

赵鑫, 秦江涛. 基于 LSTM-DTW-GAT 混合模型的加密货币预测研究[J]. 智能计算机与应用, 2024, 14(12): 96-103. DOI: 10.20169/j.issn.2095-2163.241213

基于 LSTM-DTW-GAT 混合模型的加密货币预测研究

赵鑫, 秦江涛

(上海理工大学 管理学院, 上海 200093)

摘要: 本文旨在开发一种能够有效融合时间序列特征和加密货币之间关系信息的新型预测模型, 以提升加密货币市场预测的精度。因此, 本文提出了一种基于长短期记忆网络 (LSTM)、动态时间规整 (DTW) 算法和图注意力网络 (GAT) 的 LSTM-DTW-GAT 混合预测模型。通过 LSTM 捕捉时间序列依赖关系, 结合 DTW 算法动态衡量加密货币间相似度, 并利用 GAT 学习加密货币间的复杂关系, 模型能够融合时序特征和关系特征共同预测加密货币回报率。实验结果表明, 与现有模型相比, LSTM-DTW-GAT 模型在各项性能评估指标上均有显著提升, 尤其是关系信息的引入大幅提高了预测精度, 验证了该模型在加密货币预测领域的有效性和优越性。

关键词: 长短期记忆网络; 动态时间规整; 图注意力网络; 加密货币关系; 加密货币预测

中图分类号: F821; TP181

文献标志码: A

文章编号: 2095-2163(2024)12-0096-08

Research on cryptocurrency prediction based on the LSTM-DTW-GAT hybrid model

ZHAO Xin, QIN Jiangtao

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The aim of this paper is to develop a novel forecasting model that can effectively fuse time series features and relationship information between cryptocurrencies to enhance the accuracy of cryptocurrency market forecasting. Therefore, this paper proposes a hybrid LSTM-DTW-GAT forecasting model based on the Long Short-Term Memory Network (LSTM), Dynamic Time Warping (DTW) algorithm and Graph Attention Network (GAT). By capturing time series dependencies through LSTM, combining the DTW algorithm to dynamically measure the similarity among cryptocurrencies, and using GAT to learn the complex relationships among cryptocurrencies, the model is able to fuse temporal and relational features to jointly predict cryptocurrency returns. The experimental results show that compared with the existing models, the LSTM-DTW-GAT model is significantly improved in all performance evaluation indexes, especially the introduction of relational information substantially improves the prediction accuracy, which verifies the effectiveness and superiority of the model in the field of cryptocurrency prediction.

Key words: Long Short-Term Memory networks; Dynamic Time Warping; Graph Attention Networks; cryptocurrency relationships; cryptocurrency forecast

0 引言

自比特币 2008 年诞生以来, 加密货币市场迅猛发展, 截至 2023 年 12 月, 已拥有超过 8 000 种加密货币, 总市值达到 1.68 万亿美元^[1]。由于加密货币的价值波动极端, 使其被视为纯粹的投机性资产, 呈现出泡沫化的价格变动。预测加密货币市场表现颇具挑战性, 已成为加密货币研究的一个重要分支^[2]。

近年来, 加密货币的预测方法从早期的统计学模型^[3-4]逐渐过渡到应用机器学习模型^[5], 例如, 支持向量机 (SVM)^[6]、决策树 (DT)^[7] 和随机森林 (RF)^[8] 等机器学习方法在一定程度上改进了预测精度。然而, 这些方法虽然提升了预测精度, 但在处理加密货币市场的非线性和高波动性时常遇到过拟合和适应性问题^[9]。

深度学习技术为此领域带来了重大突破, 能有效识别复杂数据模式和非线性关系, 使其成为理想

作者简介: 赵鑫 (1998—), 男, 硕士研究生, 主要研究方向: 数据挖掘, 人工智能。

通信作者: 秦江涛 (1966—), 男, 博士, 副教授, 主要研究方向: 制造系统分析。Email: zhaoxin9822@163.com。

收稿日期: 2023-06-16

哈尔滨工业大学主办 ◆ 系统开发与应用

的工具^[10]。例如, Lahmiri 等学者^[11]采用长短期记忆神经网络(LSTM)来预测加密货币价格,研究发现 LSTM 预测效果优于传统机器学习模型。Dutta 等学者^[12]将 LSTM 网络与传统的 RNN 网络进行对比,发现 LSTM 更适合处理加密货币预测。由于 LSTM 独特的细胞状态设计,可以存储长期记忆,有效捕捉序列中的长期依赖性,从而解决长序列数据处理中常见的梯度消失问题^[13]。

然而,尽管深度学习模型在单一加密货币预测方面表现出色,却往往忽略了加密货币之间的相互关联性^[14]。实际上,一个加密货币的价格变动常受到其他相关加密货币的影响^[15]。近年来,一些研究开始探索利用加密货币之间的关系来提升预测模型的性能。例如, Li 等学者^[16]采用 AttentiveLSTM 和嵌入网络(ALEN)分析比特币及相关加密货币的时间序列数据,发现加密货币间的关系对预测比特币价格至关重要。

