

文章编号: 1671-251X(2023)07-0092-07

DOI: 10.13272/j.issn.1671-251x.18142

基于可变时序移位 Transformer-LSTM 的 集成学习矿压预测方法

李泽西

(西安科技大学 通信与信息工程学院, 陕西 西安 710054)

摘要: 现有的矿压预测模型多为依赖固定长度时序序列特征的单一预测模型, 难以准确捕捉矿压时序数据的复合特征, 影响矿压预测的准确度。针对该问题, 提出一种基于可变时序移位 Transformer-长短期记忆(LSTM)的集成学习矿压预测方法。基于拉依达准则和拉格朗日插值法, 剔除矿压监测数据中的异常值, 插入缺失值, 并进行归一化预处理; 提出可变时序移位策略, 划分不同尺度的矿压时序数据, 避免固定长度时序序列可能存在的数据偏移问题; 在此基础上, 构建基于 Transformer-LSTM 的集成学习矿压预测模型, 通过结合注意力机制和准确的时序特征表示能力, 多层次捕捉矿压变化规律的动态特征, 采用集成学习的投票算法, 联合预测矿压数据, 克服单一预测模型的局限性。实验结果表明: 采用集成学习的投票算法可降低矿压预测平均绝对误差(MAE)和均方根误差(RMSE)的波动性, 有效减小不同尺度特征序列对矿压预测结果的敏感性影响; Transformer-LSTM 模型在 2 个综采工作面顶板矿压数据集上预测结果的 MAE 较 Transformer 模型分别提高了 8.9% 和 9.5%, RMSE 分别提高了 12.7% 和 16.5%, 且高于反向传播(BP)神经网络模型和 LSTM 模型, 有效提升了矿压预测准确度。

关键词: 矿压预测; 可变时序移位; Transformer-LSTM 模型; 集成学习; 投票算法

中图分类号: TD323

文献标志码: A

Ensemble learning mine pressure prediction method based on
variable time series shift Transformer-LSTM

LI Zexi

(College of Communication and Information Technology, Xi'an University of
Science and Technology, Xi'an 710054, China)

Abstract: The existing mine pressure prediction models are mostly single prediction models that rely on fixed length time series features. It is difficult to accurately capture the composite features of mine pressure time series data, which affects the accuracy of mine pressure prediction. To solve this problem, an ensemble learning mine pressure prediction method based on variable time series shift Transformer-long short-time memory (LSTM) is proposed. Based on the Laida criterion and Lagrange polynomial method, the outlier values in the mine pressure monitoring data are eliminated, and the missing values are inserted. Then normalized preprocessing is performed. The paper proposes a variable time series shift strategy to divide mine pressure time series data at different scales. It avoids potential data shift issues that may exist in fixed length time series. On this basis, the ensemble learning mine pressure prediction model based on Transformer-LSTM is constructed. By combining the attention mechanism and the accurate time series feature representation capability, the dynamic features of the mine pressure change law are captured at multiple levels. The voting algorithm of ensemble learning is used to jointly predict the mine pressure data to overcome the limitations of a single prediction model. The experimental results show that the voting algorithm of ensemble learning can reduce the volatility of mean absolute error (MAE) and

收稿日期: 2023-05-10; 修回日期: 2023-07-10; 责任编辑: 李明。

作者简介: 李泽西(2001—), 男, 湖北孝感人, 研究方向为人工智能、电子信息技术在煤矿领域的应用, E-mail: 1632982226@qq.com。

引用格式: 李泽西. 基于可变时序移位 Transformer-LSTM 的集成学习矿压预测方法[J]. 工矿自动化, 2023, 49(7): 92-98.

LI Zexi. Ensemble learning mine pressure prediction method based on variable time series shift Transformer-LSTM[J]. Journal of Mine Automation, 2023, 49(7): 92-98.



扫码移动阅读

root mean square error (RMSE) of mine pressure prediction. It effectively reduces the sensitivity impact of different scale feature series to the mine pressure prediction results. Compared with the Transformer model, the MAE of the Transformer-LSTM model's prediction results on two roof mine pressure datasets of fully mechanized working faces improves by 8.9% and 9.5% respectively, and the RMSE has increased by 12.7% and 16.5% respectively. The above indexes are also higher than those of back propagation (BP) neural network model and LSTM model. The method proposed in the paper effectively improves the accuracy of mine pressure prediction.

