中，每个面板在我们的三个精确指标之一中显示了预测性能的度量。我们分别报告了每种股票指数的六年业绩和六年的平均值。

Year	CSI 300 Index							Nikkei 225 Index						
	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Average	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Average
WSAEs-LSTM	0.025	0.014	0.016	0.011	0.033	0.016	0.019	0.024	0.019	0.019	0.019	0.019	0.017	0.019
WLSTM	0.025	0.029	0.021	0.020	0.020	0.023	0.023	0.024	0.026	0.026	0.025	0.025	0.025	0.029
LSTM	0.027	0.027	0.027	0.026	0.023	0.020	0.023	0.024	0.024	0.024	0.024	0.024	0.024	0.024
RNN	0.022	0.027	0.022	0.020	0.020	0.019	0.021	0.021	0.021	0.021	0.021	0.021	0.021	0.021

Panel B: HK

WSAEs-LSTM	0.001	0.003	0.005	0.007	0.010	0.007	0.004	0.009	0.004	0.006	0.006	0.007	0.004	0.004
WLSTM	0.001	0.001	0.010	0.004	0.007	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
LSTM	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
RNN	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Panel C: Thai

WSAEs-LSTM	0.017	0.009	0.011	0.007	0.020	0.011	0.013	0.016	0.013	0.013	0.013	0.013	0.012	0.013
WLSTM	0.016	0.019	0.014	0.013	0.004	0.001	0.016	0.020	0.020	0.020	0.020	0.020	0.020	0.020
LSTM	0.042	0.049	0.036	0.024	0.031	0.036	0.036	0.036	0.037	0.035	0.034	0.034	0.034	0.034
RNN	0.044	0.052	0.037	0.041	0.036	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046	0.046

<https://doi.org/10.1371/journal.pone.0180944.t003>

表3. 发展中国家的预测准确性。
<https://doi.org/10.1371/journal.pone.0180944.t003>

Hong Kong Index							Nikkei 225 Index						
Year	Year 1	Year 2	Year 3	Year 4	Year 5	Average	Year 1	Year 2	Year 3	Year 4	Year 5	Average	
Panel A. MAPE													
WSAEs-LSTM	0.016	0.017	0.012	0.011	0.021	0.013	0.015	0.020	0.016	0.017	0.014	0.016	0.017
WLSTM	0.020	0.027	0.017	0.018	0.026	0.021	0.020	0.025	0.022	0.020	0.020	0.027	0.028
LSTM	0.025	0.027	0.024	0.026	0.023	0.024	0.026	0.030	0.021	0.020	0.026	0.030	0.031
RNN	0.026	0.042	0.020	0.021	0.032	0.023	0.041	0.038	0.026	0.024	0.034	0.030	0.036
Panel B. R													
WSAEs-LSTM	0.944	0.926	0.920	0.927	0.904	0.926	0.890	0.927	0.940	0.945	0.976	0.961	0.937
WLSTM	0.930	0.910	0.898	0.933	0.900	0.917	0.876	0.748	0.836	0.873	0.796	0.951	0.906
LSTM	0.895	0.872	0.727	0.812	0.802	0.801	0.842	0.779	0.759	0.873	0.706	0.916	0.814
RNN	0.805	0.875	0.684	0.700	0.870	0.840	0.770	0.684	0.825	0.865	0.575	0.893	0.846
Panel C. TeIU													
WSAEs-LSTM	0.011	0.010	0.008	0.007	0.018	0.009	0.011	0.013	0.010	0.009	0.010	0.011	0.011
WLSTM	0.012	0.017	0.011	0.011	0.013	0.014	0.021	0.027	0.021	0.015	0.013	0.017	0.018
LSTM	0.015	0.017	0.016	0.019	0.014	0.015	0.022	0.020	0.018	0.014	0.016	0.020	0.020
RNN	0.022	0.026	0.018	0.021	0.020	0.021	0.028	0.020	0.021	0.021	0.021	0.021	0.023

