

Bitcoin price forecasting: A perspective of underlying blockchain transactions

摘要

加密货币价格预测在金融市场中发挥着重要作用。传统方法面临两个挑战：(1)难以确定影响价格预测的相关因素；(2)由于24/7的交易政策，加密货币的价格面临很大的波动，从而削弱了传统模型的预测能力。为了解决这些问题，我们将重点放在比特币上，并从底层区块链交易的角度确定与其价格预测相关的影响因素。然后，我们提出了一个价格预测模型WT-catcn，它利用小波变换(WT)和随机多头注意力机制(CA)时间卷积网络(TCN)来预测加密货币的价格。该模型可以捕获输入序列的重要位置，并对不同数据特征之间的相关性进行建模。使用真实世界的比特币交易数据，我们测试并比较了WT-CATCN与其他最先进的价格预测模型。实验结果表明，该模型的价格预测性能提高了25%。

关键词： 加密货币；区块链；比特币；价格预测；深度学习

1 引言

选题背景：

加密货币是一种新型数字资产，利用基于区块链技术和密码学的去中心化网络来促进、保护和验证交易。比特币是代表性的加密货币，采用区块链技术和点对点交易的支付系统。由于比特币支付协议的独特性和日益增长的流行度，全球企业、消费者和投资者都对它非常关注。然而，比特币价格波动性很高，因此需要进行准确的价格预测以促进投资决策。此外，在决定是否采用比特币作为支付方式之前，零售商和大型企业需要了解其价格走势。研究人员对比特币及其基础区块链技术也表现出了极大的兴趣。因此，需要确定影响新金融技术（尤其是比特币）传播的重要因素。

选题依据：

- 1、文章调查了加密货币市场和传统市场中的价格预测方法，对于我来说，这是一个有价值的研究领域。
- 2、文章提出了一个基于深度学习模型的价格预测模型，并对其进行了评估和可扩展性研究，这能为我进一步研究提供基础和参考。
- 3、文章还探讨了多源数据的应用，如情感数据和社交媒体数据，以提高价格预测的准确性，值得我探讨和学习。

选题意义：

本文章通过分析区块链交易的特征，发现大型和小型交易所之间的交易量差异与比特币价格存在相关性，并利用这一相关性来预测比特币价格。同时，本文章还提出了一种名为WT-CATCN的新型高维预测模型，它结合了WT-CATCN模型能够清晰地检测比特币的价格走势，这在传统方法中几乎是不可能实现的。本文章的意义在于实现更有效的价格预测，从而使金融公司能够提高处理与高波动价格预测相关问题的能力，并获得持续的竞争优势。

2 相关工作

在本节中，不仅调查了加密货币市场的价格预测，还调查了传统市场的价格预测。

2.1 加密货币市场的价格预测

已经提出了几种方法来预测比特币和其他加密货币的价格。例如，Katsiampa 等人[1]使用每日收盘价评估 GARCH 型模型并解释比特币价格的波动性。他们发现，根据拟合优度，AR-GARCH 是最好的模型之一。Mallqui 等人[2]提出了预测比特币汇率的机器学习集合算法。结合人工神经网络和自回归条件异方差，从 12 种不同的模型组合中选择了 ANN-GARCH 框架，并根据比特币价格序列预测比特币价格波动[3]。McNally 等人[4]基于价格数据使用贝叶斯优化的循环神经网络（RNN）和长短期记忆（LSTM）预测价格变化的三种状态（价格上涨、下跌和不变），发现深度学习模型在时间序列预测方面比传统的 ARIMA 模型有更好的表现。Alonso-Monsalve 等人[5]研究了不同的神经网络架构，认为 CNN 可用于比特币预测。上述仅使用过去价格数据的方法无法在比特币价格预测中实现高性能。

最近的研究已纳入了多种数据源，如基于情感的数据源，以实现更好的预测性能。Kim 等人 [6] 分析了用户在加密货币在线社区的用户评论进行情感标记，并使用平均单依赖估计器（AODE）来预测价格。Stenqvist 等人[7] 首先将 Twitter 上有关比特币的情绪分为三类：积极、中性和消极。然后，他们利用这三类情绪来预测价格的三种状态，结果表明所提出的预测模型达到了很高的预测准确率。同样，Li 等人[8] 利用 Twitter 上的情绪数据预测比特币的价格波动。他们认为，使用语义进行预测对性能的提高是有限的。因为有些语言模式（如讽刺）很难分类。Mohapatra 等人[9] 基于 Twitter 情绪建立了一个加密货币价格预测平台，支持实时预测。他们利用 Twitter 中的观点信息预测比特币价格[10]。Jain 等人[11]认为，采矿成本和经济因素等其他特征也可能影响比特币的价格波动。

除了情感数据，还引入了其他数据源。Rebane [12]的研究表明，在使用谷歌趋势、Altcoin 数据和比特币价格等其他输入源时，Seq2Seq 的预测性能比 ARIMA 更好。但是，如果比特币价格大幅下跌，Seq2Seq 模型的性能就会显著下降。Jang 等人[13]使用贝叶斯神经网络（BNN）根据从区块链中提取的特征预测比特币价格，他们发现了区块链特征与宏观经济变量之间的共线性问题。他们的实验表明，BNN 在波动性和波动率方面的预测性能均优于线性回归（LR）和支持向量回归（SVR）模型。Amjad 等人[14]在提取了比特币的买入价和卖出价等一般属性后，使用简单的实时机器学习模型预测了价格的三种状态并模拟了交易策略，其结果优于 EC 模型和 ARIMA 模型。Guo 等人[15]提出了基于比特币价格和订单数据的概率时间混合模型来预测短期波动率。订单数据包括买入订单（买入的最高价）和卖出订单（卖出的最低价），这些数据存储在交易所数据库中。

