

基于深度可分离卷积混合网络模型的地浸采铀注液量预测研究

刘志锋^{1,2,3}, 唐俊贤^{1,3}, 林芝宁^{1,3}, 周义朋^{1,2,3}

(1. 东华理工大学 核资源与环境国家重点实验室, 江西 南昌 330013;

2. 铀资源探采与核遥感全国重点实验室, 江西 南昌 330013;

3. 东华理工大学, 江西 南昌 330013)

摘要: 地浸采铀作为铀矿的绿色开采技术, 在生产运行中产生海量数据, 利用这些海量数据进行大数据分析和趋势预测, 能够提升技术人员制定生产计划的可靠性。目前采用的基于编码器-解码器结构的时序预测模型, 由于存在注意力机制, 导致计算复杂、内存消耗大。本研究提出深度可分离卷积混合模型, 通过动态序列分割模块降低固定分割带来的语义破坏, 通过深度可分离卷积混合模块降低模型运行时间并捕获局部和全局特征。结果表明, 深度可分离卷积混合网络模型的均方误差 (Mean Square Error, MSE) 与平均绝对误差 (Mean Absolute Error, MAE) 相较于时间序列分块自注意力模型 (Patch Time Series Transformer, PatchTST) 分别降低了 1.04% 和 4.13%, 提出的动态序列分割模块的 MSE 与 MAE 相较于原有模型分别降低了 7.32% 和 5.03%; 在性能对比分析上, 深度可分离卷积混合模型的训练速度相较于趋势季节分解线性模型 (Decomposition Linear, DLinear) 提高了 59.91%。建立的模型能够准确预测采区生产运行中硫酸注液量的变化趋势, 改善了现有预测模型针对地浸铀矿数据集存在的运行时间长、运行内存大、数据拟合差的问题, 可为地浸铀矿生产决策提供理论和实践参考。

关键词: 地浸采铀; 注液量预测; 深度可分离卷积; 预测模型

中图分类号: P631; TP39; TL212 **文献标志码:** A **文章编号:** 1000-8063(2025)01-0009-09

DOI: 10.13426/j.cnki.yky.2024.10.12

随着地浸采铀技术的快速发展, 铀矿地质勘查信息化建设逐步完善, 对地浸过程中产生的生产、化验分析等数据与数字化、信息化、人工智能算法进行的融合, 大幅提高了矿业技术人员工作效率^[1-3]。目前, 随着数字化矿山建设的深入以及信息化技术的发展, 生产数据得到有效收集; 但对收集数据的分析利用不足, 往往只对部分数据进行了人工分析^[4]。面对收集的海量数据, 由于缺乏有效的数据建模手段, 技术人员难以挖掘数据之间的内在规律, 也无法充分利用现有数据来制定未来的生产计划。

在地浸采铀中, 注液量受多种因素影响, 包括

矿石的矿物组成、矿石的粒度、浸出剂的性质, 以及含矿含水层厚度等。注液量与地浸铀矿山生产效率关系密切, 注液量过多可能导致资源浪费和环境污染, 而注液量不足则会影响铀的浸出率^[5]。为提高铀浸出率, 需要准确预测注液量的未来变化趋势, 进而合理安排铀矿资源的开采, 降低矿业公司开采成本。

近年来, 人工智能模型主要用于处理复杂非线性关系, 在铀矿领域已有诸多成果。Kuchin Y 等人利用地浸铀矿山勘探阶段的水文地质研究数据与测井数据, 并使用基于支持向量算法、梯度提升树等的回归模型, 对其母岩的过滤特性进行了

收稿日期: 2024-10-29

基金项目: 中国铀业有限公司—东华理工大学核资源与环境国家重点实验室联合创新基金(2022NRE-LH-14); 国家国防科技工业局核能开发项目“铀裂变瞬发 n-γ 融合测井及航空监测关键技术研究”; 江西省自然科学基金(20242BAB25084)。

第一作者简介: 刘志锋(1979—), 男, 内蒙古赤峰人, 博士, 副教授, 主要研究方向为大数据分析可视化。

预测^[6]。贾皓等人使用人工神经网络对某矿床采区日金属量进行了预测^[7]。余东原等人利用地浸铀矿生产数据,采用机器学习方法,筛选出多层感知机(Multilayer Perceptron, MLP)和随机森林(Random Forest, RF)模型来预测铀浸出金属量^[8]。张宇等人使用多目标优化算法为“水平井-直井”抽注系统提供经济高效的抽注液量决策方案选择^[9]。近年来,已有部分学者采用深度学习对铀矿领域进行相关分析,如贾明滔等人基于碱法地浸单元数据,提出了集成多种深度学习方法来有效预测浸采单元铀浓度^[10];周渊凯等人采用深度学习,对纳岭沟地区测井解析岩性获得90%以上的分类精度^[11]。然而,目前的研究主要是通过机器学习、数理统计等方法对相关酸法地浸铀矿数据进行预测处理,探寻特征间的关系;但无法对未来趋势做出准确判断,没有把时间作为协变量进行相关时序序列的预测分析。