Key words: mine pressure prediction; variable time series shift; Transformer-LSTM model; ensemble learning; voting algorithm

0 引言

随着煤炭开采深度的增加, 矿压显现规律的研究仍是煤炭开采的重要课题。矿压是煤矿生产中不可避免的问题, 其变化规律的研究能够揭示煤炭开采的物理本质和变形机制^[1], 为煤矿安全生产提供科学依据。因此, 针对矿压预测问题的研究持续受到关注^[2]。目前, 矿压预测方法已由早期的物理力学、统计学等方法^[3]逐步转变为以浅层神经网络预测模型^[4-5]、深度神经网络预测模型和其他类别机器学习模型^[6-8]为主的预测方法。

国内外学者针对综采工作面矿压预测问题进行了大量研究工作, 主要分为 3 类: ① 基于物理理论的矿压预测方法。惠艳荔^[9]利用电磁辐射技术预测采煤工作面的矿压, 通过对采煤机推进过程中产生的采动应力分布情况进行分析, 成功预测了矿压发生的情况。李静等^[10]提出了一种基于 CT 技术的矿压灾害预测方法, 利用 CT 技术反演待开采工作面的应力场, 预测可能发生矿压事故的区域范围, 并实时监测和分析矿震特征, 提前预报矿压事故的发生。② 基于模糊数学的矿压预测方法。Li Ye 等^[11]利用模糊认知图技术构建了一种矿压预测模型, 在引入遗漏正则化和权重正则化基础上, 提出了实数编码遗传算法, 用于矿压预测。霍丙杰等^[12]建立了矿压显现分级预测指标体系, 并利用模糊数学理论构建了矿压显现强度分级预测模型, 能够准确刻画坚硬顶板厚煤层的矿压显现特征。③ 基于机器学习的矿压预测方法。赵铭生等^[13]利用遗传算法优化反馈神经网络, 建立了矿压毁坏高度的预测模型, 在矿压破坏深度预测方面具有更高的预测精度。赵毅鑫等^[14]基于实测矿压数据, 利用长短时记忆 (Long Short-Term Memory, LSTM) 模型对工作面支架工作阻力、初次来压与周期来压等矿压显现规律进行分析, 实验结果表明该方法较基于反馈神经网络的预测方法具有更高的准确性。曾庆田等^[15]提出了融合相邻多架影响的 Prophet+LSTM 矿压数据预测模型, 通过数字信号处理技术减少噪声和随机误差, 结合 Prophet

和 LSTM 模型特性, 通过添加额外回归变量融合相邻支架矿压数据, 构建组合模型, 实现工作面矿压预测, 通过实验验证了该模型优于反馈神经网络和自回归积分滑动平均模型 (Autoregressive Integrated Moving Average Model, ARIMA), 具有更强的稳定性和准确性。李泽萌^[16]研究了 LSTM 模型时间窗设置对综采工作面矿压预测精度的影响, 通过优化历史数据和预测数据长度, 显著提高了预测结果的准确性。

目前, 基于机器学习的矿压预测方法已广泛应用于采矿领域。矿压数据具有时序性和非线性特性, 其变化受多种因素 (采煤工艺、地质条件等) 的影响。而基于机器学习的矿压预测方法大多采用单一预测模型, 较依赖固定长度时序序列特征, 难以有效捕捉其复合特征, 影响矿压预测的准确度。为解决该问题, 本文以综采工作面实测的顶板矿压数据为研究对象, 提出了一种基于可变时序移位 Transformer-LSTM 的集成学习矿压预测方法: 基于拉依达准则和拉格朗日插值法, 剔除矿压监测数据中的异常值, 插入缺失值, 并进行归一化数据预处理; 提出可变时序移位策略, 通过划分不同尺度的矿压时序数据, 避免固定长度时序序列可能存在的数据偏移问题; 构建基于 Transformer-LSTM 的集成学习矿压预测模型, 通过多层次捕捉矿压变化规律的动态特征, 联合预测矿压数据, 克服了单一预测模型的局限性。

1 矿压预测方法

基于可变时序移位 Transformer-LSTM 的集成学习矿压预测方法主要包括矿压数据预处理、基于可变时序移位的矿压数据划分、基于 Transformer-LSTM 模型的集成学习矿压预测 3 个部分, 如图 1 所示。

1.1 矿压数据预处理

现代化煤矿企业在井下施工时通常会配备监测系统, 可实时记录开采环境及设备状态参数。但当监测设备发生系统故障或受到其他因素干扰时, 监测数据可能会受到显著的扰动, 产生与真实值存在