尽管如此,现有研究通常依赖于传统深度学习架构如 LSTM 和 RNN,这些架构并非专门设计来处理图形数据结构,因此在捕捉和理解加密货币间复杂网络动态方面存在局限。Zhao 等学者^[17]使用的图卷积网络(GCN)在股票预测中超越了传统时间序列模型,展示了其处理复杂网络数据的潜力。然而,上述的方法依赖于静态关系图,基于先验知识构建,无法适应关系随时间的变化^[18]。相比之下,图注意力网络(GAT)^[19]通过引入注意力机制,能够动态地学习和调整节点间的相对重要性,为实时处理加密货币间的动态关系提供了新途径。

为了在提取时间序列信息的基础上进一步融合关系信息,并提高模型预测的精度,本文在长短期记忆神经网络(LSTM)基础上,提出了一种结合动态时间规整(DTW)算法和图注意力网络(GAT)的 LSTM-DTW-GAT 预测模型。该模型的核心创新在于,不仅考虑了时间序列数据中的时序特征,通过 LSTM 处理时间序列数据中的长期依赖性,解决梯度消失问题,同时 DTW 算法提供时间序列之间的弹性匹配,捕捉不同时间尺度上的相似性,而 GAT 则动态学习节点间的相对重要性,处理加密货币间的复杂关系。通过结合这些技术,模型不仅能处理时间序列数据,还能利用加密货币间的关系信息,全面提高对加密货币市场表现的预测精度。实验结果表明,该模型在预测加密货币未来回报率方面显示出了优异的精度,并通过实验证明了其在加密货币预测领域的实用性和有效性。

1 加密货币间关系的时间序列分析

加密货币之间的多种关系源于共享的技术基础、行业应用和市场事件,可以通过时间序列数据来进行揭示和分析^[20]。技术基础,如比特币现金和比特币 SV 都采用 SHA-256 算法;行业应用,如以太坊及其代币在去中心化金融中的关联性;市场事件,如特斯拉购买比特币对多个加密货币价格的影响。分析这些关系,可以更全面地理解市场动态。时间序列分析可以帮助识别和量化不同加密货币之间的关系。

1.1 时间序列趋势分析

通过对不同加密货币的日收盘价格进行归一化处理,可以观察到具有相似技术基础或行业应用的加密货币,其价格趋势高度一致。例如,图 1(a)展示了 BCH 和 BSV 的价格趋势,采用的都是 SHA-256 算法;图 1(b)展示了 ETH 和 LEO,使用的都是 Ethereum 平台;图 1(c)展示了 ETH 和 MKR,都归属于 DeFi 行业;图 1(d)展示了 USDT 和 USDC,则都是稳定币。此外,加密货币价格时间序列趋势如图 2 所示。这些也都表明市场重大事件对加密货币价格有直接影响,展示了加密货币间因事件驱动的短期联系。

1.2 时间序列协整分析

通过对归一化处理后的加密货币时间序列进行观察,可以看出加密货币之间的价格趋势存在一致性。为了定量分析这种关系,对图 1 中的 4 对加密货币进行了协整检验。协整分析检测多个时间序列间的长期平衡关系。结果见表 1。BCH-BSV、ETH-LEO、ETH-MKR 和 USDT-USDC 的 Trace 统计量均显著高于临界值,表明各统计量之间存在显著的协整关系,说明这些加密货币在长期内共同移动,验证了加密货币间关系可以通过时间序列识别和量化。

趋势分析和协整分析表明,时间序列分析是识别并量化加密货币间关系的有效工具。加密货币间的相互关系作用于交易行为,形成价格时间序列并表现出相关性。加密货币时间序列的相似性反映了序列间的综合关系,二者间存在相互映射关系。这揭示了加密货币价格时间序列相关性与其复杂关系之间的互为反映关系。本文使用如下公式描述了这种关系:

$$R_{ij} = \rho(P_i, P_j) \quad (1)$$

其中, R_{ij} 表示加密货币间的关系, $\rho(P_i, P_j)$ 表示加密货币价格时间序列间的相关系数。通过此方式,从而定量捕捉并解释加密货币之间的复杂关系。

