

用于金融时间序列预测的神经网络集成模型

张 晗¹, 王维国²

(1. 东北财经大学 数据科学与人工智能学院, 大连 116025; 2. 东北财经大学 经济学院, 大连 116025)

摘要: 准确预测金融时间序列数据对金融市场的运行和管理起着重要作用。基于神经网络和集成学习思想, 将卷积神经网络(CNN)、长短期记忆(LSTM)网络以及自回归移动平均(ARMA)模型在集成框架中进行组合, 提出一种新的用于预测金融时间序列数据的 ARMA-CNN-LSTM 模型。该模型通过 CNN-LSTM 模型对数据中时空特征进行建模, 同时利用 ARMA 模型对数据的自相关特征进行建模, 实现对金融时间序列数据中线性和非线性特征的混合建模。实验结果表明, 与基准个体模型相比, 所提模型在预测金融时间序列数据的精度和鲁棒性方面都显示出优异的性能。

关 键 词: 金融时间序列; 卷积神经网络; 长短期记忆; 自回归移动平均; 集成预测模型

中图分类号: TP183

文献标志码: A

Neural Network Ensemble Models for Financial Time Series Forecasting

ZHANG Han¹, WANG Weiguo²

(1. School of Data Science and Artificial Intelligence, Dongbei University of Finance and Economics, Dalian 116025, China;

2. School of Economics, Dongbei University of Finance and Economics, Dalian 116025, China)

Abstract: The accurate forecasting of financial time series data holds a significant position in the operation and management of financial markets. Grounded in the notions of neural networks and ensemble learning, the convolutional neural network (CNN), long short-term memory (LSTM) network, and autoregressive moving average (ARMA) model are integrated within an ensemble framework to put forward a novel ARMA-CNN-LSTM model for forecasting financial time series data. The spatiotemporal characteristics within the data are modeled by the CNN-LSTM model, while the autocorrelation features of the data are simultaneously modeled by the ARMA model, achieving hybrid modeling of linear and nonlinear traits in financial time series data. The experimental results show that, compared with the baseline individual models, the proposed model demonstrates excellent performance in both the accuracy and robustness on predicting financial time series data.

Key words: financial time series; convolutional neural network; long short-term memory; autoregressive integrated moving average; ensemble forecasting models

金融市场预测是金融时间序列预测领域中重要的研究主题之一, 而各种金融市场的金融产品价格通常以金融时间序列的形式呈现。大量实证

研究表明, 金融时间序列数据会表现出不同的特征^[1]。除了利用均衡分析法对金融市场进行建模外, 通过数据分析和建模对金融市场进行预测也

是研究热点之一。

时间序列模型和人工智能模型等的应用近年来在上述领域都得到了长足发展,每种模型因其特定的假设而被设计用于捕获不同的数据特征^[2]。例如,自回归移动平均(ARMA, autoregressive moving average)旨在对数据中的自相关性进行建模;卷积神经网络(CNN, convolutional neural network)、长短时记忆(LSTM, long short-term memory)和支持向量回归(SVR, support vector regression)等人工智能模型对数据中的非线性特征的建模,对金融时间序列的准确预测起到了非常重要的作用。其中,深度学习模型是目前较先进的模型,该类模型的独特之处在于它针对特定的时间和空间数据特征建模。而结合了不同模型的混合方法在很大程度上提高了模型的拟合效果与预测精度^[3]。

笔者在 ARMA 和 CNN-LSTM 模型的基础上提出一种新的集成预测模型 ARMA-CNN-LSTM。不同于现有的依靠人工神经网络和 SVR 等对非线性数据特征进行建模的集成模型,所提模型利用 CNN-LSTM 模型对相邻时间点观测值之间的非线性空间特征和数据中的长时依赖关系同时建模,并通过 ARMA 模型对线性自相关数据特征建模。因此,所提模型的最终预测值由线性模型 ARMA 的线性预测值和非线性模型 CNN-LSTM 的非线性预测值两部分组成。研究发现,在集成预测框架下,通过利用深度学习模型能很好地提取时间序列数据时空特征的优势,实现对金融时间序列更准确地建模和预测。将 ARMA-CNN-LSTM 与 ARMA 等多个个体模型在预测性能方面进行了综合评价。实验结果表明,与其他所评估的各模型相比,ARMA-CNN-LSTM 在金融时间序列数据预测方面具有最佳的预测效果。

