

DOI: 10.13382/j.jemi.B2407585

基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型*

陈 薇^{1,2} 解俊哲^{1,2} 康志伟^{1,2} 赵 军^{1,2} 刘海军^{1,2} 褚 彪³ 张宏图³

(1. 合肥工业大学电气与自动化工程学院 合肥 230009; 2. 安徽省工业自动化工程技术研究中心 合肥 230009;
3. 合肥水泥研究设计院有限公司 合肥 230051)

摘 要:在水泥生产中,熟料是关键成分,其质量直接影响水泥的整体性能。水泥熟料中游离氧化钙(f-CaO)的含量是评估熟料质量的重要参数之一。为了弥补传统实验室化验方法在时效性上的不足,本文构建了一种高效准确的水泥熟料 f-CaO 含量软测量模型,该模型结合了通道注意力机制和长短时记忆网络。利用融合了注意力机制的特征提取方法对数据集进行特征提取;然后将特征输入到 LSTM 网络进行学习,使模型能够以自适应方式聚焦于核心的特征通道;由于 LSTM 在预测波动性较大的数据上预测效果较差,针对此问题对上述软测量模型进行改进,在原有模型基础上引入分类模块与加权模块对 LSTM 网络的预测结果进行修正,使得模型可以更加灵活地适应不同类别之间的差异,通过优化各类别之间的权重,提高了模型预测的准确性。实验结果表明,改进 SE-LSTM 的水泥熟料 f-CaO 预测模型的决定系数(R²)、均方根误差(RMSE)和平均绝对误差(MAE)比普通 LSTM 和 SE-LSTM 预测模型都有明显提升,因此在水泥熟料 f-CaO 的含量预测上所提的改进模型提高了预测精度,达到了预期的预测效果。

关键词:水泥熟料;通道注意力机制;时序匹配;分类加权;LSTM;f-CaO;预测精度

中图分类号: TP183;TN081 **文献标识码:** A **国家标准学科分类代码:** 120.3030

Prediction model of cement clinker f-CaO based on the improved SE-LSTM

Chen Wei^{1,2} Xie Junzhe^{1,2} Kang Zhiwei^{1,2} Zhao Jun^{1,2} Liu Haijun^{1,2} Chu Biao³ Zhang Hongtu³

(1. School of Electrical and Automation Engineering, Hefei University of Technology, Hefei 230009, China;
2. Anhui Engineering Technology Research Center of Industrial Automation, Hefei 230009, China;
3. Hefei Cement Research & Design Institute Co., Ltd, Hefei 230051, China)

Abstract: In cement production, clinker is the key component, and its quality directly affects the overall performance of cement. The content of free calcium oxide (f-CaO) in cement clinker is one of the important parameters to evaluate the quality of cement clinker. In order to make up for the lack of timeliness of traditional laboratory methods, this paper constructs an efficient and accurate soft sensor model for cement clink f-CaO content, which combines channel attention mechanism and long short-term memory network. The feature extraction method combined with attention mechanism was used to extract the features of the data set. Then, the features were input into the LSTM network for learning, so that the model could focus on the core feature channels in an adaptive manner. Due to the poor prediction effect of LSTM on data with large volatility, the above soft sensor model is improved. Based on the original model, the classification module and weighting module are introduced to modify the prediction results of the LSTM network, so that the model can be more flexible to adapt to the differences between different categories. The accuracy of the model prediction is improved. The experimental results show that the coefficient of determination (R²), root mean square error (RMSE) and mean absolute error (MAE) of the improved SE-LSTM prediction model for cement clink f-CaO are significantly improved compared with the ordinary LSTM and SE-LSTM prediction models. Therefore, the improved model proposed in the prediction of cement clinker f-CaO content improves the prediction accuracy and achieves the expected prediction effect.

收稿日期: 2024-06-08 Received Date: 2024-06-08

* 基金项目:安徽省重点研发计划项目(202104a05020054)资助

Keywords: cement clinker; channel attention mechanisms; time series matching; classification weighting; LSTM; f-CaO; prediction accuracy

0 引言

水泥作为一种主要的建筑材料,在我国工业、民用、交通运输、水利、军事和工业建筑等工业工程领域得到了广泛的应用。在水泥生产中,熟料作为核心组成部分,其质量对水泥的整体性能具有决定性影响。其中,游离氧化钙(free calcium oxide, f-CaO)是熟料质量评估的关键指标之一^[1],f-CaO 指的是在生料烧结过程中,未与氧化硅、氧化铝、氧化铁等酸性氧化物发生反应而残留的氧化钙,它以游离状态存在于熟料之中^[2-3]。f-CaO 的含量对于熟料的稳定性和水泥的整体性能具有显著影响。若其含量过高,将严重影响熟料的安定性,甚至可能影响熟料与混凝土外加剂之间的相容性^[4],进而削弱水泥的整体强度和其他物理性能。相反,若 f-CaO 含量过低,则可能降低熟料的品质,包括其易磨性的降低,同时还可能增加生产过程中的能耗和成本^[5]。因此,准确测量熟料中游离氧化钙的含量在水泥生产过程中显得尤为重要^[6]。

目前,传统的 f-CaO 含量检测方法是将冷却后的熟料送到实验室进行离线检测,包括乙二醇萃取后配合乙二胺四乙酸滴定法^[7]、甘油酒精法^[7]以及乙二醇法^[8]。这些方法虽然有效,但通常需要间隔 1~2 h 才能获取结果,导致对生产过程的指导存在明显的滞后性,不利于实时地调控生产流程^[9]。为了解决这一问题,研究者们提出了 f-CaO 在线测量方法,以实现生产过程的更精准、更及时地监控与调整。

