

Multivariate Time Series Forecasting Based on Pre-trained Temporal Encoders and Graph Neural Networks

Yu Li, Xiaoxiao Wu

College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China

Abstract

Multivariate time series widely exist in many real-world systems such as finance, transportation, and social networks. Forecasting such data requires modeling both the temporal dynamics within individual sequences and the dependency relationships among different entities. In this paper, we propose a multimodal multivariate forecasting framework (MMPF) that jointly models temporal and relational information. In the temporal dimension, temporal neural networks (TNNs) are employed to capture short-term fluctuations and long-term trends of time series. In the spatial dimension, a graph structure is constructed to represent the relationships among entities, and graph neural networks (GNNs) are used to learn the structural dependencies. In addition, external news text information is incorporated to provide complementary semantic signals. The extracted multimodal features are finally fused by a multilayer perceptron (MLP) for prediction. Experimental results demonstrate that multimodal feature fusion can effectively improve the forecasting performance of multivariate time series.

Keywords: *Multivariate Time Series; Graph Neural Networks; Multimodal Feature Fusion; Time Series Forecasting; Cryptocurrency*

基于预训练时间编码器与图神经网络的多变量时间序列预测方法

李煜, 吴晓晓

深圳大学电子与信息工程学院, 中国深圳

摘要: 多变量时间序列广泛存在于金融、交通和社交网络等系统中, 其预测任务需要同时建模序列内部的时间动态以及不同实体之间的依赖关系。本文提出一种多模态多变量预测框架(MMPF), 从时间与空间两个维度进行联合建模。在时间维度上, 利用时间神经网络(TNN)提取序列的短期波动与长期趋势; 在空间维度上, 通过构建实体关系图并利用图神经网络(GNN)捕捉序列之间的结构依赖。同时引入新闻文本信息以补充潜在语义信号, 并通过多层感知器(MLP)融合多模态特征进行预测。实验结果表明, 多模态特征融合能够有效提升多变量时间序列的预测性能。

关键字: 多变量时间序列; 图神经网络; 多模态特征融合; 时间序列预测; 加密货币

1 引言

许多现实应用涉及多个相关时间序列的预测问题, 其中各序列不仅具有复杂的时间动态特征, 还通过隐式或显式的关系结构相互影响。这类问题广泛存在于金融、交通、能源系统和社交网络等领域^{[1][2][3][4]}。例如, 在

交通预测中，道路流量受到邻近道路的影响；在金融市场中，不同资产之间也存在行业关联与市场情绪等交互关系。从建模角度看，多变量时间序列通常表现出非平稳性、结构异质性以及动态依赖关系等特点^[5]。宏观事件、政策变化或群体行为等因素进一步增加了系统复杂性^{[6][7]}。因此，不同时间序列之间的关系往往随时间变化而变化^[8]。

传统方法主要依赖统计模型，例如 GARCH^[9]、VAR^[10]以及动态条件相关模型^[11]。尽管这些方法具有一定解释性，但在处理复杂非线性关系时存在局限。随着深度学习的发展，神经网络模型逐渐成为时间序列预测的重要工具，例如 LSTM^[12]和 TCN^[13]。近年来，预训练时间序列模型（如 TimesFM^[14]）能够利用大规模时间序列数据学习通用表示。同时，外部文本信息（如新闻或社交媒体）能够提供额外上下文信息，从而提升预测性能^[15]。此外，图神经网络（GNN）被广泛用于建模实体之间的关系结构^{[16][17][18]}。

基于上述背景，本文提出一种多模态多变量预测框架（MMPF），在统一架构下联合建模时间动态、实体关系以及文本语义信息。时间特征通过 TNN 模型提取，文本语义由 BERT^[22]编码器获取，实体之间的关系通过 GCN^[23]建模，并通过多模态融合实现预测。实验结果表明，该方法能够有效提升加密货币价格预测性能。

2 方法

2.1 问题建模

假设系统中包含 N 种加密货币，其节点集合为 $V=\{1,...,N\}$ ，并在 T 个时间步上进行观测。记货币 i 在时间 t 的价格为 $x_{i,t}$ 。为了减小不同资产价格尺度的影响，我们采用对数收益率表示价格变化：

$$r_{i,t} = \log x_{i,t} - \log x_{i,t-1}$$

在预测任务中，模型使用长度为 L 的历史窗口作为输入，并预测未来 h 个时间步的价格变化。

除了价格时间序列之外，我们还收集与资产相关的新闻文本信息，并构建资产之间的关系图 $G=(V,E,A)$ ，其中 $A \in \mathbb{R}^{N \times N}$ 为邻接矩阵，可由资产收益率相关系数计算得到，用于刻画资产之间的结构依赖关系。