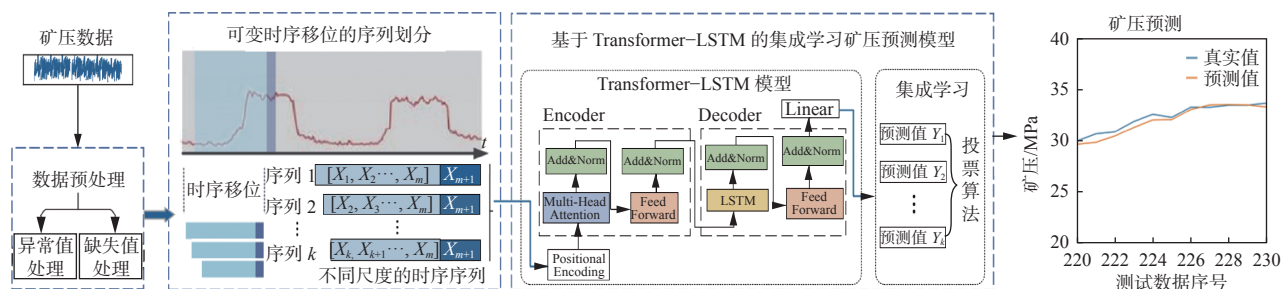


图 1 基于可变时序移位 Transformer-LSTM 的集成学习矿压预测方法

Fig. 1 Ensemble learning mine pressure prediction method based on variable time series shift Transformer-LSTM

较大差异的异常数据,影响监测数据后续应用。因此,需对监测数据进行针对性的预处理和数据清洗。采用拉依达准则^[17]检测矿压数据中的异常值,其对于大量的矿压监测数据具有显著效果。该方法通过计算数据的标准偏差,并根据一定的概率设置范围实现异常数据检测。超过该范围的误差称为严重误差,达到这种严重误差的数据即为异常数据。异常值检测公式为

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (1)$$

$$|x_i - \bar{x}| > 3\sigma \quad (2)$$

式中: σ 为标准偏差; n 为矿压数据总数; x_i 为第 i 个矿压数据; \bar{x} 为矿压数据的算术平均值。

煤矿井下监测设备易受到干扰,引起系统操作故障,甚至使传感器脱离系统,导致矿压监测数据缺失。采用拉格朗日插值方法^[18]插值缺失的矿压数据,从而减小误差,提高数据的完整性和准确性。

在此基础上,考虑矿压为多变量数据,存在不同量纲和数值的差异,直接应用会对模型训练和测试造成影响。因此,采用数据归一化方法(式(3))将矿压样本数据映射到 $[0, 1]$ 区间,保证不同变量之间的数值范围相同,减小对于数据的偏好和权重分配,保证模型预测的精度和稳定性。

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (3)$$

式中: x_i^* 为第 i 个矿压数据的归一化值; x_{\min} 为矿压数据最小值; x_{\max} 为矿压数据最大值。

1.2 基于可变时序移位的矿压数据划分

考虑矿压数据通常包含复杂的时间序列模式和周期性变化,提出一种可变时序移位策略,避免固定长度时序序列可能存在的数据偏移问题。

假设综采工作面某液压支架监测数据含 n 个按时间顺序排列的数据,时间序列为 $P = \{X_1, X_2, \dots, X_n\}$, X_m 为该支架在 m 时刻的矿压监测值, $m=1, 2, \dots$,

n 。若以 m 时刻为当前时刻,则矿压数据 P 前 t 时刻的时间序列可表示为 $P(m) = \{X_{m-t+1}, X_{m-t+2}, \dots, X_{m-1}, X_m\}$ 。为了划分可变时序移位的矿压数据,以 m 时刻监测数据为基础,通过移动非固定长度的窗口,获取不同尺度的时序序列:

$$\begin{cases} y_1 = \{X_1, X_2, \dots, X_m\} \\ y_2 = \{X_2, X_3, \dots, X_m\} \\ \vdots \\ y_k = \{X_k, X_{k+1}, \dots, X_m\} \end{cases} \quad (4)$$

式中 y_k 为尺度 k 的时序序列。

进行矿压预测时,将不同尺度的时序序列依次注入基于 Transformer-LSTM 的集成学习矿压预测模型。

1.3 基于 Transformer-LSTM 的集成学习矿压预测模型

Transformer 模型是一种基于自注意力机制的序列建模模型,由编码块和解码块组成,每个块都由多个独立的编码层(Encoder)或解码层(Decoder)堆叠构成^[19]。每个编码层包含多头注意力层(Multi-Head Attention)、全连接层(FFN)和正则化层(Add & Norm),而每个解码层则包含 2 个多头注意力层。Transformer 模型采用位置编码(Positional Encoding)策略获取输入序列中的相对位置信息,通过多头注意力机制实现在不同子空间上关注不同的细节信息。