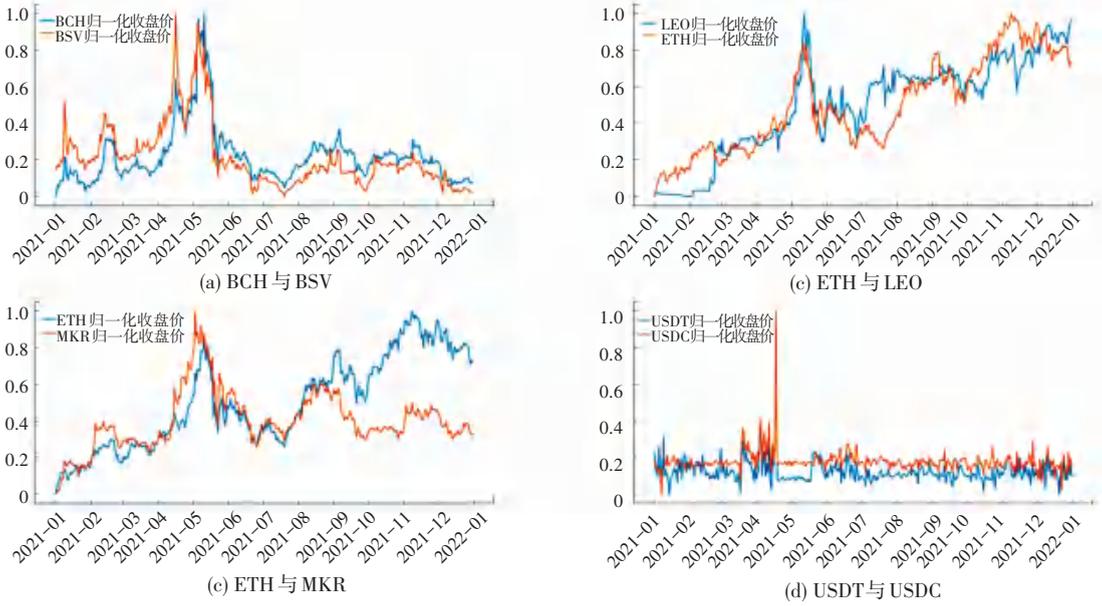


图1 加密货币价格时间序列趋势图

Fig. 1 Time series trend graph of cryptocurrency prices

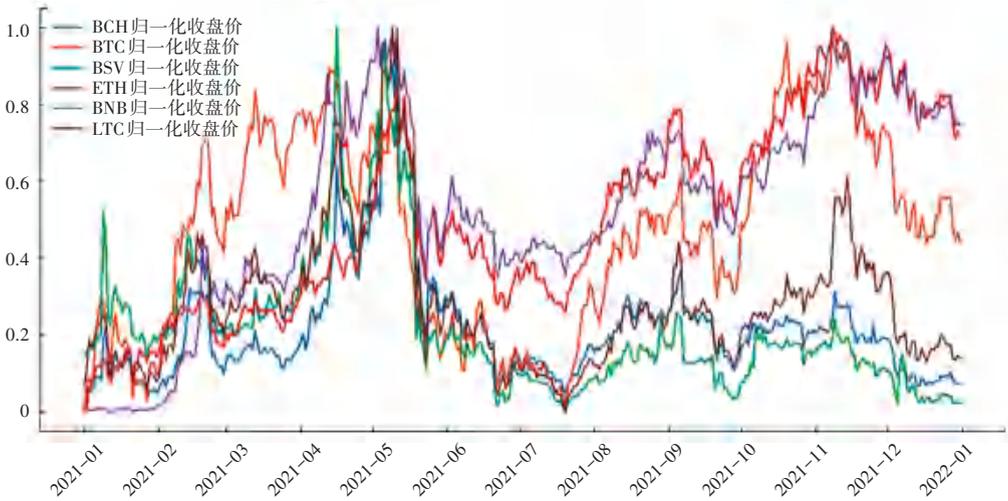


图2 加密货币价格时间序列趋势图

Fig. 2 Time series trend graph of cryptocurrency prices

表1 协整检验分析表

Table 1 Cointegration test analysis table

加密货币对	Trace 统计量	90%临界值	95%临界值	99%临界值
BCH-BSV	29.553 1	13.429 4	15.494 3	19.934 9
ETH-LEO	26.078 0	13.429 4	15.494 3	19.934 9
ETH-MKR	23.143 9	13.429 4	15.494 3	19.934 9
USDT-USDC	170.428 2	13.429 4	15.494 3	19.934 9

2 相关理论及模型构建

2.1 长短期记忆网络

长短期记忆网络(LSTM)是一种专为序列数据设计的深度学习模型,广泛应用于自然语言处理和时间序列预测等领域。作为循环神经网络(RNN)的改进版,LSTM旨在克服传统RNN在学习过程中

遇到的梯度消失或爆炸问题。其核心特点是内部设有一系列单元,由“细胞”组成,通过时间上的循环连接。这些细胞具有输入门、输出门和遗忘门,协同调控信息的存储、更新和遗忘。相关的数学公式如下所示:

$$\begin{cases}
 \dot{f}_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\
 \dot{i}_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\
 \dot{\tilde{C}}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \\
 \dot{C}_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\
 \dot{o}_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\
 \dot{h}_t = o_t \cdot \tanh(C_t)
 \end{cases} \quad (2)$$

尽管 LSTM 在处理时间序列数据方面表现出色,但在面对加密货币间复杂的相互关系时,LSTM 无法充分捕捉这些关系的动态变化。因此,需要将 LSTM 与其它技术相结合,以弥补其在处理加密货币复杂关系方面的不足。