为了解决上述问题,针对地浸铀矿生产数据集的自身特点,以深度可分离卷积神经网络(Depthwise Separable Convolution, DS-Conv)为基础时序预测模型,提出一种深度可分离卷积混合网络模型(Depthwise Separable Convolution Mixer Net, DS-Conv-Mixer-Net),并用其进行研究。

1 研究基础

1.1 数据获取与处理

本地浸铀矿数据集A1来源于某矿业公司采区的生产运行数据。采区生产运行数据参数包括日期、浸出液、浸出剂、液固比、浸出率;自2017年7月至2023年8月,记录时长为6年,采样时间间隔为1天。

为了提高模型预测能力以及训练速度,减少冗余信息的干扰,结合专家知识筛选出合适的地浸铀矿数据集,并构成相应时序序列,即 $d = \{d_{cjs}, d_{zjs}, d_{zH_2SO_4}, d_{dqu}, d_{zyl}\}$,其中 d_{cjs} 、 d_{zjs} 、 $d_{zH_2SO_4}$ 、 d_{dqu} 、 d_{zyl} 分别表示 t 时刻历史抽液金属量、浸出剂金属量、注液硫酸浓度、当前矿山剩余金属量、注液量。以注液量为输出预测数据,记为 x_i ;同时,根据滑动窗口形式生成每个时间序列实例,记为 i (i 代表第 i 个实例)。由于本研究将时间作为地浸铀矿数据集中的协变量;而这些时间序列数据

在实际应用中存在不稳定性,测试数据还可能面临分布漂移问题,这会影响模型的预测性能和准确性^[12-13]。为提高预测准确性通过用零均值和单位标准差在序列分割之前对每个 x_i 进行归一化,并且在结果输出后将平均值和偏差添加到输出上,然后经过全连接预测层将结果输出。模型流程见图1。

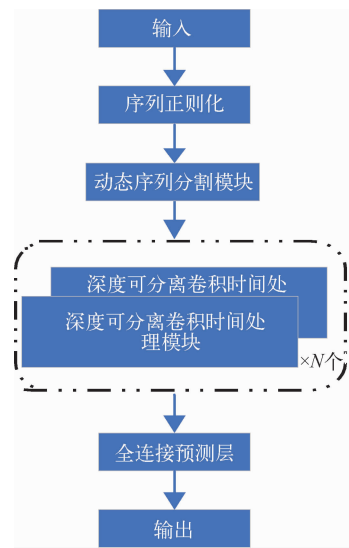


图1 深度可分离卷积混合网络模型架构
Fig. 1 Architecture of depthwise separable convolutional mixer net model

1.2 动态序列分割模块

动态序列分割模块(Dynamic Sequence Segmentation Module, DSSM)的主要目标是辨别不同时间步长的数据点之间的相关性。在大多数现有时序预测研究中,保留原始逐点时间序列数据作为输入几乎是默认做法。然而,单个时间步骤所带来的语义不能像句子中的单词那么多,因此提高每个输入单元的信息密度对于提高最终结果的准确性至关重要。近年来,基于迁移(Transformer)结构的时序预测方法将相邻时间点聚合成子序列^[14],然后将每个子序列分组到局部区域中,进而耦合它们的依赖性。但是固定大小的序列分割有可能破坏序列之间的语义性,导致性能下降。因此,本研究根据Dai、Chen等人的思想^[15-16],建立动态序列分割模块(图2)。