<https://doi.org/10.1371/journal.pone.0180944.t004>

表4.相对发达市场的预测精度。
<https://doi.org/10.1371/journal.pone.0180944.t004>

S&P 500 Index							DAX Index						
Year	Year 1	Year 2	Year 3	Year 4	Year 5	Average	Year 1	Year 2	Year 3	Year 4	Year 5	Average	
Panel A. MAPE													
WSAEs-LSTM	0.012	0.014	0.012	0.008	0.011	0.010	0.016	0.013	0.008	0.008	0.008	0.010	0.011
WLSTM	0.015	0.040	0.012	0.010	0.015	0.015	0.015	0.018	0.011	0.011	0.017	0.012	0.014
LSTM	0.021	0.018	0.013	0.012	0.017	0.022	0.017	0.019	0.026	0.016	0.020	0.020	0.020
RNN	0.017	0.019	0.013	0.019	0.020	0.019	0.019	0.040	0.020	0.024	0.028	0.026	0.023
Panel B. R													
WSAEs-LSTM	0.944	0.944	0.944	0.973	0.880	0.902	0.940	0.920	0.948	0.952	0.953	0.952	0.940
WLSTM	0.917	0.880	0.871	0.927	0.770	0.860	0.846	0.815	0.891	0.911	0.817	0.927	0.861
LSTM	0.873	0.900	0.968	0.903	0.790	0.756	0.875	0.880	0.791	0.848	0.751	0.718	0.786
RNN	0.880	0.900	0.880	0.887	0.816	0.827	0.847	0.884	0.929	0.871	0.869	0.266	0.827
Panel C. TeIU													
WSAEs-LSTM	0.009	0.010	0.006	0.006	0.008	0.006	0.007	0.010	0.009	0.006	0.006	0.006	0.007
WLSTM	0.011	0.011	0.008	0.002	0.011	0.011	0.010	0.012	0.008	0.007	0.011	0.008	0.008
LSTM	0.014	0.012	0.009	0.007	0.011	0.016	0.011	0.013	0.016	0.014	0.014	0.014	0.013
RNN	0.012	0.012	0.009	0.010	0.010	0.010	0.010	0.025	0.010	0.018	0.020	0.017	0.021

<https://doi.org/10.1371/journal.pone.0180944.t005>

表5.发达市场的预测准确性。
<https://doi.org/10.1371/journal.pone.0180944.t005>

表3记录预测CSI 300和Nifty 50的模型性能。从表中可以看出，WSAEs-LSTM在预测两种股票指数方面的表现要好于其他三种模型。例如，在预测CSI 300指数时，WSAEs-LSTM的MAPE和TeIU平均值分别达到0.019和0.013，远远小于其他三个模型。此外，指标R的平均值为0.944，是四种模型中最高的一种。事实上，WSAEs-LSTM不仅在平均水平上而且在每年的表现上都优于其他三家。为了验证我们发现的稳健性，我们检验了WSAEs-LSTM与其他三种模型之间差异的统计意义。具体来说，我们比较了24个季度的WSAEs-LSTM和三个模型的每一个准确性指标的结果。这些比较采用统计方法T检验.最后，统计证据证明，WSAEs-LSTM模型与其余三种模型在两种股指的5%水平上的差异均具有统计学意义。

表4和5介绍模型在其余四种股票指数中的表现：表4展示香港及东京市场的示范表现表5报告标准普尔500指数和DJIA指数的结果。类似于我们在表3无论从平均值角度还是从年度结果来看，WSAEs-LSTM仍然具有最低的MAPE和TeIU，以及最高的R值。尽管如此，我们提出的模型与其他三种模型的这些差异在5%的显著性水平上通过了统计检验。与其他三种模型相比，无论市场条件如何，WSAEs-LSTM模型都能稳定地获得较低的预测误差和较高的预测精度。