在本文中，研究了从底层区块链上交易所之间的交易中生成的一组新特征，并探讨了它们在比特币价格预测中的作用。据我们所知，本文是利用此类信息预测比特币价格的先驱作品之一。

2.2 传统市场的价格预测

加密货币的价格预测问题与外汇、股票、商品等传统金融市场的价格预测问题类似。本文研究了一些可用于预测加密货币价格变化的方法。例如，自回归模型被用于研究可预测性和股票收益[16]。由于股票价格的非线性和非平稳性特征，传统的自回归模型在进行准确预测方面的能力有限。为了解决股票价格的非线性问题，许多机器学习技术得到了应用。Hassan 等人[17]提出了一种基于隐马尔可夫模型（HMM）的自适应模糊推理系统来分析价

格趋势。Nayak 等人[18]提出了一种混合框架，它结合了支持向量机（SVM）和 K-最近邻方法，以平衡每个模型的复杂性和误差，实现了对股票市场指数的良好预测性能。

深度学习是一种功能强大的方法，它给包括金融行业在内的各行各业带来了革命性的变化[19]。多项研究表明，人工神经网络（ANN）模型比 ARIMA 模型具有更好的性能。Jiao 等人[20]使用标准机器学习模型（包括 ANN 模型）预测了标准普尔 500 指数中 463 只股票的股价变化。Faccini 等人[21]提出了一个新的预测指标，即代表性投资者的隐含相对风险厌恶程度。Neely 等人[22]的研究表明，技术指标（依赖于过去的价格和成交量模式）和宏观经济变量（如股息价格比）可以提高预测性能。为了提取多频交易模式，受离散傅里叶变换的启发，开发了状态频率记忆（SFM）[23]。Sul 等人[24]利用社会情绪和新闻来预测股票价格。但是，他们只是将向量连接成一个超级特征向量，因此这种方法会失去特征之间的内在联系。因此，Li 等人[25]利用股票数据比较了基于张量的方法与支持向量回归（SVR）、PCA+SCR 和 ISOMAP+SVR 的效果，结果表明基于张量的方法优于其他模型。

与传统金融市场不同，比特币作为一种去中心化的数字货币，可以全天候不打烊地进行交易。此外，任何离线信息和事件都可能立即影响加密货币的价格，而不是在市场（如股票市场）开盘时。我们分析了作品[23]中的股票数据集，发现 50 种股票价格标准偏差的平均值为 23:97。然而，比特币价格的标准偏差为 3923:15。因此，比特币比大多数传统金融资产面临更大的波动，这表明迫切需要一个新的预测模型来应对如此高的价格波动。

3 本文方法

在本节中对本文将要复现的工作进行概述。

3.1 系统概述

为了解决加密货币价格预测问题，文章提出了一个加密货币价格预测系统，如图 1 所示。该系统包括两大部分：特征提取模块和预测模型。系统的第一部分涉及从三类数据集中提取代表性特征(X_d^T, M_d^T, B_d^T)，详情将在第 3.2 节中讨论。在第二部分中，文章提出了一个名为 WT-CATCN 的预测模型，该模型基于小波变换（WT）和随性多头注意力时空卷积网络（CATCN）。小波变换用于减轻非线性和非平稳输入数据的影响，而 CATCN 则通过在传统 TCN 中添加随意多头注意力来处理输入数据之间的关系。详情见第 3.3 节。

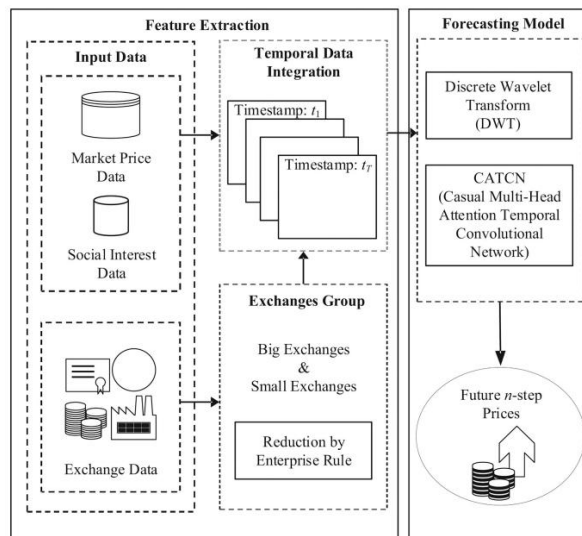


图 1. 新的加密货币价格预测系统

3.2 特征提取

本节将介绍从输入数据中提取特征的方法。根据上一节的定义，我们的输入数据包括交易所间交易数据 X_d^T 、交易所内市场价格数据 M_d^T 和社会兴趣数据 B_d^T 。

大小交易所之间的交易量差异可能会对比特币价格产生潜在影响。因此，我们首先清理数据并生成交易所间交易量 τ_i 。然后，我们将 τ_i 分成两个交易集：大交易所和小交易所。在此，我们采用行业规则，以交易量排名前 k 的交易所作为区分大交易所和小交易所的阈值。具体细节如下。

我们从 WalletExplorer 收集并预处理加密货币交易 ξ 。由于收集的数据来自底层区块链，因此这些底层交易数据是可靠的。由于交易有多个发送方和接收方，我们根据发送方和接收方构建了区块链交易的卡方交易集。然后还记录所有交易的交易费。最后一步是从卡方交易集中选择交换间交易 τ_i 。WalletExplorer 对地址进行了分组和标记，因此我们可以排除发送方和接收方没有标记的交易。

