

基于 SMoE 模型和电池融合数据的 SOH 和 RUL 联合预测

常 伟, 胡志超, 潘多昭, 师继文

(南通乐创新能源有限公司, 江苏 南通 226000)

摘要: 电池的健康状况(SOH)和电池的剩余使用寿命(RUL)是衡量电池性能衰减和剩余使用时间的关键指标。预测电池 SOH 和 RUL 在实际应用中具有重要意义。通常会借助电池运行数据来训练机器学习算法,如神经网络或深度学习,以此来捕捉电池 SOH 和 RUL 的变化规律并进行预测。传统的机器学习模型往往采用单个模型来适配整个数据集,这在面对复杂且具有高度异质性的数据时显得力不从心,每一个预测目标构建一个模型,模型训练和维护成本较高。使用稀疏混合专家模型(SMoE),构建电池 SOH 和 RUL 的联合预测模型,使用电池融合数据,同时预测电池 SOH 和 RUL。在 NASA(美国国家航空和宇宙航行局)公开数据集上测试效果。结果表明,提出的联合预测模型能够很好地预测电池 SOH 和 RUL,SOH 预测值的均方误差为 0.069,RUL 预测值的均方误差为 2.042。在电化学阻抗谱(EIS)公开数据集上测试效果,联合预测模型的 SOH 预测值的均方误差为 0.118,RUL 预测值的均方误差为 3.072,准确性均有大幅度提升。

关键词: MoE(混合专家模型); SMoE(稀疏混合专家模型); SOH(健康状况); RUL(剩余使用寿命)

中图分类号: TM912.9 **文献标志码:** A **文章编号:** 1671-1807(2025)11-0091-09

电池健康状态(state of health, SOH)是衡量电池相对于新状态的性能水平的指标,用于监测电池性能、预测寿命、评估安全性、指导维护计划,并帮助用户做出更换决策,对电池的优化使用和安全管理具有重要作用。

电池剩余使用寿命(remaining useful life, RUL)是预测电池在当前状态下还能使用多久的指标,它有助于进行预防性维护、优化成本、保障安全、规划资源、提升用户体验、确保系统可靠性、改进产品设计、遵从法规要求,并对电池的回收和再利用提供指导。常见的电池 RUL 为电池剩余循环次数。

预测电池 SOH 和 RUL 具有很高的应用价值。电池的运行数据,如电流、电压和温度,与电池的健康状态(SOH)紧密相关,因此,利用这些数据来预测 SOH 是一条可行路径。同时,电池的剩余使用寿命(RUL)与电流、电阻和阻抗等参数有着显著的联系,这促使研究者将这些参数作为估算 RUL 的关键指标。通常,研究者会利用电流、电压、温度、

电阻和阻抗的数据来训练机器学习模型,如支持向量机、神经网络和深度学习等^[1-4],以便捕捉并预测电池 SOH 和 RUL 的变化趋势。

电化学阻抗谱(electrochemical impedance spectroscopy, EIS)是一种用于表征电池内部电化学过程的测试方法。如果电池活性材料减少,使用性能变差,EIS 曲线也会相应地改变,因此使用 EIS 预测 SOH 和 RUL 也是一条可行的路径^[5]。

传统机器学习模型通常依赖单一模型来拟合所有数据,难以处理复杂、异质性高的数据集。单一模型的泛化能力有限,容易过拟合。单一模型在处理大规模数据时效率较低。

混合专家模型(mixture of experts, MoE)是一种机器学习模型,它结合了多个不同的专家模型,每个专家模型在数据的某个子空间上具有较好的预测能力^[6]。

MoE 模型的基本思想是将输入数据通过一个门控网络(gating network)分配给不同的专家模型,每个专家模型负责学习数据的一个特定部分。然

收稿日期: 2024-10-16

作者简介: 常伟(1976—),男,湖南长沙人,硕士,高级工程师,研究方向为电池工业;胡志超(1984—),男,湖北黄梅人,硕士,高级工程师,研究方向为电池热失控、EIS、SOH 估计;潘多昭(1991—),男,壮族,新疆阿勒泰人,硕士,高级工程师,研究方向为储能与能源工程、AI 算法、智慧能源;师继文(1990—),男,陕西宝鸡人,硕士,中级工程师,研究方向为新能源材料。

后,这些专家模型的输出被组合起来,形成最终的预测结果。

MoE 模型通过多个专家模型(子模型)的组合,每个专家模型专注于数据的不同部分,能够更好地捕捉数据的多样性和复杂性,提高模型的整体表达能力。通过多个专家模型的组合,能够减少单一模型过拟合的风险,提高整体模型的泛化能力和鲁棒性。专家模型可以并行训练和预测,提高计算效率,尤其在处理大规模数据和复杂任务时具有优势。

稀疏混合专家模型(sparse mixture of experts, SMoE)是一种针对复杂预测问题设计的机器学习架构。该模型的核心在于采用多个“专家”子模型,每个子模型专门针对数据集中的特定部分或特定类型的输入进行处理。这些子模型的预测结果通过一个门控网络进行加权汇总,从而得出最终的预测。

SMoE 的“稀疏”特性体现在,在每次预测时,并非所有的专家子模型都会被激活,而是只有一部分被选中参与计算。这种选择性的激活机制大幅降低了计算成本,提升了模型的运算效率。通常,这种选择是通过 Top-K 策略实现的,即只挑选出权重最高的 K 个子模型来参与预测过程。

