



# 基于深度学习的 300 kA 铝电解槽 阳极效应预测

尹刚<sup>1</sup>, 陈根<sup>1</sup>, 何文<sup>2</sup>, 颜非亚<sup>3</sup>, 罗斌<sup>4</sup>, 李锐<sup>2</sup>

- (1. 重庆大学 资源与安全学院, 煤矿灾害动力学与控制国家重点实验室, 重庆 400044;
2. 博眉启明星铝业有限公司, 眉山 620010;
3. 贵阳铝镁设计研究院有限公司, 贵阳 550000;
4. 四川省四维环保设备有限公司, 遂宁 629000)

**摘要:** 本文对铝电解槽阳极效应机理和故障参数进行研究, 提出了一种基于深度学习的阳极效应预测方法, 能适应不同维度、不同数据特征的槽况参数, 直接从海量原始数据中挖掘故障特征信息, 大幅缩减效应响应时间, 具有很好的鲁棒性和抗干扰能力, 同时在模型调试优化上, 采用 Batch normalization 算法和梯度检验, 提高了模型收敛速度和稳定性。结果表明: 该模型效应预测准确率和 F1 分数分别达到 94.65% 和 0.9317, 提前预报时间可达 16 min, 并通过现场实验验证, 达到实际生产要求。

**关键词:** 铝电解; 300 kA; 阳极效应预测; 深度学习; 算法优化

文章编号: 1004-0609(2021)-01-0161-10

中图分类号: TF821

文献标志码: A

**引文格式:** 尹刚, 陈根, 何文, 等. 基于深度学习的 300 kA 铝电解槽阳极效应预测[J]. 中国有色金属学报, 2021, 31(1): 161–170. DOI: 10.11817/j.ysxb.1004.0609.2021-36526

YIN Gang, CHEN Gen, HE Wen, et al. Anode effect prediction of 300 kA aluminium electrolysis cell based on deep learning[J]. The Chinese Journal of Nonferrous Metals, 2021, 31(1): 161–170. DOI: 10.11817/j.ysxb.1004.0609.2021-36526

阳极效应发生时, 电压短时间内急剧上升, 严重增加铝电解的吨铝能耗, 降低铝电解电流效率, 缩短电解槽寿命, 严重的还会造成设备损坏, 同时生成大量有害气体, 造成环境污染<sup>[1-2]</sup>, 因此对阳极效应进行提前且准确的预报具有重要意义。目前, 铝电解行业对阳极效应进行预报的方法较多, 并且我国对于阳极效应的抑制已达到比较高的水平, 但因铝电解高温强腐蚀环境导致在线采集参数少、离线测量不连续且不稳定, 使得对阳极效应进行超前、精准的预报仍存在极大的困难<sup>[3]</sup>。在实际生产过程中, 对阳极效应的诊断预测仍只能凭借电压这个单一参数, 其预测准确率及预报提前量均有待提高。

近年来国内外学者对阳极效应预测的研究成果基于阳极电流分布和数据驱动方法的较多<sup>[4-5]</sup>。ZHOU 等<sup>[6]</sup>结合了支持向量机和 K 最邻近算法, 从实时生产数据中提取样本, 结果显示, 该混合方法预测精度达到 89%。田晔非等<sup>[7]</sup>设计了一种基于广义回归神经网络的阳极效应预测模型, 通过实验证实了该模型预测准确度高于传统预测方法。李翹楚等<sup>[8]</sup>使用改进的交叉熵算法与支持向量机的组合优化算法, 通过每次迭代过程中的最优样本, 优化了参数更新策略, 提高了阳极效应预测准确率, 传统阳极效应预测方法存在不完全信息的问题, ZHOU 等<sup>[9]</sup>利用奇异值阈值算法对所采集到的样本进行数据填充, 再通过极端梯度提升机对阳极效应进行分

基金项目: 国家自然科学基金资助项目(50728504)

收稿日期: 2020-02-25; 修订日期: 2020-06-26

通信作者: 尹刚, 教授, 博士; 电话: 023-65104511; E-mail: yk115@cqu.edu.cn

类。CHEN 等<sup>[10]</sup>将电阻斜率和电阻累计斜率的变化趋势以及通过相似搜索技术获得的有用知识作为评判阳极效应发生可信度的主要指标,将物料累计偏差作为辅助评判标准,以此来判断阳极效应是否发生。PAN 等<sup>[11]</sup>通过修正邻域互信息计算相关参数特征与阳极效应之间的相关性,预先选择出较大相关性的特征再利用轻量梯度提升机算法对其进行分类。目前,前述阳极效应预测方法仅考虑部分槽况参数,忽略了部分关键效应特征信息,因而预测准确率仍不够理想,同时,这些方法仍依赖基于专家经验的预处理,降低了预测方法的通用性和灵活性。对于效应提前预报量,目前的方法只是通过效应前某一时刻的状态参数来提前做出判断,而忽视了槽况参数的连续变化性及其变化趋势会对效应的预测结果产生影响。

深度学习作为故障诊断领域的一种新兴方法,相比于浅层神经网络,其强大的特征提取能力和非线性映射能力能更好的表征相关参数与槽况之间复杂的映射关系,减少对专家故障诊断经验和信号处理技术的依赖,在航空发动机、电力变压器、风力发电机组、晶圆片等众多故障诊断领域都有了成功的应用<sup>[12-18]</sup>。魏立新等<sup>[19]</sup>提出了一种基于深度学习的多层感知器铝热轧轧制力预测模型,该模型预测结果与实测结果误差降低到 3%以内,实现了轧制力的高精度预报。循环神经网络适用于处理时序数据或前后关联数据<sup>[20]</sup>,王鑫等<sup>[21]</sup>提出了一种基于