除了上述发现外，我们还发现了一个基于我们的数据的有趣的模式：如果两种股票指数在发展状态相似的市场上交易，那么一个特定模型预测两种股票指数的可预测性之间的差别很小，而如果这两种股票指数在不同发展状态的市场上进行交易，差别就会增大。例如，WSAEs-LSTM预测标准普尔500和DJIA的MAPE为0.011，而当预测CSI 300和Nifty 50时则增加到0.019。其他三种模型也存在类似的模式。尽管所有的模型都表现出这种模式，但不同的模型对市场条件的影响程度是不同的。例如，市场状况似乎对RNN有相当大的影响。在六种股票指数中，其平均MAPE值在0.018到0.066之间。这意味着RNN最糟糕的可预测性仅为其最佳可预测性的四分之一左右。WLSTM和LSTM都具有与RNN相似的模式。相比之下，与这三种模型相比，WSAEs-LSTM在市场上的表现相当稳定.这可能是因为SAE在处理噪声数据方面比其他三个方面更强大。这些发现的含义是，我们的模型可能比其他模型更有价值的预测系统是不成熟的，具有较高的波动性。

盈利能力测试

盈利能力测试结果载于表6。同样，我们报告了六年的年度回报和平均回报。每个面板描述模型在特定市场条件下获得的交易回报。特别是，每个面板中的最后一行报告了在特定股票指数交易中买入和持有策略的回报。

Year	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Average	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Average
Panel A. Developing market														
CSI 300 Index								Nifty 50 Index						
WSAEs-LSTM	46.426	55.580	65.685	71.897	63.951	70.613	63.026	58.911	51.710	66.785	57.107	51.882	28.134	45.418
WLSTM	26.780	21.942	22.502	25.264	18.964	42.263	26.820	18.244	13.675	22.465	44.026	28.488	22.017	
LSTM	-15.802	-16.802	24.082	1.345	17.803	53.479	17.368	26.247	16.509	0.921	-4.459	37.190	18.279	15.113
RNN	-3.510	1.750	-6.827	-0.747	17.306	-8.430	7.550	-7.737	-10.558	18.891	22.885	29.482	-7.435	7.686
Buy-and-hold	-18.880	-13.880		0.202	16.610	3.402	-1.687	20.564	13.880	30.880	-0.767	4.471	3.738	
Panel B. Relatively developed market														
Hong Kong Index								Nikkei 225 Index						
WSAEs-LSTM	75.844	81.890	54.685	49.715	64.727	59.949	64.468	63.662	37.846	81.022	68.632	59.419	76.020	59.427
WLSTM	29.087	19.221	27.410	23.148	19.020	51.220	26.905	27.774	10.196	43.143	32.411	31.262	62.569	34.558
LSTM	16.361	29.730	27.882	1.545	15.860	62.717	25.263	14.632	31.827	50.261	-15.605	8.774	28.272	19.581
RNN	5.287	40.810	-7.134	-17.024	27.845	48.463	19.378	-7.320	-8.967	8.310	-15.389	42.258	50.950	11.974
Buy-and-hold	-27.321	-16.261	1.470	-0.821	-0.246	9.453	-1.025	-19.271	-1.588	46.280	3.790	6.265	-4.728	-6.827
Panel C. Developed market														
S&P 500 Index								DAX Index						
WSAEs-LSTM	71.316	48.551	59.229	59.940	61.187	45.950	45.907	76.742	55.433	41.790	67.034	82.523	82.522	63.989
WLSTM	40.240	25.728	21.864	15.506	13.820	18.168	25.641	47.180	38.174	20.682	28.532	30.605	54.885	38.020
LSTM	-7.630	23.136	39.710	-2.978	2.010	54.264	11.630	30.886	22.851	17.917	-4.902	4.493	37.036	18.476
RNN	18.319	11.930	11.261	-8.475	12.260	3.807	8.102	4.787	6.763	-2.096	8.730	37.388	17.814	12.228
Buy-and-hold	-12.271	22.761	12.512	13.747	-6.306	9.819	-6.966	-7.860	19.270	9.720	6.620	-7.144	9.506	-5.467