下一步是对 τ_i 进行分割；我们消除了 $\{\tau_i(p), \tau_i(r) | \tau_i \in X_d^T\}$ ，然后形成一个唯一的元素集 u ，它由 τ_i 的接收者或发送者。接下来，我们将从时间 $t-1$ 到 t 期间产生的交易所 u_j 的交易集表示为 $H_t^{u_j} = \{\tau_i | \tau_i(p) = u_j, t-1 < \tau_i(\delta) \leq t\}$ ，其中 $t = \{d-T+1, \dots, d\}$ 。为了选出前 k 个交易所，我们按照广受认可的 CoinMarketCap 网站的建议，计算每个交易所的交易量，并据此对交易所进行排名。对于交换 u_j ，我们将交换的 vol_{u_j} 定义为：

$$\text{vol}_{u_j} = \sum_{t=d-T+1}^d \sum_{\tau_i \in H_t^{u_j}} (\tau_i(v) + \tau_i(s)) v_t \quad (1)$$

我们将交易量排名前 k 的交易所定义为大型交易所，它们的交易量占总交易量的近一半，而且交易量的交易量总是排在前 k 位。然后，我们将其他交易所定义为小交易所。如图 2 所示，排名前 4 位的交易所约占市场交易量的 50%。占市场交易量的 50% 左右。根据行业规则，我们将 k 设为 4。

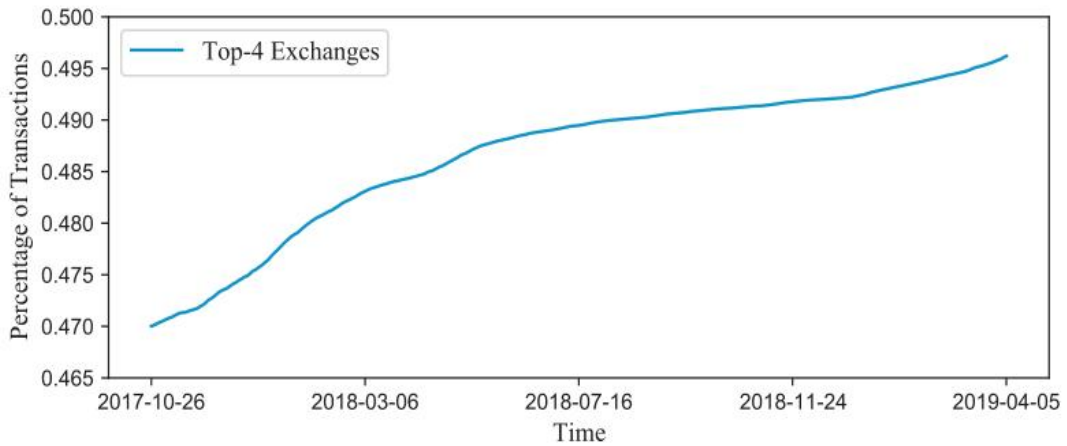


图 2. 排名前4的交易所主导着整个比特币交易

为了探索大型和小型交易所之间的交易行为，我们记录了它们之间的所有交易。我们忽略了大型交易所之间的交易，因为它们被视为内部交易。同样，小型交易所之间的交易也不被考虑。

小交易所的流入就是大交易所的流出，反之亦然。对于这些对称链接，我们只需要计算一次交易的流出 $\tau_i(s)$ 和相应的交易费用。

我们将提取的特征定义为 x_t^z ，其中 $t=\{d-t+1, \dots, d\}$ ， z 是由表1所示特征组成的向量。最后四个特征是真实世界中转换成美元的比特币价格。

The features of \mathcal{X}_t^z .		
Data source		Features
Market prices data	z_m	The market price m_t as defined in Definition 3.
Social interest data	z_g	The social behavior data g_t^r as defined in Section 3.2.
Inter-exchange transaction data	z_1	The average number of Bitcoin a_s flowing from big exchanges to small exchanges.
	z_2	The average transaction fee e_s flowing from big exchanges to small exchanges.
	z_3	The average number of Bitcoin a_l flowing from small exchanges to big exchanges.
	z_4	The average transaction fee e_b flowing from small exchanges to big exchanges.
	z_5	The average number of USD $a_s^p = a_s p_t$ flowing from big exchanges to small exchanges.
	z_6	The average transaction fee (USD) $e_s^p = e_s p_t$ flowing from big exchanges to small exchanges.
	z_7	The average number of USD $a_b^p = a_b p_t$ flowing to from small exchanges to big exchanges.
	z_8	The average transaction fee (USD) $e_b^p = e_b p_t$ flowing from small exchanges to big exchanges.

表 1

3.3 预测模型

在本节中，我们旨在使用时间序列特征预测加密货币的 n 阶价格 \widehat{P}_d^n 用 x_t^z 表示。我们的预测模型WT-CATCN基于小波变换（WT）和随机多头注意力时间卷积网络（CATCN）。为了解决输入数据的非线性和非平稳问题，我们利用离散小波变换（DWT）的小波分解来分解 x_t^z 转换为时频特征（详见第3.3.1节）。

接下来，我们利用CATCN，该CATCN堆叠多个CATCN块来预测 \widehat{P}_d^n 的时频特征。每个CATCN块包含偶然的多头注意和TCN层。

随机多头注意力（详见第3.3.2节）可以使区块集中在数据的重要位置，同时确保因果关系。至于TCN层（详见第3.3.3节），其卷积是因果的，具有类似于LSTM的长期记忆。每个CATCN块（详见第3.3.4节）通过叠加因果自我关注层和TCN层来定义。最后，我们通过小波重构将时频价格 \widehat{P}_d^n 重构为最终价格（详见第3.3.5节）。

