

# HS-MoE 和 MH-MoE 模型在电池 SOH 和 RUL 预测中的对比研究

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:** 电池的健康状态(SOH)和剩余使用寿命(RUL)是评估电池性能退化和潜在使用期限的核心指标。在实际应用中,准确预测电池的 SOH 和 RUL 至关重要。为了捕捉电池性能的变化并做出预测,通常依赖于电池的运行数据来训练机器学习算法,如神经网络或深度学习算法。然而,传统的机器学习模型往往采用单一架构来适应整个数据集,这在处理复杂且异质性强的大数据时显得不足,这类模型普遍存在泛化能力不足和过拟合的风险,且在大数据处理上效率低下。为此,采用多层稀疏混合专家模型(HS-MoE)和多头混合专家模型(MH-MoE)分别构建电池 SOH 和 RUL 的预测模型。在 NASA(美国国家航空和宇宙航行局)和 EIS(电化学阻抗谱)公开的数据集上进行对比实验,结果显示在两个数据集上,MH-MoE 模型在预测 SOH 和 RUL 方面的表现均优于 HS-MoE 模型。

**关键词:** 混合专家模型(MoE); 多层稀疏混合专家模型(HS-MoE); 多头混合专家模型(MH-MoE); 电池的健康状态(SOH); 剩余使用寿命(RUL)

**中图分类号:** TM912.9 **文献标志码:** A **文章编号:** 1671-1807(2025)13-0122-11

电池的操作数据包括电流、电压和温度,与电池的健康状况(state of health, SOH)密切相关,故而,通过这些数据来预测 SOH 是一种有效的方法。此外,电池的剩余使用寿命(remaining useful life, RUL)与电流、电阻和阻抗等因素显著相关,研究者因此将这些因素视为预测 RUL 的重要指标。研究者通常采用电流、电压、温度、电阻和阻抗等数据来训练机器学习算法,如支持向量机、神经网络和深度学习等<sup>[1-12]</sup>,目的是捕捉并预测电池 SOH 和 RUL 的变化。传统的机器学习模型往往采用单个模型来适应全部数据,这在面对复杂且异质性强的大数据集时显得力不从心。这样的单一模型在泛化能力上存在局限,容易导致过拟合问题,并且在处理大规模数据时的效率不高。

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<sup>[13]</sup>。

MoE 模型的核心概念是将输入数据通过一个门控网络分配至不同的专家模型,每个专家模型专

攻数据的一个独特方面。随后,这些专家模型的输出被汇总,以生成最终的预测结果。MoE 模型通过集成多个专家模型,每个模型专注于数据的不同方面,从而更有效地捕捉数据的复杂性和多样性,增强了模型的整体表现力。这种多模型组合的方式降低了单一模型过拟合的风险,提升了模型的泛化能力和鲁棒性。专家模型支持并行训练和预测,提高了计算效率,尤其在处理大规模和复杂任务时更显优势。基于 MoE 模型,研究者们发展出了多种变体,其中包括多层稀疏混合专家模型和多头混合专家模型。

稀疏混合专家模型(sparse mixture of experts, SMOE)是为了解决复杂预测问题而设计的机器学习架构,它利用多个专家子模型,每个子模型针对数据集的特定部分或特定输入类型进行处理。这些子模型的预测通过门控网络进行加权合并,以得出最终预测。SMOE 的“稀疏”特性表现在,每次预测时只有部分专家子模型被激活,这种机制显著降低了计算成本,提高了效率。通常,这种激活是通过 top-K 策略完成的,即仅选择权重最高的 K 个子

**收稿日期:** 2024-10-24

**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,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。

模型进行预测。

多层稀疏混合专家模型(hierarchical sparse mixture of experts, HS-MoE)是 SMoE 模型的进一步发展,它通过引入多层级的专家和门控网络,增强了模型处理多层次、复杂任务的能力。HS-MoE 的结构类似于分级树状结构,每一层都包含多个专家网络和门控网络。