SMoE 模型的优点主要包括:①计算效率提升。通过激活少量专家网络,减少计算资源消耗。②泛化能力增强。降低模型复杂度,有效防止过拟合。③特征重要性识别。突出关键特征,帮助理解数据驱动因素。④噪声抵抗性。稀疏激活减少噪声影响,提高预测准确性。⑤动态适应性。根据输入数据调整激活专家,适应数据变化。

另外,在很多情况下,获取电池的电流、电压、温度、阻抗以及对应的 SOH 和 RUL 或者测试 EIS 以及对应的 SOH 和 RUL 时,是分别采集两套不同的数据集,虽然这些数据存在一定的关联性,但是数据的数量和维度往往不尽相同,因此每个预测目标都会训练一个单独的预测模型,数据使用效率不高,模型训练和维护成本也较高。实现 SOH 和 RUL 的联合估计可以实现数据的有效利用,降低维护成本^[7-9]。

本文提出一种基于 SMoE 模型,使用融合不同情况下的电池运行数据或者 EIS 数据以及对应的 SOH 和 RUL 数据,构建电池 SOH 和 RUL 的联合预测模型,在一个模型中使用电池融合数据,并联合预测电池 SOH 和 RUL,使得模型训练和维护成本大大降低。另外,为了提升模型的预测精度,本

文使用了多种深度学习模型,来构建专家子模型。

本文使用的深度学习模型及其作用有:①卷积神经网络(convolution neural network, CNN)。CNN 可以用于提取输入数据的空间特征。②双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)。BiLSTM 可以从正向序列和方向序列同时学习电流、电压、温度和阻抗的时间序列变化规律。

综上所述, CNN 和 BiLSTM 的组合可以学习输入数据与 SOH 和 RUL 之间的复杂的映射关系。本文使用以上两种模型的组合,构建联合预测模型,使用电池融合数据联合预测电池 SOH 和 RUL。

本文的亮点主要是:①多目标联合训练。创新性地同时对电池的健康状态(SOH)和剩余使用寿命(RUL)进行预测,这有助于模型更全面地理解电池性能的衰退过程。②多源数据融合。通过融合不同条件下获取的电池运行数据或者电化学阻抗谱(EIS)数据,能够更准确地描绘出电池的整体性能状况。

本文的创新点如下。

(1)多目标联合预测框架创新。①目标关联性建模。传统的预测模型通常单独预测 SOH 或 RUL,而该模型通过联合训练,能够同时考虑 SOH 和 RUL 之间的内在联系,这种关联性建模有助于更深入地理解电池性能退化的动态过程。②互信息最大化。在训练过程中,模型可能采用互信息最大化等技术来确保 SOH 和 RUL 预测之间的相关性得到充分利用,从而提高预测的准确性。

(2)多源数据融合技术创新。①异构数据处理:该方法分别融合了传统的电池运行数据(如电池的充放电数据)和其他电池测试数据(如 EIS 谱图),通过先进的数据预处理和特征提取技术,在不同类型的数据中均实现了数据的有效融合。②条件自适应融合。该方法可能采用条件生成模型或其他自适应机制来处理不同条件下获取的数据,确保模型在不同环境下的预测性能。

(3)SMoE 模型结构创新。①专家模型的选择与权重分配。SMoE 模型通过软选择机制动态地分配不同专家模型的权重,这种结构允许模型根据输入数据的特性自动选择最合适的专家模型进行预测。②模型容错性与多样性。通过组合多个专家模型,SMoE 结构增强了模型的容错性,即使某些专家模型在某些情况下表现不佳,其他模型也能补充其不足,保持整体预测的稳定性。

(4) CNN-BiLSTM 子专家模型组合创新。

①时空特征提取。CNN-BiLSTM 组合模型在特征提取方面具有创新性,CNN 能够有效地从数据中提取空间特征,而 BiLSTM 则能够捕捉时间序列数据的长距离依赖关系,两者的结合为模型提供了更丰富的特征表示。②端到端学习。该组合模型实现了端到端学习,从原始数据直接学习到预测输出,减少了传统特征工程的需求,提高了学习效率和模型的泛化能力。

这些创新点共同构成了一个高效、精确且鲁棒的电池 SOH 和 RUL 预测模型,为电池健康管理提供了强有力的技术支持。

综上所述,SMoE 联合预测模型兼顾了多专家、多目标、多数据源、精度、效率和鲁棒性等综合优势,是一个高性能高效率的电池 SOH 和 RUL 预测模型。本文的实用性在于,提供了一种电池 SOH 和 RUL 的联合预测方法。

1 电池 SOH 和 RUL 预测流程

电池 SOH 和 RUL 预测流程如图 1 所示。

(1)测试并采集电池数据,电池运行数据包含但不限于电流、电压、温度、电阻和 EIS 等,以及对应的 SOH 和 RUL。

(2)对不同的电池运行数据以及对应的 SOH 和 RUL 数据进行融合。

(3)构建 SMoE 模型,SMoE 模型包括多个专家子模型和一个门控网络模型;

(4)训练和校验 SMoE 模型,训练数据按 8:2 的比例分成训练集合和测试集合,在训练集合上训练模型,在测试集合上校验模型;

(5)预测电池 SOH 和 RUL,在 SMoE 模型训练完毕之后,在模型效果良好的前提下,可以部署 SMoE 模型,并使用 SMoE 对电池的 SOH 和 RUL 进行预测。