LSTM 模型是一种循环神经网络(Recurrent Neural Network, RNN)的变体,专门设计用于处理和建模具有时序特征的序列数据^[20]。相较于其他类型的 RNN 模型, LSTM 模型能够更好地捕捉时序数据中的长期依赖关系,保证模型的准确性,目前已广泛应用于矿压预测。

为了更好地捕捉时序数据中潜在的矿压变化规律,基于传统 Transformer 模型结构,将解码器修改为全连接层,并在解码器中引入 1 个 LSTM 模块,构建了 Transformer-LSTM 模型,如图 2 所示。

在此基础上,将不同尺度的矿压数据分别注入 Transformer 模型的多头注意力机制中,然后计算不

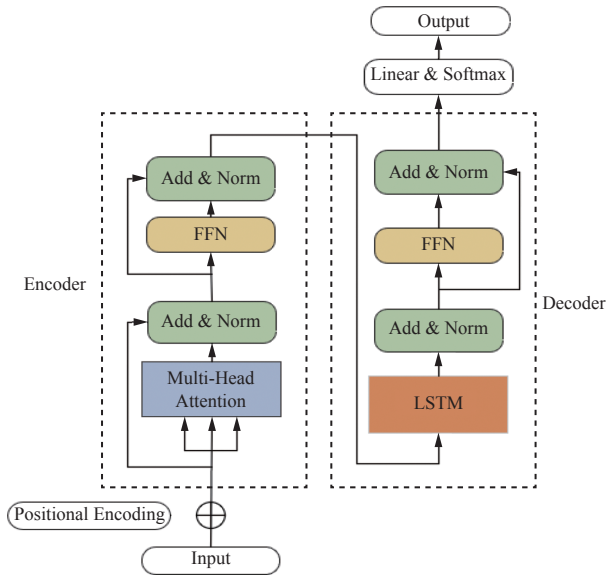


图2 Transformer-LSTM 模型结构

Fig. 2 Transformer-LSTM model structure

同时序序列的特征权重。这些权重用于局部注意力机制,以便模型在每个时间点仅关注该尺度内时序数据的局部矿压变化特征,最终捕捉到不同尺度下的多层次矿压变化特征。在上述过程中,通过采用相对位置编码,使 Transformer 模型能够对时序数据进行编码。由此,可将不同尺度的时序数据特征通过 Transformer 模型输出,并利用拼接方式实现矿压变化特征的向量表示,以便后续矿压预测。具体过程如下。

1) 对于输入的时序数据 X_m , 将其分解成不同尺度的时序序列片段, 每个片段在时间维度上关注某个局部时序序列特征, 由此得到时序序列矩阵 X'_m 。将 X'_m 注入 Transformer 模型中的多头注意力机制, 得到局部特征表示 A_m 。

$$A_m = \text{MultiHeadAttention}(X'_m) \quad (5)$$

2) 分别计算不同时序序列的特征权重 S_m , 用于计算局部自注意力分数 $M^{(m)}$ 。

$$S_m = \text{Softmax}(\text{ReLU}((W_q A_m)^T W_k A_m)) \quad (6)$$

$$M^{(m)} = \sum_{m=1}^n S_m W_v A_m \quad (7)$$

式中 W_q, W_k, W_v 为权重矩阵。

3) 将 Transformer 模型输出的各特征向量进行拼接, 构成不同尺度下矿压数据的特征向量表示 H 。

4) 将特征向量表示 H 注入 LSTM 模型, 进一步抽取时序特征。

$$Y'_m = \text{LSTM}(H, W_m, Y_m^{t-1}, \theta_m) \quad (8)$$

式中: Y'_m 为 t 时刻的隐藏状态; $\text{LSTM}(\cdot)$ 为 LSTM 模型的时序特征提取函数; W_m 为 LSTM 模型训练过程中学习的参数矩阵; θ_m 为 LSTM 模型的超参数。

通过上述过程, 针对不同尺度的矿压序列可得预测值 p'_1, p'_2, \dots, p'_m 。引入集成学习方式, 通过投票算法计算出最终矿压预测结果:

$$p' = \frac{1}{m} \sum_{j=1}^m p'_j \quad (9)$$

采用集成学习的投票算法可克服单一预测模型自身的局限性, 提高矿压预测的准确性和稳定性。

2 实验及结果分析

2.1 数据集

采用某煤矿 2 个不同位置的综采工作面上液压支架压力采集器监测数据进行实验。2 个综采工作面均布置 121 台液压支架。从 1 号液压支架开始, 每间隔 5 台液压支架装设 1 台压力采集器, 监测液压支架左右柱压力, 并以此作为 1 个矿压测站。1 个工作面上共设置 25 个测站, 局部如图 3 所示。