2.2 动态时间规整算法

动态时间规整算法 (DTW)^[21] 由 Itakura 首次提出,最早应用于语音识别领域。DTW 通过计算不同长度的时间序列间相似性,克服了欧氏距离在序列维度不同情况下的局限,应用于各种相似性计算场景中^[22]。对任意 2 种加密货币序列 s 和 q , 对其归一化后的收盘价格序列匹配路径表示为 $w = \{w_1, w_2, \dots, w_k\}$ 。函数 $d(s_i, q_j)$ 表示序列 s 中的点 s_i 与序列 q 中的点 q_j 之间的欧氏距离。DTW 的核心是通过建立最优匹配路径来最小化加密货币价格时间序列间的整体距离。具体计算过程如下:

$$d(w_l) = d(s_i, q_j) = |s_i - q_j|, \\ i \in [1, n], j \in [1, n], l \in [1, k] \quad (3)$$

$$DTW(s, q) = \min \sum_{l=1}^k d(w_l) \quad (4)$$

$$\gamma(i, j) = d(s_i, q_j) + \min(\gamma(s_i, q_{j-1}), \gamma(s_{i-1}, q_j), \gamma(s_{i-1}, q_{j-1})) \quad (5)$$

$$\nu_{sq} = DTW(s, q) \quad (6)$$

其中, $\gamma(0, 0) = 0, \gamma(i, 0) = \gamma(0, j) \geq 0$ 。本文使用 DTW 算法,利用归一化后的收盘价时间序列确定 2 种加密货币的相似度值 ν_{sq} 。

2.3 图注意力神经网络

图注意力神经网络 (GAT)^[23] 是一种增强版的图神经网络,通过引入注意力机制,动态识别和量化图中节点间的相对重要性。GAT 通过计算注意力系数 α_{ij} 来衡量节点 i 和 j 之间的关系,公式如下:

$$e_{ij} = \text{LeakyReLU}(\mathbf{a}^T [W h_i P W h_j]) \times A_{ij} \quad (7)$$

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in N(i)} \exp(e_{ik})} \quad (8)$$

其中, e_{ij} 表示节点 i 和 j 之间关系的原始注意力分数,加权邻接矩阵 \mathbf{A} 中的元素表示节点 i 与节点 j 之间的边权重。

利用计算得到的注意力系数 α_{ij} , GAT 对每个节点的邻居节点特征进行加权聚合,以得到该节点的更新特征表示:

$$r_i = \sigma \left(\sum_{j \in N(i)} \alpha_{ij} W h_j \right) \quad (9)$$

GAT 通过引入注意力机制,能够动态识别和学习图中节点间的相对重要性,精确捕捉节点之间的

复杂关系。

2.4 LSTM-DTW-GAT 预测模型框架

由于加密货币市场受多种因素影响,时间序列表现出非线性和高波动性。本文结合 LSTM、DTW 和 GAT,构建 LSTM-DTW-GAT 混合预测模型,以提高预测精度。LSTM-DTW-GAT 混合结构如图 3 所示。首先,利用 LSTM 处理时间序列的长期依赖性,解决长序列数据的梯度消失问题。接着,通过 DTW 算法弹性匹配时间序列,捕捉不同时间尺度的相似性。然后,采用 GAT 动态学习和调整节点间的相对重要性,充分利用加密货币间的复杂关系。最终,结合提取的时间序列特征和节点关系信息进行市场预测。多层次结构设计使模型在处理非线性和高波动性数据时表现出优异的预测精度和适用性,解决了传统模型的过拟合和适应性问题。

(1) 特征提取层。在特征提取层,LSTM-DTW-GAT 模型使用长短期记忆网络 (LSTM) 独立处理每种加密货币的时间序列数据 \mathbf{x}^s , 以捕获价格波动和时间依赖性。具体来说,LSTM 通过对时间序列进行深度学习,提取出代表历史价格趋势和时间依赖关系的特征向量 \mathbf{h}^s 。这些特征不仅直接用于预测未来趋势,也会作为关系提取层的输入,推得的公式如下:

$$\mathbf{h}^s = \text{LSTM}(\mathbf{x}^s) \quad (10)$$

其中, S 表示加密货币数量。时间序列特征 $H = \{h^1, h^2, \dots, h^s\}$ 表示所有待预测加密货币的综合价格趋势特征。利用这些特征,LSTM-DTW-GAT 模型能够为每种货币生成详尽的时间序列特征,为识别市场行为和进行准确预测奠定了基础。

(2) 关系提取层。在关系提取层中,LSTM-DTW-GAT 模型首先使用动态时间规整 (DTW) 算法计算不同加密货币时间序列之间的相似度 ν , 以此构建表示加密货币相互关系的邻接矩阵 \mathbf{A} , 具体公式为:

$$\mathbf{A} = \begin{matrix} \hat{e}_{11} & \cdots & \nu_{1s} \\ \vdots & \ddots & \vdots \\ \hat{e}_{s1} & \cdots & \nu_{ss} \end{matrix} \hat{u} \quad (11)$$

得到加密货币间的邻接矩阵 \mathbf{A} 后,模型进一步运用图注意力网络 (GAT) 学习节点间的相互关系。GAT 优化了节点间 (即加密货币间) 的权重分配,从而根据节点的相互关系动态计算其重要性。GAT 的输入包括 LSTM 层输出的时序特征 H (图的节点信息) 以及由 DTW 算法得出的邻接矩阵 \mathbf{A} (图的结构信息):

$$\tilde{\mathbf{A}} = \mathbf{A} + \lambda \mathbf{I}_N \quad (12)$$

$$\mathbf{R} = \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H} \mathbf{W}_r^1) \mathbf{W}_r^2) \quad (13)$$