2 电池 SOH 和 RUL 预测详细过程

2.1 采集电池数据

当采集电池运行数据时,记录电池的电流、电压和温度数据,以及电池容量数据 SOH。另外,还要记录电池的电流、电阻和阻抗数据,以及电池 RUL 数据。

当采集电池 EIS 数据时,需要使用专业的阻抗谱测试仪器,并且需要覆盖一定的频率范围,频率范围可以设置为 $10^{-2} \sim 10^5$ Hz。在频率范围中间按一定间隔选取多个频率值,需要记录每个频率对应的阻抗的实部和虚部数据,以及需要预测的目标值,这里是相对应的电池 SOH 和 RUL。

2.2 数据融合

在联合预测模型中,不同来源的数据需要进行融合,以确保模型输入数据具有相同的数量。在收集电池运行数据或者 EIS 以及相应的 SOH 和 RUL 时,通常这些数据是在不同的条件下产生的,导致所得数据的数量和维度存在差异。为了构建一个联合预测模型,需要确保输入数据的总量一致,同时每个训练批次的数据量也需保持一致。因此,必须对来自不同测试环境的电池运行数据或者 EIS 数据及其相关的 SOH 和 RUL 数据进行融合。

数据融合的过程涉及以下步骤:①计算不同数据集数量的最小公倍数;②通过重复的方法将每个数据集的数量扩展至这个最小公倍数;③将不同数据集的扩展数据集按列进行拼接,得到训练数据。

2.3 构建 SMoE 模型

2.3.1 SMoE 联合预测模型

基于 SMoE 模型的使用电池融合数据预测 SOH 和 RUL 的联合预测模型结构如图 2 所示。

输入数据是电池融合数据,数据来自不同的数据集,图 2 中以 2 个输入数据集为例,也可以拓展到多个数据集。接下来是专家层和门控网络层,以 3 个 CNN-BiLSTM 专家子模型为例,以基于多层感知机(MLP)的门控网络为例。

数据分别输入到 3 个专家子模型中,从不同角度提取相应的特征。3 个专家子模型的输出分别是 output1、output2 和 output3,这 3 个输出均是高维向量,以 32 维为例,分别表示每个子专家模型对于输入数据的特征表示。同时,输入数据输入到门控网络中,用于计算每个专家子模型的权重,门控网络的输出是一个三维向量,分别对应 3 个专家子模型的权重,3 个权重之和为 1。然后,选取门控网络输出中,对应概率最高的 2 个专家子模型,将其输出的特征表示与门控网络对应输出的专家权重进行

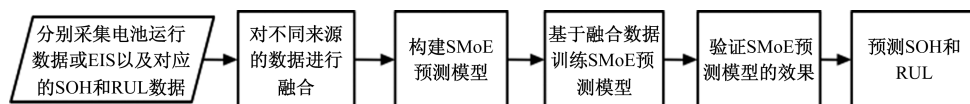
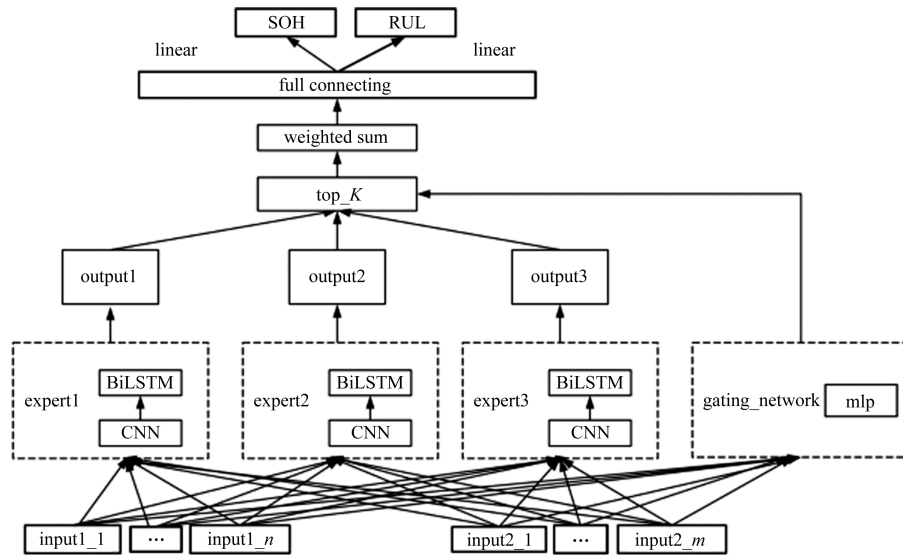


图 1 基于 SMoE 模型联合预测 SOH 和 RUL 整体流程



input₁, ..., input₂_m 表示输入数据; expert1、expert2 和 expert3 表示专家子模型; CNN 表示卷积神经网络; BiLSTM 表示双向长短期记忆网络; gating_network 表示门控网络模型; mlp 表示多层感知机模型; output1、output2 和 output3 表示专家子模型的输出值; top_K 表示激活前 K 个专家子模型; weighted_sum 表示加权求和; full_connecting 表示全连接层; linear 表示线性回归

