

基于 LSTM 混合模型的比特币价格预测

张 宁 方靖雯 赵雨宣

中央财经大学金融学院 北京 100081

(nzhang@amss.ac.cn)

摘 要 聚焦于具有极度非线性、非平稳性等特征的比特币价格预测问题,在长短时记忆网络(Long Short-Term Memory, LSTM)基础上构建了 4 个混合预测模型,利用小波变换(Wavelet Transform, WT)以及自适应噪声的完备经验模态分解(Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, CEEMDAN)对序列进行分解与重构,并引入了样本熵(Sample Entropy, SE)进行重构优化,使用 LSTM 对重构以后的子序列分别进行预测,最后将其叠加得到最终的预测结果。在预测结果的评判上,使用均方根误差、平均绝对百分误差以及希尔不等系数来进行拟合评价,并将结果与单一 LSTM 模型进行比较。研究发现混合模型的预测准确性均优于单一模型,且样本熵的引入可有效降低预测误差。

关键词: 比特币价格;长短时记忆网络;小波变换;自适应噪声完备经验模态分解;样本熵

中图法分类号 TP183

Bitcoin Price Forecast Based on Mixed LSTM Model

ZHANG Ning, FANG Jing-wen and ZHAO Yu-xuan

School of Finance, Central University of Finance and Economics, Beijing 100081, China

Abstract For the reason that Bitcoin price is highly nonlinear and non-stationary, this paper proposes four mixed forecasting model based on Long Short-Term Memory (LSTM) model to get better prediction performance. Firstly, we use Wavelet Transform (WT) and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to decompose and reconstruct the original sequence. Then, we introduce Sample Entropy (SE) to optimize the reconstruction. Finally, we predict the reconstructed sub-sequences respectively using LSTM and superpose the outcomes yields of prediction results. To evaluate the prediction performance, three evaluation functions are used, which are RMSE, MAPE and TIC. Besides, the results are compared with single LSTM model prediction result. The research shows that the prediction accuracy of mixed model is better than that of the single model, and the introduction of Sample Entropy can effectively reduce prediction error.

Keywords Bitcoin price, Long short-term memory network, Wavelet transform, Complete ensemble empirical mode decomposition with adaptive noise, Sample entropy

1 引言

自 2008 年诞生以来,比特币以其稀缺性和独特的发行机制,加之多种因素助推,表现出价格的极端波动性:2013 年起,比特币价格急剧上升,从起初的 13 美元攀升至 1 200 美元,涨幅近 80 倍;而 2017 年比特币则经历了站上 20 000 美元高点后跌至 10 000 美元以下的大起大落;转眼至 2021 年,比特币价格已成功突破 60 000 美元大关,市值已达 1.2 万亿美元,超过两个特斯拉市值之和。比特币市场的跌宕以及比特币价格呈现出的非平稳性,为比特币价格预测提出了挑战。

而在众多预测模型工具中,深度学习以其强大的学习和表现能力脱颖而出,成为金融领域时间序列预测的首选工具。早在 20 世纪末,Donaldson 就利用人工神经网络(ANNs)对

S&P500 股票价格进行预测,并交叉验证了神经网络相比于传统方法如加权最小二乘方法的优势^[1]。Takeuchi 和 Lee 使用一个基于堆栈限制玻尔兹曼机的自动编码器提取股票价格特征信息,预测哪些股票会有比中值更高或更低的月度收益,最终得到 53% 的准确率以及 45.93% 的年度回报^[2]。国内学者方面也有相关探索。吴微利用 BP 神经网络对股票市场进行预测,取得了 70% 的预测准确率^[3];宋晓勇建立了一个金融参数预测系统,该系统基于神经网络与遗传算法的耦合,利用后者对前者的拓扑结构层权值进行调整,并证实了耦合能够提高系统的预测准确性^[4]。

LSTM 得到重视并在金融领域得到广泛应用始于 2014 年 LSTM 算法在机器翻译领域的巨大成功。Sutskever 等使用多层长短时记忆网络进行 WMT-14 数据集的英法翻译,最

基金项目:中央高校基本科研业务费专项基金;中央财经大学新兴交叉学科建设项目

This work was supported by the Fundamental Research Funds for the Central Universities and Emerging Interdisciplinary Construction Project of Central University of Finance and Economics.

通信作者:方靖雯(stb16fjw@126.com)

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

终整个测试集的翻译结果 BLEU 分数为 34.8, 优于基于短语的 SMT 系统, 且在长句翻译上更有优势^[5]。在语音识别领域, Graves 也利用 LSTM 模型来改进 Google DeepMind 公司的语音识别系统, 将准确率提高到了 93.3%^[6]。随后在处理金融时间序列的应用上, Murtaza 利用 2 层 LSTM 网络对 NIFTY50 股票价格进行预测, 获得 RMSE 为 0.00859 的测试集结果, 预测准确度显著高于计量模型^[7]。LSTM 模型的改进方面, Bao 等首次将小波变换、栈式编码与 LSTM 结合, 利用 6 个股票市场指数期货来验证模型性能, 结果发现其在预测精度与盈利能力上均优于同类模型^[8]。