3.3.1 离散小波分解

采用小波变换来减轻非线性和非平稳输入数据的影响。它的主要特性是可变的时频分辨率，这允许小波变换具有具有较低时间解的较高频率解（局部行为），或具有较高时间解的较低频率解（一般行为）。WT的使用也方便了训练神经网络[26]。因此，我们分解 x_t^z 转换为 z 的多个时频特征，以探索输入数据中的不同隐藏模式。详情如下所示。

小波用于将输入序列分解成不同的尺度分量。离散小波[27] ψ 如下：

$$\Psi_{n_1, n_2}(t) = a^{-\frac{n_1}{2}} \Psi(a^{-n_1} t - b n_2) \quad (2)$$

其中 a 是比例因子， b 是平移因子。比例因子 a 控制小波的拉伸或收缩。平移因子 b 可以移动小波以在输入序列上移动 x_t^z 表示。 \mathbb{Z} 表示整数集。

离散小波变换定义为：

$$DWT_{n_1, n_2} = a^{-\frac{n_1}{2}} \sum_{t=d-T+1}^d \chi_t^z \Psi^*(a^{-n_1} t - b n_2) \quad (3)$$

其中*表示复共轭。重建 x_t^z 只能在离散小波满足一定条件时执行。

重建的常用方法是构建如下的离散二进小波变换：

$$DWT_{i,j} = 2^{-\frac{i}{2}} \sum_{t=d-T+1}^d \chi_t^z \Psi^*(2^{-i} t - j) \quad n_1, n_2 \in \mathbb{Z} \quad (4)$$

其中 n_1 和 n_2 设置为1， $a=2^i$ ， $b=j2^i$ 。

然后 x_t^z 可以通过使用Stephane Mallat的多分辨率分析（MRS）来分解和重建[46]。MRS分解后，离散小波分解的函数可以定义为：

$$\Gamma(\chi_t^z) = \{A_j^z, D_1^z, D_2^z, \dots, D_J^z\} \quad (5)$$

其中J是分解级别， A_j^z 是z的每个特征的近似分量（即，用于捕获一般行为的低频分量），而 $\{D_j^z | j=1, 2, \dots, J\}$ 是z的每一个特征的细节分量（即用于捕获局部行为的高频分量） x_t^z 可以分解为 A_j^z 、 D_1^z 、 D_2^z 、...、 D_J^z ，如图3所示。

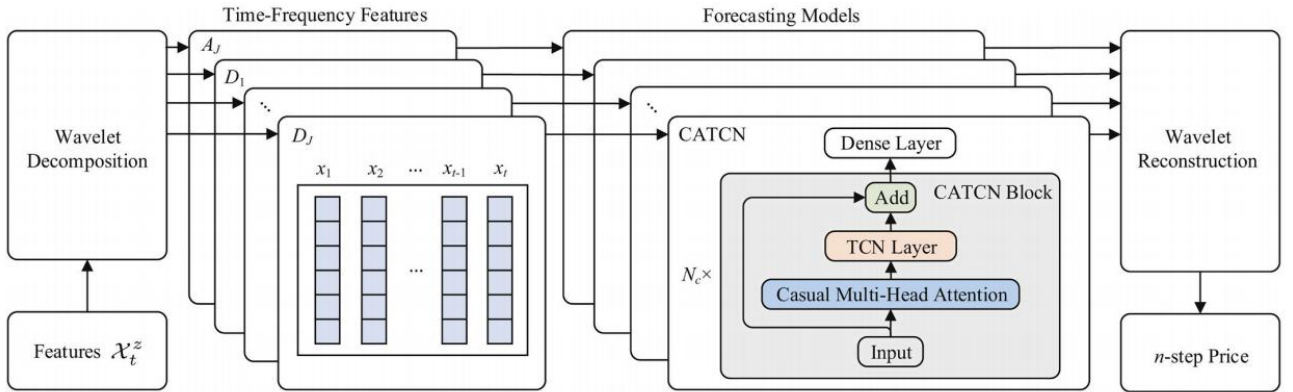


图 3. WT-CATCN架构

3.3.2 随意的多头注意力

注意机制[18]可以从大量输入序列中选择性地筛选出重要信息。多头注意力[18]用于动态计算输入序列中每个位置的注意力值，使得不同的输入位置具有不同的重要性级别。由于我们的输入时间序列数据在不同的时间戳相互关联，因此我们使用随意的多头注意力来检查输入序列的注意力值，并在数据中保持因果关系。

根据我们的设计，随意的多头注意力是每个CATCN块的第一层，在我们的系统中有多CATCN块。因此，有两种类型的随意多头注意力输入：

(1) 如果层在第一CATCN块中，则输入包括相应的时间-频率特征 $\{A_j^z, D_1^z, D_2^z, \dots, D_J^z\}$ 。在不失一般性的情况下，我们将 x 定义为 A_j^z 或 D_i^z ， $i \in [1, J]$ 。

(2) 如果该层位于CATCN块的其余部分，则输入为其上一个CATCN块中的输出。

随意的多头注意力层结合了多个因果自我注意力层，类似于[28]，如图4所示。每个因果自我注意力层都可以计算 x 的注意值。多个因果自我关注层可以计算多个注意值，并从输入数据中提取更多信息。我们将因果自我注意力层的数量表示为 N_m 。