2.2 多模态预测框架

如 Fig.1 所示，所提出框架主要包括时间编码、文本编码、特征融合和图结构建模四个模块。

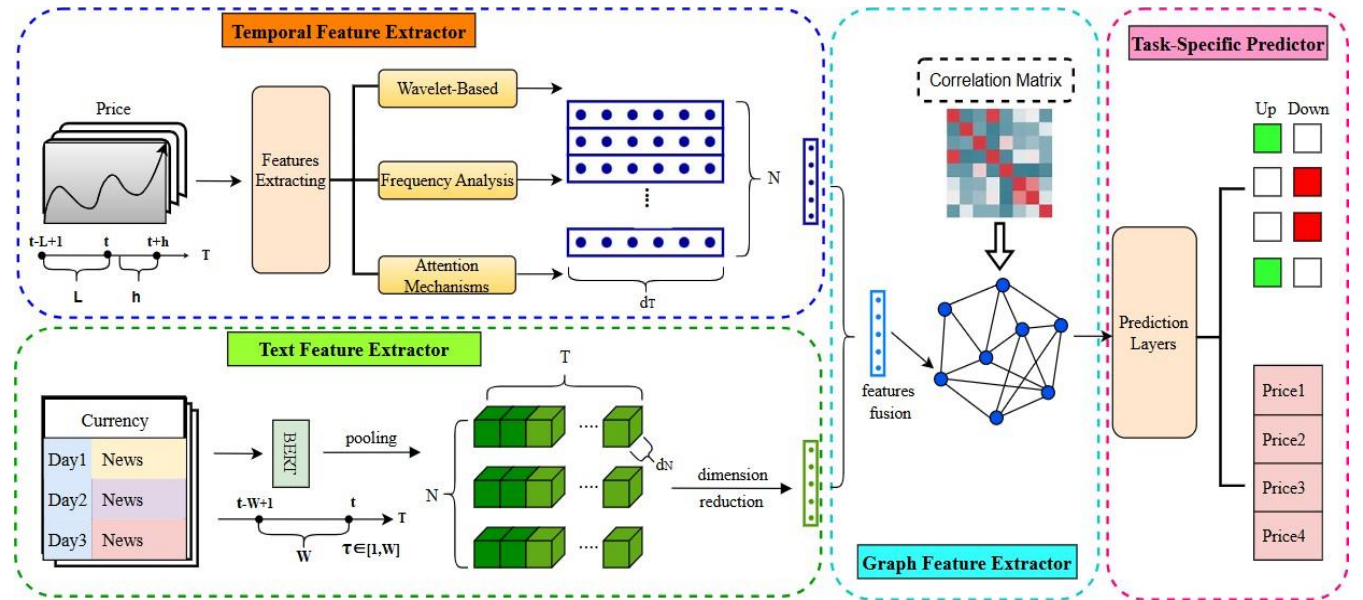


图 1 多模态多货币预测框架（MMPF）

(1) 时间编码器

时间编码器 $f_T(\cdot)$ 用于从历史价格序列中提取时间特征：

$$F_{i,t}^T = f_T(X_{i,t} - L + 1:t) \in R^{d_T}$$

编码器可以采用 LSTM、TCN 或预训练时间序列模型（如 WPMixer、TimeMixer），以捕捉序列中的长期依赖与短期波动。

(2) 文本编码器

新闻文本通过 BERT 转换为语义向量，并考虑新闻影响随时间衰减。我们采用加权时间池化聚合新闻特征：

$$F_{i,t}^N = \frac{\sum_{\tau=1}^W \sum_k w_{t,\tau} e_{i,t-\tau+1}^{(k)}}{\sum_{\tau=1}^W \sum_k w_{t,\tau}}$$