以数据驱动的在线软测量中,利用如支持向量机、极限学习机和神经网络等模型,通过分析数据来实时监测和预测水泥熟料中的 f-CaO 含量。例如,赵朋程^[10]结合机理分析与灰色关联分析,对极限学习机进行了创新性的改造,提出了基于变形 Cholesky 分解的在线多核超限学习机算法(COS-MKELM),以应对复杂的工业数据;Zhao 等^[11]选取生料配比及回转窑的相关运行参数建立了一种基于多元时间序列分析的卷积神经网络(MVTS-CNN)预测模型;Li 等^[12]则选取喂煤量、窑主机电流等常规变量外还加入了火焰图像的特征,改进了多源建模技术。杨黎明^[13]选取了生料成分配比及回转窑相关参数对深度残差网络模型进行改进,并引入了主成分分析(PCA)和长短时记忆网络(LSTM),提出了一种综合性的机理分析方法和辅助变量选取策略,为水泥熟料 f-CaO 的预测和控制提供了新思路;文献[14]介绍了一种将 CNN、LSTM 和 AM 动态集成的预测模型来实现对电站风机状态的预测;刘文彪等^[15]将 CNN 与 BiLSTM 相结合

来实现对轴承性能退化的预测。

上述各类预测模型仅仅将辅助变量数据作为 f-CaO 预测的简单参考,却忽视了不同辅助变量对 f-CaO 影响力差异的重要性。其次,尽管 LSTM 模型在时序数据预测方面较其他模型来说表现出色,但在面对波动性强的数据时,其预测误差仍然显著。这是因为数据的大幅波动使得模型难以全面捕捉其背后的复杂上下文信息,从而影响了预测的准确性。

综上所述,为了解决当前研究存在的问题,实验提出了一种新的水泥熟料 f-CaO 预测模型。针对各辅助变量对 f-CaO 影响程度不同的问题,实验引入了通道注意力模块^[16](squeeze-and-excitation block, SE block),对辅助变量进行加权处理。此举旨在赋予对 f-CaO 影响较大的辅助变量更高的权值,而对影响较小的辅助变量则分配较低的权值^[17],从而显著提升模型的预测精度。针对 LSTM 在预测波动性较强数据时出现的显著误差问题,考虑通过集成分组模块与加权模块对 LSTM 模型进行改进,使模型能够更好地适应复杂多变的数据关系,尤其是在处理波动性较强数据时,展现了出色的适应性和鲁棒性。这一优化不仅增强了模型的灵活性,同时显著提高了其在复杂数据环境下的预测精度,为相关领域的研究和应用提供了新的思路和方法。

1 水泥熟料 f-CaO 机理分析及辅助变量选取

1.1 水泥熟料 f-CaO 机理分析

水泥熟料中的 f-CaO 含量,作为衡量水泥品质的关键指标,特指那些未与二氧化硅、三氧化二铁和三氧化二铝等氧化物发生反应的游离态氧化钙。此含量的多少直接关系到水泥的强度及稳定性。因此,在水泥的生产流程中,对 f-CaO 含量的准确预测及时调整生产参数显得尤为关键,它能够有效提升水泥的最终质量。图 1 为新型干法水泥烧成系统的工艺流程图。

新式干法水泥烧成系统涵盖了多个关键阶段,即预热器预热、分解炉分解、回转窑煅烧和篦冷机冷却。这些不同的生产环节在各自特定的环境条件下,为碳酸盐的分解和氧化钙的反应提供了所需的条件。因此,在整个生产流程中,每个环节的过程参数都会对水泥熟料中 f-CaO 的含量产生有着一定的影响。

新式干法水泥烧成系统步骤具体为:在生料准备工序完成后,生料粉末经斗式提升机输送至五级旋风筒预热器。在预热器内部完成预加热过程后,因重力的作用落入分解炉中进行预分解步骤,分解炉内的温度主要通