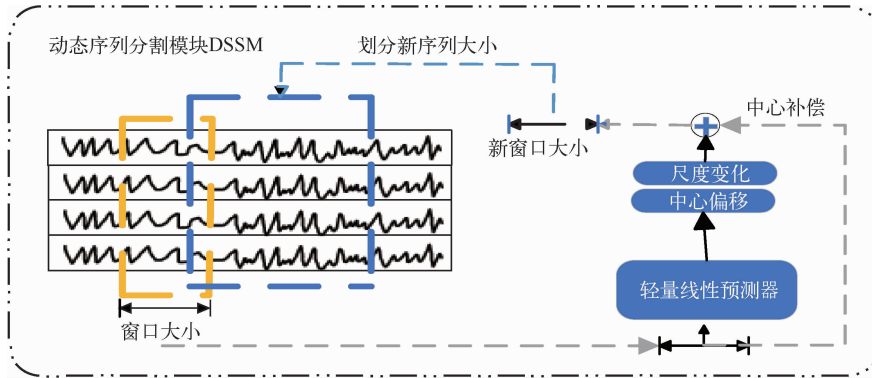


图 2 DSSM 模块框架

Fig. 2 DSSM modular framework

对于 DSSM 模块来说,时序序列的输入为 $X \in R^{C \times T}$,其中 C 和 T 分别表示变量(通道)的数量和序列长度。首先,在进行 DSSM 模块之前,序列数据固定分割大小为 $N = (T - P)/S + 2$,其中每个子序列(子序列简称 patch)的形状用 $X \in R^{P \times C}$ 表示, N 代表 patch 的数量, P 为每个相应 patch 的大小, T 表示输出通道数量, S 表示每个 patch 窗口滑动大小。同时,为了保证均匀分割序列,在 patch 分割之前将最后一个 $X \in R^{P \times C}$ 使用 S 个重复数填充到原始序列末尾。其次,每个 patch 设定 3 个关键控制变量,即中心位置 x_c 、中心偏移 δ_c 和尺度变化 δ_p 。 x_c 由 patch 分割决定其大小, δ_c 和 δ_p 通过式(1)~式(2)计算得到。

$$\delta_c = \text{Gelu}[\mathbf{W}_{\text{off}} \cdot f(x_{\text{in}})], \quad (1)$$

$$\delta_p = \text{Relu}[\mathbf{W}_{\text{dss}} \cdot f(x_{\text{in}})], \quad (2)$$

式中: $f(x_{\text{in}})$ 为轻量预测线性函数;Gelu、Relu 为激活函数; \mathbf{W}_{off} 、 \mathbf{W}_{dss} 为权重矩阵,分别用于学习中心偏移和尺度变化。基于以上中心位置 x_c 、中心偏移 δ_c 和尺度变化 δ_p 结果,可以计算出新的 patch 序列的中心位置 x_c^{i+1} 和长度 p^{i+1} ,以及左边界位置 L 和右边界位置 R 。

$$x_c^{i+1} = x_c + \delta_c, p^{i+1} = P + \lambda \cdot \delta_p; \quad (3)$$

$$L = x_c^i - \frac{p^i}{2}, R = x_c^{i+1} + \frac{p^{i+1}}{2}; \quad (4)$$

式中: λ 为缩放因子; P 为初始化 patch 的大小; x_c^i 为当前 patch 的中心位置; p^i 为当前 patch 的长度大小; x_c^{i+1} 为下一个 patch 的中心位置; p^{i+1}

为下一个 patch 的长度大小; L 为左边界位置; R 为右边界位置。根据这些控制变量,DSSM 模块能够自适应调整每个 patch 的位置和尺度,减少固定大小的序列分割可能引起的语义信息丢失。最后,通过轻量级线性预测函数将新的 patch 变换到嵌入空间中,从而使得模型能够更加灵活地捕捉时间序列数据中的局部模式和长期依赖关系。

1.3 深度可分离卷积混合模块

为了更有效地处理 DSSM 生成的自适应 patch,采用深度可分离卷积混合模块(Conv-Mixer-Net)对多个分支的时间序列数据进行建模。该网络首先使用分组卷积,即分组卷积中分组数等于 DSSM 模块中 patch 长度。同时,对于深度卷积层(Depthwise Convolution)来说, N 个 patch 同时采用相同的卷积核大小(均为 patch 的长度 P),分组数量等于 patch 的数量 N ;而对于逐点卷积层(Pointwise Convolution)来说, N 个 patch 同时采用相同的卷积核大小(均为 1),分组数量等于 1。这种设计通过将输入数据划为多个“patch”,并进行混合捕捉全局上下文和局部特征,使得每个 patch 在深度卷积和逐点卷积上通过一个专用的卷积核进行处理,从而增强模型对不同时间序列特征的捕捉能力。

同时,为了扩展模型的感受野并捕捉时间序列数据中的长期依赖关系,本研究参考了时间序列分块自注意力模型(Patch Time Series Trans-

former, PatchTST)研究中的策略,将卷积操作的步长设置为 S ,针对步长 S 设置卷积核大小为 8。在这一步骤中,输入特征图中的每个 patch 都独立地与一个卷积核进行卷积操作。这种操作生成了 N 个特征图,每个特征图代表了对应 patch 的卷积结果。随后,这些特征图按照生成的顺序进行连接,形成一个具有 N 个通道的输出特征图。深度卷积网络通过使用组卷积核,对共享相同空间位置的块应用相同的卷积操作,这有助于模型捕捉时间序列中 patch 间的潜在周期性模式。