<https://doi.org/10.1371/journal.pone.0180944.t006>

表6.各模型的盈利绩效。
<https://doi.org/10.1371/journal.pone.0180944.t006>

小组A展示了每种模式在发展中市场的盈利能力。左边部分是基于CSI 300的预测数据的交易业绩，而右边部分是基于Nifty 50的预测数据的交易业绩。结果表明，WSAE-LSTM比其他三种模型获得的利润要高得多.例如，在中国大陆和印度市场，该模型的年平均收益可分别达到63.026%和45.418%，而其他三种模式的年收益几乎都低于40%。对于每一个年度回报，WSAEs-LSTM也优于其他模式。它几乎可以稳定地每年获得超过40%的收益，这对其他三种模式来说是非常困难的。

B和C小组分别报告相对发达和发达市场的交易回报。类似于A小组的研究结果，WSAEs-LSTM可以在每年获得稳定的收益，而其他模型在交易收益上面临较大的差异。此外，从样本期内的平均收益来看，根据B和C小组的结果，我们提出的模型仍能获得最高的利润。

此外，为了得到一个稳健的结论，我们还检验了WSAEs-LSTM与其余三个模型之间的收益差异是否具有统计学意义。再一次，我们比较了24个季度在模型之间的回报。t检验结果表明，WSAEs-LSTM模型与其它三种模型的回归差异均在5%的水平上通过了显著性检验。因此，我们的研究结果支持WSAEs-LSTM在四种模型中具有最佳的可预测性。

结语

本文建立了一种新的预测框架，用以预测在不同金融市场上交易的六种流行股票指数的一步收盘价。该预测框架的建立过程如下：第一，利用Haar小波进行离散小波变换生成去噪时间序列；第二，通过无监督的方法提取深度的日常特征；第三，利用长-短期记忆以监督的方式生成一步输出。我们的输入变量包括每日OHLC变量、技术指标和宏观经济变量。这项工作对社会的主要贡献是首次尝试引入SAES方法来提取金融时间序列的深度不变日特征。此外，本文还提出了一套完整的去噪、深度特征提取、特征选择和金融时间序列拟合的学习框架。在该框架内，可以通过在去噪、深度特征提取或时间序列拟合等领域中，用一种最先进的方法代替各个模块来建立预测模型。

通过与其它三种模型的比较，验证了本文提出的模型的预测精度和盈利能力。研究结果表明，无论选择哪种股票指数进行检验，其预测精度和盈利能力均优于其他三种指标。虽然该集成系统具有较好的预测性能，但仍存在一些不足。例如，可以在系统中嵌入一种更先进的超参数选择方案，以进一步优化所提出的深度学习框架。此外，深度学习方法费时，基于GPU和异构计算的深度学习方法需要更多的关注。所有这些都可以通过今后的研究加以加强。

辅助信息

图S1.Displays the actual data and the predicted data from the four models for each stock index in Year 2 from 2011.10.01 to 2012.09.30.

<https://doi.org/10.1371/journal.pone.0180944.s001>
(TIF)

图S2.Displays the actual data and the predicted data from the four models for each stock index in Year 3 from 2012.10.01 to 2013.09.30.

<https://doi.org/10.1371/journal.pone.0180944.s002>
(TIF)

图S3.Displays the actual data and the predicted data from the four models for each stock index in Year 4 from 2013.10.01 to 2014.09.30.

<https://doi.org/10.1371/journal.pone.0180944.s003>
(TIF)

图4. Displays the actual data and the predicted data from the four models for each stock index in Year 5 from 2014.10.01 to 2015.09.30.

<https://doi.org/10.1371/journal.pone.0180944.s004>
(TIF)

S5图5.Displays the actual data and the predicted data from the four models for each stock index in Year 6 from 2015.10.01 to 2016.09.30.

<https://doi.org/10.1371/journal.pone.0180944.s005>
(TIF)

S1表指数与相应指数期货的关系。

<https://doi.org/10.1371/journal.pone.0180944.s006>
(PDF)

作者贡献

概念化：WB JY.