多层网格搜索的长短时记忆网络预测模型参数优选算法,其在故障时间序列分析中相比于其他时间序列预测模型具有更高的准确性。YIN 等<sup>[22]</sup>利用递归神经网络对网络攻击进行识别分类,其分类准确率优于传统机器学习分类方法。DE BRUIN 等<sup>[23]</sup>利用长短时记忆神经网络通过对区域铁路具有时空相关性的多轨道电路信号进行分析,学习数据间的依赖关系,实现了对铁路轨道电路故障的检测。由此可见,深度学习非常适合于大数据背景下多样性、非线性、高维监测数据诊断分析需求,但其在铝电解槽阳极效应诊断领域的应用还未见报道。

本文将深度学习应用于铝电解槽阳极效应的诊断预测中,提出基于长短时记忆网络的阳极效应预测方法,以解决目前铝电解槽阳极效应预测方法关键特征信息遗漏、高度依赖专家经验、缺乏可推广性等不足,其完整预报和处理过程如图 1 所示。

## 1 方法理论

### 1.1 长短时记忆网络

长短时记忆网络充分考虑了前后样本对当前样本的影响,适用于复杂设备或系统的实时故障诊断,善于处理时序数据,学习数据间的依赖关系。为了避免循环神经网络产生梯度消失或梯度爆炸等问题,构建门控单元,实现较远距离的信息传递。其基本单元结构如图 2 所示。

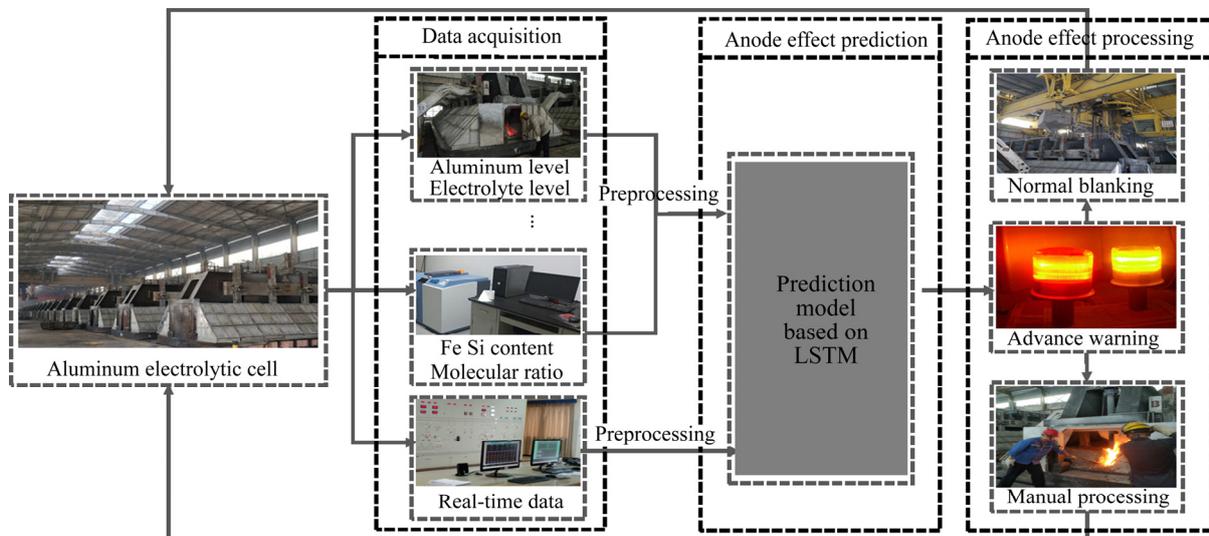


图 1 铝电解槽阳极效应预测及处理过程

Fig. 1 Anode effect prediction and treatment process of aluminium electrolysis cell

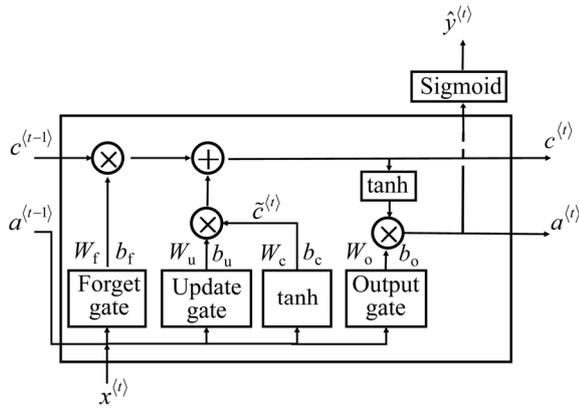


图 2 LSTM 基本单元结构图

Fig. 2 Basic unit of LSTM

对于 LSTM 结构单元, 其增加了遗忘门  $\Gamma_f^{(t)}$ , 它决定了过去输出信息对当前输入信息的重要程度, 控制继承过往多大程度的信息:

$$\Gamma_f^{(t)} = \sigma(W_f [a^{(t-1)}, x^{(t)}] + b_f) \quad (1)$$

通过更新门  $\Gamma_u^{(t)}$  控制更新后的信息所占比例,  $\tilde{c}^{(t)}$  为临时记忆单元, 由上一时刻的输出以及这一时刻的输入共同决定其临时记忆的内容, 记忆单元  $c^{(t)}$  所携带的信息最终由遗忘门和更新门共同决定:

$$\Gamma_u^{(t)} = \sigma(W_u [a^{(t-1)}, x^{(t)}] + b_u) \quad (2)$$

$$\tilde{c}^{(t)} = \tanh(W_c [a^{(t-1)}, x^{(t)}] + b_c) \quad (3)$$

$$c^{(t)} = \Gamma_f^{(t)} \cdot c^{(t-1)} + \Gamma_u^{(t)} \cdot \tilde{c}^{(t)} \quad (4)$$