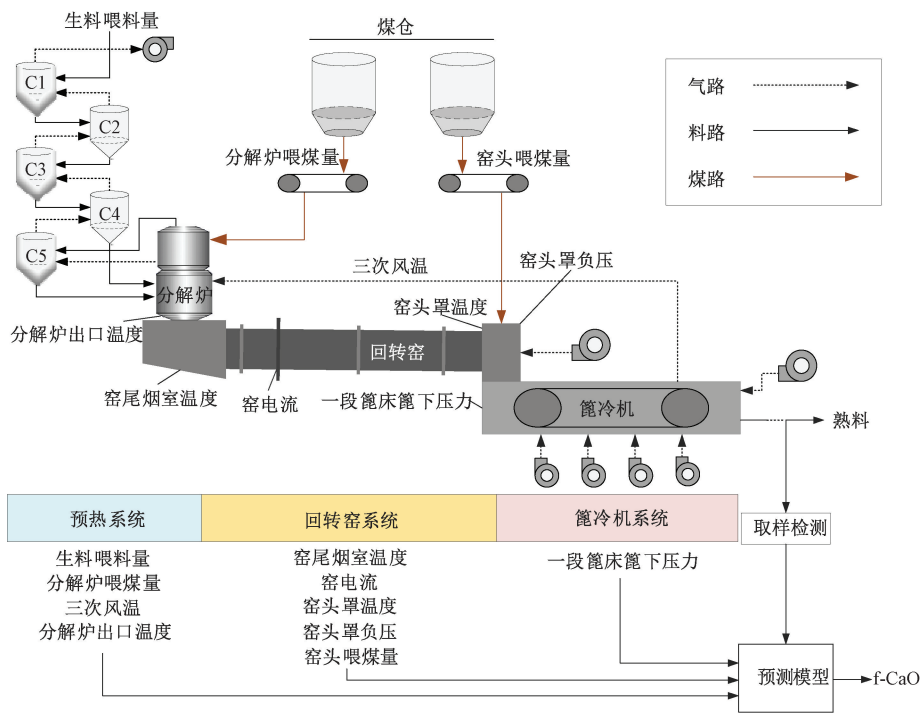


图 1 新型干法水泥烧成工艺流程图

Fig. 1 Flow chart of new dry process cement firing process

过喂煤量的多少和三次风温的大小进行控制。在预分解完成之后,水泥生料随即进入回转窑中进行高温煅烧。在这个过程中,生料与燃煤发生复杂的物理化学反应,最终从窑尾逐渐滑移至窑头,并落入篦冷机中,窑内的温度控制主要依赖于窑头喂煤量的调整以及篦冷机回收的二次风温的调节,同时回转窑主电机的持续旋转保证了熟料在煅烧过程中的均匀性。随后,经过高温煅烧的物料从窑头滑落至篦冷机后,在篦冷机的多个风室中经过充分的冷却降温处理后使物料达到合适的温度,为了提高能源利用效率,从窑下料区产生的二次风被重新回收至回转窑,而来自热回收区的三次风则被送回分解炉进行能量回收,形成了高效的能源循环体系。最终冷却后的物料通过熟料破碎机进行破碎,转化为水泥熟料。上述每个环节都会对水泥熟料中 f-CaO 的含量造成不同程度的影响。

1.2 辅助变量的确定

根据上述机理分析,影响水泥熟料 f-CaO 含量的环节多,因素多,初步选择分解炉出口温度、分解炉喂煤量、三次风温、生料喂料量、窑尾烟室温度、窑头喂煤量、窑头罩温度、窑主机电流、窑头罩负压、一段篦床篦下压力作为影响 f-CaO 含量大小的主要辅助变量。为了进一步确定辅助变量,考虑采用皮尔逊相关系数 (pearson correlation coefficient) 来度量辅助变量与 f-CaO 含量的相关程度^[18],其公式如式(1)所示。

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{1}$$

式中: σ_X 为样本标准差, $Cov(X,Y)$ 为辅助变量与 f-CaO 之间的协方差。通过计算水泥烧成系统中各个过程参量与 f-CaO 之间的相关系数,结果如表(1)所示。

表 1 各辅助变量与 f-CaO 相关系数

Table 1 Correlation coefficients of each auxiliary variable with f-CaO

变量名	PEARSON 相关系数
分解炉出口温度	0.150 61
分解炉喂煤量	-0.101 24
三次风温	0.217 33
生料喂料量	0.249 93
窑尾烟室温度	-0.151 07
窑头喂煤量	-0.371 44
窑头罩温度	0.445 86
窑主机电流	-0.006 66
窑头罩负压	-0.045 97
一段篦床篦下压力	-0.172 29

最终选择相关系数绝对值大于 0.1 的变量作为 f-CaO 含量预测的辅助变量,其中包括:分解炉出口温度、分解炉喂煤量、三次风温、生料喂料量、窑尾烟室温度、窑头喂煤量、窑头罩温度、一段篦床篦下压力。

2 融合通道注意力 SE block 的特征提取方法

2.1 数据预处理

在通过现场 DCS 进行数据采集的过程中会出现两种数据异常情况,第 1 种情况是由于设备故障或需要停窑检修时,导致辅助变量和 f-CaO 数据都为零值。第 2 种情况是某个设备中出现异常工况时导致对应设备下的辅助变量数据出现大幅度跳动后又迅速恢复正常。针对于这两种异常,将异常时间段内的所有辅助变量和对应时间段内的 f-CaO 数据同时进行删除。

由于不同变量之间量纲的不同,例如:窑主机电流、三次风温和生料喂料量等数据量级有着很大的差异,因此采用 python 的 sklearn 库中的 MinMaxScaler 归一化函数来降低输入数据量纲上的差异,增强数据之间的可比性,表达式如公式(2)所示。

$$x_t' = \frac{x_t^i - x_{\min}^i}{x_{\max}^i - x_{\min}^i} \times (x_{\max}^i - x_{\min}^i) + x_{\min}^i \quad (2)$$

式中: $x_{\max}^i$ 和 $x_{\min}^i$ 为第 i 个辅助变量 x^i 的最大值与最小值, x_t^i 为第 i 个辅助变量第 t 时刻的数值, x_t' 为 t 时刻归一化后的值。

2.2 辅助变量与 f-CaO 的时序匹配

新式干法水泥烧成工艺整个流程大约需要 120 min,由于物料在不同设备移动速度和设备长度不同,导致物料在各设备中停留时间不同。根据水泥工艺流程可知,物料在各设备停留时间如表 2 所示。

表 2 物料在各设备中停留时间

设备名称	物料在设备中停留时间
预热系统(分解炉+回转窑)	53:00~55:00
回转窑	58:00~78:00
篦冷机	84:00~114:00

各辅助变量数据和水泥熟料 f-CaO 含量的数据的采样间隔分别为 5 s 和 2 h,在一个 f-CaO 的采样周期中,物料移动到某个设备中,该设备中的辅助变量对 f-CaO 产生的影响最大,如当物料移动到预热系统时,分解炉出口温度、分解炉喂煤量等辅助变量的大小对 f-CaO 产生的影响最大,但是当物料进入回转窑时,预热系统中的变量造成的影响几乎为零。

因此,将上述各辅助变量根据设备划分,预热系统:分解炉出口温度、分解炉喂煤量、三次风温和生料喂料量;回转窑:窑尾烟室温度、窑头喂煤量、窑头罩温度;篦冷机:一段篦床篦下压力。在一个水泥熟料 f-CaO 的采样周期中,选择物料在各设备中停留时间区域数据作为各辅助变量的参考数据。

2.3 融合注意力机制 SE block 的特征提取方法

SE block 是一种灵活且高效的通道注意力模块,其核心理念在于实施了压缩(Squeeze)与激励(Excitation)两大策略。在压缩阶段,利用全局平均池化技术,将各个通道的特征图精妙地压缩为单一的数值,以此彰显该通道全局特征的关键性。进入激励阶段,凭借一个或多个全连接层,深入挖掘并学习每个通道的权重,从而精细调整各通道的响应强度。最终,这一模块赋能网络以自适应方式聚焦于核心的特征通道,实现网络性能的显著提升。