图2 基于 SMoE 模型联合预测 SOH 和 RUL 的模型结构

加权求和,得到综合特征表示。之后,将综合特征连接到一个全连接层。最后,通过线性回归层,同时输出 SOH 和 RUL。

2.3.2 SMoE 专家子模型和门控网络模型

SMoE 模型包括多个专家子模型和一个门控网络模型。

1) 专家子模型

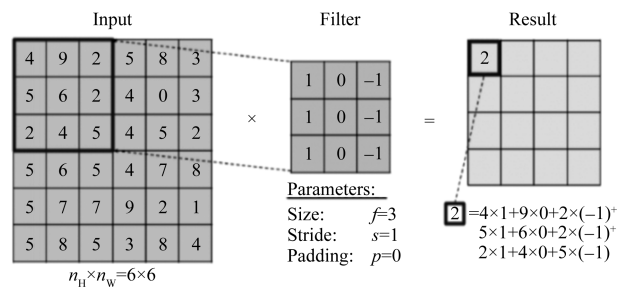
每个子模型均是 CNN-BiLSTM 模型。CNN 可以自动学习输入数据的空间相关性,提取高层抽象特征。本文的 CNN 结构包含 3 个卷积层,卷积层的激活函数均为 relu 函数,过滤器数量设置为 32 个,核的尺寸为 1。

图3的卷积层中,有一个 $1 \times 6 \times 6$ 的矩形作为输入数据,称为输入特征图 Input,它经过一个 3×3 的滤波器 Filter(也被称为卷积核)进行卷积运算,即从 Input 左上角 3×3 的受野区开始计算其张量积,每完成一次计算向右移动,步幅为 1,完成此计算后,就会得出一个 $1 \times 4 \times 4$ 的输出特征图,即 Result。

BiLSTM 由前向 LSTM 和后向 LSTM 组成。LSTM 是一种特殊的 RNN,可以学习长期依赖关系,捕捉时间序列数据的动态变化。

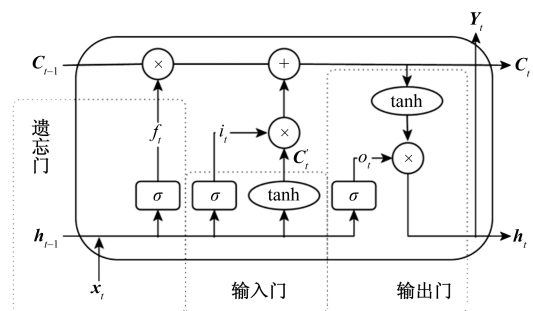
LSTM 模型的结构如图4所示,组成包含:

①输入门。决定当前时刻输入信息中的哪些部分允许进入状态单元。②遗忘门。决定状态单元中的哪些信息需要被遗忘或者移除。③状态单元。用



Input 表示输入数据; Filter 表示滤波器; Result 表示输出结果

图3 卷积计算示意图



C_t 表示记忆单元,通过门控机制选择性保留或遗忘信息; f_t 表示遗忘门输出,控制旧信息的保留程度; i_t 表示输入门输出,控制新信息的添加程度; C'_t 表示候选细胞状态,存储当前时刻的临时记忆; h_t 表示 LSTM 的输出向量,用于传递到下一时间步或作为最终输出; O_t 表示输出门输出,控制隐藏状态的生成; σ 表示 Sigmoid 函数,输出值在 (0,1) 之间; tanh 表示双曲正切函数,输出值在 (-1,1) 之间; $\oplus$ 表示求和; $\otimes$ 表示乘积

图4 LSTM 模型结构

于存储历史状态信息的核心组件。④输出门。决定输出到下一个隐层的信息。⑤激活函数 (Sigmoid、

tanh)。用于进行非线性转换。⑥输出层。与输入数据相对应的 LSTM 输出数据。

LSTM 通过输入门、遗忘门及输出门的开闭,来控制信息进入、保留和输出的流动,实现时序信息的记忆及处理。相比简单的 RNN,LSTM 更能捕捉时序数据中的长期依赖,处理较长的时间序列,拥有更强大的建模能力。

BiLSTM 由前向 LSTM 和后向 LSTM 组成,包含前向隐状态和后向隐状态。在同一时刻,前向 LSTM 处理正向序列,后向 LSTM 处理反向序列,最后通过拼接或者添加两者的隐状态作为该时刻的输出状态。

图 5 中 X 表示输入量,BiLSTM 由前向 LSTM 与后向 LSTM 组合而成,最终每个时间步 i 输出 $h_i = [h_{fi} \cdot h_{bi} - i]$, $[\cdot]$ 表示拼接操作,即如果 LSTM 的隐含层维度是 16,那么 BiLSTM 的维度则为 32。

BiLSTM 的优点:①可以访问更多上下文信息,对序列模式理解更全面;②理论上可以处理任意长度的远距离依赖关系;③训练效果一般优于单向 LSTM,解决梯度消失问题;④对序列数据具有很强的特征提取能力;⑤能够有效学习数据的时间上下文;⑥对异构时间序列具有更好的拟合性能。

本文的 BiLSTM 层的神经元个数为 32,BiLSTM 层可以学习输入数据的长期依赖关系,识别时间序列的动态变化模式。

2) 门控网络模型

门控网络模型采用多层感知机模型,门控网络模型结构如图 6 所示。

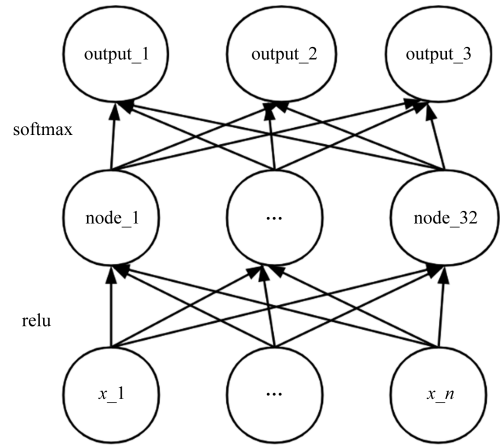
门控网络模型采用多层感知机模型,为一个 3 层的网络结构,输入层接收输入数据,中间层为