最后利用输出门  $\Gamma_o^{(t)}$  实现信息输出, 通过  $\tanh$  函数决定记忆单元的输出内容并与输出门相乘:

$$\Gamma_o^{(t)} = \sigma(W_o [a^{(t-1)}, x^{(t)}] + b_o) \quad (5)$$

$$a^{(t)} = \Gamma_o^{(t)} \cdot \tanh(c^{(t)}) \quad (6)$$

式中:  $W_f$ 、 $b_f$ 、 $W_u$ 、 $b_u$ 、 $W_o$ 、 $b_o$  分别为遗忘门、更新门、输出门的权值矩阵和阈值。

## 1.2 Adam 优化算法

本文采用 Adam 优化算法, 相比于传统的优化算法, Adam 算法通过计算梯度的一阶矩估计和二阶矩估计而为不同的参数设计独立的自适应性学习率, 该算法计算了梯度的指数移动平均值并通过

$\beta_1$  和  $\beta_2$  两个超参数控制移动均值的衰减率, 从而在非稳态和在线问题上有很好的表现。

$$V_{dw} = \beta_1 V_{dw} + (1 - \beta_1) dw \quad (7)$$

$$V_{db} = \beta_1 V_{db} + (1 - \beta_1) db \quad (8)$$

$$S_{dw} = \beta_2 S_{dw} + (1 - \beta_2) (dw)^2 \quad (9)$$

$$S_{db} = \beta_2 S_{db} + (1 - \beta_2) (db)^2 \quad (10)$$

式中:  $V_{dw}$ 、 $V_{db}$ 、 $S_{dw}$ 、 $S_{db}$  为指数加权移动平均数, 均初始化为 0;  $\beta_1$ 、 $\beta_2$  为超参数, 分别取值 0.9 和 0.999。同时引入了偏差修正:

$$V_{dw}^{\text{correct}} = \frac{V_{dw}}{1 - \beta_1^t} \quad (11)$$

$$V_{db}^{\text{correct}} = \frac{V_{db}}{1 - \beta_1^t} \quad (12)$$

$$S_{dw}^{\text{correct}} = \frac{S_{dw}}{1 - \beta_2^t} \quad (13)$$

$$S_{db}^{\text{correct}} = \frac{S_{db}}{1 - \beta_2^t} \quad (14)$$

式中:  $t$  为迭代次数; 最后对  $w$ 、 $b$  进行更新:

$$w = w - \partial \frac{V_{dw}^{\text{correct}}}{\sqrt{S_{dw}^{\text{correct}} + \epsilon}} \quad (15)$$

$$b = b - \partial \frac{V_{db}^{\text{correct}}}{\sqrt{S_{db}^{\text{correct}} + \epsilon}} \quad (16)$$

式中:  $\epsilon$  是为了防止分母为 0, 一般取  $1 \times 10^{-7}$ , 本文在 Adam 算法中, 学习率  $\partial$  取 0.001, 为了在保证模型精度的基础上提高模型的迭代速度。

## 2 阳极效应预测模型

### 2.1 模型框架

本文提出了一种基于深度学习长短时记忆网络的阳极效应预测方法, 所选择的样本特征既考虑了铝电解生产过程中长期积累的槽况参数如分子比、电解质水平、硅含量、铁含量等, 也包含了能够反映短时生产操作影响以及效应发生前短时特征的实时监测参数如槽电压、分布式阳极电流。为了尽可能避免关键特征的遗漏, 达到更好的预测效果, 本文选取所有能采集到的原始数据作为样本数据直接输入, 详细输入特征参数见表 1。其中电解温度、分子比、铝水平、电解质水平、硅含量、铁

表1 阳极效应模型输入参数

Table 1 Input parameters of proposed model

Feature	Name	Description
X1	NB times	Total times of NB during one day
X2	Average voltage	Average value of voltage during one day
X3	Electrolysis temperature	Internal temperature of aluminum electrolytic cell
X4	Temperature change	Variation of temperature compared with previous day
X5	Molecular ratio	Ratio molecules number of NaF to AlF <sub>3</sub> in electrolyte
X6	Molecular ratio change	Variation of molecular ratio compared with previous day
X7	Aluminum liquid level	Height of aluminum liquid in aluminum electrolytic cell
X8	Aluminum liquid level change	Variation of aluminum liquid level compared with previous day
X9	Electrolyte level	Height of electrolyte in aluminum electrolytic cell
X10	Electrolyte level change	Variation of electrolyte level compared with previous day
X11	TAP volume	Aluminum tapping volume of one day
X12	TAP volume change	Variation of TAP volume compared with previous day
X13	Silicon content	Percentage of silicon impurities in aluminum liquid
X14	Silicon content change	Variation of silicon content compared with previous day
X15	Iron content	Percentage of iron impurities in aluminum liquid
X16	Iron content change	Variation of iron content compared with previous day
X17	Cell voltage	Cell voltage between anode and cathode
X18	Cell current	Electrolyte current