深度卷积混合模块使用卷积核为 $k = P$ 的分

组卷积操作,通过对输入的通道(input channels)进行卷积处理。这种操作将生成与输入通道数量相同的特征图(feature maps),每个特征图都对应一个特定的输入通道;逐点卷积使用 1×1 的卷积核将深度卷积得到的特征图进行组合,生成已经定义的 N 个特征图,具体卷积操作见图 3。上述模块同时结合激活函数 Gelu 以及残差网络来弥补梯度消失、简化网络结构,Conv-Mixer-Net 实现跨 patch 通道混合信息,生成最终输出特征图,然后通过全连接层将预测结果输出。Conv-Mixer-Net 模块具体流程见图 4。

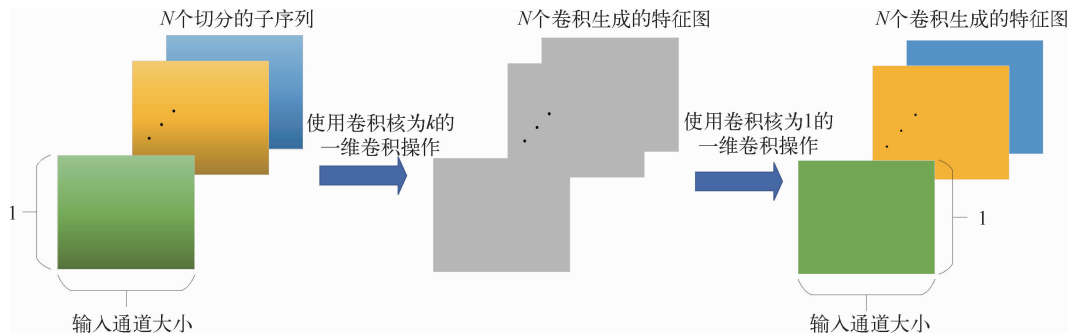


图 3 深度可分离卷积示意图

Fig. 3 Schematic diagram of depthwise separable convolution

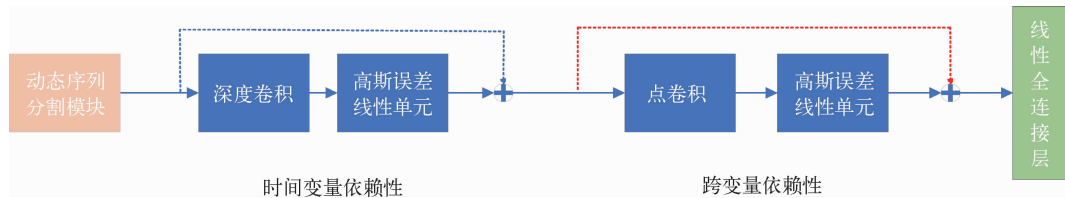


图 4 Conv-Mixer-Net 模块框架

Fig. 4 Conv-Mixer-Net modular framework

2 结果与分析

2.1 实验设置

网络训练优化均以 Tesla V100 服务器为硬件平台,运行内存为 16 GB,采用 PyTorch 深度学习框架构建 DS-Conv-Mixer-Net 模型。模型训练采用损失函数和 Adam 优化器,数据集划分比例为 7:2:1,分别对应训练集、测试集和验证集。模型输入时间序列长度设置为 $seq_len = 96$,用于预测未来 24、36、48、60 天的生产数据,分别对

应不同的预测步长 $pre_len = \{24, 36, 48, 60\}$ 。迭代次数设定为 30 次,学习率设定为 0.000 1。在模型的 DSSM 模块中,patch 长度默认为 16,步长默认为 8。Conv-Mixer-Net 的层数默认设置为 $N=1$ 。

模型性能评价指标以平均绝对误差 (Mean Absolute Error, MAE) 和均方误差 (Mean Square Error, MSE) 为准,误差指标计算公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (5)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

式中: y_i 为第 i 个观测值的实际输出; \hat{y}_i 为模型预测的输出; n 为观测值的总数。

2.2 实验分析

2.2.1 评价指标结果分析

对当前时间序列预测领域的多个先进模型进行了比较分析。这些模型包括基于 Transformer 架构的最新研究成果,例如 PatchTST^[14]、Crossformer^[17] 以及 Informer^[18],以及基于线性网络结构的 TiDE^[19]和 DLinear^[20]模型。为了确保实验的一致性,所有参与比较的模型在实验中均采