1 相关工作

1.1 ARMA 模型

ARMA 等传统线性时间序列模型广泛应用于分析金融时间序列中的自相关性等线性数据特征,并常作为基准模型与其他模型的预测结果进行比较。Chevallier 等^[4]提出了一个非参数模型分别对碳现货和碳期货价格进行预测,该模型与线性自回归(AR, autoregression)模型相比有效地降低了预测误差。Zhao 等^[5]提出了一种新的碳期货

价格预测模型,其预测效果优于 AR、移动平均(MA, moving average)和自回归条件异方差等基准模型。

在 ARMA 模型中,时间序列数据的当前值是之前的值和白噪声项的线性组合:

$$\phi(\mathbf{L})\mathbf{r}_t = c + \theta(\mathbf{L})\boldsymbol{\mu} \quad (1)$$

其中: $\mathbf{r}_t (t=1, 2, \dots, n)$ 为时刻 t 的金融时间序列; $\boldsymbol{\mu}$ 为独立同分布的随机变量(即白噪声项); c 为常数,通常假设为 0; \mathbf{L} 为滞后算子,是将 \mathbf{r}_t 转换为 \mathbf{r}_{t-1} 的转换矩阵;设 p 和 q 为整数, $\phi(\mathbf{L}) = 1 - \phi_1\mathbf{L} - \phi_2\mathbf{L}^2 - \dots - \phi_p\mathbf{L}^p$ 为 AR 的 p 阶滞后多项式, $\theta(\mathbf{L}) = 1 - \theta_1\mathbf{L} - \theta_2\mathbf{L}^2 - \dots - \theta_q\mathbf{L}^q$ 为 MA 的 q 阶滞后多项式。ARMA 等传统线性模型擅长对金融时间序列的线性特征进行建模,但在对非线性特征进行建模时会面临困难。

1.2 人工智能模型

近年来,人工智能模型因其强大的非线性建模能力而备受关注,不同人工智能模型的组合使用在提高时间序列的预测效果上展现了不凡实力。Wen 等^[6]提出一种结合了奇异谱分析(SSA, singular spectrum analysis)与支持向量机(SVM, support vector machine)的混合预测模型 SSA-SVM,在预测股票价格方面比基于集合经验模态分解和 SVM 等模型具有更高的预测精度;Shen 等^[7]将门控循环单元(GRU, gated recurrent unit)和 SVM 模型相结合提出的 GRU-SVM 与 GRU、SVM 和深度神经网络相比,拥有更好的股票交易信号预测效果。

CNN 和 LSTM 作为两种典型的深度学习模型,因其对数据的长时依赖性和时间序列中局部相邻特征方面的强大建模优势^[8]。

1.2.1 CNN 模型

CNN 模型通过在卷积层引入卷积操作对输入数据的局部区域之间的非线性关系进行建模。CNN 一般由输入层、卷积层、池化层、展平层、全连接层和输出层组成。卷积层中的卷积操作为

$$\mathbf{F}_t = \sigma \left(\sum \mathbf{w} * \mathbf{x}_t + \mathbf{b} \right) \quad (2)$$

其中: \mathbf{F}_t 为卷积层中滤波器所输出的特征图, \mathbf{x}_t 为输入矩阵, \mathbf{w} 和 \mathbf{b} 分别为滤波器的权重向量和偏置向量, $\sigma(\cdot)$ 为激活函数,* 为卷积操作。

池化层用于对前一个卷积层生成的特征图进行

下采样。Flatten 层用于将多维特征图的形状展平为一维形状,而全连接层根据变换后的特征图计算最终的预测结果。

1.2.2 LSTM 模型

LSTM 模型是传统循环神经网络 (RNN, recurrent neural network) 模型的特殊变体,其旨在解决在 RNN 中,对时间序列的长期依赖关系进行建模时可能会发生梯度消失或梯度爆炸等问题。LSTM 的隐藏层由内循环系统和外循环机制两部分组成。外循环机制使得 LSTM 能够递归地馈送网络输入中先前时间点的隐藏状态,并对最终预测值产生影响。内循环系统的主要部分是记忆单元,它是一个自连接单元,其主要作用是存储长时间序列的时间状态信息,起到防止梯度爆炸和梯度消失的作用。LSTM 隐藏层的基本结构是以记忆单元为关键部分的储存块,其包括 3 个特殊的门控单元,其中,输入门用于决定当前输入中的哪些信息可以馈送到记忆单元中以更新单元状态;遗忘门的功能是确定前一个单元状态中的哪些信息可以存储在记忆单元中,哪些应该被丢弃,遗忘门还可以防止单元状态的值无限增长;输出门的作用是控制在计算当前时刻预测值时,应该过滤掉存储单元中的哪些信息。