含量等参数通过每天离线采集,可称之为长周期变量,这类特征可反映阳极效应的长期积累效应,用于对铝电解槽状态进行预测评价;而对于能实时监测获取的数据如槽电压和电解槽系列电流等瞬时变化的量,作为反映阳极效应的短时突发性的关键特征参数,充分考虑其时序性,在效应发生前一段时间,每隔 2 min 采集一次数据,共采集 10 个时间步的数据,与反映阳极效应累积性的长周期变量相结合,利用长短时记忆网络学习阳极效应与数据间的依赖关系,实现对阳极效应的预报,同时实现对阳极效应提前预报时间的控制。最终建立的模型结构如图 3 所示。本文模型中涉及到的超参数较多,且深度学习缺少超参数调试的理论指导,本文在大量实验的基础上,以模型在训练集/验证集/测试集中的准确率为粗略衡量模型性能指标,对模型中的关键性参数进行了确定,模型调试过程如图 4 所示,模型的整体优化算法采用 Adam 算法,模型最终输出层采用 Sigmoid 激活函数,具体超参数设置

如表 2 所示。

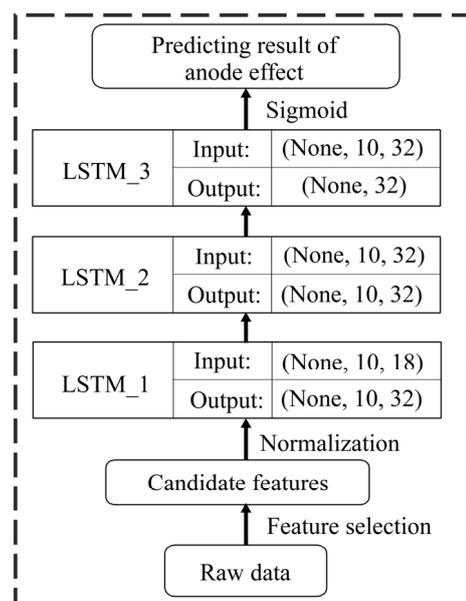


图3 长短时记忆网络阳极效应预测模型

Fig. 3 LSTM model for anode effect prediction

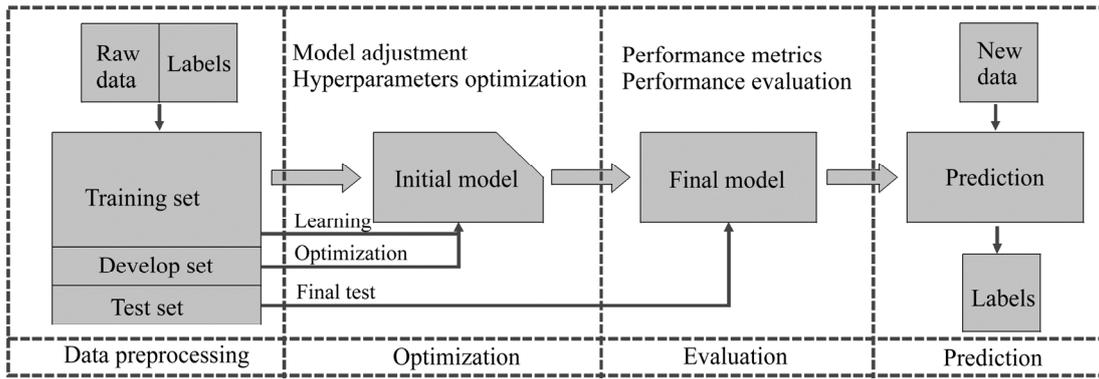


图 4 模型训练、优化和测试过程

Fig. 4 Model training, optimization and testing process

表 2 LSTM 模型超参数设置

Table 2 Hyperparameters of LSTM model

Hyperparameter	Value
Time step	10
Numbers of layer	3
Dropout	0.1
Recurrent dropout	0.5

## 2.2 模型优化

本文所用数据样本量较大, 传统 Batch 梯度下降法迭代速度慢, 随机梯度下降法参数更新不稳定, 本文采用 Mini-batch 梯度下降法, 经过实验, 每个 Mini-batch 样本容量设置为 64 时最为理想, 算法 1 细节见表 3。

表 3 Mini-batch 梯度下降法

Table 3 Mini-batch gradient descent

Algorithm 1. Mini-batch gradient descent

**Input:** Parameters with a mini-batch**Output:** Parameters have been updated**For**  $t=1, 2, \dots$ Forward propagation on  $X^{[24]}$ 

.....

$$C^{[1]} = \Gamma_f^{[1]} \cdot C^{[0]} + \Gamma_u^{[L]} \cdot \tilde{C}^{[1]}$$

$$A^{[1]} = \Gamma_o^{[1]} \cdot \tanh(C^{[1]})$$

.....

$$A^{[L]} = \Gamma_o^{[L]} \cdot \tanh(C^{[L]})$$

$$\text{Compute cost } J^{(t)} = \frac{1}{m} \sum_{i=1}^m l(\hat{y}^{(i)}, y^{(i)})$$

Back propagation to compute gradients of  $J^{(t)}$  using  $(X^{(t)}, Y^{(t)})$ 

Using Adam algorithm to update parameters

深度学习模型包含多隐层网络结构, 在训练过程中, 每个隐含层都会出现 Covariate-shift 问题, 即每个隐层的输入分布不断变化, 整体分布逐渐往非线性函数的取值区间上下限两端靠近, 导致反向传播时出现梯度消失, Batch normalization(简称 BN)算法可使输入值落入激活函数的线性激活区域, 以此避免梯度消失的问题, 本文将 BN 算法引入长短时记忆网络, 有效提升了模型收敛速度, 一定程度上提高了模型性能, 且降低了模型超参数调试难度。在模型训练过程中:

$$\mu = \frac{1}{m} \sum_i c^{(i)} \quad (17)$$

$$\sigma^2 = \frac{1}{m} \sum_i (c^{(i)} - \mu)^2 \quad (18)$$

$$c_{\text{normal}}^{(i)} = \frac{c^{(i)} - \mu}{\sqrt{\sigma^2 + \varepsilon}} \quad (19)$$