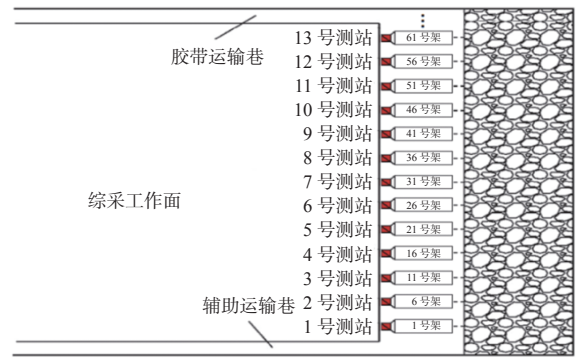


图3 综采工作面矿压测站局部布置

Fig. 3 Local layout of mine pressure measuring stations in fully mechanized working face

通过分析综采工作面各液压支架压力实测数据发现, 工作面两侧液压支架压力监测数据明显小于中间数据。这与文献[21]中“顶板岩层‘O-X’型破断规律对液压支架工作阻力的影响”相符, 即工作面中部液压支架的工作阻力明显比工作面上部和下部液压支架的工作阻力大。因此, 分别将 2 个综采工作面中部 13 号测站的监测数据作为时序矿压数据预测的研究对象, 构建数据集 1 和数据集 2。对矿压数据进行预处理之后, 数据集 1 共有 15 640 条矿压数据, 数据集 2 共有 20 360 条矿压数据。基于上述 2 个数据集, 验证本文方法的预测效果。

部分矿压监测数据见表 1。实测矿压数据由 6 个字段组成: 时间、支架编号、采集器编号、左柱压力、右柱压力、整架压力。其中, 时间表示压力采集器采集数据的时间点; 支架编号表示液压支架序号; 采集器编号表示该液压支架上压力采集器的编号; 左柱压力表示在同一时间下压力采集器所采集的液

压支架左柱油缸压力;右柱压力表示采集的液压支架右柱油缸压力;整架压力表示在当前时间点该液压支架所受的工作阻力,可根据左右柱压力来计算,即

$$p_{\text{all}} = \frac{p_l + p_r}{2} \quad (10)$$

式中 p_{all} , p_l , p_r 分别为整架压力、左柱压力、右柱压力。

表 1 13 号测站部分矿压监测数据

Table 1 Part of measured mine pressure data of No.13 station

时间	支架编号	采集器编号	左柱压力/MPa	右柱压力/MPa	整架压力/MPa
2020-05-06T12:25:00	61	13	22.40	19.30	20.85
2020-05-06T12:30:00	61	13	22.10	18.60	20.35
2020-05-06T12:35:00	61	13	22.10	18.10	20.10
2020-05-06T12:40:00	61	13	21.90	17.20	19.55
2020-05-06T12:45:00	61	13	21.90	16.70	19.30

在每条矿压数据中, p_{all} , p_l , p_r 与检测时间相对应,因此均为时序矿压数据。整架压力表示液压支架在不同时刻承受的负荷,将其作为预测指标。

2.2 评价指标

选择平均绝对误差 (Mean Absolute Error, MAE) 和均方根误差 (Root Mean Square Error, RMSE) 对模型预测精度进行评价。

$$E_{\text{MAE}} = \frac{1}{n} \sum_{t=1}^n |p_t - \tilde{p}_t| \quad (11)$$

$$E_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n (p_t - \tilde{p}_t)^2} \quad (12)$$

式中: E_{MAE} , E_{RMSE} 分别为预测结果的 MAE 和 RMSE; p_t 为 t 时刻矿压真实值; \tilde{p}_t 为 t 时刻矿压预测值。

通常来说, MAE 和 RMSE 越小, 表示模型预测精度越高, 模型预测效果越好。

2.3 对比实验

对不同矿压序列下的预测指标进行可视化分析, 结果如图 4 所示。可看出针对可变尺度的矿压数据, 矿压预测结果的 MAE 和 RMSE 波动明显, 说明不同尺度的时序特征序列对于矿压预测的准确度有较大的敏感性。因此, 采用集成学习的投票算法对多个预测结果进行融合平均, 以降低不同尺度的特征序列因噪声等因素对矿压预测结果产生的干扰。基于该算法得到的矿压预测结果总体趋势与未采用该算法时近似, 但波动明显减小, 有效降低了不同尺度对矿压预测结果的敏感性影响, 保证矿压预测准确度。