因果自我关注层由分别表示查询、关键字和值的三个矩阵 Q 、 K 和 V 计算[28]。 Q 通过将输入 x 乘以具有相同大小的三个参数矩阵来计算 K 和 V 。它们的值在矩阵训练后更新。注意值通过取 Q 和 K 的点积来计算。

由于未来价格数据事先不可用，我们需要设计 $Mask$ 来将相应的关注值替换为负无穷大，并在 $softmax$ 函数之后使与未来数据相关的关注值为零。因此，通过加上面具，自我关注就变成了因果关系。自我关注层的具体计算过程定义如下。

$$Mask = \begin{cases} X_{(i,j)} = X_{(i,j)}, & \text{if } i \leq j \\ X_{(i,j)} = -inf, & \text{otherwise} \end{cases} \quad (6)$$

其中 $X_{(i,j)} = QK^T$ 是方阵。

$$Attention(Q, K, V) = softmax\left(\frac{Mask(QK^T)}{\sqrt{d_x}}\right)V \quad (7)$$

其中 d_x 是 x 的维数， $softmax$ 函数用于计算注意力值的分布。在我们计算注意力值 $Attention(Q, K, V)$ 之后，因果多头注意力可以定义为：

$$MultiHead(x) = Concat(\{head_i\})W_m \quad (8)$$

其中 $head_i = Attention(Q, K, V), i = \{1, 2, \dots, N_m\}$ 。 Q_i 、 K_i 和 V_i 通过将输入 x 乘以不同的参数矩阵来生成。 W_m 是一个参数矩阵，用于投射不同的注意力值，最终的注意力值具有相同的维度 x 。

接下来，我们计算 x 的多头注意力值 $MultiHead(x)$ ，TCN层的输入（在下一小节中介绍）计算为 $x' = MultiHead(x) = \{x_i | i = 1, 2, \dots, l_b\}$ ，其中 l_b 是输入 x 的长度。

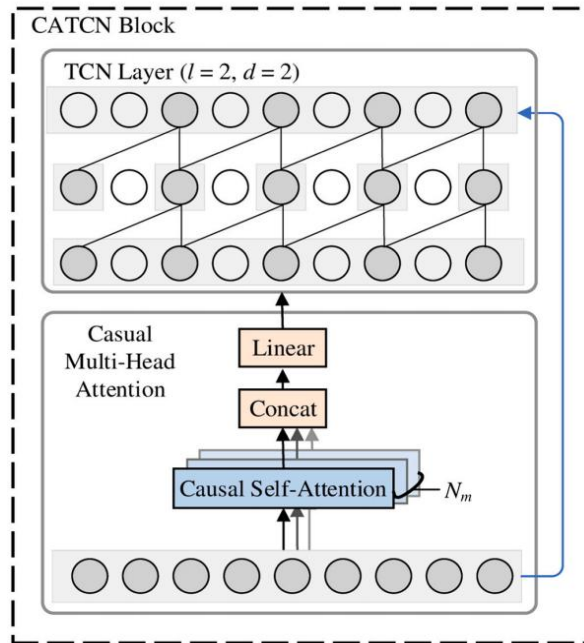


图 4. CATCN 块由 TCN 层和多个因果自注意层组成。图示内核大小 l 为 2，扩张因子 d 为 2

3.3.3 时间卷积网络层

时间卷积网络（TCN）可用于实现长记忆并简化神经网络。它不仅可以缓解消失梯度和爆炸梯度的问题，而且可以用并行方法计算。TCN层。TCN层用于构建第3.3.4节中的CATCN块。TCN层由两个相同的子层组成。为了实现长内存，每个TCN层的膨胀系数是前

一层的两倍（稍后将在膨胀卷积部分详细介绍）。扩张因子随着 TCN 层数的增加而呈指数增长，这使得 CATCN 的内存呈指数增长。以这种方式构建 TCN 层，保证了我们的模型能更高效地处理长序列数据。TCN 层的子层由因果卷积和扩张卷积组成，具体实现如下。

因果卷积：具有因果卷积的模型可以形成一个函数，这与 RNN 的递归连接类似。

因果卷积需要对时间 x'_i 之前的数据进行卷积运算。因果卷积可定义为：

$$y'_i = (x' * W)(i) = \sum_{\eta=0}^{l_w-1} x'_{i-\eta} W(\eta) \quad (9)$$

其中 $i = \{1, 2, \dots, l_b\}$ ， W 表示因果卷积的矩阵（通常称为核）， l_w 是 W 的核大小。随意卷积模型是随意而高效的，但它的感受野不够大，无法处理长序列数据集。为了解决这个问题，我们采用了以下扩张卷积。

扩张卷积：稀释卷积用于扩大内核，以增加感受野并保持内核的大小，这意味着在由稀释因子 d 控制的相同间隔内向内核添加孔：

$$y'_i = (x' *_d W)(i) = \sum_{\eta=0}^{l_w-1} v_{i-d\cdot\eta} W(\eta) \quad (10)$$

其中 W 表示扩张卷积的矩阵乘法。每增加一层，感受野就会以 2 倍 d 的指数增长。

3.3.4 CATCN块

通过堆叠因果多头注意力和 TCN 层，拟议的 CATCN（Casual Multi-Head Attention TCN）区块如图 5 所示。CATCN 模块具有以下特点：

因果关系.为了保证时间序列每一步的顺序，我们使用 CATCN 模块对输入数据进行综合因果处理。在因果多头注意力层中，我们使用偶然掩码使涉及未来数据的注意力值无效。同时，TCN 层使用因果卷积来处理因果多头注意力层的数据。将这两层叠加在一起，CATCN 块就能在以因果关系方法处理数据的同时，将注意力集中在重要位置上。