时间 t 时,输入门、遗忘门、输出门的值以及记忆单元的值分别为

$$i_t = \sigma(\mathbf{w}_i[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_i) \quad (3)$$

$$g_t = \sigma(\mathbf{w}_g[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_g) \quad (4)$$

$$o_t = \sigma(\mathbf{w}_o[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_o) \quad (5)$$

$$\hat{C}_t = \sigma(\mathbf{w}_c[\mathbf{x}_t, \mathbf{h}_{t-1}] + \mathbf{b}_c) \quad (6)$$

其中: $\mathbf{w}_i, \mathbf{w}_g, \mathbf{w}_o$ 和 \mathbf{w}_c 分别为对应单元的权重矩阵, $\mathbf{b}_i, \mathbf{b}_g, \mathbf{b}_o$ 和 \mathbf{b}_c 分别为对应单元的偏置向量。

根据上述输入门、遗忘门、记忆单元以及之前记忆单元在 $t-1$ 时刻的状态值,更新当前时间点 t 的记忆单元:

$$C_t = g_t \circ C_{t-1} + i_t * \hat{C}_t \quad (7)$$

其中: \circ 为按元素乘积, \hat{C}_t 为当前时间点 t 的单元状态的更新值。

输出门从记忆单元中筛选出所需信息,并根据从记忆块中计算出的隐状态信息得到 LSTM 的最终输出值。LSTM 的隐输出状态和最终预测值分别为

$$h_t = o_t * \tanh C_t \quad (8)$$

$$y_t = a(\mathbf{w}_y h_t + \mathbf{b}_y) \quad (9)$$

其中: \mathbf{w}_y 和 \mathbf{b}_y 分别为 LSTM 模型输出层的权重矩阵

和偏置向量, $a()$ 为 LSTM 输出层中的激活函数。在训练过程中,使用基于梯度下降的反向传播算法最小化损失函数来优化 LSTM 模型的参数。

1.3 集成框架

很多学者尝试将深度学习模型和传统线性时间序列模型在集成框架中进行组合,用于构建具有更好预测效果的金融时间序列数据模型。尽管在某些特定场景中,ARMA 等单一预测模型对金融时间序列表现出了良好的预测性能,但随着时间序列数据的激增,数据中存在更大的随机波动性,单一模型不再适合用来处理复杂的时间序列数据。如何选择不同模型,通过集成的方式构建具有更高挖掘金融时间序列复杂数据特征能力的预测模型,是一个具有现实意义及研究价值的问题。目前,有关金融时间序列的混合预测模型和方法的研究与应用取得了很多成果,但利用人工智能和传统时间序列模型对金融时间序列中不同数据特征进行同时建模的研究却少之又少^[9]。因此,在集成框架下,笔者通过传统时间序列模型 ARMA 与深度学习模型 CNN-LSTM 的组合建模,同时获取金融时间序列数据的线性和非线性特征,在金融时间序列混合模型的构建方法上做了有益尝试,为提高金融时间序列模型的预测效果提供一条有效的途径。

2 集成预测模型的构建

2.1 个体模型

现有研究中已存在众多为预测金融时间序列而构建的模型,需要从中选择合适的个体模型来进行预测,从而构建一个能够对更为广泛数据特征进行建模的集成预测模型。

在众多线性模型中,选择最具代表性的 ARMA 模型对数据中的自相关特征进行建模。ARMA 模型是 AR 以及 MA 两种模型的结合体,对时间序列数据具有良好的短期预测性能和解释能力。

近年来,学术界已涌现出很多针对时间序列数据的不同特征进行建模的深度学习模型,CNN 和 LSTM 是其中最主流的 2 个模型。CNN 能够有效地提取时间序列数据中的空间局部特征,LSTM 可以很好地表达输入中的长期依赖的信息。此外,根据 Karim 等^[10]的研究,LSTM 和全卷积网络以联合的方式用于预测时间序列数据时都能表现出非常好的效果。因此,使用全卷积网络作为 CNN 的结构,