其中, \mathbf{I}_N 表示单位矩阵; λ 表示调整项, 加权自连接的调整项(本研究中设为 1), 以保持节点自信

息; \mathbf{D} 表示度矩阵, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$; $\mathbf{R} = \{r^1, r^2, \dots, r^s\}$ 表示 GAT 输出的加密货币间关系特征集合, 这里 r^s 表示单一加密货币的关系特征向量。

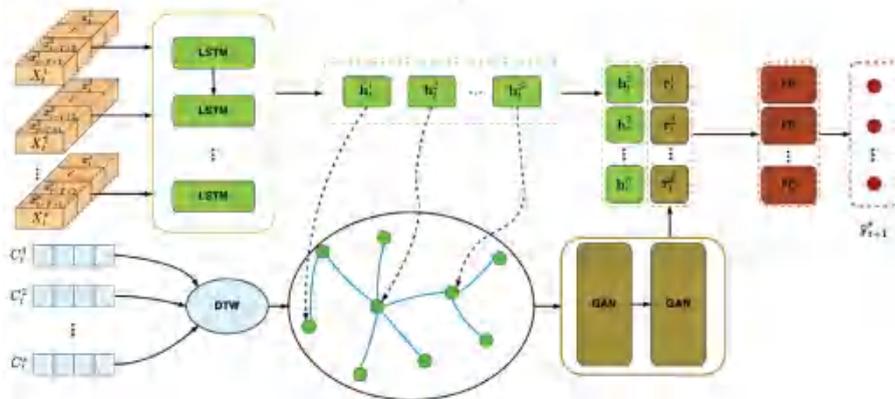


图3 LSTM-DTW-GAT 混合模型结构

Fig. 3 Structure of LSTM-DTW-GAT hybrid model

(3) 预测层。在预测层中, LSTM-DTW-GAT 模型整合从特征提取层获得的时序特征 h^s 和关系提取层得到的关系特征 r^s 。通过拼接这 2 种特征, 模型构建一个综合特征集, 为每种加密货币及其相互作用提供全面视角。对加密货币 s 在时间 $t+1$ 的预测回报率 \hat{y}_{t+1}^s , 计算公式如下:

$$\hat{y}_{t+1}^s = \sigma(\mathbf{W}_p[r^s \parallel h^s] + b_p) \quad (14)$$

为了精确预测每种加密货币的回报率, 本研究的目标损失函数结合了逐点回归损失和成对排序感知损失:

$$L(\hat{y}_{t+1}, y_{t+1}) = \|\hat{y}_{t+1} - y_{t+1}\|^2 + \alpha \sum_i \sum_j \max\{0, -(\hat{y}_{t+1}^i - \hat{y}_{t+1}^j)(y_{t+1}^i - y_{t+1}^j)\} \quad (15)$$

其中, \hat{y}_{t+1} 和 y_{t+1} 分别表示第 $t+1$ 天的预测回报率和回报率; \hat{y}_{t+1}^i 和 y_{t+1}^i 分别表示加密货币 i 在第 $t+1$ 天的预测回报率和实际回报率; α 表示调节 2 个损失项的超参数。式(15)中, 第一项损失针对预测回报率与实际回报率的差异, 第二项损失针对回报率的相对顺序, 确保预测结果的准确性和排序合理性。

通过这种结构设计, LSTM-DTW-GAT 模型不仅分析单个加密货币的时间序列数据, 还综合其关系数据, 有效提取加密货币间的关系, 从而提升市场表现的预测准确性。

3 实验设计和结果分析

3.1 数据准备与预处理

本文使用的加密货币历史时间序列数据来自

CoinMarketCap.com。根据市值排名和数据质量, 共选取了 46 种加密货币, 包括 BTC、ETH、USDT、DOGE 等。数据覆盖 2020 年 1 月 1 日至 2023 年 12 月 1 日的 1 431 天, 分为训练集、验证集和测试集, 比例为 8:1:1。特征包括开盘价、收盘价、最高价、最低价、成交量和市值。为了消除量纲和加快模型训练, 数据进行了最小-最大归一化处理, 处理公式如下:

$$X_n = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (16)$$

其中, X_{\min} 和 X_{\max} 分别为数据集中的最小值和最大值。

3.2 模型训练策略及参数设置

本文实验环境配置如下: 操作系统为 Windows 11, 内存 16 GB, GPU 为 RTX3060, 编程语言使用 Python 3.9, 深度学习框架为 Pytorch。滑动窗口周期为 15 天, 即通过前 15 天的数据预测后一天的加密货币回报率。训练配置中, 使用随机梯度下降法优化损失函数, batch size 设为 16, 模型 epoch 设为 50, 采用 Adam 优化器, 初始学习率为 0.001, α 设为 10。具体模型参数设置见表 2。