其中权重为：

$$w_{t,\tau} = 1 - \frac{\tau - 1}{W}$$

(3) 特征融合

时间特征与文本特征通过融合函数 $\phi(\cdot)$ 进行整合：

$$H_{i,t} = \phi(F_{i,t}^T, F_{i,t}^N)$$

融合可以通过特征拼接与多层感知器（MLP）实现，也可以采用交叉注意力机制。

(4) 图结构建模

为了建模资产之间的依赖关系，我们利用图神经网络进行信息传播。首先对邻接矩阵进行归一化：

$$\hat{A} = \tilde{D}^{-1/2}(A + I)\tilde{D}^{-1/2}$$

随后通过图卷积更新节点表示：

$$S_t^{(\ell+1)} = \sigma(\hat{A} S_t^{(\ell)} W^{(\ell)})$$

其中 $S^{(0)}=H_t$ ，最终得到图增强表示 S_t 。

(5) 预测层

基于最终特征 $S_{i,t}$ ，模型输出未来价格趋势或回归预测结果。

2.3 训练目标

整体模型可表示为

$$\hat{Y}_{t+1:t+h} = f_{\theta}(H_t, G)$$

训练过程中通过最小化预测误差进行优化，并采用正则化项 $\Omega(\theta)$ 防止模型过拟合。

3 实验

3.1 数据集与实验设置

实验选取 BTC、ETH、XRP、BNB、SOL、DOGE、LTC 和 AVAX 八种加密货币。数据来自 OKX 交易所，包含 2022 年 12 月至 2025 年 6 月的小时级收盘价，共 21,422 个样本。数据按 80%:20% 划分为训练集和测试集，并进行归一化处理。

新闻数据来自 CoinDesk，使用过去 10 天新闻作为输入。

时间特征采用 WPMixer^[19]、TimeMixer^[20]、TimeXer^[21]、LSTM 和 TCN。文本特征通过 BERT 提取。空间

关系通过三层 LE-GCN 建模，隐藏维度为 128、64 和 32。

预测任务包括：（1）分类任务：预测价格涨跌，评价指标为 ACC^[24]；（2）回归任务：预测价格趋势，评价指标为 RE^{[25][26]}。

3.2 不同 TNN 模型的影响

表 1 给出了不同时间神经网络模型在 LE-GCN 框架下的预测结果。结果表明，引入图结构能够显著提升整体性能。其中 TimeMixer 在融合 GCN 与新闻特征后取得最佳表现，ACC 达到 54.01%。

表 1 不同 TNN 模型在 LE-GCN 框架下的性能比较

Methods	WPMixer		TimeMixer	
	ACC(%)	RE(%)	ACC(%)	RE(%)
Baseline	51.73%	0.19%	52.67%	0.14%
+News	52.27%	0.08%	52.91%	0.09%
+LE-GCN	52.45%	0.01%	53.73%	0.03%
+LE-GCN and News	52.72%	0.17%	54.01%	0.05%
	TimeXer LSTM		LSTM	
	ACC(%)	RE(%)	ACC(%)	RE(%)
Baseline	52.13%	2.04%	51.88%	1.05%
+News	52.23%	2.18%	51.63%	1.82%
+LE-GCN	52.43%	0.88%	52.59%	0.90%
+LE-GCN and News	53.05%	0.32%	52.70%	0.45%
	TCN			
	ACC(%)	RE(%)		
Baseline	49.89%	1.26%		
+News	50.28%	1.42%		
+LE-GCN	50.01%	0.90%		
+LE-GCN and News	52.22%	1.22%		

3.3 不同 GNN 模型的影响

为了进一步评估图结构建模的重要性，我们固定 WPMixer 作为时间模型，并比较不同 GNN 架构的性能，如表 2 所示。结果表明，简单堆叠 GCN 可能导致性能下降，而改进的图模型（如 GAT）能够显著降低预测误差。

表 2 不同 GNN 模型在 WPMixer 框架下的性能比较

Configuration	ACC(%)	RE(%)
Baseline	51.59%	1.48%
Basic GCN	51.34%	13.22%
LE-GCN	52.38%	1.05%
GELU-GCN	52.02%	0.71%
GAT	52.16%	0.05%

4 结论

本文提出了一种融合多模态信息的多变量时间序列预测框架，通过联合建模价格时间序列、新闻文本信息

以及资产之间的关系结构，实现对加密货币价格变化的预测。模型利用时间神经网络（TNN）提取时间依赖特征，并通过图神经网络（GNN）建模资产之间的结构关系，同时引入新闻文本信息以补充外部语义信号。

实验结果表明，多模态信息融合能够有效提升价格趋势分类与价格回归任务的预测性能，并有助于揭示不同加密货币之间的潜在关联关系。未来工作将进一步探索利用大语言模型（LLMs）增强文本语义表示，并研究更加动态的关系图建模方法。