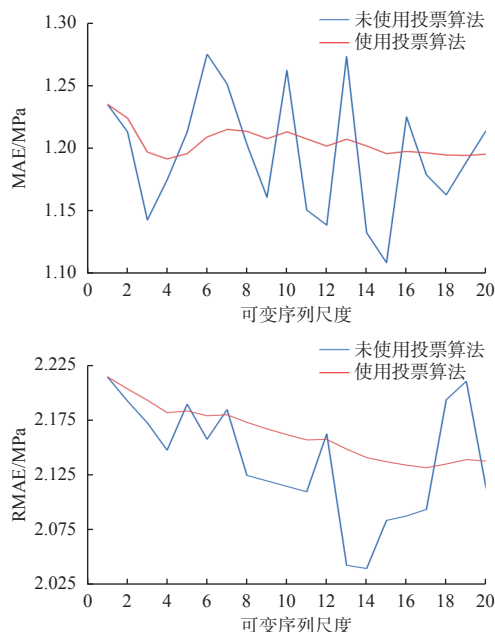


图 4 不同矿压序列的预测结果

Fig. 4 Prediction results of different mine pressure series

选择反向传播 (Back Propagation, BP) 神经网络^[13]、LSTM^[14]、Transformer^[22]作为对比模型, 与 Transformer-LSTM 模型的矿压预测结果进行对比。为了避免模型在训练过程中发生过拟合或欠拟合等情况, 根据数据集规模、数据特征等因素设置各模型参数, 使用网格搜索方法和交叉验证法进行参数调优^[23]。BP 神经网络模型和 LSTM 模型的初始化参数: epochs 为 100, batch_size 为 64, 选择 ReLU 激活函数, 以 Adam 优化算法为模型优化器。Transformer 模型和 Transformer-LSTM 模型的初始化参数: epochs 为 100, batch_size 为 64, n_heads 为 10, 选择 Gelu 激活函数。在 2 个数据集上进行模型训练和测试时, 每个数据集均选取 90% 数据作为训练集, 剩余 10% 作为测试集。

采用训练集训练各模型, 之后采用测试集数据对模型的预测性能进行评价。以区间在 200~400 的 200 个测试数据为例, 其预测结果如图 5 所示, 评价指标见表 2。

从图 5 可看出: BP 神经网络模型的预测值与真实值之间误差较大; Transformer 模型因具有自注意力机制, 可同时关注序列中的所有位置, 更好地捕捉全局信息, 使模型更加全面地理解序列, 对于时序矿压数据具有较好的预测精度; LSTM 模型具有较好的时间序列预测能力, 对于矿压预测的准确率较高; Transformer-LSTM 模型利用 Transformer 模型的自注意力机制来获取时序数据的潜在变化规律, 并利用 LSTM 模型实现长期时序数据依赖关系的特征, 使得预测值最接近真实值, 预测效果最优。

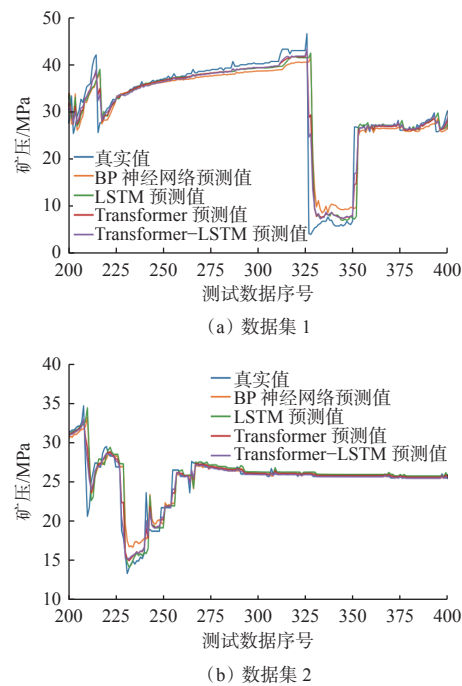


图 5 不同模型的部分矿压预测结果

Fig. 5 Part prediction results of mine pressure by different models

表 2 不同模型预测结果的评价指标

Table 2 Evaluation index comparison of prediction results of different models

数据集	模型	MAE/MPa	RMSE/MPa
数据集1	BP神经网络	2.214	3.756
	LSTM	1.847	3.632
	Transformer	1.313	2.447
	Transformer-LSTM	1.196	2.136
数据集2	BP神经网络	1.333	2.755
	LSTM	1.205	2.764
	Transformer	0.820	1.810
	Transformer-LSTM	0.742	1.512