数据管理：世行。

获得资金：耶。

方法：WB JY.

项目管理：Jy WB Yr.

软件：Jy WB

书面-原始草案：WB JY.

写作-审查和编辑：WB JY.

参考文献

1. 王B, 黄H, 王X金融时间序列预测的一种新的文本挖掘方法。神经计算机。2012年; 83(6): 136-45。
查看文章 • 谷歌学者
2. 郭Z, 王宏, 刘清, 杨杰。基于特征融合的金融时间序列预测模型。PLOS一号。2014年; 第9(6)段: 172-200。
查看文章 • 谷歌学者
3. Cherkassky V.统计学习理论的性质: Springer; 1997。
查看文章 • 谷歌学者
4. 用神经网络对An, Zapranis A, Francis G.股票的性能建模: 与回归模型的比较研究。神经网络1994年; 7(2): 375-88。
查看文章 • 谷歌学者
5. 音Y, 马加维奥TM.判别分析与人工神经网络的比较。运筹学会学报。1993年; 44(1): 51-60。

[查看文章](#) • [谷歌学者](#)

6. 支持向量机预测印度股市期货价格。国际计算机应用杂志。2012年; 41(3): 22-6。
[查看文章](#) • [谷歌学者](#)
7. 陆CJ, 李TS, 邱CC. 基于独立成分分析和支持向量回归的金融时间序列预测。决策支持系统2009年; 47(2): 115-25。
[查看文章](#) • [谷歌学者](#)
8. 黄伟, 中森Y, 王茜. 用支持向量机预测股市走势。计算机与运筹学。2005年; 32(10): 2513-22。
[查看文章](#) • [谷歌学者](#)
9. Hinton GE, Salakhutdinov RR. 用神经网络降低数据维数。科学。2006年; 313(5786): 504-7。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
10. 古维尔A, 文森特P. 代表性学习: 回顾和新的观点。模式分析和机器智能的IEEE事务。2013年; 35(8): 1798-828。
[查看文章](#) • [谷歌学者](#)
11. Cavalcante RC, Brasileiro RC, Souza VLF, Nobrega JP, Oliveira Ali. 计算情报和金融市场: 概览和未来的方向。有应用的专家系统。2016年; 55: 194-211。
[查看文章](#) • [谷歌学者](#)
12. Krizhevsky A, Sutskever I, Hinton GE, 编辑. 基于深度卷积神经网络的ImageNet分类。神经信息处理系统的进展; 2012年。
13. Hinton GE, Osindero S, Tyw. 一种深度信念网的快速学习算法。神经计算2006年; 18(7): 1527-54。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
14. 贝尼吉奥Y, 兰布林P, 波波维西D, 拉罗谢尔H. 贪婪层-深入网络的培训。神经信息处理系统的研究进展。2007年; 19: 153。
[查看文章](#) • [谷歌学者](#)