用统一的时间步长设置,其中序列长度为 $seq_len = 96$,预测步长 pre_len 则分别设定为 24、36、48 和 60。

不同模型在相同实验环境下,通过 MSE 与 MAE 来评价模型预测性能,结果见表 1。从归一化后的数值来看,与表现最好的基于 Transformer 结构的 PatchTST 模型相比,本模型 (DS-Conv-Mixer-Net) 在 MSE 与 MAE 上分别降低了 1.04% 与 4.13%。与模型简单且表现较好的基于多层感知机结构的 TiDE 与 DLinear 模型相比,本模型在 MSE 上分别降低 10.80% 和 38.31%,在 MAE 上分别降低 16.69% 和 49.83%。

表 1 不同模型对比实验

Table 1 Comparison of experimental results by different models

| 预测 步长 | DS-Conv-Mixer-Net | | PatchTST | | TiDE | | DLinear | | Crossformer | | Informer | |
|----------|-------------------|----------------|----------------|----------------|---------|---------|---------|---------|-------------|---------|----------|---------|
| | MSE | MAE | MSE | MAE | MSE | MAE | MSE | MAE | MSE | MAE | MSE | MAE |
| 24 | 0.016 2 | 0.053 0 | <u>0.016 5</u> | <u>0.055 9</u> | 0.019 5 | 0.067 5 | 0.020 5 | 0.088 7 | 0.038 3 | 0.148 2 | 0.160 6 | 0.286 4 |
| 36 | 0.018 3 | 0.058 7 | 0.018 3 | <u>0.060 7</u> | 0.020 4 | 0.070 3 | 0.029 1 | 0.113 5 | 0.055 5 | 0.182 4 | 0.153 2 | 0.263 0 |
| 48 | 0.020 4 | 0.063 9 | <u>0.020 5</u> | <u>0.066 3</u> | 0.022 2 | 0.074 6 | 0.034 8 | 0.132 8 | 0.064 5 | 0.181 4 | 0.183 6 | 0.336 1 |
| 60 | 0.020 9 | 0.065 9 | <u>0.021 3</u> | <u>0.069 1</u> | 0.022 9 | 0.077 6 | 0.038 7 | 0.146 7 | 0.066 0 | 0.195 2 | 0.260 8 | 0.374 5 |
| Avg | 0.019 0 | 0.060 4 | <u>0.019 2</u> | <u>0.063 0</u> | 0.021 3 | 0.072 5 | 0.030 8 | 0.120 4 | 0.056 1 | 0.176 8 | 0.189 6 | 0.315 0 |
| 性能/% | 0 | 0 | -1.04 | -4.13 | -10.80 | -16.69 | -38.31 | -49.83 | -66.13 | -65.84 | -89.98 | -80.83 |

注: Avg 表示所有时间步长下的平均评价指标;加粗数值表示对比其他模型最优的结果;下划线的数值表示对比其他模型次优的结果。性能栏中“MSE”“MAE”所在列的值表示 DS-Conv-Mixer-Net 模型相比于所在列模型降低百分比。

2.2.2 模型拟合能力评估

2022 年 6 月 26 日至 2023 年 6 月 26 日时间跨度内测试集的拟合效果见图 5。可以看出,在 $seq_len = 96$ 与 $pre_len = 60$ 的情况下,DS-Conv-Mixer-Net 与 PatchTST 模型与真实值拟合较好;但根据表 1 结果,通过计算得出,本模型的 MSE 与 MAE 较 PatchTST 模型分别下降了 1.88% 与 4.63%,优于 Patch 模型。同样, TiDE 与 DLinear 模型的 MSE 与 MAE 低于本模型且测试集拟合较差。Crossformer 和 Informer 模型由于自身注意力机制比较依赖于大型数据集等特点,其 MSE 和 MAE 值较大,导致与真实值拟合偏差较大。综上,本模型在 $seq_len = 96$ 与 $pre_len = 60$ 的测试集情况下拟合能力优于其他 5 种模型。

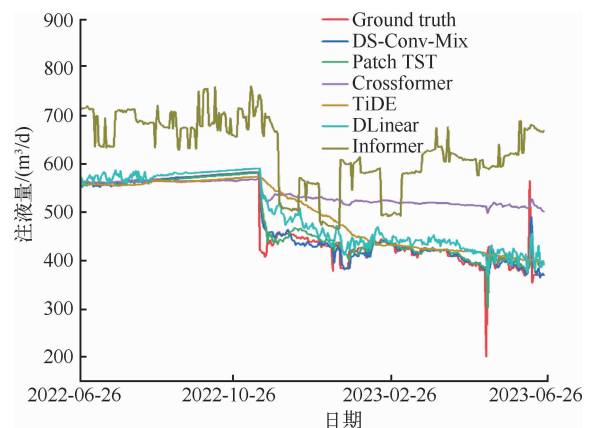


图 5 不同模型在 $seq_len = 96$ 与 $pre_len = 60$ 下的测试集表现

Fig. 5 Test set performance of different models with $seq_len = 96$ and $pre_len = 60$