在对比特币价格预测的研究上, 深度学习与 LSTM 的相关模型也逐渐占据主流。比特币进入大众视野的前期, Shah 等将贝叶斯回归中的潜在源模型应用到比特币的预测中, 通过二元分类方法实现了不错的比特币投资收益^[9]。Yang 等通过研究比特币交易流网络的复杂测度建立了这些复杂测度与比特币收益率与波动率的联合动态关系, 并使用其中一个特殊复杂性变量, 即比特币网络流量的剩余多样性变量, 来提高比特币的波动率可预测性^[10]。随后, McNally 等通过贝叶斯优化递归神经网络与长短时记忆网络进行比特币价格指数的预测, 并将预测结果与常用预测方法 ARIMA 进行比较, 发现非线性深度学习性能明显优于后者^[11]。Stenqvist 和 Lönnö 通过分析 272 万条与比特币有关的评论来对价格进行预测, 预测准确率高达 83%^[12]。Ravi 则从比特币价格数据中提取出了双指数移动平均、简单移动平均、移动平均收敛发散等指标, 采用基于 FuzzyCSA 的 Deep-LSTM 进行预测, 并在公开数据集上进行实验, 预测结果的 MAE 与 RMSE 分别为 0.4811 和 0.3905^[13]。

总结上述文献, 深度学习方法被越来越广泛地应用到比特币价格的预测中, 而长短时记忆网络对长期依赖问题的有效解决使得其在预测中具有一定优势。从具体效果来看, 比特币价格的高度复杂性使得单一预测模型很难取得理想的效果, 利用多种模型集成可以有效提取比特币价格时间序列的不同维度的特征, 提高预测准确性。由此, 本文以 LSTM 预测模型为核心, 构建 WT-LSTM 模型、CEEMDAN-LSTM 模型、CEEMDAN-SE-LSTM 模型以及 WT-SE-LSTM 模型用于比特币价格预测, 并对各模型预测结果进行评估。

本文第 2 节对循环神经网络以及长短时记忆网络的原理与特点进行介绍; 第 3 节具体介绍本文 4 个模型的构建过程以及评价指标; 第 4 节得出预测实证结果并对结果进行分析; 最后总结全文并展望未来。

2 循环神经网络与长短时记忆网络

2.1 循环神经网络

循环神经网络(Recursive Neural Network, RNN)通过构建同一隐含层的神经元之间的联系, 获取数据的前后关联信息, 以此发挥其短期记忆优势。循环神经网络展开图如图 1 所示。从图 1 中可以看出, 除当前时刻输入信息外, 隐含层输入信息还包括上一时刻隐含层的输出信息, 并以此形成时序依赖。凭借这种不同时刻隐含节点的连接结构, RNN 可以实

现对历史时刻信息的记忆并应用于当前输出的计算。

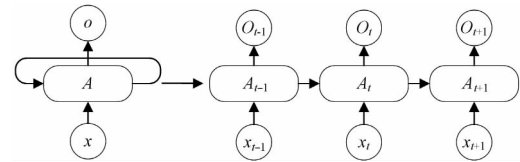


图 1 RNN 展开图

Fig. 1 RNN structure

但当对 RNN 进行参数优化并将其应用于长时间跨度序列处理时, 由于各层共享相同权值参数以及激活函数导数乘数不断累积等, 常常发生梯度消失以及梯度爆炸导致记忆值过小, 故 RNN 在处理长期时序问题时往往效果不佳。

2.2 长短时记忆网络

为了解决前述 RNN 产生的问题, 长短时记忆网络(LSTM)应运而生。作为一种特殊的循环神经网络, LSTM 在 RNN 的基础上引入了门控单元系统, 包括输入门、遗忘门以及输出门, 这 3 类门分别控制当前输入数据、历史数据对记忆单元状态值的更新以及输出, 由不同的门对信息通过进行选择控制, 使得网络学习可以适当地遗忘历史信息并依据新信息更新细胞状态。其基本结构如图 2 所示。

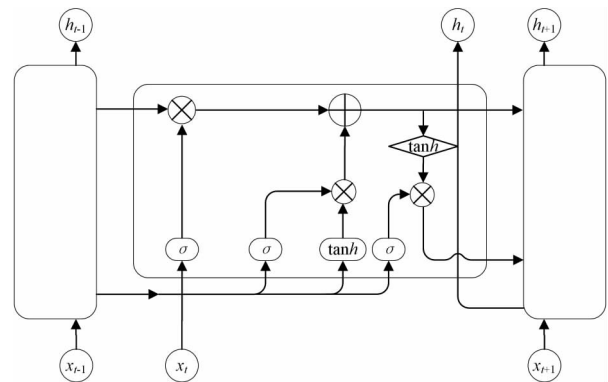


图 2 LSTM 单元内部结构

Fig. 2 LSTM unit internal structure

LSTM 单元的计算过程分为以下几步。

(1) 分别对 t 时刻的候选记忆单元值 \tilde{c}_t 、输入门的值 i_t 以及遗忘门的值 f_t 进行计算, 计算公式如下:

$$\tilde{c}_t = \tanh(\omega_c[h_{t-1}, x_t] + b_c) \quad (1)$$

$$i_t = \sigma(\omega_i[h_{t-1}, x_t] + b_i) \quad (2)$$

$$f_t = \sigma(\omega_f[h_{t-1}, x_t] + b_f) \quad (3)$$

其中, $\omega_c, \omega_i, \omega_f$ 分别为对应的权重矩阵; b_c, b_i, b_f 分别为对应的偏置; h_{t-1} 为 t 时刻前一时刻的 LSTM 单元输出; x_t 为 t 时刻记忆单元值; σ 为 sigmoid 函数。