15. 作者声明: DingX, Zhang Y, Liu T, Duan J, 编者. 深度学习事件驱动的股票预测。国际人工智能会议; 2015年。
16. Yoshihara A, Fujikawa K, Seki K, Uehara K. PRICAI 2014: 人工智能趋势。2014年; 759-69。
17. 深信度网络和共轭梯度法预测汇率。神经计算机。2015年; 167(C): 243-53。
[查看文章](#) • [谷歌学者](#)
18. Kuremoto T, Kimura S, Kobayashi K, Obayashi M. 时间序列预测使用有限制的Boltzmann机器的深度信念网络。神经计算机。2014年; 137(15): 47-56。
[查看文章](#) • [谷歌学者](#)
19. 陈勇, 林Z, 赵欣, 王庚, 顾永. 基于深度学习的高光谱数据分类. IEEE J Sel 顶层应用地球观测远程传感器. 2014; 7(6): 2094-107。
[查看文章](#) • [谷歌学者](#)
20. Chu KL, Sahari KSM. 基于长时记忆的类人机器人行为识别。2016年; 13(6): 172988141666336。
[查看文章](#) • [谷歌学者](#)
21. 帕兰吉H, 邓L, 沈玉龙, 高JF, 何旭德, 陈锦生, 等. 基于长时记忆网络的深句子嵌入: 分析及其在信息检索中的应用. IEEE-ACM 跨音频语音朗. 2016年; 24(4): 694-707。
[查看文章](#) • [谷歌学者](#)
22. 帕兰吉H, 沃德R, 邓L. 分布式压缩传感: 一种深度学习的方法. IEEE Transaction 关于信号处理. 2016年; 64(17): 4504-18。
[查看文章](#) • [谷歌学者](#)
23. 谢天杰, 小HF, 叶WC. 基于小波变换和递归神经网络的股票市场预测: 一种基于人工蜂群算法的集成系统. 应用软件计算2011年; 11(2): 2510-25。
[查看文章](#) • [谷歌学者](#)
24. Aminghafari M, Cheze N, Poggi JM. 小波与主成分分析相结合的多元去噪方法。计算统计与数据分析。2006年; 50(9): 2381-98。
[查看文章](#) • [谷歌学者](#)
25. 多诺霍·DL. 软阈值去噪。信息理论上的IEEE交易。1995年; 41(3): 613-27。
[查看文章](#) • [谷歌学者](#)
26. 张P, 王勇, 杨W. 台湾股票价格变化的混合预测模型研究。中国工业工程师学会学报。2004年; 21(4): 358-68。
[查看文章](#) • [谷歌学者](#)
27. 陈AS, 梁振英, 道克H. 神经网络在新兴金融市场的应用: 台湾股指的预测与交易。计算机与运筹学。2003年; 30(6): 901-23。
[查看文章](#) • [谷歌学者](#)
28. 文章题目: Parasuraman K. 小波网络: 传统神经网络的一种替代。IEEE神经网络国际联席会议, 2005年IJCNN'05会议记录; 2005年。