$$\tilde{c}^{(i)} = \gamma z_{\text{normal}}^{(i)} + \beta \quad (20)$$

$$a^{(t)} = \Gamma_o^{(t)} \cdot \tanh(\tilde{c}^{(t)}) \quad (21)$$

式中:  $c^{(i)}$  为 LSTM 记忆单元携带的信息;  $m$  为 Mini-batch 的大小。

在模型测试过程中, 无法对测试集采用 Mini-batch 来计算均值  $\mu$  和方差  $\sigma^2$ , 为了将模型用于测试过程, 本文利用指数加权平均来估算测试样本的均值和方差, 即通过训练集每个 Mini-batch 的均值和方差, 利用式计算出其指数加权移动平均值, 以此来替代测试样本中的均值和方差。相比其他均值估计算法, 指数加权平均可以直接将数据带入公式, 不断覆盖旧数据, 极大的减少内存, 提高效率。

$$v_i = \beta v_{i-1} + (1 - \beta) \theta_i \quad (22)$$

### 2.3 梯度检验

深度学习模型的训练、验证、预测过程包含大量的计算，无法保证模型在反向传播过程中的所有细节都准确无误。在对模型进行反复训练之后，通过梯度检验用来验证模型的反向传播过程是否正确，把所有参数转换成一个向量数据，采用双边误差，对梯度做数值逼近，计算其欧几里得范数并进行归一化与设定的误差值相比较。算法 2 细节见表 4。

表 4 梯度检验

Table 4 Gradient checking

Algorithm 2. Gradient checking

**Input:** Vectors with parameters

**Output:** credible or questionable

Take  $W^{[1]}, b^{[1]}, \dots, W^{[L]}, b^{[L]}$  and reshape into a big vector  $\theta$ .

$$J(\theta) = J(W^{[1]}, b^{[1]}, \dots, W^{[L]}, b^{[L]})$$

Take  $dW^{[1]}, db^{[1]}, \dots, dW^{[L]}, db^{[L]}$  and reshape into a big vector  $d\theta$ .

$$\varepsilon = 10^{-7}$$

For  $i$  in epochs:

$$d\theta_{\text{approx}}^{[i]} = \frac{J(\theta_1, \theta_2, \theta_i + \varepsilon, \dots) - J(\theta_1, \theta_2, \theta_i - \varepsilon, \dots)}{2\varepsilon}$$

$$d\theta^{[i]} = \frac{\partial J}{\partial \theta_i}$$

**Check**

$$\frac{\|d\theta_{\text{approx}} - d\theta\|_2}{\|d\theta_{\text{approx}}\|_2 + \|d\theta\|_2} \text{ and } \varepsilon$$

**If** both values are approximately equal

The calculation process is credible

**Else**

The calculation process is questionable

### 2.4 模型评价方法

为了更好地评价模型性能，除了计算训练集/

验证集/测试集的预测准确率来衡量模型的表现，本文还采取了 F1 分数来作为分类模型评估指标，它是精确率和召回率的调和平均数，最大为 1，最小为 0。

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (23)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (24)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (25)$$

式中：TP(True positive)表示预测答案正确；FP(False positive)表示错将他类预测为本类；FN(False negative)表示将本类标签预测为其他类标签。最终将求得各个类别下的 F1 分数求均值，得到最后的评测结果。

## 3 实验验证

本文数据来源于眉山市博眉启明星铝业有限公司 300 kA 系列电流铝电解槽。采用 X 射线荧光光谱仪测定硅含量、铁含量和分子比。采用热电偶温度计测得铝电解槽槽内温度并通过人工标记硅含量、铁含量、分子比、温度等参数相比于前一天的变化量。从铝电解生产线槽控系统中采集加料次数及实时电压数据。实验数据的获取时间为 2018 年 8 月 1 日至 2019 年 8 月 1 日。具体实验设置如表 5 所示。

表 5 实验设置

Table 5 Experimental settings

Item	Value
Time range of data acquisition	From 1 August 2018 to 1 August 2019
Number of cells	21
Number of features	17
Experimental samples	3671
Normal samples	3036
Anode effect samples	635
Training set:Development set: Test set ratio	6:2:2

### 3.1 模型性能验证

本文共采集了 3671 个样本, 其中正常样本 3036 个, 效应样本 635 个, 所有样本均匀分布。按照 6:2:2 的比例分为训练集、验证集和测试集, 首先将训练集用于模型训练, 验证集用于模型性能验证, 并对模型进行修正, 其过程如图 5 和 6 所示, 在初始迭代时, 训练集和验证集的预测准确率相差较大, 主要由于模型初次接触样本时未习得数据特征, 所以训练集的初始预测准确率较差。经过 130 次左右迭代损失和准确率趋于稳定, 达到了较低的损失和较高的预测准确率。将测试集作为模型的最终的性能评估, 部分预测结果如图 7 所示, 模型分类准确率达到 94.65%, F1 分数达到 0.9317。