32 维向量,激活函数为 relu 函数,输出层是一个 3 维向量,激活函数为 softmax,分别对应 3 个专家子模型的权重,3 个权重之和为 1。

3 个专家子模型的输出分别是 output1、output2 和 output3,这 3 个输出均是 32 维向量,分别表示每个子专家模型对于输入数据的特征表示。

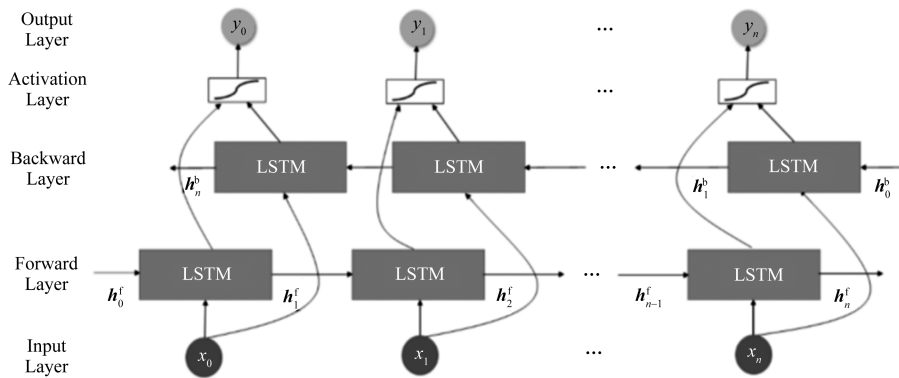
选取概率最高的 2 个专家子模型与门控网络模型对应的权重进行加权求和,得到综合特征表示。

在综合特征表示之后是一个全连接层,激活函数为 linear 函数,全连接层之后是两个线性回归模型,预测值分别是电池容量即 SOH 和剩余循环次数即 RUL,激活函数均为 linear 函数。模型的损失函数为均方误差(mean square error,MSE);优化器 optimizer 为 rmsprop;评估指标 metrics 为平均绝对误差(mean absolute error,MAE),即目标值和预测值之差的绝对值之和。



$x_1, \dots, x_n$ 表示输入数据;relu 表示激活函数;node_1, ..., node_32 表示隐藏节点;softmax 表示激活函数;output_1, output_2 和 output_3 表示输出值

图 6 门控网络模型结构



Input Layer 表示输入层; $x_0, x_1, \dots, x_n$ 表示输入数据;Forward Layer 表示前向层, h_i^f 表示前向 LSTM 的第 i 步输出向量;Backward Layer 表示后向层; h_{n-i}^b 表示后向 LSTM 的第 $n-i$ 步输出向量;Activation Layer 表示激活层;Output Layer 表示输出层; $y_0, y_1, \dots, y_n$ 表示输出数据

图 5 BiLSTM 模型结构

2.4 训练和校验 SMoE 模型

本文在 Keras 平台实现 SMoE 模型代码并训练,模型训练的优化器 optimizer 为 adam 算法,损失函数为均方误差(MSE)。

训练数据按 8:2 的比例分成训练集合和测试集合,在训练集合上训练模型,在测试集合上校验模型。

2.5 预测电池 SOH 和 RUL

在 SMoE 模型训练完毕之后,在模型效果良好的前提下,可以部署 SMoE 模型,并使用 SMoE 模型对电池的 SOH 和 RUL 进行预测。

3 预测效果展示

3.1 基于电池运行数据的电池 SOH 和 RUL 联合预测效果

3.1.1 电池运行数据

本文的测试数据来源有两个:一是 NASA 电池数据集中的 B0018 电池的放电(type=discharge)阶段的数据,使用其中的电池电流、电压、温度和 SOH;二是 NASA 电池数据集中的 B0018 电池的阻抗状态下(type=impedance)的数据,使用其中的电池电流、电阻、阻抗和 RUL。其中,电池电流、电压、温度和 SOH 数据有 34 866 条,电池电流、电阻、阻抗和 RUL 数据有 2 544 条。电池电流、电压、温度和 SOH 数据共有 5 个特征和 1 个预测目标,特征和预测目标的名称和含义见表 1。预测目标 Capacity 表示电池容量即 SOH。具体的数据样例见表 2。

电池电流、电阻、阻抗和 RUL 测试数据共有 7 个特征和 1 个预测目标,名称和含义见表 3。具体的数据样例见表 4。

表 1 电池电流电压温度和 SOH 数据的特征名称和含义

特征名称	含义	单位
Voltage_measured	测量的电压	V
Current_measured	测量的电流	A
Temperature_measured	测量的温度	℃
Current_charge	在负载下测量的电流	A
Voltage_charge	在负载下测量的电压	V
Capacity	放电至 2.7 V 的电池容量	A·h

表 2 电池运行数据和 SOH 样例

Voltage_measured	Current_measured	Temperature_measured
4.188 108 651	0.000 130 667	23.819 520 25
4.188 195 943	0.001 459 081	23.828 807 16
...	...	...
3.477 276 739	-0.001 940 069	34.581 659 54
Current_charge	Voltage_charge	Capacity
0.0006	0	1.855 004 521
0.0006	4.203	1.855 004 521
...	...	...
0.0006	0	1.341 051 441

表 3 电池电流电阻阻抗和 RUL 数据的特征名称和含义

特征名称	含义
Sense_current	传感器支路电流
Battery_current	电池支路电流
Current_ratio	传感器支路电流与电池支路电流的比率
Battery_impedance	根据原始数据计算的电池阻抗
Rectified_impedance	校准和平滑的电池阻抗
Re	估计电解液电阻
Rct	估计充电转移电阻
RUL	循环次数