通过传感器收集上述 8 个辅助变量的原始数据构成完整的时间序列数据集 $X(t) = [X_1(t), X_2(t), \dots, X_8(t)]$ 。为了应对卷积池化环节中特征图通道间重要性差异所带来的性能损耗问题,实验引入了通道注意力机制 SE block,对这一时间序列数据进行精细化特征提取,具体步骤如图 2 所示。

步骤 1)通过数据预处理方法对原始数据 $X(t)$ 进行异常数据删除和归一化处理,得到时序数据 $X'(t)$;

步骤 2)将时间序列数据 $X'(t)$ 按照上述时序匹配方法截取出各辅助变量对应的有效数据,并根据 f-CaO 采样周期 T 进行均值压缩,得到新的时间序列数据 $X_T(t)$;

步骤 3)由于 SE block 的输入格式为三维数据,设置切片长度 n-step,将 $X_T(t)$ 进行整合后并按照时间步做切片处理得到三维数据 $X_R(t)$,其维度为 $H \times W \times C$;

步骤 4)设置 SE block 中的激活函数、优化器、通道缩放比例;

步骤 5)将 $X_R(t)$ 输入 SE block 中学习其各通道间的依赖关系,自适应地选择特征的表达,最终计算得出加权后的数据特征 $\tilde{X}(t) = [\tilde{X}_1(t), \tilde{X}_2(t), \dots, \tilde{X}_8(t)]$ 。

3 基于改进的 SE-LSTM 水泥熟料 f-Cao 预测模型

3.1 基于 SE-LSTM 的水泥熟料 f-CaO 预测模型

长短时记忆网络(long short-term memory, LSTM)作为循环神经网络(recurrent neural network, RNN)的一种特殊变体,其主要设计目标在于克服传统 RNN 在处理序列数据时容易出现的梯度消失或梯度爆炸问题^[19],可以有效地传递和表达时间序列中的信息并且不会因为时间过长而遗忘之前的有用信息。

每个 LSTM 单元由记忆细胞、输入门、遗忘门、输出门组成,其中 3 个门通常使用 Sigmoid 激活函数进行激活。通过记忆细胞来存储信息,通过这些门控单元来控制模型学习何时输入、遗忘和输出信息。图 3 为 LSTM 模型的网络结构图。

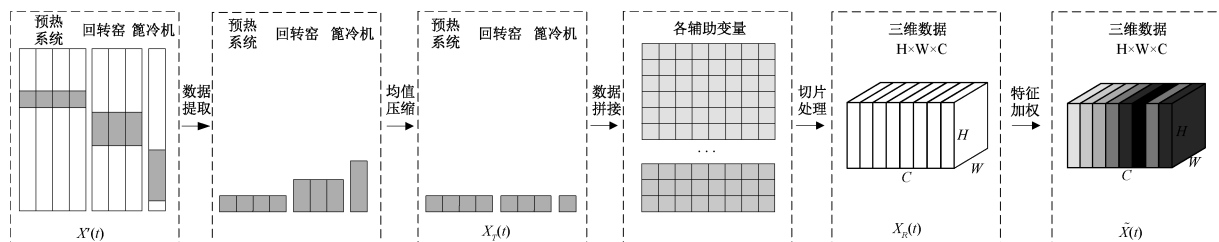


图 2 融合注意力机制 SE block 的特征提取方法

Fig. 2 The feature extraction method incorporating the attention mechanism of SE block

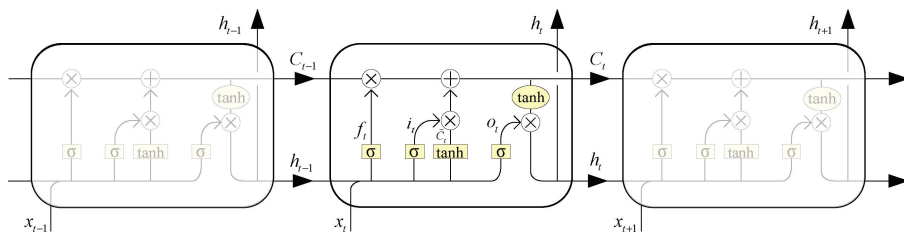


图 3 LSTM 的网络结构图

Fig. 3 Network structure diagram of LSTM

其中, x_t 为当前时刻的输入, h_{t-1} 为上一时刻的输出, C_t 为当前时刻的细胞状态, LSTM 的计算公式如式(3)~(8)所示。

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (3)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (4)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (5)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (6)$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (7)$$

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

其中, W_i, W_o, W_f, W_c 分别为输入门、更新门、遗忘门和输出门所对应的权重向量, b_i, b_o, b_f, b_c 分别为对应门的偏置向量, σ 与 $\tanh$ 分别为 Sigmoid 和 tanh 激活函数, 通过上述公式来计算出当前时刻的输出 h_t 和细胞状态 C_t 。

通过巧妙地结合长短时记忆网络 (LSTM) 与通道注意力模块 (SE block), 构建了融合注意力机制的 LSTM 的水泥熟料 f-CaO 预测模型 (下文统一称为 SE-LSTM 预测模型)。该模型采用了结合注意力机制的特征提取技术, 以精确捕捉关键特征, 并通过 SE-LSTM 网络进行深度训练, 构建了一个高效的预测模型。利用通道注意力模块, 模型显著增强了重要特征的影响力, 同时降低了非关键特征的干扰, 从而成功解决了模型在识别特征变量重要性差异方面的难题。