2.2.3 模型性能分析

为了更好地突出 DS-Conv-Mixer-Net 模型的特点,对不同模型在相同实验环境下,设定 $seq_len = 96$ 、 $pre_len = 60$ 进行性能分析,结果见表 2。实验表明,在评价指标为 MSE 的情况下,DS-Conv-Mixer-Net 模型训练的总时长最小

且对应实验评价指标 MSE 最小,分别为 27.762 2 s 和 0.020 9;DLinear 模型的结果次之,训练的总时长和 MSE 分别为 44.395 s 和 0.038 7;DS-Conv-Mixer-Net 模型比次优模型(DLinear)的训练速度提高 59.91%。

表 2 不同模型在 $seq_len = 96$ 与 $pre_len = 60$ 下的性能分析

Table 2 Performance analysis of different models with $seq_len = 96$ and $pre_len = 60$

| 项目 | DS-Conv-Mixer-Net | | PatchTST | | TiDE | | DLinear | | Crossformer | | Informer | |
|-------------|-------------------|----------------|-----------|----------------|-----------|----------------|----------|---------|-------------|---------|----------|---------|
| | time/s | MSE | time/s | MSE | time/s | MSE | time/s | MSE | time/s | MSE | time/s | MSE |
| 预测步 长 60 | 27.762 2 | 0.020 9 | 131.179 0 | <u>0.021 3</u> | 182.766 7 | <u>0.022 9</u> | 44.395 0 | 0.038 7 | 79.725 8 | 0.066 0 | 93.331 0 | 0.260 8 |
| 性能/% | 0 | 0 | 372.51 | -1.88 | 558.33 | -8.73 | 59.91 | -45.99 | 187.17 | -68.33 | 236.18 | -91.99 |

注:1)加粗数值表示对比其他模型最优的结果;2)下划线的数值表示对比其他模型次优的结果;3)性能栏中,“time/s”所在列的值表示 DS-Conv-Mixer-Net 模型相比于所在列模型训练速度提高百分比,“MSE”所在列的值表示 DS-Conv-Mixer-Net 模型相比于所在列模型降低百分比。

2.2.4 消融实验

为了证明 DSSM 模块的有效性,进行了相关消融实验,结果见表 3。可以看出,在相同实验环境下,在研究的预测步长范围内,未使用 DSSM 的模型的平均 MSE 和 MAE 性能下降 7.32%和

5.03%。实验证明,相较于常规固定式分割 patch 来说,DSSM 可根据数据特性动态分割成“非固定式”patch,而动态变化的 patch 可增强模型对局部和全局上下文的理解,减少由于固定分割造成的语义信息丢失。

表 3 消融实验结果

Table 3 Results of ablation experiments

| 预测步长 | DS-Conv-Mixer-Net | | w/o DSSM | |
|---------------------|-------------------|---------|--------------|--------------|
| | MSE | MAE | MSE | MAE |
| 24 | 0.016 2 | 0.053 0 | 0.016 8 | 0.056 8 |
| 36 | 0.018 3 | 0.058 7 | 0.018 5 | 0.061 2 |
| 48 | 0.020 4 | 0.063 9 | 0.023 0 | 0.067 3 |
| 60 | 0.020 9 | 0.065 9 | 0.023 7 | 0.069 2 |
| Avg ^① | 0.019 0 | 0.060 4 | 0.020 5 | 0.063 6 |
| Avg ^② /% | 0 | 0 | <u>-7.32</u> | <u>-5.03</u> |

注:1)Avg^①表示所有时间步长下的平均评价指标;2)Avg^②表示所有时间步长下性能平均降低指标;3)w/o 表示在相同实验环境下,未使用 DSSM 模块;4)下划线数值表示 DS-Conv-Mixer-Net 模型相比于未使用 DSSM 模块的平均 MSE 与 MAE 下降百分比。

3 结论

收集某矿业公司采区生产运行数据,将时间作为协变量,同时结合多个影响地浸采铀注液量的特征参数,形成地浸采铀数据集,以深度卷积神

经网络为架构,改进常用的序列分割模块,构建了一种深度可分离卷积混合网络模型(DS-Conv-Mixer-Net),并用于预测注液量。该模型能够有效降低运行时间,提高数据的拟合能力。选取约 6 年生产运行数据,制作地浸采铀数据集,用于注