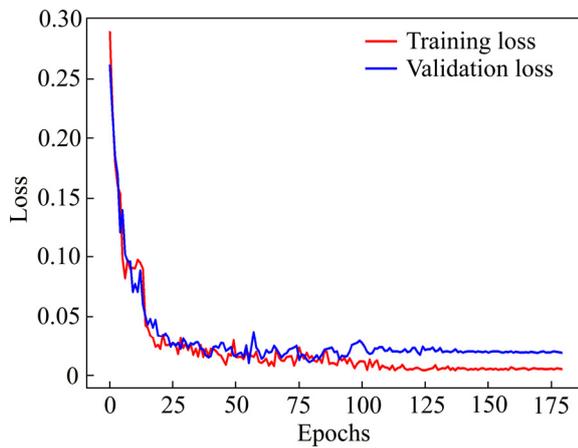


图 5 模型训练集、验证集损失下降过程  
Fig. 5 Model training set/development set loss changes

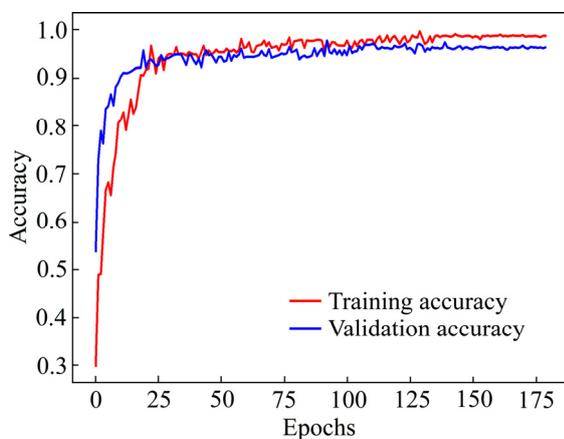


图 6 模型训练集、验证集预测准确率变化过程  
Fig. 6 Model training set/development set accuracy changes

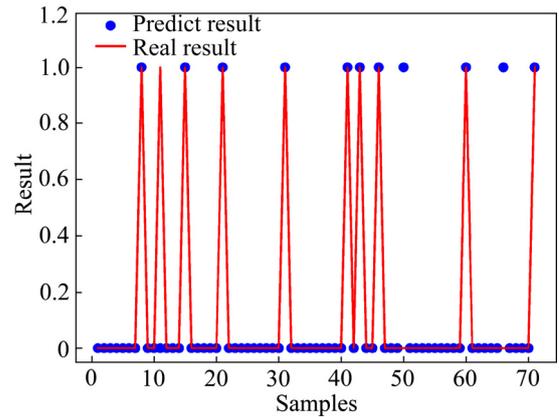


图 7 模型部分预测结果  
Fig. 7 Partial prediction results of LSTM model

### 3.2 不同方法对比

传统阳极效应预测模型都采用基于数据驱动思想的机器学习方法, 且模型输入特征参数较少, 一定程度上造成了关键特征信息的损失, 为了验证基于深度学习的长短时记忆网络模型的性能, 本文另外单独构建了 BP 神经网络模型和广义回归网络模型, 输入同样的数据对三个模型性能进行对比验证, 采用测试集准确率及 F1 分数作为评估指标, 实验结果表明, 长短时记忆网络模型不论在预测准确率还是在 F1 分数均远高于传统机器学习模型, 如图 8 所示。

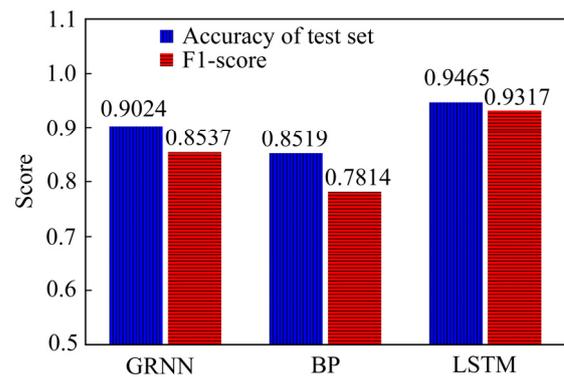


图 8 LSTM 模型与传统模型性能对比  
Fig. 8 Performance comparison between LSTM model and traditional models

### 3.3 效应提前预测时间

为了确定提前预报时间, 本文分别研究了效应前 4、8、12、16、20、24、28 和 32 min 所采集的样本数据, 对于实时监测数据, 每隔 2 min 采集一

次电压数据,共采集10次,如研究效应前10 min的阳极效应预报准确率,就需要从效应前半小时的电压开始采集。同样采用预报准确率和F1分数作为评价指标,实验结果如图9所示。可以看出随着预测时间增加,预报准确率和F1分数均随之下降,在前16 min内模型性能相差不大,但更多的提前预报时间更有利于现场工人提前处理,所以16 min的提前预报时间更为合理。

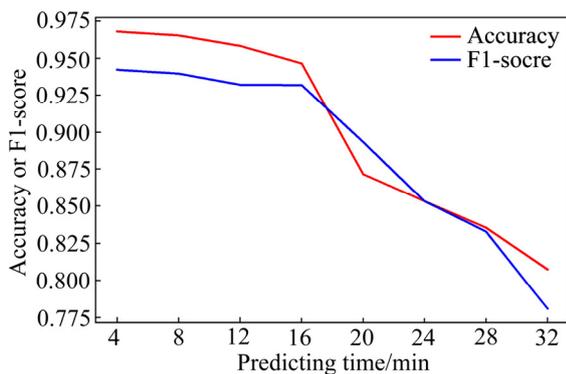


图9 预测准确率和F1分数随提前预测时间变化

Fig. 9 Prediction accuracy and F1-score change with prediction time

### 3.4 现场实验

为了进一步验证基于深度学习的长短时记忆网络在阳极效应预报任务中的有效性,本文选取眉山市眉眉启明星铝业有限公司300 kA系列铝电解生产线中的4个未用于模型训练、验证、测试的铝电解槽进行实验,如图10所示。其中单个电解槽现场抽取效应样本和正常样本共计50份,具体样本情况及预测结果如表6所示。现场实验结果显示,模型性能与仿真实验结果相符。