从表 2 可看出：Transformer-LSTM 模型在数据集 1 上预测结果的 MAE 和 RMSE 分别为 1.196, 2.136 MPa, 较 Transformer 模型分别减小 8.9%, 12.7%, 在数据集 2 上的 MAE 和 RMSE 分别为 0.742, 1.512 MPa, 较 Transformer 模型分别减小 9.5%, 16.5%; 与其他 3 种模型相比, Transformer-LSTM 模型矿压预测结果的 MAE 和 RMSE 最低, 表明该模型预测精度最高。

3 结论

1) 构建可变时序移位策略对矿压时序数据进行不同尺度的划分, 避免了固定长度时序序列可能存在的数据偏移问题, 可适应不同时间段内矿压数据的变化速度、幅度和模式, 为准确捕捉矿压变化规律

特征提供数据支持。

2) 通过融合 Transformer 和 LSTM 模型, 学习矿压时序数据中的时序特征和前后矿压数据之间的依赖关系, 解决了单一矿压预测模型较难准确刻画时序序列复合特征的问题, 提高了矿压预测模型的准确度。

3) 采用现场数据建立了 2 个数据集, 验证基于可变时序移位 Transformer-LSTM 模型的矿压预测性能。2 个数据集预测结果的 MAE 分别较 Transformer 模型提高了 8.9% 和 9.5%, RMSE 分别提高了 12.7% 和 16.5%, 有效提升了矿压预测精度。

4) 未来将重点研究矿压的分阶段预测模型, 并构建高效的预测机制, 实现在准确预测矿压的同时, 实时进行矿压灾害预警。

参考文献(References):

[1] 钱鸣高, 许家林, 王家臣. 再论煤炭的科学开采[J]. 煤炭学报, 2018, 43(1): 1-13.
QIAN Minggao, XU Jialin, WANG Jiachen. Further on the sustainable mining of coal[J]. Journal of China Coal Society, 2018, 43(1): 1-13.

[2] 冀汶莉, 刘艺欣, 柴敬, 等. 基于随机森林的矿压预测方法[J]. 采矿与岩层控制工程学报, 2021, 3(3): 71-81.
JI Wenli, LIU Yixin, CHAI Jing, et al. Mine pressure prediction method based on random forest[J]. Journal of Mining and Strata Control Engineering, 2021, 3(3): 71-81.

[3] 曹富荣, 吴学松, 传金平, 等. 冲击危险区域震动波 CT 反演探测技术应用[J]. 煤炭科技, 2021, 42(2): 67-71, 76.
CAO Furong, WU Xuesong, CHUAN Jinping, et al. Application of CT inversion detection technology for shock wave in hazardous area of rock burst[J]. Coal Science & Technology Magazine, 2021, 42(2): 67-71, 76.

[4] 常峰. 基于 GA-BP 神经网络的工作面顶板矿压预测模型应用研究[D]. 徐州: 中国矿业大学, 2019.
CHANG Feng. Application research of the prediction model for the coal working face roof pressure based on GA-BP neural networks[D]. Xuzhou: China University of Mining and Technology, 2019.

[5] FENG Xiating, WANG Yongjia, YAO Jianguo. A neural network model for real-time roof pressure prediction in coal mines[J]. International Journal of Rock Mechanics and Mining Science & Geomechanics Abstracts, 1996, 33(6): 647-653.

[6] 柴敬, 王润沛, 杜文刚, 等. 基于 XGBoost 的光纤监测矿压时序预测研究[J]. 采矿与岩层控制工程学报, 2020, 2(4): 64-71.
CHAI Jing, WANG Runpei, DU Wengang, et al. Study