3.2 基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型

经过多次实验测试, 发现 SE-LSTM 预测模型在处理波动较大的 f-CaO 数据时表现欠佳。尽管 SE-LSTM 的设

计在一定程度上缓解了梯度消失问题^[20-21], 但在面对大幅度波动数据时, 梯度问题依然突出。这种数据波动会导致反向传播过程中的梯度值变得极大或极小, 进而造成权重更新的困难。同时, 大幅度的数据波动也使得模型难以捕捉到足够的上下文信息, 严重降低了预测精度。为了改善上述提出的问题, 提出了基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型, 其结构图如图 4 所示。

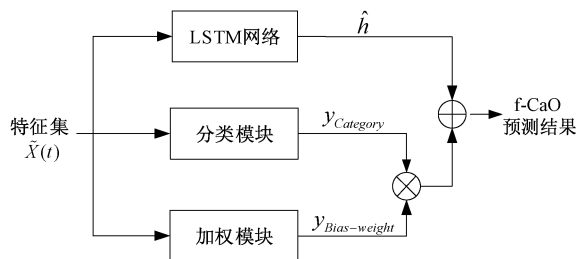


图 4 基于改进的 SE-LSTM 预测模型结构图

Fig. 4 Structure of the improved SE-LSTM based prediction model

模型引入分类模块和加权模块, 分类模块能够提供对输入数据进行更细粒度的分类信息, 使得模型能够自适应地理解不同类别之间的关系, 加权模块中引入可学习的偏置项, 使得模型能够自适应地为每个类别调整偏置, 最终通过分类和加权结果对 LSTM 预测结果进行调整。该方法增加了模型的表达能力, 使其能够更好地适应复杂的数据关系, 同时提高了模型的灵活性和精确度。

分类模块是由扁平层和多个全连接层堆叠而成, 如图 5 所示。使用全连接层执行多类别分类任务, 将输入

数据映射到不同类别的概率分布,提供了不同类别的预测,使模型能够更好地理解输入数据的类别特征,增强模型捕捉上下文信息的能力。

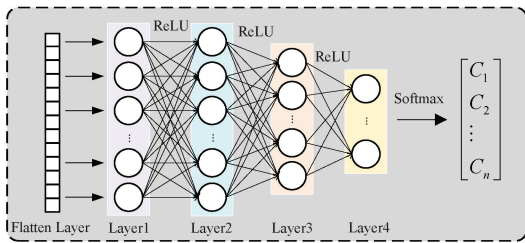


图 5 分类模块网络结构图

Fig. 5 Classification module network diagram

分类模块通过扁平层 (flatten layer) 对上述提取到的特征数据 $X_{(H, W, C)}$ 进行展开, 形成一个一维向量 $X(t)$, 使得全连接层 (dense layer) 能够对整个序列的特征进行全连接操作。然后将 $X(t)$ 传递给全连接层, 对其进行线性变换并通过激活函数 ReLU 引入非线性, 计算公式如式 (9) 所示。

$$y_i = \text{ReLU}(W_i x + b_i) \quad (9)$$

其中, W_i 是对应层的权重矩阵, b_i 为偏置向量, ReLU 为激活函数; 将最后一个全连接层的输出通过 softmax 激活函数进行规范化处理如式 (10) 所示。

$$\text{softmax}(x) = \frac{e^{x_i}}{\sum_i e^{x_i}} \quad (10)$$

得到包含对每个类别概率的向量 $y_{\text{Category}} = [C_1, C_2, C_3, \dots, C_n]$, 用于表示输入数据属于每个类别的可能性。

加权模块也是由扁平层和全连接层构成, 如图 8 所示。模型通过引入可学习的偏置项, 使得模型能够自适应地为每个类别调整偏置。这使得模型在序列中自适应地调整每个类别的重要性, 提高了模型在复杂序列任务上的性能, 提高了模型的灵活性。

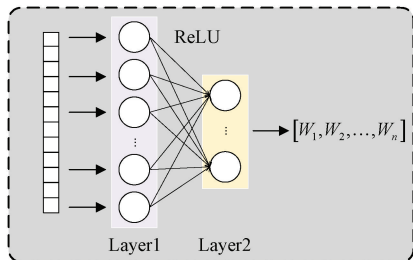


图 6 加权模块网络结构图

Fig. 6 Weighted module network diagram

加权模块包括前向传播过程和反向传播过程, 前向传播过程与上述分类模块相同, 通过扁平层与全连接层学习输入数据的非线性特征, 但是不同之处在于其最后

一层全连接层未使用激活函数, 使得最后一层为线性层, 利用其线性特性计算出含有每个类别的偏置权重的向量 $y_{\text{Bias-weight}} = [W_1, W_2, W_3, \dots, W_n]$ 。在反向传播过程中, 首先定义均方差损失函数, 如式 (11) 所示。

$$\text{loss} = \frac{1}{2n} \sum_{i=1}^n (Y_{\text{hat}_i} - Y_i)^2 \quad (11)$$

其中, Y_{hat} 与 Y 分别为模型的预测值和真实值; 其次用链式法则计算损失对于各类的偏置权重的梯度, 其计算如式 (12) 所示。

$$\frac{\partial \text{loss}}{\partial W_i} = \frac{1}{n} (Y_{\text{hat}_i} - Y_i) \quad (12)$$

其中, W_i 给某个类别的偏置权重; 通过梯度下降法更新各类别的偏置权重, 其计算如式 (13) 所示。

$$W'_i = W_i - \eta \frac{\partial \text{loss}}{\partial W_i} \quad (13)$$

其中, η 为神经网络学习率, W_i 与 W'_i 分别更新前后的权重。在训练集训练过程中, 通过多次迭代调整权重, 使得损失逐渐减小, 得到合适的偏置权重 $y_{\text{Bias-weight}}$ 。

最后将分类模块结果 y_{Category} 和权重模块结果 $y_{\text{Bias-weight}}$ 进行点积得到每个类别的偏置 y_B , 并与 LSTM 部分的输出结果 $\hat{h}$ 进行相加得到最终预测结果 $\hat{y}$, 其计算如式 (14) 所示。

$$\hat{y} = \hat{h} + (y_{\text{Category}} \times y_{\text{Bias-weight}}) \quad (14)$$