图10 阳极效应预测模型现场实验

Fig. 10 Field experiment of anode effect prediction model

表6 现场实验样本详情及预测结果

Table 6 Sample details and prediction results

Number of cell	Effect sample	Normal sample	Accuracy	F1-score
1	9	41	0.92	0.9583
2	11	39	0.88	0.9362
3	7	43	0.92	0.9583
4	9	41	0.90	0.9474

## 4 讨论与分析

本文研究了基于深度学习方法的长短时记忆网络模型在阳极效应预测任务上的表现,实验结果表明,相比于传统机器学习方法,深度学习模型在预测准确率及提前预测时间上的表现更为出色,机器学习方法在模型调试过程中缺少统一的方法同时降低模型训练过程中产生的偏差和方差,导致模型训练、验证、测试效果不理想,从而直接影响模型的性能表现,而深度学习方法具有更大的网络结构以及更多的数据,可显著减少模型的调试优化的工作量,并且能很好的平衡偏差和方差问题。此外,阳极效应故障特征参数具有累积性和前后关联性,而长短时记忆网络擅长于处理时序数据和前后关联数据,能很好的找出槽况参数变化趋势和阳极效应之间的联系。因此,模型在阳极效应预测任务上具有很好的表现。

同时本文在长短时记忆网络模型中引入了Batch normalization算法,提高了模型的稳定性和收敛速度,本文采用了Mini-batch优化算法,由于每个Mini-batch所计算出的标准偏差和方差带有一部分噪声,BN算法在对标准偏差进行缩放和减去均值的同时也带入了部分噪音,从而使每层隐含层的激活值增加了噪音,使得后层网络的单元不会过分依赖任何一个隐藏层单元,这在一定程度上对模型有正则化作用。

对于效应提前预测时间,本文研究发现,随着预测时间的增加,预测准确率和F1分数都会随之下降,但在ZHOU等<sup>[9]</sup>采用极端梯度提升算法得到的研究结果中,效应预测准确率随着时间变化先增加后减少,在第10 min达到峰值,而在PAN等<sup>[11]</sup>的研究中,预测准确率在50 min内处于稳定波动的状态,维持在90%左右的准确率。这主要因为不同

算法所提取的关键特征不同,从而导致数据在不同时间节点与阳极效应的关联性不同。

## 5 结论

1) 长短时记忆网络阳极效应预测模型(LSTM)在阳极效应预测任务上实现 94.65%的预测准确率和 0.9317 的 F1 分数,相比于传统机器学习预测方法,其准确率和 F1 分数分别提高了 4.9%和 9.1%。

2) 在提前预测时间上,本文研究了模型在不同预报提前量的预测准确率和 F1 分数,综合考虑模型性能和现场工人处理时间两个因素,最终确定模型提前预测时间为 16 min。

3) 此外,本文在模型调试优化过程中引入了 Batch Normalization 算法和梯度检验,提高了模型的稳定性和收敛速度。

4) 最后通过现场实验,选取 4 个未用于模型训练、验证、测试的铝电解槽进行实验,进一步验证了方法的有效性和可行性。

## REFERENCES

- [1] DION L, MARKS J, KISS L I, et al. Quantification of perfluorocarbons emissions during high voltage anode effects using non-linear approach[J]. *Journal of Cleaner Production*, 2017, 164: 357–366.
- [2] 陈建军, 陈善永. 降低铝电解槽阳极效应生产实践[J]. *中国有色金属*, 2017, 27(S2): 387–389.  
CHEN Jian-jun, CHEN Shan-yong. Production practice of reducing anode effect in aluminum electrolyzer[J]. *China Nonferrous Metals*, 2017, 27(S2): 387–389.
- [3] KONG L S, YU C J, TEO K L, et al. Robust real-time optimization for blending operation of alumina production[J]. *Journal of Industrial and Management Optimization*, 2017, 13(3): 1149–1167.
- [4] BAZHIN V Y, VLASOV A A, LUPENKOV A V. Controlling the anode effect in an aluminum reduction cell[J]. *Metallurgist*, 2011, 55: 463–468.
- [5] HARLEY J B, MOURA J M F. Data-driven matched field processing for Lamb wave structural health monitoring[J]. *Journal of the Acoustical Society of America*, 2014, 135(3): 1231–1244.
- [6] ZHOU K B, XU G F, GUO S H, et al. Anode effect prediction based on support vector machine and K nearest neighbor[C]//2017 Chinese Automation Congress, 2017: 341–345.
- [7] 田晔非, 翟渊. 基于广义回归神经网络的阳极效应自动预测[J]. *电子器件*, 2018, 41(5): 1297–1295.  
TIAN Ye-fei, ZHAI Yuan. Automatic prediction of anode effect based on generalized regression neural network[J]. *Journal of Electron Devices*, 2018, 41(5): 1297–1295.
- [8] 李翘楚, 潘浩, 陈晓冉. 基于改进交叉熵算法的阳极效应预测方法[J]. *国外电子测量技术*, 2019, 38(9): 142–146.  
LI Qiao-chu, PAN Hao, CHEN Xiao-ran. Anode effect prediction method based on improved cross entropy algorithm[J]. *Foreign Electronic Measurement Technology*, 2019, 38(9): 142–146.
- [9] ZHOU K B, ZHANG Z X, LIU J, et al. Anode effect prediction based on a singular value thresholding and extreme gradient boosting approach[J]. *Measurement Science and Technology*, 2019, 30(1): 015104.
- [10] CHEN Z G, LI Y G, CHEN X F, et al. Anode effect prediction based on collaborative two-dimensional forecast model in aluminum electrolysis production[J]. *Journal of Industrial and Management Optimization*, 2019, 15(2): 595–618.
- [11] PAN H, KONG L, CHEN X R, et al. A modified neighborhood mutual information and light gradient boosting machine-based long-term prediction approach for anode effect[J]. *Measurement Science and Technology*, 2019, 30(11): 115105.
- [12] 任浩, 屈剑锋, 柴毅, 等. 深度学习在故障诊断领域中的研究现状与挑战[J]. *控制与决策*, 2017, 32(8): 1345–1358.  
REN Hao, QU Jian-feng, CHAI Yi, et al. Research status and challenges of deep learning in the field of fault diagnosis[J]. *Control and Decision*, 2017, 32(8): 1345–1358.
- [13] TAMILSELVAN P, WANG P F. Failure diagnosis using deep belief learning based health state classification[J]. *Reliability Engineering & System Safety*, 2013, 115(6): 124–135.
- [14] YANG Z X, WANG X B, ZHONG J H. Representational learning for fault diagnosis of wind turbine equipment: A multi-layered extreme learning machines approach[J]. *Energies*, 2016, 9(6): 397.
- [15] LEE H, KIM Y, KIM C O. A deep learning model for robust wafer fault monitoring with sensor measurement noise[J]. *IEEE Transactions on Semiconductor Manufacturing*, 2017, 30(1): 23–31.
- [16] 唐赛, 何苻兮, 张家悦, 等. 基于长短期记忆网络的轴承故障识别[J]. *汽车工程学报*, 2018, 8(4): 297–303.