29. 作者声明: [by]张YQ, Akkaladevi S, Vachtsevanos G, Lin Ty.用于股票预测的颗粒神经网络代理.软计算2002年; 6(5): 406-13。
查看文章 • 谷歌学者
30. 月J, 赵伟, 毛S, 刘H.用深卷积神经网络对高光谱图像进行空间分类.遥感信件。2015年; 6(6): 468-77。
查看文章 • 谷歌学者
31. 一个基于空间金字塔池的高光谱图像分类的深度学习框架。遥感信件。2016年; 第7(9)段: 875-84。
查看文章 • 谷歌学者
32. Chan PMJ, Mehralizadeh M.通过一种新的小波-PCA去噪和人工神经网络模型预测东亚指数期货。PLOS一号。2016年; 第11(6)段。
查看文章 • 谷歌学者
33. 阿尔泰·E, 萨特曼·MH.股票市场预测: 人工神经网络与新兴市场线性回归比较。财务管理与分析杂志。2005年; 18(2): 18-33。
查看文章 • 谷歌学者
34. 埃克欧。尼日利亚证券交易所收益预测: 自回归综合移动平均(ARIMA)模型的证据。SSRN电子杂志2010年。
查看文章 • 谷歌学者
35. [2]姚杰, 谭克伦, 波赫.用于技术分析的神经网络: KLCI的研究.国际理论与应用金融杂志。1999年; 2(2): 1590-603。
查看文章 • 谷歌学者
36. Chan PMJ, Rozaimah Z.使用东盟-5货币评估可调移动平均线的有效性。PLOS一号。2016年; 11(8): e0160931。
查看文章 • PubMed/NCBI • 谷歌学者
37. 魏丽英, 程青.基于综合特征的混合递归神经网络模型对台湾股市的预测。国际创新计算机信息与控制杂志。2012年; 第8(8)段: 5559-71。
查看文章 • 谷歌学者
38. 而是AM, Agarwal A, Sastry VN.递归神经网络与股票收益预测的混合模型。有应用的专家系统。2015年; 42(6): 3234-41。
查看文章 • 谷歌学者
39. Farabet C, Couprie C, Najman L, LeCun Y.学习场景标记的层次特征。模式分析和机器学习, IEEE事务。2013年; 35(8): 1915-29。
查看文章 • 谷歌学者
40. 题名/责任者: Rev.J, J.J, Jain A, LeCun Y, Bregler C.人体姿态估计的卷积网络和图形模型的联合训练.EPrint Arxiv.2014年; 1799-807。
查看文章 • 谷歌学者
41. 塞格迪C, 刘伟, 贾英, Sermanet P.2015年: 1-9。
查看文章 • 谷歌学者
42. 书名/作者责任者: by J.大规模神经网络语言模型的训练策略。自动语音识别和理解; 2011年。
43. Hinton G, 邓L, Yu D, Dahl GE, Mohamed A, Jaitly N, 等。语音识别中声学建模的深层神经网络IEEE信号处理杂志。2012年; 29(6): 82-97。
查看文章 • 谷歌学者
44. Sainath TN, Mohamed AR, Kingsbury B, Ramabhadran B, 编辑。2013年LVCSR的深卷积神经网络
45. 简S, CHO K, Memisevic R, Bengio Y.关于使用非常大的目标词汇进行神经机器翻译。计算机科学。2014年。
查看文章 • 谷歌学者
46. Sutskever I, Vinyals O, Le QV, Sutskever I, Vinyals O, Le QV.基于神经网络的序列学习。神经信息处理系统的研究进展。2014年; 4: 3104-12。
查看文章 • 谷歌学者
47. [中英文摘要]Ma J, Sheridan RP, Liaw A, Dahl Ge, Svetnik V.化学信息与建模杂志。2015年; 55(2): 263-74。
查看文章 • 谷歌学者
48. 梁MK, 熊HY, 李LJ, 弗雷·BJ.深度学习组织调控的剪接代码。生物信息学。2014年; 30(12): 121-9。
查看文章 • 谷歌学者
49. 熊HY, Alipanahi B, Lee LJ, Bretschneider H, Merico D, 袁RKC, 等.人类剪接密码揭示了疾病的基因决定因素的新见解。科学。2015年; 347(6218): 1254806-PMID: 25525159
查看文章 • PubMed/NCBI • 谷歌学者
50. 拉姆齐JB.小波对经济和金融数据分析的贡献。皇家生物科学学会的哲学交易。1999年; 357(357): 2593-606。
查看文章 • 谷歌学者
51. 波波拉A, 艾哈迈德K, 编辑。检验小波预处理对TSK模糊模型的适用性。IEEE模糊系统国际会议; 2006年。
52. 阿布拉莫维奇F, 贝斯比亚斯P, 萨帕蒂纳斯T经验Bayes方法块小波函数估计。计算统计与数据分析。2002年; 39(4): 435-51。
查看文章 • 谷歌学者