(2) 将旧状态与遗忘门信息相乘并丢弃部分信息, 加上输入门与候选记忆单元值, 得到当前时刻记忆单元的值 c_t , 公式为:

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (4)$$

(3) 最后通过输出门确认输出门的值 o_t , 并确定输出部分 h_t :

$$o_t = \sigma(\omega_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t \tanh(c_t) \quad (5)$$

其中, ω_0, b_0 为其对应的权重矩阵与偏置。

通过建立以上控制门与记忆单元结构,可以实现更长期数据保存与处理,而梯度爆炸等问题也能通过建立输入反馈的长时期延迟得到解决。

3 混合模型构建与评价

3.1 WT-LSTM 模型

小波变换本身具有处理非平稳金融时间序列数据的能力,通过对原始信号进行多层分解重构可以分离趋势信息与波动信息,分解过程如下。

(1) 首先通过快速二进正交小波变换(Mallat 算法)分解:

$$D_n = H A_{n-1} \quad (6)$$

$$A_n = G A_{n-1} \quad (7)$$

其中, H 和 G 分别为低通滤波器与高通滤波器, n 为分解次数, A_0 为初始时间序列。原始数据第一次分解得到 D_1 与 A_1 分量,下一次分解对近似信号 A_1 进行,得到 A_2 与 D_2 ,如此持续进行 n 次,直至得到 $n+1$ 个信号序列。

(2) 二进制采样导致的数据缺失问题使用插值法来进行恢复重构:

$$a_{n-1} = H^* A_n + G^* D_n \quad (8)$$

其中, H^* 和 G^* 为 H 和 G 的对偶算子,并使得重构序列之和等于原序列。

本文在母小波的选取上选择适用范围最大的 Daubechies4,分解层数选择 4 层。通过小波变换对比特币价格时间序列进行多层分解,分解首先得到低频近似序列 A_4 以及不同频率的高频细节序列 D_1, D_2, D_3, D_4 ,插值法恢复重构得到近似序列 a_4 以及细节序列 d_1, d_2, d_3, d_4 。

重构后的子序列分别用 LSTM 进行预测,最后对预测得到的子序列进行加总,得到最终预测结果。预测流程如图 3 所示。

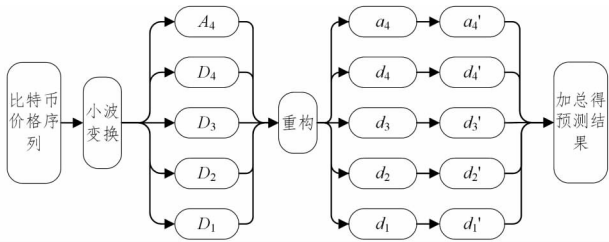


图3 WT-LSTM模型预测流程

Fig. 3 WT-LSTM forecasting procedure

LSTM 模型训练参数设置为:LSTM 层隐含单元个数为 200,训练最大迭代次数为 200,梯度阈值设置为 1,初始学习率为 0.005,迭代 125 次后通过乘因子 0.2 来降低学习率,预测步长为 1。下文 LSTM 模型参数设置相同。

3.2 CEEMDAN-LSTM 模型

自适应噪声完备经验模态分解(CEEMDAN)是基于经验模态分解(EMD)的改进,后者也是一种分析非线性、非平稳数据的方法,通过将序列分解为一系列本征模态函数(imf)分量,来代表不同时间尺度上的数据特征。但是未被改进的 EMD 存在模态混叠缺陷,于是 CEEMDAN 基于集合经验模态分解(EEMD)的白噪声方法提出了在原始数据中加入独立

分布的高斯白噪声的改进措施,自适应的噪声添加同时解决了模态混叠与噪声过量残留问题,提高了分解效率。CEEMDAN 算法过程如下。

(1) 在原始时间序列基础上加入正态分布的高斯白噪声:

$$y_i(n) = y(n) + \epsilon_0 \omega^i(n), i = 1, 2, 3, \dots, t \quad (9)$$

其中, $y(n)$ 为原始序列, $\omega^i(n)$ 为高斯白噪声, ϵ_0 为噪声的标准差, t 为噪声的添加次数。

(2) 根据 EMD 方法得到一阶模态分量 $imf_1(n)$,取均值作为第一个 imf 分量,并计算出第一阶段后的残差值:

$$imf_1(n) = \frac{1}{t} \sum_{i=1}^t imf_1^i(n) \quad (10)$$

$$res_1(n) = y(n) - imf_1(n) \quad (11)$$

(3) 同理,对将残差项作为原始时间序列,重复步骤(1)、步骤(2),加入自适应的高斯白噪声后进行 EMD 分解,得到 $imf_2(n)$ 与相应残差值,重复上述步骤直至残差量不可继续分解,即残差项已经是一个单调函数或常数,或振幅低于既定阈值无法继续提取下一模态函数时,分解过程结束。最终得到 k 个正交的 imf 函数与最终趋势项 res_k :

$$y(n) = \sum_{i=1}^k imf_i(n) + res_k \quad (12)$$

基于 CEEMDAN 在序列分解上的优点,本文构建了 CEEMDAN-LSTM 模型来进行比特币价格预测。图 4 为预测的流程:首先将比特币价格序列作为输入数据,经 CEEMDAN 分解后得到各 imf 分量以及趋势项,通过对 imf 分量的重构得到 3 个预测分量,即高频分量、低频分量以及趋势项,将以上三大分量输入 LSTM 模型进行预测并加总得到最终的预测结果。

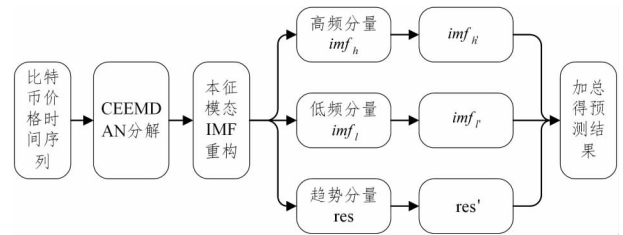


图4 CEEMDAN-LSTM模型预测流程

Fig. 4 CEEMDAN-LSTM forecasting procedure

重构方法参照 Zhang 等^[14]的方法。由于高频的分量反映短期波动,对原序列影响较小,均值接近零,故可先通过 t 检验对高频分量做出筛选,然后合并为一个新序列 imf_h ,反映随机波动,除残差项以外的剩余分量合并为低频分量 imf_l ,反映各种因素,如基本面、技术调整等对序列的影响,趋势项则反映原序列的变化趋势。具体操作如下。

(1) 计算分解出的各本征模态 $imf_i(n)$ 的均值, $i = 1, 2, 3, \dots, k$;

(2) 对 $imf_i(n)$ 依次进行显著水平为 0.05、均值不为零的 t 检验;

(3) 顺次检验后,得到第一个均值显著非零的分量 $imf_q(n)$,将 $imf_1(n)$ 至 $imf_{q-1}(n)$ 相加得到高频子序列 imf_h ,将 $imf_q(n)$ 至 $imf_k(n)$ 相加得到低频子序列 imf_l , res_k 继续作为趋势项。

参数设置上参考 Torres^[15] 参数设置原则, 在利用 CEEMDAN 对原始序列进行分解时, 加入标准差为 0.2 的高斯白噪声, 添加次数为 500, 最大迭代次数取 2000。

3.3 CEEMDAN-SE-LSTM 模型

在上述模型的基础上, 本文引入样本熵作为 imf 分量重构依据。样本熵从时间序列复杂度出发, 定量地描述系统复杂度以及规则度, 从而判断产生新模式概率的大小。计算所得熵值越大代表时间序列越复杂, 产生新模式的概率也就越大, 反之序列越简单, 产生新模式的概率也越小。

样本熵的计算过程如下。

(1) 对给定时间序列 $y(n)$, 按照序号排列得到一组 k 维向量 $y_k(1), \dots, y_k(n-k+1)$, 其中 $y_k(i) = \{y(i), y(i+1), \dots, y(i+k-1)\} (1 \leq i \leq n-k+1)$;

(2) 定义 $y_k(i)$ 与 $y_k(j)$ 的距离为两者对应元素中最大差值的绝对值, 记为 $d[y_k(i), y_k(j)]$;

(3) 给定阈值 h , 对每一个 i , 计算 $d[y_k(i), y_k(j)] < r$ 的数目, 记为 $N_k(i)$, 并定义

$$B_k^r(i) = \frac{N_k(i)}{n-k+1}, 1 \leq i \leq n-k+1, i \neq j \quad (13)$$

(4) 计算上述所有上述定义值的均值, 记为 B_k^r ;

$$B_k^r = \frac{1}{n-k+1} \sum_{i=1}^{n-k+1} B_k^r(i) \quad (14)$$

(5) 重复以上步骤, 得到 B_{k+1}^r , 当 n 有限时即可求出样本熵 $SampEn$ 的估计值为

$$SampEn(k, r, n) = -\ln\left(\frac{B_{k+1}^r}{B_k^r}\right) \quad (15)$$

基于样本熵在判断序列复杂度与新模式概率上的特点, 本文引入样本熵的计算作为重构依据。不同于前两个模型中以低频与高频作为重构依据, 本模型需要计算样本熵值。样本熵越接近, 代表分量之间越相似, 波动越具有一致性特征。本模型的预测流程如图 5 所示。