表 4 电池运行数据和 RUL 样例(部分)

Sense_current_real	Sense_current_imag	Battery_current_real	
832.749 572 8	−27.377 193 45	327.527 618 4	
826.062 866 2	−41.281 600 95	331.814 27	
.....			
803.516 357 4	443.114 929 2	98.716 178 89	
Battery_current_imag	Current_ratio_real	Current_ratio_imag	
−90.784 202 58	2.382 645 118	0.576 834 847	
−78.227 035 52	2.386 236 081	0.438 156 484	
.....			
158.175 430 3	4.297 792 306	−2.397 684 157	
Battery_impedance_real	Battery_impedance_imag	Rectified_impedance_real	
0.165 409 535	−0.129 054 344	0.095 314 82	
0.165 409 535	−0.129 054 344	0.095 314 82	
.....			
0.164 978 108	−0.126 231 668	0.091 522 433	
Rectified_impedance_imag	Re	Rct	RUL
−0.001 691 894	0.065 158 152	0.095 553 696	1
−0.001 691 894	0.065 158 152	0.095 553 696	2
.....			
−0.002 372 766	0.066 068 503	0.088 958 851	314

注:Sense_current、Battery_current、Current_ratio、Battery_impedance 和 Rectified_impedance 等特征均分成了实部数据和虚部数据,以“_real”结尾的特征是实部数据,以“_imag”结尾的特征是虚部数据。

3.1.2 电池运行数据融合

分别将以上两套数据集按照 8:2 的比例分成训练集合和测试集合,分别得到电池电流电压温度和 SOH 的 27 892 条训练数据和 6 974 条测试数据,得到电池电流电阻阻抗和 RUL 的 2 035 条训练数据和 509 条测试数据。

为了保证训练数据和测试数据不存在交集,因此,分别在两套训练集合和测试集合上进行数据整合。

电池电流电压温度和 SOH 的 27 892 条训练数据与电池电流、电阻、阻抗和 RUL 的 2 035 条训练数据的最小公倍数是 56 760 220 条,因此分别对训练集合进行重复,使得各自的训练集合数量增加至

56 760 220 条。

电池电流、电压、温度和 SOH 的 6 974 条测试数据与电池电流、电阻、阻抗和 RUL 的 509 条测试数据的最小公倍数是 3 549 766 条,因此分别对测试集合进行重复,使得各自的测试集合数量增加至 3 549 766 条。

3.1.3 SOH 预测结果

电池容量的测试集合的均值为 1.584 5,标准方差为 0.156 3,预测集合的均值为 1.589 3,标准方差为 0.138 4,这表明预测结果的数据分布与真实数据的数据分布基本一致。

电池 SOH 的预测集合和测试集合的均方根误差只有 0.069 33,说明模型的预测效果良好。预测结果见表 5。预测结果如图 7 所示。从图 7 中可以看出,预测效果良好。

3.1.4 RUL 预测结果

电池循环次数的测试集合的均值为 142.369 7,标准方差为 98.558 1,预测集合的均值为 141.544 1,标准方差为 98.873 9,预测结果的数据分布与真实数据的数据分布基本一致。

电池 RUL 的预测集合和测试集合的均方根误差为 2.042,说明模型在预测 RUL 时,预测效果良好。预测结果见表 6。预测结果如图 8 所示。从图 8 中可以看出,预测效果良好。

表 5 基于 SMoE 模型和电池融合数据的电池 SOH 预测结果

真实值	预测值	误差
1.771 2	1.806 0	-0.034 8
1.415 5	1.390 8	0.024 7
1.694 0	1.710 3	-0.016 2
1.501 2	1.579 0	-0.077 8
1.695 8	1.692 1	0.003 7
1.393 5	1.440 6	-0.047 1