剩余连接.如果输入是一个很长的序列，模型的层数就会增加。这可能会导致模型性能下降。

受 LSTM 栅极控制的启发，He 等人[29]提出了残差块，它在输入的非线性变换中增加了一个输入。使用残差块后，层数越多的模型改善越明显。因此，在我们的区块中实现了残差连接。残差区块的定义如下：

$$y'' = A(x + F(x)) \quad (11)$$

其中， $i = \{1, 2, \dots, l_b\}$ ， x 是残差块的输入， A 是激活函数， F 是非线性变换，由因果多头注意力层和 TCN 层组成。

CATCN 块可以利用扩张卷积来扩大感受野；因此，CATCN 的感受野会随着 CATCN 块数量的增加而呈指数增长。为了防止感受野过大，需要控制 CATCN 块的数量。我们用 N_C 表示 CATCN 块的数量。最后一个区块的感受野 l_r 的计算公式为：

$$l_r = 1 + 2(l_w - 1)(2^{N_C} - 1) \quad (12)$$

一般来说，感受野应覆盖所有可用数据；因此， l_r 应覆盖第一个区块的输入。通过固定 l_r ，可以根据公式计算出 N_C 。

3.3.5 小波重构

对于每个CATCN，其输出的分量可以集成为 $\{A_j^p, D_1^p, \dots, D_j^p\}$ ，是 P_d^n 的分量。我们可以利用公式 (5) 的逆变换重构 P_d^n ，其定义为：

$$\Gamma^{-1}(\{A_j^p, D_1^p, \dots, D_j^p\}) = P_d^n \quad (13)$$

4 复现细节

4.1 与已有开源代码对比

本复现没有参考任何相关源代码。主要实现了以下部分代码：（1）爬取论文数据代码实现。（2）数据预处理代码实现。（3）论文整体模型代码实现。（4）用另一数据集跑论文模型的代码实现。（5）针对另一数据集预测问题改进论文模型的代码实现。

4.2 实验环境搭建

Pycharm；Python 3.7；TensorFlow框架；pandas、numpy、scikit-learn、matplotlib、Seaborn等python库。

4.4 创新点

将论文模型用在了石油价格的预测上面，并针对该数据的特征对模型进行一定的改进。

5 实验结果分析

5.1 预测结果与实际结果对比

通过在数据集中随机选择连续14天作为测试集，将预测的比特币价格与实际价格进行对比，如图5所示。可以观察到该模型具有较高的预测准确性，能够准确预测比特币价格的走势，证明了该模型的可靠性和有效性，与论文结果基本相符。

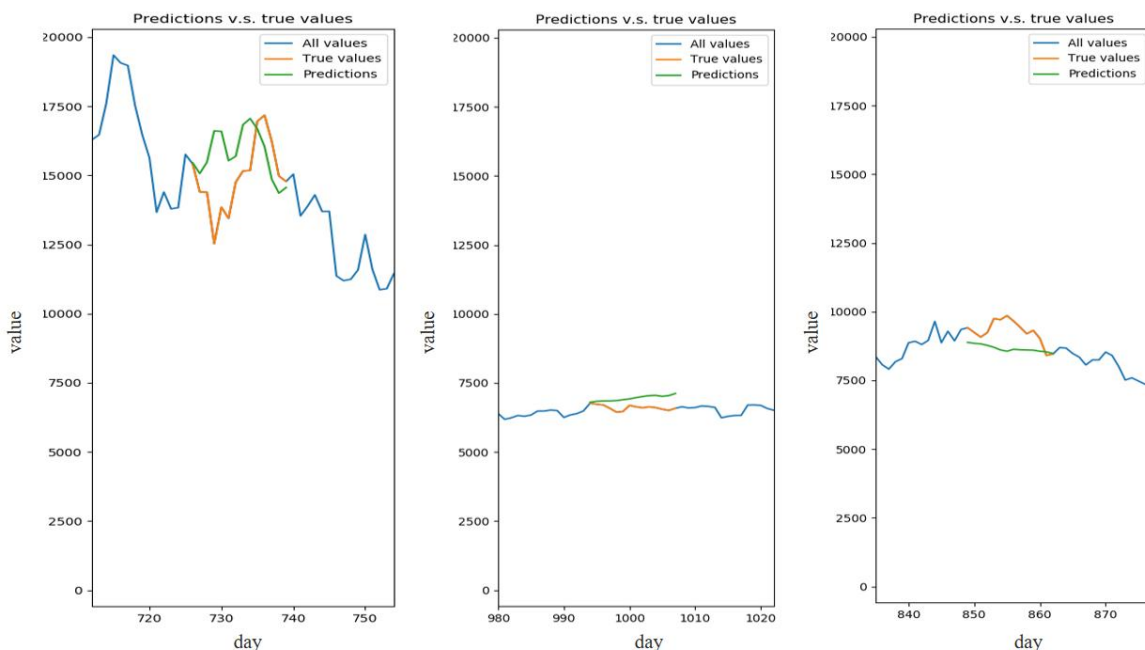


图 5. 预测价格与实际价格对比

5.2 预测趋势对比

为了清晰地可视化预测价格，将结果与基线中的两个模型(ARIMA和BNN)进行比较。各趋势的结果如图6所示。可以发现WT-CATCN可以准确地预测大部分的趋势，复现结果与论文结果基本一致。