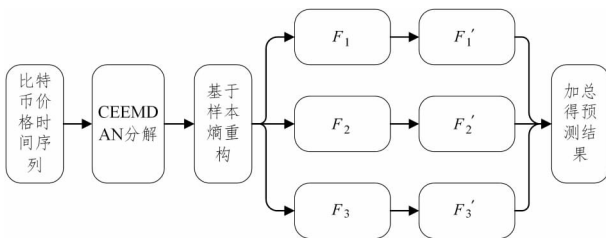


图 5 CEEMDAN-SE-LSTM 模型预测流程

Fig. 5 CEEMDAN-SE-LSTM forecasting procedure

在参数选取上, 取 $k=2, r=0.2 \text{std}(y(n))$ 。

3.4 WT-SE-LSTM 模型

本文还将样本熵重构法应用到小波变换-LSTM 模型中, 对插值法重构后得到的近似序列 a_4 以及细节序列 d_1, d_2, d_3, d_4 进行熵值计算, 按照计算结果叠加以后输入 LSTM 进行预测, 相关步骤参照前述模型即可, 此处不再赘述。

3.5 模型评价指标

衡量指标方面, 对于比特币价格的预测结果, 本文采用均方根误 (RMSE)、平均绝对百分误 (MAPE) 以及希尔不等系

数 (TIC) 来检验其精确度, 指标公式如下:

$$RMSE(y_i, \hat{y}_i) = \sqrt{\frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{n}} \quad (16)$$

$$MAPE(y_i, \hat{y}_i) = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (17)$$

$$TIC(y_i, \hat{y}_i) = \frac{\sqrt{\frac{\sum_{i=0}^n (y_i - \hat{y}_i)^2}{n}}}{\sqrt{\frac{1}{n} \sum_{i=0}^n y_i^2} + \sqrt{\frac{1}{n} \sum_{i=0}^n \hat{y}_i^2}} \quad (18)$$

其中, y_i 为实际的比特币价格, \hat{y}_i 为经模型预测后的比特币价格。

4 数据与实证结果分析

本文选取了 2019 年 1 月 1 日至 2020 年 12 月 31 日比特币兑美元价格共 731 个数据作为研究对象。数据进行归一化处理以后分为训练集与测试集两部分, 取其中 90% 的数据为训练集, 对预测模型进行训练, 余下 10% 的数据作为测试集对预测结果进行分析。

4.1 WT-LSTM 模型预测结果

比特币价格序列经小波变换得到 5 个分量, 各分量数据经 LSTM 预测以后的子序列预测结果如图 6 和图 7 所示。将子序列预测结果相加得到价格预测结果, 图 8 为预测结果与实际价格的对比图。