液量实验验证。

1) 针对地浸铀矿的生产运行数据,将时间作为重要因素纳入模型,并结合多个注液量影响因素,使用深度卷积与点卷积来处理局部特征与全局特征,提高了数据处理速度。同时,改进常规固定 patch 分割处理,提高了对采区注液量的预测精度和效率。

2) 对比实验显示,DS-Conv-Mixer-Net 模型在数据拟合和运行时间上优于其他 5 种现有预测模型,具备明显的预测精度和计算效率优势。消融实验显示,本研究提出的 DSSM 模块在一定程度上能够动态处理地浸采铀数据集,提高预测精度。

3) 尽管模型拟合程度良好,但现地浸矿山数据难以获取且数据集尚不完整,未来有望通过完整的数据集进行更有效地建模。在地浸采铀过程中,对控制流量等设备的调整可能导致数据异常,影响模型预测能力。因此,未来研究将考虑引入突变因素作为协变量,进一步提升模型的鲁棒性和适应能力。

参考文献:

- [1] 阳奕汉,苏学斌,陈梅芳,等. “数字建井”在砂岩型铀矿地浸开发中的探索与应用[J]. 中国矿业, 2022, 31(12): 64-69.
YANG Yihan, SU Xuebin, CHEN Meifang, et al. Exploration and application of "digital well construction" in in-situ leaching of sandstone uranium deposit[J]. China Mining Magazine, 2022, 31(12): 64-69(in Chinese).
- [2] 张明林,刘洋,吴建勇,等. 中国铀矿地质勘查信息化建设现状及“十四五”发展思路[J]. 世界核地质科学, 2021, 38(3): 287-294.
ZHANG Minglin, LIU Yang, WU Jianyong, et al. Construction situation of uranium exploration informatization and development idea for the 14th Five-Year Plan in China[J]. World Nuclear Geoscience, 2021, 38(3): 287-294(in Chinese).
- [3] 毛鑫磊. 信息化技术在内蒙古某地浸采铀矿山中的应用[J]. 化工管理, 2024, 694(6): 82-84.
MAO Xinlei. Application of information technology in a in-situ leaching of uranium mine in Inner Mongolia[J]. Chemical Engineering Management, 2024, 694(6): 82-84(in Chinese).
- [4] 袁志华,孙占学,周义朋,等. 绿色铀矿山研究进展[J]. 有色金属(冶炼部分), 2021(11): 85-95.
YUAN Zhihua, SUN Zhanxue, ZHOU Yipeng, et al. Research progress of green uranium mines[J]. Nonferrous Metals (Extractive Metallurgy), 2021(11): 85-95(in Chinese).
- [5] 王亮,王海峰,霍建党,等. 纳岭沟矿床地浸采铀现场试验浸出液铀浓度变化研究[C] // 中国核学会. 中国核科学技术进展报告(第三卷)——中国核学会 2013 年学术年会论文集: 第 2 册. 北京: 核工业北京化工冶金研究院, 2013: 7.
- [6] KUCHIN Y, MUKHAMEDIEV R, YUNICHEVA N, et al. Application of machine learning methods to assess filtration properties of host rocks of uranium deposits in Kazakhstan[J]. Applied Sciences, 2023, 13(19): 10958.
- [7] 贾皓,廖文胜,杜志明,等. 基于人工智能算法的地浸采区产量预测[J]. 铀矿冶, 2023, 42(3): 10-16.
JIA Hao, LIAO Wensheng, DU Zhiming, et al. Production prediction of in-situ leaching mining area based on artificial intelligence algorithm[J]. Uranium Mining and Metallurgy, 2023, 42(3): 10-16(in Chinese).
- [8] 余东原,罗跃,梁大业,等. 基于机器学习方法预测地浸过程中铀浸出金属量的变化[J]. 有色金属(冶炼部分), 2024(2): 92-98.
YU Dongyuan, LUO Yue, LIANG Daye, et al. Predicting the variation of uranium leaching metal content in ground-leaching process based on machine learning methods[J]. Nonferrous Metals (Extractive Metallurgy), 2024(2): 92-98(in Chinese).
- [9] 贾明滔,谭笑,苏学斌,等. 基于卷积和长短期记忆网络的地浸开采铀浓度预测研究[J]. 铀矿地质, 2024, 40(3): 578-586.
JIA Mingtao, TAN Xiao, SU Xuebin, et al. Uranium concentration prediction for in-situ leaching based on convolution and long short-term memory networks[J]. Uranium Geology, 2024, 40(3): 578-586(in Chinese).
- [10] 张宇,马亮,李召坤,等. 基于多目标优化算法的“水平井-直井”系统流量决策方法[J/OL]. 铀矿冶, 2024: 1-7 [2024-11-03]. <https://doi.org/10.13426/j.cnki.yky.2024.06.05>.
ZHANG Yu, MA Liang, LI Zhaokun, et al. Flow decision-making method for the "horizontal well-vertical well" system based on multi-objective optimization algorithm[J/OL]. Uranium Mining and Metallurgy, 2024: 1-7 [2024-11-03] (in Chinese).