由于在真实的加密货币交易中, 如果预测 $t+1$ 日的回报率最高的加密货币, 很难在 t 日以收盘价买入该加密货币, 并在 $t+1$ 日以收盘价卖出该加密货币。

因此本文在实验中假设: 投资者将能够利用该模型获得所有加密货币的在 $t+1$ 日预测回报排名后, 在当天(t 日)以收盘价买入预测回报率最高的加密货币, 然后在下一个交易日($t+1$ 日)以收盘价卖出, 并重复该过程。计算评价指标基于以下假设:

- (1) 投资者在每个交易日花费相同的金额。
 (2) 投资者可以在 t 日的收盘价买入,在 $t+1$ 日

的收盘价卖出。

(3) 不考虑交易成本。

表 2 模型参数设置

Table 2 Model parameter settings

模块	参数
LSTM	Layers:2;Number of neurons:{64,128};Dropout rate:0.1 Activation:Relu
GAT	Layers:2;Number of neurons:{128,128};Dropout rate:0.1 Activation:Relu
P_Dense	Layers:2;Number of neurons:{128,64};Dropout rate:0.1 Activation:LeakyRelu

本文的目的是准确预测加密货币的回报率,同时平衡投资回报和风险。因此,以下 5 个评估指标用于评价和比较模型:均方误差 (MSE)、平均绝对误差 (MAE)、累计投资回报率 (IRR)、最大回撤 (MDD) 和夏普比率 (SR)。公式如下:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (17)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (18)$$

$$IRR = \sum_i^n R_i \quad (19)$$

$$MDD = \max\left(\frac{R_{up} - R_{down}}{R_{up}}\right) \quad (20)$$

$$SR = \frac{E(R_p) - R_f}{\sigma_p} \quad (21)$$

IRR 反映累计投资回报, MDD 描述投资过程中可能发生的最坏情况, SR 从回报和风险两个维度评估投资表现。一个好的预测模型应有较小的 MSE 和 MAE ,以及较大的 IRR 和 SR 。

3.3 对比模型

为了验证模型的先进性,本文选择了以下 3 种基线模型进行实验对比。

(1) LSTM:标准的长短期记忆网络(LSTM),从历史价格序列中提取时序特征,并通过全连接层进

行回报率预测。

(2) CNN-LSTM:结合卷积神经网络(CNN)和 LSTM 的混合模型^[24]。CNN 提取局部特征,适合捕捉时间序列中的局部模式,LSTM 捕捉长期周期依赖关系。这种组合能够有效地处理时间序列数据。

(3) LSTM+GCN:结合 LSTM 和图卷积网络(GCN)的模型^[25]。LSTM 提取时序特征,GCN 捕捉节点间的图结构关系。这种组合能够同时处理时间序列特征和图结构特征。

3.4 对比实验结果分析

为了全面评估 LSTM-DTW-GAT 模型,进行了对比实验。结果见表 3。表 3 中,单一 LSTM 模型在各项指标上均低于其他模型。尽管 LSTM 能够处理序列特征信息,但其网络结构的局限导致预测误差较大。CNN-LSTM 结合 CNN 的特征提取能力与 LSTM 的序列学习能力,虽提高了预测精度,但主要关注单一货币时间序列信息,预测精度的提升有限。LSTM+GCN 模型的预测性能虽仅次于 LSTM-DTW-GAT 模型,表明结合图卷积网络(GCN)的混合模型在一定程度上能够提取加密货币间的关系信息,并有效地利用这些信息提升预测精度。然而,由于加密货币间的关系信息会随市场动态变化,而 GCN 主要提取静态的关系信息,LSTM+GCN 模型在捕捉这些动态关系方面显示出一定的局限性。

表 3 基线模型性能评价指标对比

Table 3 Comparison of baseline model performance evaluation metrics

模型	MSE	MAE	IRR	$MDD/\%$	SR
LSTM	8.69	1.958	0.23	9.8	2.35
CNN-LSTM	7.31	1.917	0.32	8.3	2.87
LSTM-GCN	6.86	1.763	0.46	7.8	3.26
LSTM-DTW-GAT	6.13	1.712	0.57	6.8	3.72

与此相比,LSTM-DTW-GAT 模型在 MSE 、 MAE 、 IRR 、 MDD 和 SR 方面表现最佳,通过 DTW 和

GAT 精准捕捉动态关系,表现出高预测精度和低风险投资潜力。实验得到的累计投资回报率对比如图

4所示。图4显示,LSTM-DTW-GAT模型在大多数测试时间段内累积回报率最高,体现出其强大性能和适应能力。这些成果验证了LSTM-DTW-GAT模型在动态市场环境中的稳健性和优越性。