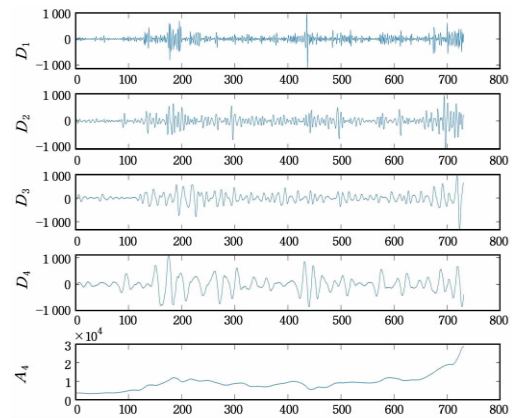


图 6 WT 分解分量

Fig. 6 WT decomposed components

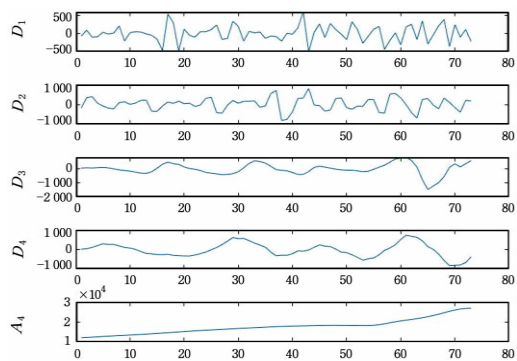


图 7 各分量预测结果

Fig. 7 Forecasting result of components

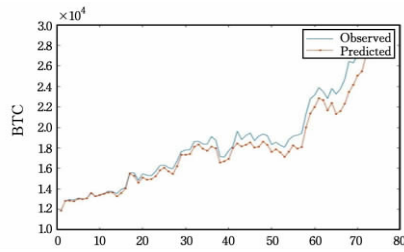


图 8 WT-LSTM 模型预测结果

Fig. 8 Forecasting result of WT-LSTM

4.2 CEEMDAN-LSTM 模型预测结果

CEEMDAN 分解得到的各 imf 分量如图 9 所示。

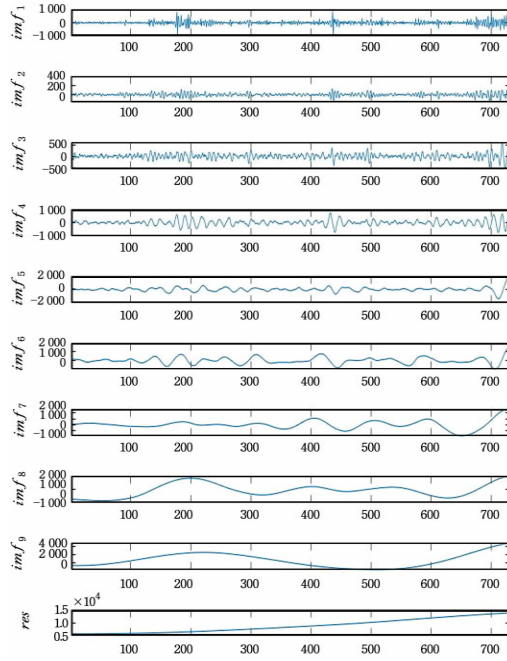


图 9 CEEMDAN 分解分量

Fig. 9 CEEMDAN decomposed components

进行重构之前,需要对各 imf 分量进行 t 检验,检验结果如表 1 所列。从表中我们可以发现, $imf_1 - imf_7$ 的显著性均大于 0.05,可以认为该分量均值与 0 在置信水平为 95%的情况下没有差异,故 $imf_1 - imf_7$ 可以被认定为高频分量。而 imf_8 与 imf_9 的显著性均小于 0.05,说明其均值与 0 存在显著差异,进而被认定为低频分量。由此我们将这 10 个分量重构为高频分量、低频分量与趋势分量,重构序列与序列预测结果如图 10 和图 11 所示。

表 1 Imf 分量与趋势项 t 检验结果

Table 1 Imf components and trends' t-test result

Component	Observations	Average	Test value=0		
			t-value	Significance	Average Difference
imf_1	731	0.8508	0.140	0.889	0.8508
imf_2	731	0.7823	0.536	0.592	0.7823
imf_3	731	1.8637	0.405	0.686	1.8637
imf_4	731	0.8907	0.110	0.912	0.8907
imf_5	731	9.7394	0.791	0.429	9.7394
imf_6	731	-4.1594	27.4	0.754	-4.1594
imf_7	731	11.0846	23.2	0.668	11.0846
imf_8	731	-124.7669	-3.710	0.000	-124.7669
imf_9	731	796.1480	12.671	0.000	796.1480
res	731	8565.7287	92.911	0.000	8565.7282

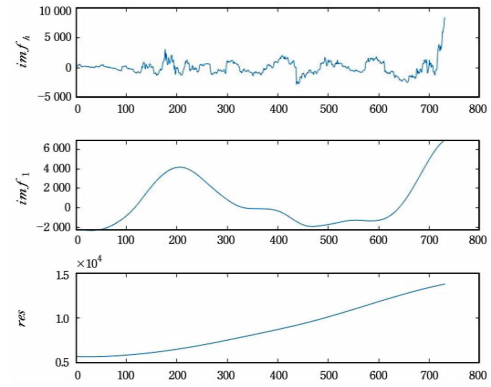


图 10 分量重构结果

Fig. 10 Components reconstruction results

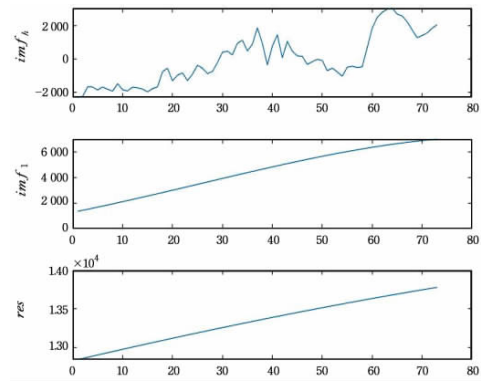


图 11 重构分量预测结果

Fig. 11 Reconstructed components forecasting result

随后将三分量预测结果进行加总,最终预测结果如图 12 所示。

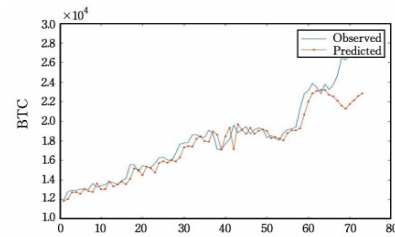


图 12 CEEMDAN-LSTM 模型预测结果

Fig. 12 Forecasting result of CEEMDAN-LSTM

4.3 CEEMDAN-SE-LSTM 模型预测结果

各 imf 分量的样本熵值如图 13 所示,可以看到各分量样本熵随着 imf 分量频率的降低整体呈递减趋势,对相邻熵值相差不大的 imf 分量进行合并重叠,即 imf_1, imf_2, imf_3 叠加为 F_1 序列, imf_4, imf_5, imf_6 叠加为 F_2 序列,剩余 imf 与 res 序列进行叠加,得到 3 个一致性序列,减少 lstm 预测分量数。熵值与重构后的 3 个分量如图 13 和 14 所示。