并利用同时包含 CNN 和 LSTM 的 CNN-LSTM 模型对时间序列数据中的特征进行建模,CNN-LSTM 模型结构如图 1 所示。

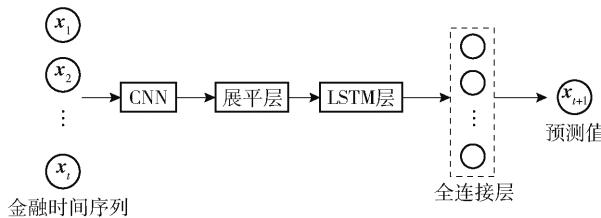


图 1 CNN-LSTM 模型的结构

2.2 集成预测模型

集成预测模型通过组合多个机器学习模型来产生一个优化的模型,从而提高模型的性能。通过集成的方式可以很容易地减少过拟合,避免模型在训练时表现更好,而在测试时不能产生良好的结果。

集成预测模型 ARMA-CNN-LSTM 是通过结合能够捕获不同数据特征的个体模型,并为每个模型分配相应权重的方式对时间序列数据进行预测。

假设总体的预测值 \hat{y}_f 与个体模型的预测值 \hat{y}_i 之间是线性关系,则若有 n 个个体模型,且每个个体模型的预测值都分配了一个可学习的权重 w_i ,那么,最终的预测值为

$$\hat{y}_f = \sum_{i=1}^n \hat{y}_i * w_i \quad (10)$$

具体来说,首先将 ARMA 和 CNN-LSTM 模型并入到集成框架中形成 ARMA-CNN-LSTM 混合模型:由 ARMA 模型捕捉金融时间序列数据的线性趋势并得出线性预测值;用 CNN 和 LSTM 所组成的 CNN-LSTM 非线性模型同时提取相邻观测值之间的非线性空间特征和整个历史观测值之间的非线性长期依赖特征,得出时间序列数据的非线性预测值。然后,将上述 2 个个体模型的预测值根据式(10)得出最终的预测值。

3 实证分析

3.1 数据集

使用 3 个不同的金融时间序列数据集对 ARMA-CNN-LSTM 模型的预测性能进行实证评估,它们分别是欧洲碳交易市场的每周欧洲碳排放交易体系(EU ETS, European union emissions trading system)数据^[11]、中国股市的每日上证综合指数

(SCI, Shanghai composite index)^[11]和每日比特币在加密货币市场的价格^[11]。其中,EU ETS 数据集涵盖了 2008 年 4 月 7 日至 2022 年 4 月 11 日期间,共 731 次每周观测;SCI 数据集覆盖 2010 年 1 月 4 日至 2022 年 1 月 23 日期间,共 3 647 个每日观测数据;比特币数据集包括了 2012 年 2 月 2 日至 2022 年 8 月 8 日期间,共 4 217 个每日观测。为了便于模型的测试和性能评估,将所有数据集按照 70:10:20 的比例划分为训练集、验证集和测试集。其中,训练集用于训练不同模型的参数,验证集用于确定最优的模型结构,测试集用于测试对比不同模型在样本外预测中的预测精度。

模型的预测性能指标包括平均绝对误差(MAE, mean absolute error)、平均绝对百分比误差(MAPE, mean absolute percentage error)、均方根误差(RMSE, root mean square error)和方向统计量 D_{st} 。 MAE 、 $MAPE$ 和 $RMSE$ 是用来衡量预测值与实际值之间的距离,其值越小表明预测值与实际值越接近,模型的预测性能越好; D_{st} 是用来衡量模型在预测金融时间序列移动方向上的准确性,其值越大表明预测模型对金融时间序列移动方向的预测越准确。

3.2 对比模型

选用 8 个个体模型作为对比模型:① ARMA 为最具代表性的线性时间序列预测模型;② CNN 能有效地对数据中的非线性特征建模;③ LSTM 网络为 RNN 的变体,能有效解决传统 RNN 模型在对时间序列的长期依赖关系进行建模时可能会遇到的梯度消失或梯度爆炸等问题;④ CNN-LSTM 是由 CNN 和 LSTM 所组成的模型,其能对数据中的时空特征建模;⑤ 长短期时间序列网络(LSTNet, long short-term time series network)^[12]是一种专门用于多元时间序列预测的深度学习网络;⑥ TimeGrad^[13]是最早使用扩散模型做时间序列预测的方法之一;⑦ RNN-GRU^[14]是具有全连接 GRU 隐藏单元的循环神经网络;⑧ STGCN^[15]模型能够提取时间序列数据中有用的空间特征,并能连贯地捕捉到最基本的时间特征。