参考文献

- [1] Corbet S, Lucey B, Urquhart A, Yarovaya L. Cryptocurrencies as a financial asset: A systematic analysis[J]. *International Review of Financial Analysis*, 2019, 62: 182–199.
- [2] Wu Z, Pan S, Long G, Jiang J, Chang X, Zhang C. Connecting the dots: multivariate time series forecasting with graph neural networks[J]. *arXiv preprint*, 2020: 753–763.
- [3] Sun L, Liu X, Liu G, Yang Z, Liu S, Li Y, Wu X. Deep learning-based multivariate load forecasting for integrated energy systems[J]. *International Journal of Low-Carbon Technologies*, 2025, 20: 957–964.
- [4] Ng K W, Mubang F, Hall L O, Skvoretz J, Iamnitchi A. Experimental evaluation of baselines for forecasting social media time series[J]. *PLoS One*, 2023, 12: 1–26.
- [5] Antonakakis N, Chatziantoniou I, Gabauer D. Cryptocurrency market contagion: market uncertainty, market complexity, and dynamic portfolios[J]. *Journal of International Financial Markets, Institutions and Money*, 2019, 61: 37–51.
- [6] Bialkowski J, Gottschalk K, Wisniewski T P. High policy uncertainty and low implied market volatility: an academic puzzle?[J]. *Journal of Financial Economics*, 2022, 143(3): 1185–1208.
- [7] Hackethal A, Inderst R, Meyer S. The characteristics and portfolio behavior of Bitcoin investors: evidence from indirect cryptocurrency investments[J]. *Review of Finance*, 2021, 26(4): 855–898.
- [8] Liu J, Serletis A. Volatility and dependence in cryptocurrency and financial markets: a copula approach[J]. *Studies in Nonlinear Dynamics Econometrics*, 2024, 28(1): 119–149.
- [9] Fung K, et al. More to cryptos than Bitcoin: A GARCH modelling of heterogeneous cryptocurrencies[J]. *Finance Research Letters*, 2022, 47.
- [10] Jlassi N B, et al. Subsample analysis of stock market–cryptocurrency returns tail dependence: A copula approach for the tails[J]. *Finance Research Letters*, 2023, 58.
- [11] Zhang J, He Q Z. Dynamic cross-market volatility spillover based on MSV model: Evidence from Bitcoin, gold, crude oil, and stock markets[J]. *Complexity*, 2021: 1–8.
- [12] Seabe P L, Madzivhandila T, Odhiambo N M. Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: A deep learning approach[J]. *Fractal and Fractional*, 2023, 7(2).
- [13] Guo H, Li X, Li Z, Chen Y. Bitcoin price forecasting: A perspective of underlying blockchain transactions[J]. *Decision Support Systems*, 2021, 151.
- [14] Das A, Kong W, Sen R, Zhou Y. A decoder-only foundation model for time-series forecasting[C]// *Proceedings of the 41st International Conference on Machine Learning (ICML)*. 2024: 10148–10167.
- [15] Ye Z, et al. A stacking ensemble deep learning model for Bitcoin price prediction using Twitter comments on Bitcoin[J]. *Mathematics*, 2022, 10(8).
- [16] Tanwar S, et al. Deep learning-based cryptocurrency price prediction scheme with inter-dependent relations[J]. *IEEE Access*, 2021, 9: 138633–138646.
- [17] Zhou Y, et al. Forecasting cryptocurrency volatility: A novel framework based on the evolving multiscale graph neural network[J].

Financial Innovation, 2025, 11(1): 1–52.

- [18] Zhao L, Chen Q, Huang Z, Xu J. Long-term, short-term and sudden event: Trading volume movement prediction with graph-based multi-view modeling[EB/OL]. arXiv:2106.03764, 2021.
- [19] Murad M M N, et al. WPMixer: Efficient multi-resolution mixing for long-term time series forecasting[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2025, 39(18): 19581–19588.
- [20] Wang S, et al. TimeMixer: Decomposable multiscale mixing for time series forecasting[C]// Proceedings of the International Conference on Learning Representations (ICLR). 2024.
- [21] Wang Y, et al. TimeXer: Empowering transformers for time series forecasting with exogenous variables[C]// Advances in Neural Information Processing Systems (NeurIPS). 2024.
- [22] Pota M, et al. Multilingual evaluation of pre-processing for BERT-based sentiment analysis of tweets[J]. Expert Systems with Applications, 2021, 181: 115101.
- [23] Yin W, Liu Z, Zhang L. Forecasting cryptocurrencies'price with the financial stress index: A graph neural network prediction strategy[EB/OL]. SSRN Preprint, 2022.
- [24] Henrique B M, Sobreiro V A, Kimura H. Literature review: Machine learning techniques applied to financial market prediction[J]. Expert Systems with Applications, 2019, 124: 226–251.
- [25] Flores B E, Wichern D W. Evaluating forecasts: A look at aggregate bias and accuracy measures[J]. Journal of Forecasting, 2005, 24(6): 433–451.
- [26] Li Z, Yang D, Gao B, Jiao Y, Hong Y, Xu T. Multiscale hydrologic applications of the latest satellite precipitation products in the Yangtze River Basin using a distributed hydrologic model[J]. Hydrology and Earth System Sciences, 2015, 16: 407–426.