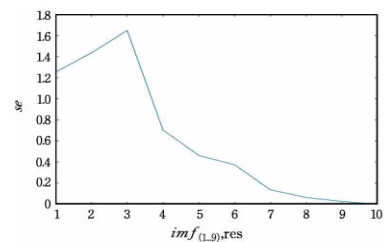


图 13 imf 样本熵值计算结果

Fig. 13 imf sample entropy result

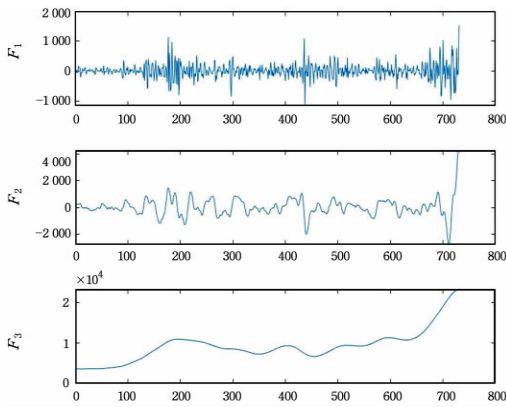


图 14 根据样本熵叠加重构结果

Fig. 14 Reconstruction result based on sample entropy

随后进行分量预测与加总。三分量预测结果以及加总结果如图 15 和图 16 所示。

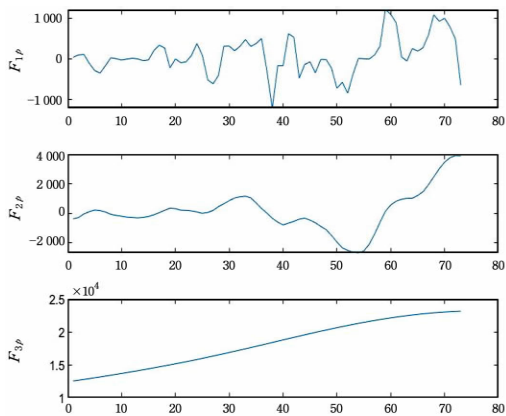


图 15 重构三分量预测结果

Fig. 15 Forecasting result of reconstructed components

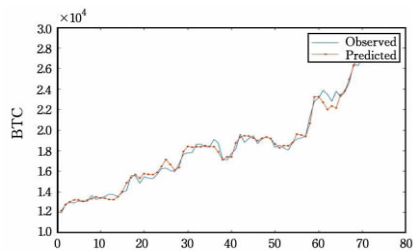


图 16 CEEMDAN-SE-LSTM 模型预测结果

Fig. 16 Forecasting result of CEEMDAN-SE-LSTM

4.4 WT-SE-LSTM 模型预测结果

在对小波变换的分量进行熵值求解后,基于熵值变化(见图 17),将 d_1, d_2, d_3, d_4, a_4 划分为: $F_1(d_1), F_2(d_2, d_3, d_4), F_3(a_4)$, 叠加结果如图 18 所示。

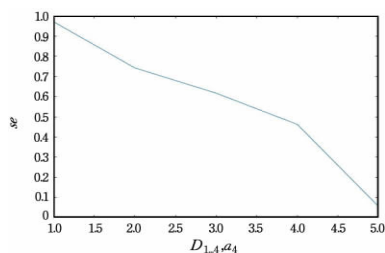


图 17 小波变换分量熵值计算结果

Fig. 17 Entropy of wavelet transform component

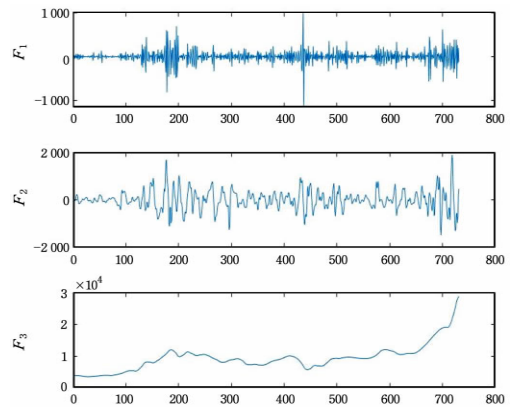


图 18 样本熵叠加重构结果

Fig. 18 Reconstruction result based on entropy

分量预测结果以及加总结果如图 19 和图 20 所示。

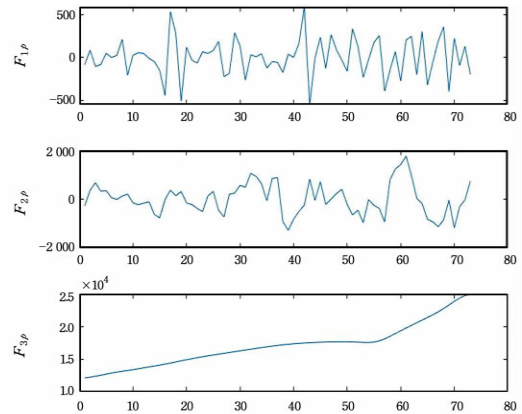


图 19 重构分量预测结果

Fig. 19 Forecasting result of reconstructed components

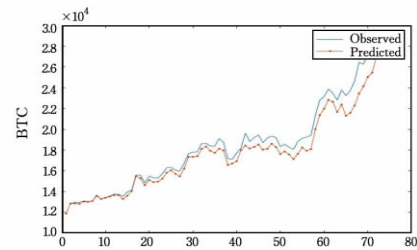


图 20 WT-SE-LSTM 模型预测结果

Fig. 20 Forecasting result of WT-SE-LSTM

4.5 拟合评价

对 4 个模型的预测结果以及实际价格进行误差估计,以评价模型与真实值的拟合程度。经对比发现,与单一 LSTM 模型预测相比,本文的 4 个模型均有明显优势。各模型的预测误差如表 2 所列。