通过对 LSTM 结果的修正, 使得模型可以同时考虑分类信息和权重信息, 更加灵活地适应不同类别之间的差异, 优化各类别之间的权重, 提高预测精度。

3.3 基于改进的 SE-LSTM 预测算法步骤

表 3 基于改进的 SE-LSTM 预测算法步骤

Table 3 Improved SE-LSTM based prediction algorithm steps

主要步骤:

- (1) 对原始数据集进行异常数据剔除、归一化等数据预处理得到数据集 $X(t)$ 。
- (2) 设置切片长度, 激活函数以及通道缩放比例, 对数据集 $X(t)$ 根据上述方法进行特征提取, 得到数据特征集 $\tilde{X}(t)$ 。
- (3) 设置 LSTM 节点数与激活函数
- (4) 设置分类模块与加权模块中的全连接层数、输出维度以及各层所使用的激活函数。
- (5) 设置优化器, 损失函数, 划分数据集比例得到训练集、测试集, 并根据 $\tilde{X}(t)$ 与 f-CaO 数据训练得到改进 LSTM 预测模型^[22]。
- (6) 将测试集输入预测模型中进行预测, 将结果通过反归一化得到最终 f-CaO 预测结果与模型预测性能指标。

4 基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型仿真

某水泥生产线每隔 2 h 化验一次水泥熟料样本得到

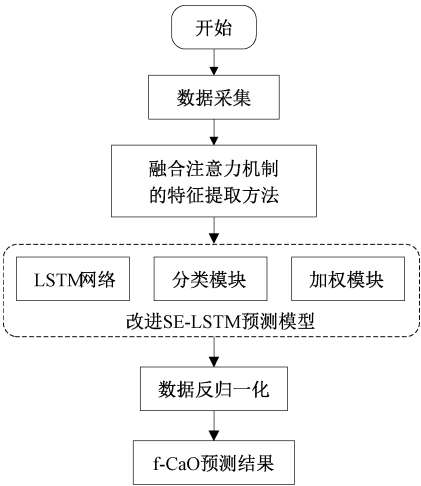


图 7 基于改进的 SE-LSTM 水泥熟料 f-CaO 预测流程

Fig. 7 Improved SE-LSTM based f-CaO prediction process for cement clinker

对应 f-CaO 数据,每隔 5 s 通过 DCS 采集一组各辅助变量的值,故一个水泥熟料 f-CaO 值对应 1 440 组辅助变量,且每个 f-CaO 数据对应前两个小时的辅助变量数据。通过连续获取 19 个月的样本数据,得到共 6 864 个 f-CaO 数据和对应 9 884 160 组辅助变量数据,并用上述方法对异常值进行处理后还有 3 705 个 f-CaO 数据和对应 5 335 200 组辅助变量,利用上述辅助变量的时序匹配和均值压缩处理后得到 3 705 个 f-CaO 数据和对应 3 705 组辅助变量作为最终的输入输出。选择前 3 223 组数据作为训练集,中间 370 组数据作为验证集,最后 112 组数据作为测试集。采用均方根误差 (RMSE)、平均绝对误差 (MAE)、决定系数 (R2) 来评估预测模型的回归精度。RMSE、MAE 的值越小,R2 值越接近 1,则模型的预测效果越好,精度越高。

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \tag{15}$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |(y_i - \hat{y}_i)| \tag{16}$$

$$R2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \tag{17}$$

实验以 TensorFlow 作为底层库,利用 Keras 框架来搭建网络结构;选择全局平均池化作为注意力模块中池化层的方式,以确保模型能够有效地捕获全局特征;对于全连接层,采用了 Sigmoid 作为激活函数,以提供非线性的转换能力。为了提升模型的泛化能力并减少过拟合的风险,将 LSTM 层数设置为 80 层,并在其后添加了一个参数为 0.3 的 Dropout 层;在模型训练中,选用 Adam 算法来加快梯度下降的速度,确保优化的高效性。设定均方

差为损失函数,并设置时间步为 40,总共进行 100 次训练迭代。然后将数据集分为训练集和测试集,并将它们分别应用于 LSTM 与 SE-LSTM 网络中进行训练和预测。通过对比预测结果与真实值,计算了均方根误差 (RMSE)、平均绝对误差 (MAE) 以及决定系数 (R2) 来量化评估模型的准确性。下图展示了普通 LSTM 与 SE-LSTM 在预测任务上的结果对比。

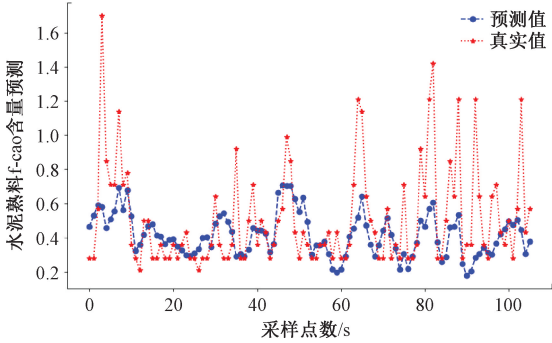


图 8 普通 LSTM 预测效果图

Fig. 8 LSTM prediction effect diagram

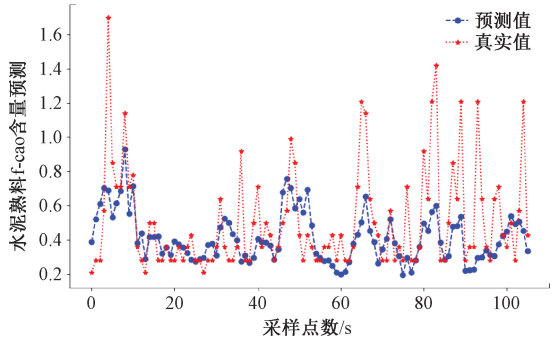


图 9 SE-LSTM 预测效果图

Fig. 9 SE-LSTM prediction effect diagram

表 4 两种预测模型预测结果比较			
Table 4 Comparison of the forecasting results of the two forecasting models			
模型	MAE	RMSE	R2
普通 LSTM	0.195	0.081 1	0.086 3
SE-LSTM	0.188	0.076 6	0.165 3