- TANG Sai, HE Xing-xi, ZHANG Jia-yue, et al. Bearing fault identification based on long short-term memory networks[J]. Chinese Journal of Automotive Engineering, 2018, 8(4): 297–303.
- [17] PANG S, YANG X Y, ZHANG X F. Aero engine component fault diagnosis using multi-hidden-layer extreme learning machine with optimized structure[J]. International Journal of Aerospace Engineering, 2016, 2016: 1329561.
- [18] CHANG W J, CHEN L B, HSU C H, et al. A deep learning-based intelligent medicine recognition system for chronic patients[J]. IEEE Access, 2019, 7(1):44441–44458.
- [19] 魏立新, 魏新宇, 孙浩, 等. 基于深度网络训练的铝热轧制力预报[J]. 中国有色金属学报, 2018, 28(10): 2070–2076.
- WEI Li-xin, WEI Xin-yu, SUN Hao, et al. Prediction of aluminum hot rolling force based on deep network[J]. The Chinese Journal of Nonferrous Metals, 2018, 28(10): 2070–2076.
- [20] YANG F, DONG H L, WANG Z D, et al. A new approach to non-fragile state estimation for continuous neural networks with time-delays[J]. Neurocomputing, 2016, 197: 205–211.
- [21] 王鑫, 吴际, 刘超, 等. 基于LSTM循环神经网络的故障时间序列预测[J]. 北京航空航天大学学报, 2018, 44(4): 772–784.
- WANG Xin, WU Ji, LIU Chao, et al. Exploring LSTM based recurrent neural network for failure time series prediction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(4): 772–784.
- [22] YIN C L, ZHU Y F, FEI J L, et al. A deep learning approach for intrusion detection using recurrent neural networks[J]. IEEE Access, 2017, 5: 21954–21961.
- [23] DE BRUIN T, VERBERT K, BABUSKA R. Railway track circuit fault diagnosis using recurrent neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(3): 523–533.
- [24] DION L, KISS L I, PONCSÁK S, et al. Prediction of low-voltage tetrafluoromethane emissions based on the operating conditions of an aluminium electrolysis cell[J]. JOM: The Journal of the Minerals, Metals & Materials Society, 2016, 68(9): 2472–2482.

## Anode effect prediction of 300 kA aluminium electrolysis cell based on deep learning

YIN Gang<sup>1</sup>, CHEN Gen<sup>1</sup>, HE Wen<sup>2</sup>, YAN Fei-ya<sup>3</sup>, LUO Bin<sup>4</sup>, LI Rui<sup>2</sup>

- (1. State Key Laboratory of Coal Mine Disaster Dynamics and Control, College of Resource and Safety Engineering, Chongqing University, Chongqing 400044, China;
2. Bomei Qimingxing Aluminum Co., Ltd., Meishan 620010, China;
3. Guiyang Aluminum Magnesium Design & Research Institute Co., Ltd., Guiyang 550000, China;
4. Sichuan Siwei Environmental Protection Equipment Co., Ltd., Suining 629000, China)

**Abstract:** The anode effect mechanism and fault parameters of aluminium electrolytic cells were studied, and a deep learning-based anode effect prediction method was proposed. It can adapt to the parameters of tank conditions in different dimensions and different data characteristics, and directly mine fault characteristic information from massive raw data. It greatly reduces the response time of the effect, has good robustness and anti-interference ability. At the same time, in the model debugging optimization, the Batch Normalization algorithm and gradient test are used to improve the model convergence speed and stability. The prediction accuracy and F1 score reach 94.65% and 0.9317, respectively. The prediction time can reach 16 min, and it is verified by field experiments to meet the actual production requirements.

**Key words:** aluminium electrolysis; 300 kA; anode effect prediction; deep learning; optimization

**Foundation item:** Project(50728504) supported by the National Natural Science Foundation of China

**Received date:** 2020-02-25; **Accepted date:** 2020-06-26

**Corresponding author:** YIN Gang; Tel: +86-23-65104511; E-mail: yk115@cqu.edu.cn

(编辑 何学锋)