53. “金融和经济学中的小波和其他滤波方法简介：学术出版社”，2002年。
54. Mallat SG.多分辨率信号分解理论：小波表示。IEEE Trans. 肛门模式。机器信息。11，674-693模式分析和机器智能的IEEE事务。1989年；第11(7)条：674-93。
[查看文章](#) • [谷歌学者](#)
55. Dghais AAA, Ismail MT.利用小波变换研究时间序列的平稳性。AIP会议记录。2014年；798-804。
[查看文章](#) • [谷歌学者](#)
56. Sch lkopf B, Platt J, Hofmann T.贪婪层-深度网络的明智培训。神经信息处理系统的研究进展。2007年；19：153-60。
[查看文章](#) • [谷歌学者](#)
57. Sutskever I, Hinton GE.深而窄的乙状结肠信念网络是通用的逼近器。神经计算2008年；20(11)：2629-36。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
58. Roux NL, Bengio Y.深信度网络是紧凑型通用近似器。神经计算2010年；22(8)：2192-207。
[查看文章](#) • [谷歌学者](#)
59. 题名/责任者：by L.机器学习研究杂志。2009年；第5(2)段：1967-2006年。
[查看文章](#) • [谷歌学者](#)
60. 辛顿·葛。训练受限玻尔兹曼机器实用指南。动力。2010年；9(1)：599-619。
[查看文章](#) • [谷歌学者](#)
61. 福岛核电站：一种不受位置移动影响的模式识别机制的自组织神经网络模型。生物控制论。1980年；36(4)：193-202。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
62. 基于单层SAE和SVM的音H、焦X、柴Y、方B场景分类.有应用的专家系统。2015年；42(7)：3368-80。
[查看文章](#) • [谷歌学者](#)
63. Dahl GE, Yu D, 邓L, Acero A.上下文相关的深层神经网络，用于大词汇量语音识别.IEEE音频语音和语言处理事务。2012年；20(1)：30-42。
[查看文章](#) • [谷歌学者](#)
64. Elman JL及时发现结构。认知科学。1990年；14(2)：179-211。
[查看文章](#) • [谷歌学者](#)
65. 罗宾逊AJ.递归网在电话概率估计中的应用。IEEE在神经网络上的交易。1994年；5(2)：298-305。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
66. 编辑Mikolov T, Karafát M, Burget L, Cernock J, Khudanpur S.基于递归神经网络的语言模型。2010年国际语音通信协会会议，日本千叶，Makuhari，2010年9月。
67. Graves A.递归神经网络序列传递。计算机科学。2012年；58(3)：235-42。
[查看文章](#) • [谷歌学者](#)
68. Bengio Y, Boulanger-Lewandowski N, Pascanu R.IEEE国际声学会议。2012年；8624-8。
[查看文章](#) • [谷歌学者](#)
69. 梅斯尼尔G, 何X, 邓L, 本吉奥Y.口语理解的递归神经网络结构和学习方法的研究.对话。2013年。
[查看文章](#) • [谷歌学者](#)
70. 题名/责任者：Reach.神经计算1997年；9(8)：1735-80。
[查看文章](#) • [PubMed/NCBI](#) • [谷歌学者](#)
71. 赵H.汇率与股价的动态关系：来自中国的证据。国际商业与金融研究。2010年；24(2)：103-12。
[查看文章](#) • [谷歌学者](#)
72. Bodyanskiy Y, Popov S.神经网络方法预测准周期金融时间序列。EUR J Oper Res.2006；175(3)：1357-66。
[查看文章](#) • [谷歌学者](#)
73. Hussain AJ, Knowles A, Lisboa PJG, El-DeredW.金融时间序列的多项式流水线神经网络预测。专家系统与应用-国际期刊。2008年；35(3)：1186-99。
[查看文章](#) • [谷歌学者](#)
74. Nourani V, Komasi M, Mano A.降雨径流模拟的多元神经网络-小波方法.水资源管理。2009年；23(14)：2877。
[查看文章](#) • [谷歌学者](#)
75. 金泰, 哦KJ, Kim C, 做JD.非平稳时间序列的人工神经网络。神经计算机。2004年；61(C)：439-47。
[查看文章](#) • [谷歌学者](#)

76. Bliemel F. Theil的预测精度系数：一个澄清。市场研究杂志。1973年；10(4)：444。
查看文章 • 谷歌学者
77. Kawaller IG, Koch PD, Koch TW. 标普500期货与标准普尔500指数之间的时间价格关系。“金融杂志”。1987年；42(5)：1309-29。
查看文章 • 谷歌学者
78. 题名/责任者：by L. 股票指数现金与期货市场关系的检验：协整法。1998年。
79. 龚CC, 纪SD, 苏林, 李SP, 任非股指与股指期货的领先滞后关系：一种热最优路径方法。Physica A: 统计力学及其应用。2016年；444：63-72。
查看文章 • 谷歌学者
80. 柴S, 郭C, 编辑。股指与期货价格的协整关系。信息和服务科学新趋势国际会议；2009年。
81. 张欣, 吕某, 编辑。优秀股指期货与现货市场的相互关系。第三次商业智能和金融工程国际会议；2010年。