从图 8、9 以及表 4 可知,基于 SE-LSTM 模型的预测效果以及模型的预测精度要略优于普通 LSTM 模型的预测效果,通过注意力机制对各辅助变量数据进行加权后,使得模型能够更加有效地提取数据中的重要信息。但是针对于波动性较大的数据预测效果较差,于是对模型做出相应改进,提出了基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型,其中分类模块采用 Flatten 层和 4 层全连接层拼接,其输出维度分别为 512、512、256、20;权重模块采用 Flatten 层和 2 层全连接层进行拼接,全连接层输出

维度为 512、20。根据上述操作步骤将训练集、验证集、测试集放入所提出的网络中进行训练和预测,其预测结果如图 10 所示,性能指标对比如表 5 所示。

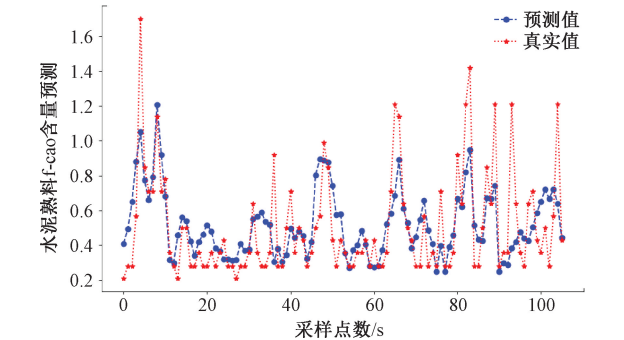


图 10 基于改进的 SE-LSTM 预测效果图
Fig. 10 Improved SE-LSTM prediction effect diagram

从图 10 和表 5 可以看出,上述所介绍的模型在预测效果及各项性能指标上均展现出相对于 SE-LSTM 预测模型的显著优势。此外,经过改进后的模型在处理波动性较大的数据时,其预测效果也得到了显著提升。基于改进的 SE-LSTM 水泥熟料 f-CaO 预测模型是通过引入通道注意力 SE block 来提高模型提取数据特征的能力,并利用分类、加权模块来挖掘数据中的分类信息和权重信息,弥补了模型在处理大幅度波动数据时可能出现的偏差,提高了水泥熟料 f-CaO 含量预测的预测精度。

表 5 两种预测模型预测结果比较

Table 5 Comparison of the forecasting results of the two forecasting models			
模型	MAE	RMSE	R2
SE-LSTM	0.188	0.076 6	0.165 3
改进的 SE-LSTM	0.177	0.058 6	0.352 3

5 结 论

为了解决传统神经网络无法辨别不同变量对预测变量的影响程度不同和针对波动性较大的数据预测精度较差的问题,论文通过结合通道注意力机制 SE block 和长短时记忆网络 LSTM 对水泥熟料 f-CaO 含量进行预测,通过与传统 LSTM 预测模型的对比试验可以看出新模型在预测精度和预测效果上得到了一定的提升。但面对波动性较大的数据时,预测效果尚存优化空间。为此,引入了分类与加权模块对模型进行改良,使模型能够综合考虑分类信息与权重信息,从而更加灵活地适应不同类别数据间的差异。实验结果表明,所提出的模型在预测精度上有明显提升。为了进一步提高模型的预测精度,未来会考虑从数据集去噪和更换预测模型方向开展进一步研究。

参考文献

[1] 赵淑彩. 熟料中游离氧化钙的控制[J]. 中国水泥, 2023(1): 94-95.
ZHAO SH C. Control of free calcium oxide in clinker[J]. China Cement, 2023(1): 94-95.

[2] 周武洲. 基于集成优化算法的水泥熟料游离氧化钙预测研究[D]. 秦皇岛:燕山大学,2019.
ZHOU W ZH. Research on prediction of free calcium oxide in cement clinker based on integrated optimization algorithm[D]. Qinhuangdao: Yanshan University, 2019.

[3] 张海博, 邓全亮, 李小青, 等. 水泥熟料中 f-CaO 预测方法研究进展[J]. 中国水泥, 2022(1): 71-78.
ZHANG H B, DENG Q L, LI X Q, et al. Research progress on prediction methods of f-caO in cement clinker[J]. China Cement, 2022(1): 71-78.

[4] 何伟, 刘洋, 郝文茹, 等. f-CaO 对水泥基材料补偿收缩和安定性影响机理研究[J]. 混凝土, 2022(4): 8-13.
HE W, LIU Y, HAO W R, et al. Study on the mechanism of the influence of f-CaO on the compensating shrinkage and stability of cement-based materials[J]. Concrete, 2022(4): 8-13.

[5] 杨利军, 刘利民. 熟料中 f-CaO 含量高的原因分析及处理措施[J]. 建材技术与应用, 2015(4): 15-17.
YANG L J, LIU L M. Analysis of the causes of high f-CaO content in clinker and treatment measures[J]. Building Materials Technology and Application, 2015(4): 15-17.

[6] 崔保华, 张成伟, 李慧霞, 等. 基于工况分类的熟料 f-CaO 含量预测方法研究[J]. 水泥工程, 2024(1): 1-5, 15.
CUI B H, ZHANG CH W, LI H X, et al. Research on clinker f-CaO content prediction method based on working condition classification[J]. Cement Engineering, 2024(1): 1-5, 15.