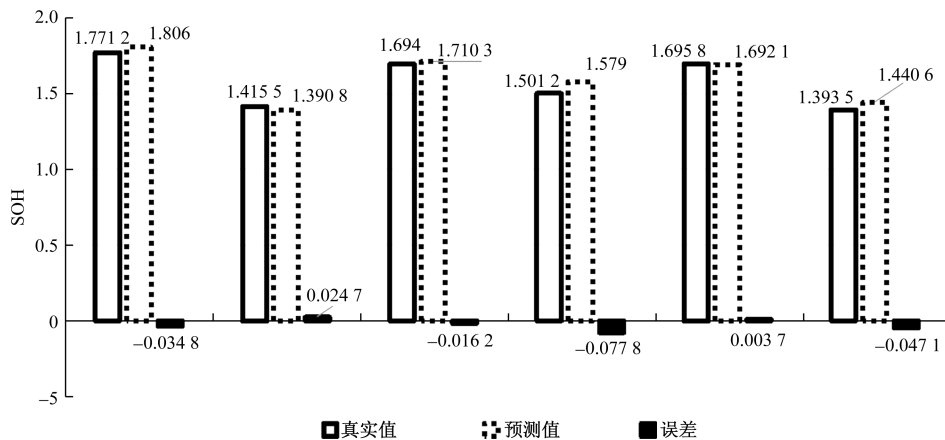


图 7 基于 SMoE 模型和电池融合数据的电池 SOH 预测结果样例

表 6 基于 SMoE 模型和电池融合数据的电池 RUL 预测结果

真实值	预测值	误差
195	197	-2
25	24	1
242	242	0
49	49	0
1	1	0
55	54	1

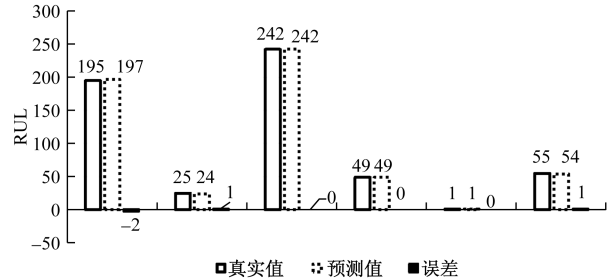


图 8 基于 SMoE 模型和电池融合数据的电池 RUL 预测结果

3.2 基于 EIS 的电池 SOH 和 RUL 联合预测效果

3.2.1 EIS 数据

本文在公开数据集合上测试效果,使用其中的阻抗数据来预测电池 SOH 和 RUL。其中,电池 EIS 和 SOH 数据有 1 358 条,电池 EIS 和 RUL 数据有 525 条。EIS 均有 120 个特征,特征序号依次为 f1~f120,前 60 个特征是频率从高到低对应的阻抗数据的实部数据,后 60 个特征是频率从高到低对应的阻抗数据的虚部数据。电池 EIS 和 SOH 数据集合、电池 EIS 和 RUL 数据集见表 7、表 8。

表 7 电池 EIS 和 SOH 数据集合(部分)

f1	f2	...	f120	SOH
0.384 7	0.391 56	...	0.327 95	37.202 7
0.388 86	0.393 43	...	0.323 00	36.223 0
0.390 38	0.396 00	...	0.329 55	35.589 3
0.391 94	0.396 43	...	0.336 73	35.108 1
0.392 25	0.397 60	...	0.325 62	34.767 3

表 8 电池 EIS 和 RUL 数据集(部分)

f1	f2	...	f120	RUL
0.384 7	0.391 56	...	0.327 95	234
0.388 86	0.393 43	...	0.323 00	232
0.390 38	0.396 00	...	0.329 55	230
⋮	⋮	⋮	⋮	⋮
0.751 14	0.757 14	...	0.296 80	4
0.753 21	0.758 79	...	0.298 27	2
0.752 11	0.760 74	...	0.291 80	0

3.2.2 EIS 数据融合

分别将以上两套数据集按照 8:2 的比例分成训练集和测试集,分别得到预测 SOH 的 1 086 条训练数据和 272 条测试数据,得到预测 RUL 的 420 条训练数据和 105 条测试数据。为了保证训练数据和测试数据不存在交集,分别在两套训练集和测试集上进行数据融合。

预测 SOH 的 1 086 条训练数据和预测 RUL 的 420 条训练数据的最小公倍数是 76 020 条,因此分别对训练集进行上采样,使得各自的训练集数量增加至 76 020 条。

预测 SOH 的 272 条测试数据和预测 RUL 的 105 条测试数据的最小公倍数是 28 560 条,因此分别对测试集进行上采样,使得各自的测试集数量增加至 28 560 条。

3.2.3 SOH 预测结果

电池 SOH 的测试集的均值为 30.593 2,标准方差为 4.128 4,预测集的均值为 30.598 7,标准方差为 4.077 5,这表明预测结果的数据分布与真实数据的数据分布基本一致。

电池 SOH 的预测集和测试集的均方根误差只有 0.118,说明模型的预测效果良好。预测结果表 9 和图 9 所示。从图 9 中可以看出,预测效果良好。

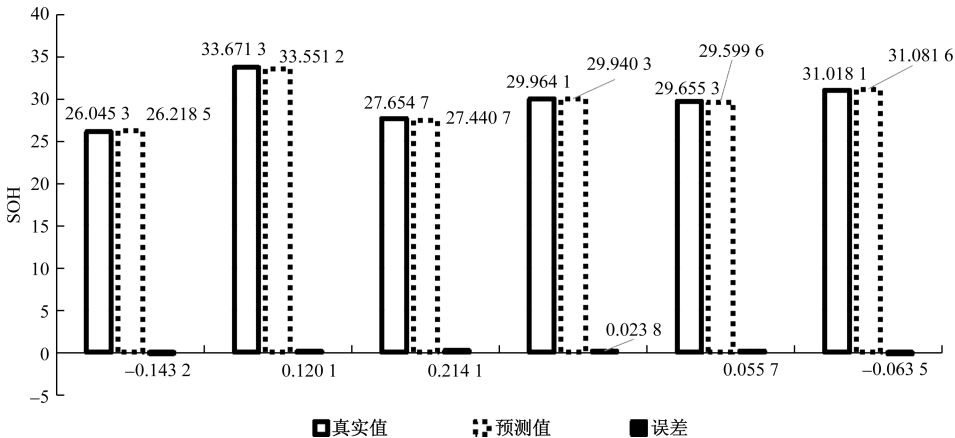


图 9 基于 SMoE 和 EIS 融合数据的电池 SOH 预测结果

表 9 基于 SMoE 模型和 EIS 融合数据的电池 SOH 预测结果

真实值	预测值	误差
26.075 3	26.218 5	-0.143 2
33.671 3	33.551 2	0.120 1
27.654 7	27.440 7	0.214 1
29.964 1	29.940 3	0.023 8
29.655 3	29.599 6	0.055 7
31.018 1	31.081 6	-0.063 5

3.2.4 RUL 预测结果

电池循环次数的测试集的均值为 149.295 2,标准方差为 105.195 3,预测集的均值为 149.995 7,标准方差为 105.782 9,这表明预测结果的数据分布与真实数据的数据分布基本一致。