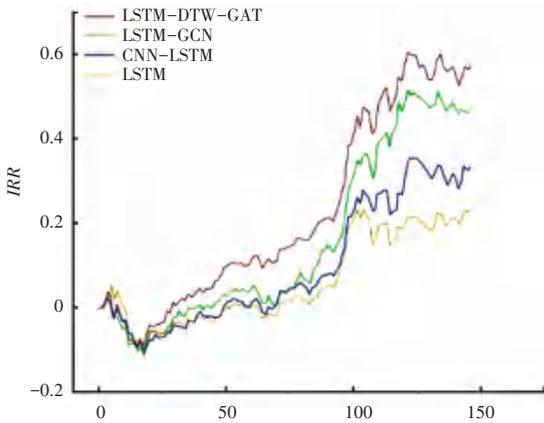


图4 对比实验累计投资回报率对比图

Fig. 4 Comparison of cumulative ROI for comparison experiments

3.5 消融实验结果分析

为了深入评估 LSTM-DTW-GAT 模型中特征提取层和关系提取层的有效性,构建了3种变体模型进行消融实验:N/T&G、N/T和N/G,并利用相同历史价格数据集验证。

(1)N/T&G:将 LSTM 替换为普通循环神经网络(RNN),并去除关系提取层,以测试 LSTM 和关系特征提取对预测性能的综合贡献。

(2)N/T:将 LSTM 替换为 RNN,保留相同的网络层数和神经元数量,探究 LSTM 对模型性能的影响。

(3)N/G:去除关系提取层,仅使用时序特征进行预测,分析加密货币之间关系特征的作用。

分别使用 N/T&G、N/T、N/G 和 LSTM-DTW-GAT 模型进行实验,实验结果见表4。

表4 消融实验评价指标对比

Table 4 Comparison of evaluation indexes of ablation experiments

模型	MSE	MAE	IRR	MDD/%	SR
N/T&G	9.14	2.149	0.18	11.9	1.97
N/T	7.13	1.817	0.42	7.7	2.98
N/G	7.53	1.964	0.33	8.5	2.75
LSTM-DTW-GAT	6.13	1.712	0.57	6.8	3.72

表4的结果显示,当移除关系提取层和特征提取层的N/T&G模型的投资回报率(IRR)仅为0.18。相比之下,引入关系信息的N/T模型,IRR显著提升至0.42,其他指标如MSE、MAE、MDD和

SR也有改善,证明了关系信息的重要性。N/G模型的IRR提升到0.33,但仍低于N/T模型,表明关系信息和时序特征的结合更为有效。各个变体模型的累计回报IRR结果如图5所示。图5显示LSTM-DTW-GAT模型在各测试时间段的表现均优于其他模型,验证了该模型结合时序和关系数据预测的优势。这一结果突出了将时序特征提取与关系特征提取模块结合使用的重要性,并进一步验证了LSTM-DTW-GAT模型结合时间序列数据和关系数据进行预测的明显优势。

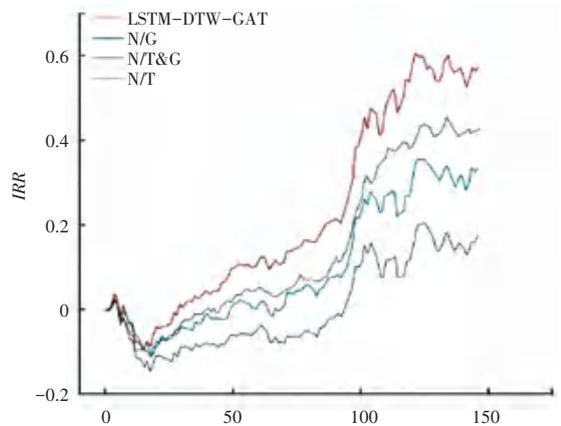


图5 消融实验累计投资回报率对比图

Fig. 5 Comparison of cumulative ROI of ablation experiments

4 结束语

本文提出了一种基于长短期记忆网络(LSTM)、动态时间规整(DTW)算法和图注意力网络(GAT)相结合的加密货币预测模型,通过融合时间序列特征和关系特征,提高模型的预测精度。主要研究结论如下:

(1)使用 LSTM 捕捉时间序列依赖关系,有效解决了长时间依赖问题,提高了模型的时间序列预测能力。

(2)采用 DTW 算法动态衡量加密货币间的相似度,使得模型能够在不同时间尺度上进行匹配,增强了对不同时间序列的适应性。

(3)引入 GAT 学习加密货币间的复杂关系,通过注意力机制动态调整节点间的权重,从而准确捕捉节点间的关系特征。

(4)实验结果表明,LSTM-DTW-GAT模型在均方误差(MSE)、平均绝对误差(MAE)、累计投资回报率(IRR)、最大回撤(MDD)和夏普比率(SR)等评估指标上均优于 LSTM、CNN-LSTM 和 LSTM+GCN 等基线模型。特别是,关系信息的引入显著提

升了模型的预测效果,验证了该模型在加密货币预测领域的有效性和优越性。

LSTM-DTW-GAT 模型在加密货币预测中表现出色,但仅考虑了时间序列特征和关系特征的融合,忽略了市场其他潜在影响因素。在未来的研究中,可以引入更多元化的数据源,例如新闻情绪分析、宏观经济指标等,以进一步提升模型的预测精度和适用性。此外,本研究主要针对市值排名前 100 的加密货币,未来可以扩展到更大范围的加密货币市场,以验证该模型的广泛适用性和稳定性。