表 2 各模型比较
Table 2 Model comparison

Model	Evaluation		
	MAPE/%	RMSE	TIC/%
LSTM	28.23	5744.1	17.69
WT-LSTM	4.55	958.466	2.61
CEEMDAN-LSTM	5.47	1769.12	4.83
CEEMDAN-SE-LSTM	0.04	519.63	1.38
WT-SE-LSTM	3.07	700.03	1.89

由表 2 可知,CEEMDAN-SE-LSTM 模型的预测效果最好,其 MAPE 值、RMSE 值以及 TIC 值分别为 0.04%,

519.63 和 1.38%,与 CEEMDAN-LSTM 模型相比均明显下降,说明采用熵值法进行分量重构比单用低高频特性重构效果更好。除此之外,加入样本熵重构后的 WT-SE-LSTM 模型预测准确率也有所提高,但优化程度不及 CEEMDAN-SE-LSTM 模型,这体现了样本熵方法在较高分量数情况下具有更强的适用性。

结束语 由于比特币价格具有高度非线性、非平稳性特点,LSTM 与传统预测模型相比具有优势,但仅使用单一 LSTM 模型进行预测时,因多尺度的周期特征捕捉困难,最终准确度方面仍有较大进步空间,故本文构建 LSTM 混合模型来进行预测。本文的最终结果表明:混合预测模型的预测效果明显优于单一模型,MAPE,RSME,TIC 指标均有所下降,而引入样本熵重构优化以后的模型可以明显提升预测结果的准确度。

本文建立的 4 个混合模型提供了对比特币价格的预测思路,运用该思路进行实验能得出比较精确的预测结果。该结果可以支持各国政府监管部门的决策,也可以帮助比特币投资者降低投资风险,实现收益平稳化。特别地,上述模型在比特币价格预测上的效果验证说明其在其他金融时间序列预测上具有应用空间,例如大宗商品现货期货价格、房地产买卖租赁价格等,并能够支持资产证券化的风险管理。

参 考 文 献

- [1] DONALDSON R G,KAMSTRA M. Neural network forecast combining with interaction effects[J]. Journal of the Franklin Institute, 1999, 336(2): 227-236.
- [2] TAKEUCHI L,LEE Y Y A. Applying deep learning to enhance momentum trading strategies in stocks[R]. Technical Report, Stanford University, 2013.
- [3] WU W,CHEN W Q,LIU B. Prediction of ups and downs of stock market by BP neural networks [J]. Journal of Dalian University of Technology, 2001, 41(1): 9-15.
- [4] SONGX Y,CHEN N S. Financial prediction system based on combination of genetic algorithm with BP neural network [J]. Journal of Shanghai Jiao Tong University, 2016, 50(2): 313-316.
- [5] SUTSKEVER I,VINYALS O,LE Q V. Sequence to sequence learning with neural networks[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems-Volume 2, 2014:3104-3112.
- [6] GRAVES A,JAITLEY N. Towards end-to-end speech recognition with recurrent neural networks[C]// International Conference on Machine Learning. PMLR, 2014: 1764-1772.
- [7] ROONDIWALA M,PATEL H,VARMA S. Predicting stock prices using LSTM[J]. International Journal of Science and Research, 2017, 6(4): 1754-1756.
- [8] BAO W,YUE J,RAO Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory[J]. PloS one, 2017, 12(7): e0180944.
- [9] SHAH D,ZHANG K. Bayesian regression and Bitcoin[C]// 2014 52nd annual Allerton conference on communication, control, and computing (Allerton). IEEE, 2014: 409-414.
- [10] YANG S Y,KIM J. Bitcoin Market Return and Volatility Forecasting Using Transaction Network Flow Properties[C]// 2015 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2015.
- [11] MCNALLY S, ROCHE J, CATON S. Predicting the price of bitcoin using machine learning[C]// 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP). IEEE, 2018: 339-343.
- [12] STENQVIST E,LÖNNÖ J. Predicting Bitcoin price fluctuation with Twitter sentiment analysis[EB/OL]. (2017-06-16) [2018-01-13]. <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-209191>.
- [13] RAVI C. Fuzzy Crow Search Algorithm-Based Deep LSTM for Bitcoin Prediction[J]. International Journal of Distributed Systems and Technologies (IJ DST), 2020, 11(4): 53-71.
- [14] ZHANG X,LAI K K,WANG S Y. A new approach for crude oil price analysis based on empirical mode decomposition[J]. Energy Economics, 2008, 30(3): 905-918.
- [15] TORRES M E, COLOMINAS M A, SCHLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise [C] // 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2011: 4144-4147.



ZHANG Ning, born in 1978, Ph.D, professor. His main research interests include fintech and artificial intelligence.



FANG Jing-wen, born in 1999, master. Her main research interests include sequence forecasting and so on.