3.3 参数选择

通过搜索确定对比模型和 ARMA-CNN-LSTM 的最优参数。对于 CNN-LSTM,选择能使其产生最小 RMSE 的参数作为其最优参数。个体模型 CNN

和 LSTM 中的超参数通过网格搜索法确定, 并使用 RMSE 作为性能度量。其中, 滤波器的搜索空间是(1, 10); 神经元的搜索空间是(1, 100)。具有至少一个隐藏层、并使用非线性激活函数的多层次感知机是一个通用的函数逼近器^[16], 如果神经元的数量足够多, 其能够对任意函数建模。因此, 隐藏层的搜索范围为 1 至 3, 每个隐藏层中神经元的搜索范围为 1 至 200; 滚动窗口的大小设置为 5, 也就是说, 每个子周期都包含 5 个先前的观测值, 这相当于一周有 5 个交易日。滚动窗口每次向前滚动一步, 通过贪婪搜索方法评估不同模型结构在验证集中的 RMSE 来确定最优的模型结构。

3.4 模型预测结果

确定最优参数后, 将个体对比模型和集成预测模型 ARMA-CNN-LSTM 应用于金融时间序列数据预测, 并对这些模型的性能进行比较和评价。表 1 至表 3 所示为各预测模型在 4 种预测性能指标下对 3 种金融时间序列数据进行 10 倍交叉验证时的实验结果。

表 1 各模型对 EU ETS 金融时间序列的预测性能

模型	RMSE	MAPE	MAE	D_{st}
ARMA	1.2379	0.0413	0.9122	0.5581
LSTM	3.9867	0.1552	3.3696	0.5504
CNN	1.7748	0.0621	1.3474	0.4884
CNN-LSTM	1.2205	0.0409	0.8978	0.5794
STGCN	1.2325	0.0411	0.9017	0.5774
LSTNet	1.2294	0.0409	0.8997	0.5803
TimeGrad	1.2287	0.0487	0.8974	0.5814
RNN-GRU	1.2275	0.0479	0.8954	0.5874
ARMA-CNN-LSTM	1.1195	0.0380	0.8537	0.5947

表 2 各模型对 SCI 金融时间序列的预测性能

模型	RMSE/ $\times 10^{-2}$	MAPE	MAE/ $\times 10^{-2}$	D_{st}
ARMA	1.2004	1.3853	0.8669	0.7416
LSTM	1.2022	1.1637	0.8644	0.7421
CNN	1.2061	1.2057	0.8679	0.7403
CNN-LSTM	1.2014	1.1632	0.8632	0.7421
STGCN	1.2154	1.1635	0.8635	0.7418
LSTNet	1.1998	1.1624	0.8624	0.7415
TimeGrad	1.1987	1.1579	0.8627	0.7419
RNN-GRU	1.1989	1.1584	0.8628	0.7421
ARMA-CNN-LSTM	1.1864	1.1429	0.8570	0.7443

所提 ARMA-CNN-LSTM 模型的 4 个预测性能指标在 3 个数据集上都取得了比个体对比模型更好的预测性能, 同时 D_{st} 高于其他所有模型, 表明该模型在预测金融时间序列方面比现有的主流模型具有更高的预测精度以及更好的方向预测精度。

表 3 各模型对比特币收盘价的预测性能

Model	RMSE	MAPE	MAE	D_{st}
ARMA	324.6788	0.0258	199.5287	0.4928
LSTM	476.8439	0.0423	327.0795	0.5395
CNN	378.6605	0.0315	243.0132	0.5298
CNN-LSTM	332.4257	0.0267	225.1432	0.5532
STGCN	375.3432	0.0325	243.1242	0.5484
LSTNet	323.9403	0.0261	215.3742	0.5501
TimeGrad	343.3278	0.0259	218.4821	0.5421
RNN-GRU	333.8942	0.0257	214.3802	0.5421
ARMA-CNN-LSTM	323.2705	0.0252	196.8432	0.5576

CNN 和 LSTM 在 3 个数据集上的预测性能都比 ARMA 的预测效果要差, 表明这 2 个深度学习模型都不能对具有复杂特征的时间序列数据提供足够的拟合, 它们只适用于对数据特征的某些方面进行建模。而金融时间序列数据通常是复杂的, 并混合了多种数据特征, 因此需要通过组合不同类型的模型才能达到较好的预测效果, 如所提的集成预测模型 ARMA-CNN-LSTM。