参考文献

- [1] CORBET S, LUCEY B, URQUHART A, et al. Cryptocurrencies as a financial asset: A systematic analysis [J]. *International Review of Financial Analysis*, 2019, 62: 182-199.
- [2] JAQUART P, DANN D, WEINHARDT C. Short-term bitcoin market prediction via machine learning [J]. *The Journal of Finance and Data Science*, 2021, 7: 45-66.
- [3] POONGODI M, VIJAYAKUMAR V, CHILAMKURTI N. Bitcoin price prediction using ARIMA model [J]. *International Journal of Internet Technology and Secured Transactions*, 2020, 10(4): 396-406.
- [4] NAIMY V Y, HAYEK M R. Modelling and predicting the Bitcoin volatility using GARCH models [J]. *International Journal of Mathematical Modelling and Numerical Optimisation*, 2018, 8(3): 197-215.
- [5] HITAM N A, ISMAIL A R, SAEED F. An optimized support vector machine (SVM) based on particle swarm optimization (PSO) for cryptocurrency forecasting [J]. *Procedia Computer Science*, 2019, 163: 427-433.
- [6] HU Zexin, ZHAO Y, KHUSHI M. A survey of forex and stock price prediction using deep learning [J]. *Applied System Innovation*, 2021, 4(1): 9.
- [7] AKYILDIRIM E, GONCU A, SENSOY A. Prediction of cryptocurrency returns using machine learning [J]. *Annals of Operations Research*, 2021, 297: 3-36.
- [8] MCNALLY S, ROCHE J, CATON S. Predicting the price of bitcoin using machine learning [C]//*Proceedings of the 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*. Piscataway, NJ:IEEE, 2018: 339-343.
- [9] REN Y S, MA C Q, KONG X L, et al. Past, present, and future of the application of machine learning in cryptocurrency research [J]. *Research in International Business and Finance*, 2022, 63: 101799.
- [10] PATEL M M, TANWAR S, GUPTA R, et al. A deep learning-based cryptocurrency price prediction scheme for financial institutions [J]. *Journal of Information Security and Applications*, 2020, 55: 102583.
- [11] LAHMIRI S, BEKIRO S. Cryptocurrency forecasting with deep learning chaotic neural networks [J]. *Chaos, Solitons & Fractals*, 2019, 118: 35-40.
- [12] DUTTA A, KUMAR S, BASU M. A gated recurrent unit approach to bitcoin price prediction [J]. *Journal of Risk and Financial Management*, 2020, 13(2): 23.
- [13] YU Yong, SI Xiaosheng, HU Changhua, et al. A review of recurrent neural networks; LSTM cells and network architectures [J]. *Neural Computation*, 2019, 31(7): 1235-1270.
- [14] GUNAY S, KASKALOGLU K. Does utilizing smart contracts induce a financial connectedness between Ethereum and non-fungible tokens? [J]. *Research in International Business and Finance*, 2022, 63: 101773.
- [15] SHI Y, TIWARI A K, GOZGOR G, et al. Correlations among cryptocurrencies: Evidence from multivariate factor stochastic volatility model [J]. *Research in International Business and Finance*, 2020, 53: 101231.
- [16] LI Panpan, GONG Shengbo, XU Shaocong, et al. Cross cryptocurrency relationship mining for bitcoin price prediction [C]//*Proceedings of International Conference on Blockchain and Trustworthy Systems*. Cham: Springer, 2022: 237-250.
- [17] ZHAO C, LIU X, ZHOU J, et al. GCN-based stock relations analysis for stock market prediction [J]. *PeerJ Computer Science*, 2022, 8: e1057.
- [18] KHEMANI B, PATIL S, KOTECHEA K, et al. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions [J]. *Journal of Big Data*, 2024, 11(1): 18.
- [19] ZHONG Chao, DU Wei, XU Wei, et al. LSTM-ReGAT: A network-centric approach for cryptocurrency price trend prediction [J]. *Decision Support Systems*, 2023, 169: 113955.
- [20] FARELL R. An analysis of the cryptocurrency industry [J]. *Wharton Research Scholars*, 2015, 130: 1-23.
- [21] MÜLLER M. Dynamic time warping [C]//*Information Retrieval for Music and Motion*. Cham: Springer, 2007: 69-84.
- [22] SHARABIANI A, DARABI H, HARFORD S, et al. Asymptotic dynamic time warping calculation with utilizing value repetition [J]. *Knowledge and Information Systems*, 2018, 57: 359-388.
- [23] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. *arXiv preprint arXiv, 1710.10903*, 2017.
- [24] 耿晶晶, 刘玉敏, 李洋, 等. 基于 CNN-LSTM 的股票指数预测模型 [J]. *统计与决策*, 2021, 37(5): 134-138.
- [25] WU W, HAN M, HU Y, et al. Optimization of financial data processing strategy based on LSTM-GCN [J]. *Journal of Education, Humanities and Social Sciences*, 2024, 30: 189-197.