电池 RUL 的预测集和测试集的均方根误差只有 3.071 6,说明模型的预测效果良好。

预测结果如表 10 和图 10 所示。从图 10 中可以看出,预测效果良好。

表 10 基于 SMoE 模型和 EIS 融合数据的电池 RUL 预测结果

真实值	预测值	误差
254	258.58	-4.58
70	67.74	2.26
50	47.57	2.43
28	31.75	-3.75
138	138.33	-0.33
274	276.28	-2.28

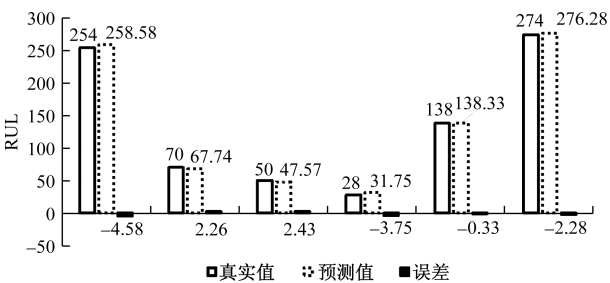


图 10 基于 SMoE 模型和 EIS 融合数据的电池 RUL 预测结果

4 结论

基于 SMoE 模型,使用电池融合数据联合预测了电池 SOH 和 RUL。在 NASA 公开数据集上测试效果,结果表明本文提出的联合预测模型能够很好地预测电池 SOH 和 RUL,SOH 预测值的均方误差为 0.069,RUL 预测值的均方误差为 2.042。在 EIS 公开数据集上测试效果,联合预测模型的 SOH 预测值的均方误差为 0.118,RUL 预测值的均方误差为 3.072。文本提出的模型和方法具有一定的参考和应用价值。

参考文献

- [1] 常伟,胡志超,潘多昭. 基于多模型组合和 EIS 的锂电池 SOH 和 RUL 预测[J]. 科技和产业, 2024, 24(2): 192-199.
- [2] 贺伟,马鸿雁,张英达,等. 基于改进门控循环单元神经网络的锂电池组荷电状态预测[J]. 科学技术与工程, 2023, 23(12): 5102-5109.
- [3] 郑永飞,文怀兴,韩昉,等. 基于电池外特征的粒子群神经网络电池健康状态预测[J]. 科学技术与工程, 2019, 19(36): 184-189.
- [4] 朱振宇. 基于深度学习的电动汽车锂电池 SOH 估计与 RUL 预测方法研究[D]. 青岛: 青岛科技大学, 2023.
- [5] 郝可青,吕志刚,邸若海,等. 基于鲸鱼算法优化长短时记忆神经网络的锂电池剩余寿命预测[J]. 科学技术与工程, 2022, 22(29): 12900-12908.
- [6] 郝可青,吕志刚,邸若海,等. 基于鲸鱼算法优化长短时记忆神经网络的锂电池剩余寿命预测[J]. 科学技术与工程, 2022, 22(29): 12900-12908.
- [7] 周雅夫,万伟东,董启超. 基于改进多频正弦激励的燃料电池电化学阻抗谱测量[J]. 科学技术与工程, 2022, 22(22): 9679-9685.
- [8] ZHANG Z G, JIANG Q L, ZHAN Y B, et al. Fault classification method of small sample rolling bearing based on VAE-GAN data enhancement algorithm[J]. Atomic Energy Science and Technology, 2023, 57(S1): 228-237.
- [9] 李伟,牛东晓. 基于灰色神经网络的短期电力负荷预测分析[J]. 科技和产业, 2008, 8(10): 57-60.
- [10] 王萍,范凌峰,程泽. 基于健康特征参数的锂离子电池 SOH 和 RUL 联合估计方法[J]. 中国电机工程学报, 2022, 42(4): 1523-1534.
- [11] 郭建康,王子赞. 基于深度学习的锂电池 SOC 和 SOH 联合估计研究[J]. 计算机科学与应用, 2023, 13(4): 672-679.

Joint Prediction of SOH and RUL Based on SMoE Model and Battery Integration Data

CHANG Wei, HU Zhichao, PAN Duozhao, SHI Jiwen

(Nantong Le Chuang New Energy Co., Ltd., Nantong 226000, Jiangsu, China)

Abstract: The state of health(SOH) and remaining useful life(RUL) of a battery are key indicators for measuring battery performance degradation and remaining useful time. Predicting battery SOH and RUL is of great importance in practical applications. Usually, battery operation data is used to train machine learning algorithms, such as neural networks or deep learning, to capture the changing patterns of battery SOH and RUL and make predictions. Traditional machine learning models often use a single model to adapt to the entire dataset, which is inadequate when dealing with complex and highly heterogeneous data. Building a model for each prediction target incurs high training and maintenance costs. Sparse mixture of experts (SMoE) was used to construct a joint prediction model for battery SOH and RUL, and battery fusion data was used to simultaneously predict battery SOH and RUL. The results of testing on NASA's public dataset show that the proposed joint prediction model can effectively predict battery SOH and RUL, with a mean square error of 0.069 for SOH prediction and 2.042 for RUL prediction. Tested on the EIS public dataset, the mean square error of the SOH prediction value of the joint prediction model was 0.118, and the mean square error of the RUL prediction value was 3.072, indicating a significant improvement in accuracy. The models and methods proposed in the text have certain reference and application value.

Keywords: MoE(mixture of experts); SMoE(sparse mixture of experts); SOH(state of health); RUL(remaining useful life)