- on time series prediction of rock pressure by XGBoost in optical fiber monitoring[J]. *Journal of Mining and Strata Control Engineering*, 2020, 2(4): 64-71.
- [7] 程敬义, 万志军, PENG Syd S, 等. 基于海量矿压监测数据的采场支架与顶板状态智能感知技术[J]. *煤炭学报*, 2020, 45(6): 2090-2103.
CHENG Jingyi, WAN Zhijun, PENG Syd S, et al. Technology of intelligent sensing of longwall shield supports status and roof strata based on massive shield pressure monitoring data[J]. *Journal of China Coal Society*, 2020, 45(6): 2090-2103.
- [8] 徐刚, 范志忠, 张春会, 等. 宏观顶板活动支架增阻类型与预测模型[J]. *煤炭学报*, 2021, 46(11): 3397-3407.
XU Gang, FAN Zhizhong, ZHANG Chunhui, et al. Support increasing resistance forms of macroscopic roof support and its prediction model[J]. *Journal of China Coal Society*, 2021, 46(11): 3397-3407.
- [9] 惠艳荔. 电磁辐射技术在采煤工作面矿压预测中的应用[J]. *煤炭技术*, 2013, 32(12): 73-75.
HUI Yanli. Application of electromagnetic radiation technology in mine pressure observation at coal face[J]. *Coal Technology*, 2013, 32(12): 73-75.
- [10] 李静, 何江, 巩思园, 等. 上覆遗留煤柱作用下冲击矿压预测预警案例研究[J]. *煤炭学报*, 2016, 41(增刊 2): 305-310.
LI Jing, HE Jiang, GONG Siyuan, et al. Case study on prediction and warning of rock burst under overlying pillar[J]. *Journal of China Coal Society*, 2016, 41(S2): 305-310.
- [11] LI Ye, SHI Xiaohu. Mine pressure prediction study based on fuzzy cognitive maps[J]. *International Journal of Computational Intelligence and Applications*, 2020, 19(3): 1-13.
- [12] 霍丙杰, 荆雪冬, 于斌, 等. 坚硬顶板厚煤层采场来压强度分级预测方法研究[J]. *岩石力学与工程学报*, 2019, 38(9): 1828-1835.
HUO Bingjie, JING Xuedong, YU Bin, et al. A classification forecasting method for the weighting intensity of stopes of thick coal seams with hard roofs[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2019, 38(9): 1828-1835.
- [13] 赵铭生, 刘守强, 纪润清, 等. 基于遗传算法优化BP神经网络的华北型煤田矿压破坏带深度预测[J]. *矿业研究与开发*, 2020, 40(6): 89-93.
ZHAO Mingsheng, LIU Shouqiang, JI Runqing, et al. Depth prediction of mining pressure failure zone in north China coalfield based on BP neural network optimized by genetic algorithm[J]. *Mining Research and Development*, 2020, 40(6): 89-93.
- [14] 赵毅鑫, 杨志良, 马斌杰, 等. 基于深度学习的大采高工作面矿压预测分析及模型泛化[J]. *煤炭学报*, 2020, 45(1): 54-65.
ZHAO Yixin, YANG Zhiliang, MA Binjie, et al. Deep learning prediction and model generalization of ground pressure for deep longwall face with large mining height[J]. *Journal of China Coal Society*, 2020, 45(1): 54-65.
- [15] 曾庆田, 吕珍珍, 石永奎, 等. 基于Prophet+LSTM模型的煤矿井下工作面矿压预测研究[J]. *煤炭科学技术*, 2021, 49(7): 16-23.
ZENG Qingtian, LYU Zhenzhen, SHI Yongkui, et al. Research on prediction of underground coal mining face pressure based on Prophet+LSTM model[J]. *Coal Science and Technology*, 2021, 49(7): 16-23.
- [16] 李泽萌. 基于LSTM的采煤工作面矿压预测方法研究[D]. 西安: 西安科技大学, 2020.
LI Zemeng. Research on prediction method of mining pressure in coal face based on LSTM[D]. Xi'an: Xi'an University of Science and Technology, 2020.
- [17] SHEN Cheng, BAO Xuejing, TAN Jiubin, et al. Two noise-robust axial scanning multi-image phase retrieval algorithms based on Pauta criterion and smoothness constraint[J]. *Optics Express*, 2017, 25(14): 16235-16249.
- [18] HIGHAM N J. The numerical stability of barycentric Lagrange interpolation[J]. *IMA Journal of Numerical Analysis*, 2004, 24(4): 547-556.
- [19] HAN Kai, XIAO An, WU Enhua, et al. Transformer in transformer[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 15908-15919.
- [20] 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.
- [21] 苏振国, 邓志刚, 李国营, 等. 顶板深孔爆破防治小煤柱冲击地压研究[J]. *矿业安全与环保*, 2019, 46(4): 21-25, 29.
SU Zhenguo, DENG Zhigang, LI Guoying, et al. Study on prevention and control of small coal pillar rock burst by roof deep hole blasting[J]. *Mining Safety & Environmental Protection*, 2019, 46(4): 21-25, 29.
- [22] 田原, 庞骁, 赵文祎, 等. 基于Transformer的滑坡短期位移预测模型[J]. *北京大学学报(自然科学版)*, 2023, 59(2): 197-210.
TIAN Yuan, PANG Xiao, ZHAO Wenyi, et al. A transformer-based model for short-term landslide displacement prediction[J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2023, 59(2): 197-210.
- [23] 文冰梅, 赵联文, 黄磊. AIC准则与留一法交叉验证渐近等价的证明[J]. *统计与决策*, 2022, 38(6): 40-43.
WEN Bingmei, ZHAO Lianwen, HUANG Lei. Proof of the asymptotic equivalence between AIC criterion and LOOCV[J]. *Statistics & Decision*, 2022, 38(6): 40-43.