平坦趋势：WT-CATCN模型能够更好地预测平坦趋势，与真实值更接近。相比之下，ARIMA模型的预测效果较差，预测曲线要么平坦，要么与真实值相差较远。

凸趋势：没有一个模型能够很好地捕捉凸模式。预测曲线要么平坦，要么与真实值相差甚远。这可能是因为凸模式的价格走势相对较复杂，难以由单一模型准确预测。

凹趋势：WT-CATCN其曲线相对WT-CATCN来说较为不平滑。ARIMA模型的曲线通常非常平坦，表明该模型在捕捉上升趋势方面通常效果较差。

下降趋势：WT-CATCN和ARIMA模型都能够较好地捕捉下降趋势，预测结果与真实值比较接近。然而，ARIMA模型的预测效果相对较差。

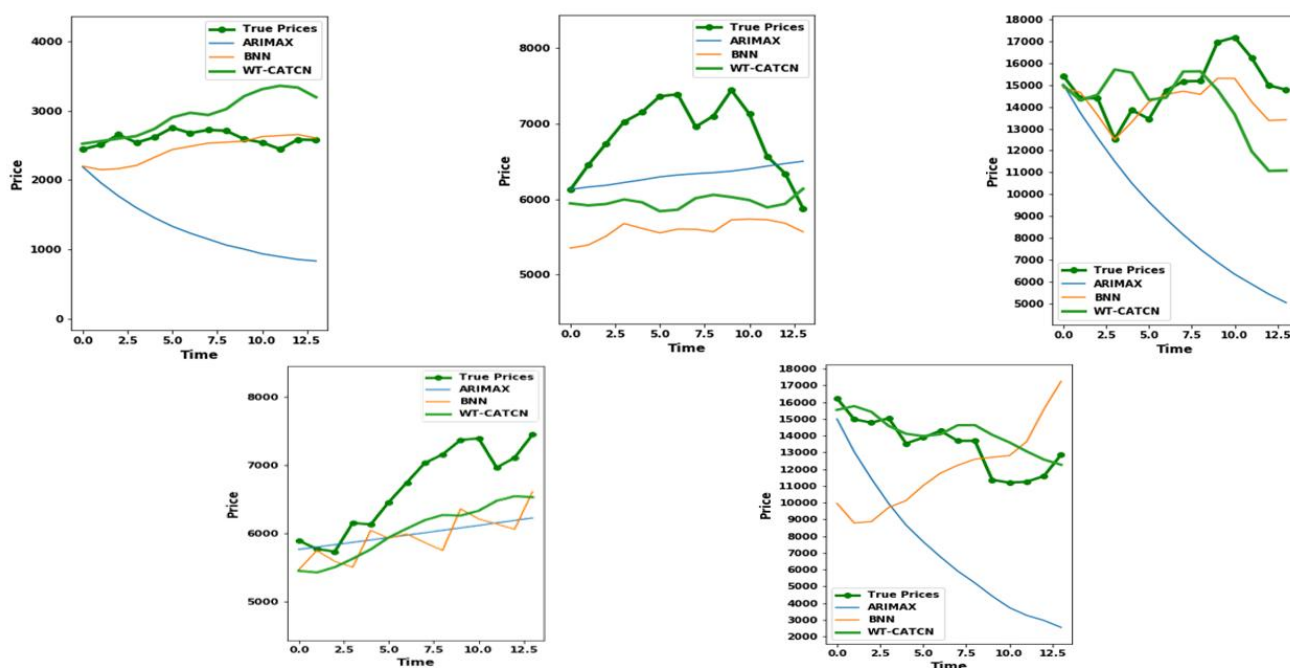


图 6. 价格走势比较:平、凸、凹、涨、跌

5.3 柴油价格预测

由于比特币价格预测问题与传统金融市场中其他资产的价格预测问题相似，因此可以将该模型扩展运用到柴油价格预测的问题上。根据图7所示的预测结果，我们可以观察到以下情况：

WT-CATCN模型在柴油价格预测方面表现较为准确。它能够捕捉到柴油价格的变化趋势，并且与真实值较为接近。这表明该模型具有一定的预测能力，并且可以为柴油市场参与者提供有价值的信息。但由于柴油价格数据与比特币价格数据在特征上存在差异，因此还可以根据柴油价格数据的特征对模型进行改进以达到更好的效果。

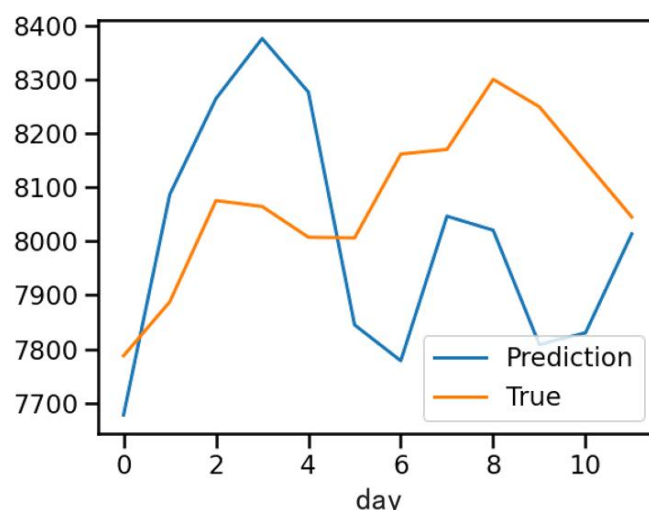


图 7. 柴油价格预测

6 总结与展望

本论文聚焦于比特币，成功地提出了一种能够准确预测高波动时间序列数据（如比特币价格）的模型。该模型能够关注输入数据的关键部分，并对数据之间的相关性进行建模。通过一系列基于真实世界数据的实验，展示了该模型能够有效预测比特币价格及其波动。