- [11] 周渊凯,刘祐.基于深度学习方法的测井岩性识别研究[J].铀矿地质,2024,40(2):336-345.
ZHOU Yuankai, LIU Hu. Deep learning based lithology recognition of well logging data[J]. Uranium Geology, 2024, 40(2): 336-345(in Chinese).
- [12] ULYANOV D, VEDALDI A, LEMPITSKY V. Instance normalization: The missing ingredient for fast stylization[EB/OL]. (2016-07-27). <https://doi.org/10.48550/arXiv.1607.08022>.
- [13] KIM T, KIM J, TAE Y, et al. Reversible instance normalization for accurate time-series forecasting against distribution shift[C]. Kigali: International Conference on Learning Representations (ICLR), 2022.
- [14] NIE Y Q, NGUYEN N H, SINTHONG P, et al. A time series is worth 64 words: Long-term forecasting with transformers[C]. Kigali: International Conference on Learning Representations (ICLR), 2023.
- [15] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]. Venice: 2017 Institute of Electrical and Electronics Engineers (IEEE) International Conference on Computer Vision, 2017: 764-773.
- [16] CHEN Z Y, ZHU Y S, ZHAO C Y, et al. DPT: Deformable patch-based transformer for visual recognition[C]. Beijing: Association for Computing Machinery, 2021: 2899-2907.
- [17] WANG W X, CHEN W, QIU Q B, et al. Cross-former++: A versatile vision transformer hinging on cross-scale attention[J]. Institute of Electrical and Electronics Engineers (IEEE) Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(5): 3123-3136.
- [18] ZHOU H Y, ZHANG S H, PENG J Q, et al. In-former: Beyond efficient transformer for long sequence time-series forecasting[J]. Association for the Advancement of Artificial Intelligence (AAAI)-21 Technical Tracks, 2021, 35(12): 11106-11115.
- [19] DAS A, KONG W, LEACH A, et al. Long-term forecasting with TiDE: Time-series dense encoder[EB/OL]. arXiv preprint arXiv:2304.08424, 2023.
- [20] ZENG A L, CHEN M X, ZHANG L, et al. Are Transformers effective for time series forecasting[J]. Proceedings of the 37-th Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence, 2023, 37(9): 11121-11128.

Injection Volume Prediction for In-situ Leaching of Uranium Based on Depthwise Separable Convolution Mixer Network Model

LIU Zhifeng^{1,2,3}, TANG Junxian^{1,3}, LIN Zhining^{1,3}, ZHOU Yipeng^{1,2,3}

(1. State Key Laboratory of Nuclear Resources and Environment, East China University of Technology, Nanchang 330013, China; 2. National Key Laboratory of Uranium Resources Exploration-Mining and Nuclear Remote Sensing, Nanchang 330013, China; 3. East China University of Technology, Nanchang 330013, China)

Abstract: In-situ leaching of uranium, as a green uranium mining technology, generates massive data in production operation, which are available for the big data analysis and trend prediction to improve the reliability of technicians in making production plans. In the current prediction algorithms, the attention mechanism in the temporal prediction model based on the encoder-decoder structure has the problems of computational complexity and high memory consumption. In this paper, we proposed a depthwise separable convolutional model, in which the semantic damage caused by fixed segmentation was reduced by the dynamic sequence segmentation module, and the depthwise separable convolutional mixer module was used to reduce the model running time and capture local features as well as global features. The results show that the Mean Square Error (MSE) and Mean Absolute Error (MAE) of the depthwise separable convolutional hybrid network model are reduced by 1.04% and 4.13% respectively, compared with Patch Time Series Transformer (PatchTST), and the proposed dynamic sequence segmentation module MSE and MAE are reduced by 7.32% and 5.03% respectively, compared to the original model; in the comparative performance analysis, the training speed of this model is 59.91% faster than the Trend Seasonal Decomposition Linear (Decomposition Linear, DLinear) model. The depthwise separable convolutional model can accurately predict the future trend of sulfuric acid injection volume in the production operation of the mining area, improve the existing prediction model for in-situ leaching data by solving the problem of long running time, large running memory, poor data fitting problems, which provide a theoretical and practical reference for the decision-making of in-situ leaching production.

Key words: in-situ leaching of uranium; injection volume prediction; depthwise separable convolution; prediction model