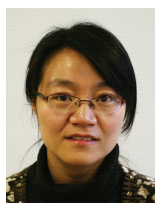
[7] 詹妮. 水泥及原料中游离氧化钙测定要点[J]. 水泥工程, 2021(6): 45-47.
ZHAN N. Key points for determination of free calcium oxide in cement and raw materials[J]. Cement Engineering, 2021(6): 45-47.

[8] 王伟, 于克孝, 刘亚民. 乙二醇萃取—EDTA 滴定法测定水泥熟料中 f-CaO[J]. 水泥, 2017(7): 40-41.
WANG W, YU K X, LIU Y M. Determination of f-CaO in cement clinker by ethylene glycol extraction—EDTA titration[J]. Cement, 2017(7): 40-41.

[9] 唐新宇. 水泥行业 f-CaO 测量技术现状与进展[J]. 居业, 2020(3): 32-33.
TANG X Y. Current status and progress of f-CaO

- measurement technology in cement industry [J]. Residential Industry, 2020(3): 32-33.
- [10] 赵朋程. 水泥熟料 fCaO 含量预测模型及烧成过程优化控制算法研究[D]. 秦皇岛:燕山大学, 2018.
ZHAO P CH. Research on prediction model of fCaO content in cement clinker and optimization control algorithm for sintering process [D]. Qinhuangdao: Yanshan University, 2018.
- [11] ZHAO Y, DING B, ZHANG Y, et al. Online cement clinker quality monitoring: A soft sensor model based on multivariate time series analysis and CNN [J]. ISA Transactions, 2021(117): 180-195.
- [12] LI W T, WANG D H, ZHOU X J, et al. An improved multi-source based soft sensor for measuring cement free lime content [J]. Information Sciences, 2015, 323: 94-105.
- [13] 杨黎明. 基于残差网络与长短时记忆网络的水泥游离钙软测量方法研究[D]. 秦皇岛:燕山大学, 2020.
YANG L M. Research on soft measurement method of cement free calcium based on residual network and long short-term memory network [D]. Qinhuangdao: Yanshan University, 2020.
- [14] 魏玮, 吕游, 齐欣宇, 等. 基于 CNN-LSTM-AM 动态集成模型的电站风机状态预测方法[J]. 仪器仪表学报, 2023, 44(4): 19-27.
WEI W, LYU Y, QI X Y, et al. Wind turbine state prediction method of power station based on CNN-LSTM-AM dynamic ensemble model [J]. Chinese Journal of Scientific Instrument, 2023, 44(4): 19-27.
- [15] 刘文彪, 殷礼祥, 耿帆, 等. 基于 CNN-BLSTM 网络的轴承性能退化预测[J]. 电子测量与仪器学报, 2021, 35(2): 80-86.
LIU W B, DUAN L X, GENG F, et al. Prediction of bearing performance degradation based on CNN-BLSTM network [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(2): 80-86.
- [16] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [17] 刘学平, 李玮乾, 刘励, 等. 嵌入 SENet 结构的改进 YOLOV3 目标识别算法[J]. 计算机工程, 2019, 45(11): 243-248.
LIU X P, LI Y Q, LIU L, et al. Improved YOLOV3 target recognition algorithm with embedded senet structure [J]. Computer Engineering, 2019, 45(11): 243-248.
- [18] 刘迪. 建筑能源系统异常数据的快速识别与定点修复方法体系[D]. 大连:大连理工大学, 2022.
LIU D. Rapid identification and targeted repair methodology for abnormal data in building energy systems [D]. Dalian: Dalian University of Technology, 2022.
- [19] 杨熙丞. 基于全卷积神经网络 OCT 内外指纹提取算法[D]. 杭州:浙江工业大学, 2019.
YANG X CH. Extraction algorithm of OCT internal and external fingerprints based on fully convolutional neural networks [D]. Hangzhou: Zhejiang University of Technology, 2019.
- [20] 陈卓, 孙龙祥. 基于深度学习 LSTM 网络的短期电力负荷预测方法[J]. 电子技术, 2018, 47(1): 39-41.
CHEN ZH, SUN L X. Short-term electric load forecasting method based on deep learning lstm network [J]. Electronic Technology, 2018, 47(1): 39-41.
- [21] 王亦斌, 孙涛, 梁雪春, 等. 基于 EMD-LSTM 模型的河流水量水位预测[J]. 水利水电科技进展, 2020, 40(6): 40-47.
WANG Y B, SUN T, LIANG X CH, et al. River flow and water level prediction based on EMD-LSTM model [J]. Advances in Science and Technology of Water Resources, 2020, 40(6): 40-47.
- [22] 胡洋, 程志江, 崔澜. 基于 LSTM 的变频太阳能-空气源热泵系统逐时负荷预测研究[J]. 可再生能源, 2022, 40(7): 866-873.
HU Y, CHENG ZH J, CUI L. Hourly load forecasting study of variable-frequency solar-air source heat pump system based on LSTM [J]. Renewable Energy Resources, 2022, 40(7): 866-873.

作者简介



陈薇(通信作者), 2002 年于中国科学技术大学获得学士学位, 2007 年于中国科学技术大学获得博士学位, 现为合肥工业大学副教授, 主要研究方向为复杂系统建模与控制和非线性预测控制。

E-mail: windy02@hfut.edu.cn

Chen Wei (Corresponding author) received her B. Sc. degree from University of Science and Technology of China in 2002, Ph. D. degree from University of Science and Technology of China in 2007, respectively. Now she is an associate professor in Hefei University of Technology. Her main research interests include Complex systems modeling and Nonlinear predictive control.



解俊哲, 2021 年于安徽信息工程学院获得学士学位, 现为合肥工业大学硕士研究生, 主要研究方向为数据驱动建模与控制。

E-mail: xjz19991117@163.com

Xie Junzhe received his B.Sc. degree from Anhui Institute of Information Technology in 2021. Now he is a M. Sc. candidate in the Hefei University of Technology. His main research interests include data-driven modeling and control.