目前的工作是基本复现出了论文的结果，并在不同趋势预测上与几种模型进行了对比。但由于部分细节没有复现出来，导致结果不是最优的。

在复现过程中，将模型扩展到与原文数据具有一定相似特性的数据集上，验证模型的泛化能力。验证结果表明，可以预测出大致趋势，但是不够精准。

参考文献

- [1] P. Katsiampa, Volatility estimation for bitcoin: a comparison of garch models, *Econ. Lett.* 158 (2017) 3–6.
- [2] D.C. Mallqui, R.A. Fernandes, Predicting the direction, maximum, minimum and closing prices of daily bitcoin exchange rate using machine learning techniques, *Appl. Soft Comput.* 75 (2019) 596–606.
- [3] W. Kristjanpoller, M.C. Minutolo, A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis, *Expert Syst. Appl.* 109 (2018) 1–11.
- [4] S. McNally, J. Roche, S. Caton, Predicting the price of bitcoin using machine learning, in: 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), IEEE, 2018, pp. 339–343.
- [5] S. Alonso-Monsalve, A.L. Su´arez-Cetrulo, A. Cervantes, D. Quintana, Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators, *Expert Syst. Appl.* 149 (2020) 113250.
- [6] Y.B. Kim, J.G. Kim, W. Kim, J.H. Im, T.H. Kim, S.J. Kang, C.H. Kim, Predicting fluctuations in cryptocurrency transactions based on user comments and replies, *PLoS One* 11 (8) (2016).

- [7] E. Stenqvist, J. Lonn " o, " Predicting Bitcoin Price Fluctuation With Twitter Sentiment Analysis, 2017.
- [8] T.R. Li, A. Chamrajnagar, X. Fong, N. Rizik, F. Fu, Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model, *Front. Phys.* 7 (2019) 98.
- [9] S. Mohapatra, N. Ahmed, P. Alencar, KryptoOracle: a real-time cryptocurrency price prediction platform using twitter sentiments, in: 2019 IEEE International Conference on Big Data (Big Data), IEEE, 2019, pp. 5544–5551.
- [10] G. Cheuque Cerda, J.L. Reutter, Bitcoin price prediction through opinion mining, in: Companion Proceedings of the 2019 World Wide Web Conference, 2019, pp. 755–762.
- [11] A. Jain, S. Tripathi, H. DharDwivedi, P. Saxena, Forecasting price of cryptocurrencies using tweets sentiment analysis, in: 2018 Eleventh International Conference on Contemporary Computing (IC3), IEEE, 2018, pp. 1–7.
- [12] J. Rebane, I. Karlsson, S. Denic, P. Papapetrou, Seq2Seq RNNs and ARIMA models for cryptocurrency prediction: a comparative study, *SIGKDD Fintech* 18 (2018).
- [13] H. Jang, J. Lee, An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information, *IEEE Access* 6 (2017) 5427–5437.
- [14] M. Amjad, D. Shah, Trading bitcoin and online time series prediction, in: NIPS 2016 Time Series Workshop, 2017, pp. 1–15.
- [15] T. Guo, A. Bifet, N. Antulov-Fantulin, Bitcoin volatility forecasting with a glimpse into buy and sell orders, in: 2018 IEEE International Conference on Data Mining (ICDM), IEEE, 2018, pp. 989–994.
- [16] P.R. Hansen, A. Lunde, A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? *J. Appl. Econom.* 20 (7) (2005) 873–889.
- [17] M.R. Hassan, K. Ramamohanarao, J. Kamruzzaman, M. Rahman, M.M. Hossain, A HMM-based adaptive fuzzy inference system for stock market forecasting, *Neurocomputing* 104 (2013) 10–25.
- [18] R.K. Nayak, D. Mishra, A.K. Rath, A naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices, *Appl. Soft Comput.* 35 (2015) 670–680.
- [19] M. Lam, Neural network techniques for financial performance prediction: integrating fundamental and technical analysis, *Decis. Support Syst.* 37 (4) (2004) 567–581.
- [20] Y. Jiao, J. Jakubowicz, Predicting stock movement direction with machine learning: an extensive study on s&p 500 stocks, in: 2017 IEEE International Conference on Big Data (Big Data), IEEE, 2017, pp. 4705–4713.
- [21] R. Faccini, E. Konstantinidi, G. Skiadopoulos, S. Sarantopoulou-Chiourea, A new predictor of us real economic activity: the s&p 500 option implied risk aversion, *Manage. Science* 65 (10) (2019) 4927–4949.
- [22] C.J. Neely, D.E. Rapach, J. Tu, G. Zhou, Forecasting the equity risk premium: the role of technical indicators, *Manage. Sci.* 60 (7) (2014) 1772–1791.
- [23] L. Zhang, C. Aggarwal, G.-J. Qi, Stock price prediction via discovering multifrequency trading patterns, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 2141–2149.
- [24] H.K. Sul, A.R. Dennis, L. Yuan, Trading on twitter: using social media sentiment to predict stock returns, *Decis. Sci.* 48 (3) (2017) 454–488.

- [25] Q. Li, L. Jiang, P. Li, H. Chen, Tensor-based learning for predicting stock movements, in: Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.
- [26] A. Aghajani, R. Kazemzadeh, A. Ebrahimi, A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm, *Energy Convers. Manage.* 121 (2016) 232–240.
- [27] O. Rioul, P. Duhamel, Fast algorithms for discrete and continuous wavelet transforms, *IEEE Trans. Inf. Theory* 38 (2) (1992) 569–586.
- [28] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need. *Advances in Neural Information Processing Systems*, 2017, pp. 5998–6008.
- [29] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.