式(10)是一种具有鲁棒性的集成方法。采用该方法对个体模型 ARMA 以及 CNN-LSTM 的预测值进行加权平均所构建的 ARMA-CNN-LSTM, 可以很好地减少预测的估计偏差, 提高预测的精度。集成预测模型 ARMA-CNN-LSTM 相对个体模型预测精度的提高不仅得益于针对不同数据特征的深度学习模型的引入, 而且在集成模式下, 实现了对深度学习模型和计量经济模型所捕捉到信息的有效融合。

4 结束语

金融时间序列数据常呈现复杂性并包含很多不同类型的特征, 单一模型已不再适合用来处理复杂的时间序列数据。利用传统 ARMA 模型能提取时间序列数据中的自相关性等线性特征的特性, 结合深度学习模型 CNN-LSTM 能提取时空数据中的非线性特征的特性, 提出集成预测模型 ARMA-CNN-

LSTM, 它在3个金融时间序列数据集上表现出良好的预测能力, 证明了该模型具有很好的鲁棒性以及较高的应用推广价值。实验结果表明, 在集成框架中引入特定的深度学习模型能对金融时间序列数据中的混合特征进行准确建模, 在降低预测误差的同时大幅提高模型的预测能力。

参考文献:

- [1] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting [J]. *Machine Learning*, 2019, 108: 1421-1441.
- [2] RANE N L, CHOUDHARY S P, RANE J. Artificial intelligence-driven corporate finance: enhancing efficiency and decision-making through machine learning, natural language processing, and robotic process automation in corporate governance and sustainability [J]. *Studies in Economics and Business Relations*, 2024, 5(2): 1-22.
- [3] LIM B, ZOHREN S. Time-series forecasting with deep learning: a survey [J]. *Philosophical Transactions of the Royal Society A*, 2021, 379(2194): 20200209.
- [4] CHEVALLIER J. Nonparametric modeling of carbon prices [J]. *Energy Economics*, 2011, 33(6): 1267-1282.
- [5] ZHAO X, HAN M, DING L L, et al. Usefulness of economic and energy data at different frequencies for carbon price forecasting in the EU ETS [J]. *Applied Energy*, 2018, 216: 132-141.
- [6] WEN F H, XIAO J H, HE Z F, et al. Stock price prediction based on SSA and SVM [J]. *Procedia Computer Science*, 2014, 31: 625-631.
- [7] SHEN G Z, TAN Q P, ZHANG H Y, et al. Deep learning with gated recurrent unit networks for financial sequence predictions [J]. *Procedia computer science*, 2018, 131: 895-903.
- [8] ZHANG G P. Time series forecasting using a hybrid ARIMA and neural network model [J]. *Neurocomputing*, 2003, 50: 159-175.
- [9] PAVLYSHENKO B M. Machine-learning models for sales time series forecasting [J]. *Data*, 2019, 4(1): 15.
- [10] KARIM F, MAJUMDAR S, DARABI H, et al. LSTM fully convolutional networks for time series classification [J]. *IEEE Access*, 2018, 6: 1662-1669.
- [11] Anon. 网易财经股票 [EB/OL]. (2016-02-17) [2023-10-13]. <https://money.163.com/stock/>.
- [12] WANG D Z, CHEN C Y. Spatiotemporal self-attention-based LSTNet for multivariate time series prediction [J]. *International Journal of Intelligent Systems*, 2023, 2023(1): 1-16.
- [13] RASUL K, SEWARD C, SCHUSTER I, et al. Autoregressive denoising diffusion models for multivariate probabilistic time series forecasting [C] // In *International Conference on Machine Learning*. Cambridge: JMLR, 2021: 8857-8868.
- [14] ZAINUDDIN Z, AKHIR P E, HASAN M. Predicting machine failure using recurrent neural network-gated recurrent unit (RNN-GRU) through time series data [J]. *Bulletin of Electrical Engineering and Informatics*, 2021, 10(2): 870-878.
- [15] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network data forecasting [C] // In *Proceedings of the AAAI Conference on Artificial Intelligence*. Cambridge: JMLR, 2020(34): 914-921.
- [16] HORNIK K. Approximation capabilities of multilayer feedforward networks [J]. *Neural Networks*, 1991, 4(2): 251-257.