HS-MoE 模型的层次化的设计允许模型在不同层次上捕捉数据的不同抽象级别,同时通过稀疏性约束保持模型的效率和可解释性。

HS-MoE 模型的优点主要包括:①高效处理复杂任务。多层结构增强了对复杂任务的处理能力。②计算效率高。稀疏激活减少了计算负担。③泛化能力强。多个专家子模型组合提高了模型泛化。④降低过拟合。稀疏性有助于减少过拟合风险。⑤动态适应性。能够根据数据特性动态选择专家子模型。

多头混合专家模型(multi-head mixture of experts, MH-MoE)是 MoE 模型另一种变体,它是一种神经网络架构,旨在提高模型的表现力和计算效率。

MH-MoE 的专家网络由多个子网络组成,各自从不同角度处理输入数据;每个门控网络被称作一个“头”,每个“头”根据输入数据选择激活相应的专家网络,多个门控网络可并行操作;多头机制允许多个门控网络同时运作,各自生成权重分布以选择不同的专家网络。

多头混合专家模型(MH-MoE)的主要优点包括:①并行处理能力。多个“头”可以同时处理信息,提高处理效率。②增强的表征能力。每个“头”学习数据的不同方面,提供更全面的信息表示。③更好的泛化能力。从多个角度学习数据,减少过拟合风险。④灵活性和可扩展性。根据任务需求调整“头”的数量。⑤提高预测准确性。组合多个“头”的预测,提供更准确的结果。

将 MoE 的变体模型用于电池 SOH 和 RUL 的预测,是一个积极的探索。将 HS-MoE 和 MH-MoE 模型分别用于电池 SOH 和 RUL 的预测中,并对比分析了两种模型的效果。另外,为了提升模型的预测精度,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。

②Transformer(深度学习模型)。Transformer 可以用多头注意力机制和前馈网络学习输入数据的特征表示。

综上所述,CNN 和 Transformer 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,使用以上两种模型的组合模型作为专家子模型,构建 HS-MoE 和 MH-MoE 模型,使用电池运行数据分别单独预测电池 SOH 和 RUL。

文本的创新点如下。

(1)HS-MoE 模型创新点:①多级特征提取。模型通过多层次结构逐层提取和组合电池运行数据特征,增强对数据多级特性的理解。②增强模型表达能力。多层次结构提升模型对电池 SOH 和 RUL 变化特征的捕捉能力。③计算效率高。稀疏选择机制减少计算量,提高训练和预测效率。④避免过拟合。稀疏选择机制防止模型依赖于所有专家输出,选择最相关专家进行组合。

(2)MH-MoE 模型创新点:①多头机制提高特征提取能力。多个门控网络并行处理输入数据,提高对电池健康状态变化的敏感度。②稀疏激活提升计算效率。减少计算量,高效处理大规模电池阻抗数据。③专家网络专注于不同模式。每个专家网络专注于特定阻抗模式,提升建模能力。④灵活适应多任务环境。模型自适应选择最合适的网络结构和参数,适应不同电池数据。

(3)CNN 与 Transformer 结合的专家子模型:空间与时间特征提取。CNN 提取空间特征,Transformer 捕捉时间序列模式,组合模型在特征提取和稳健性方面表现优异。

本文的亮点:模型效果对比分析。HS-MoE 和 MH-MoE 模型分别增加了模型的深度和宽度,对比分析了 HS-MoE 和 MH-MoE 模型在预测电池 SOH 和 RUL 中的效果。

本文的实用性在于,提供了两种电池 SOH 和 RUL 的预测方法。

## 1 电池 SOH 和 RUL 预测流程

电池 SOH 和 RUL 预测流程如图 1 所示。

首先,测试并采集电池数据,电池运行数据包括但不限于电流、电压、温度、电阻和阻抗等,以及对应的 SOH 和 RUL;其次,构建 HS-MoE 模型和 MH-MoE 模型,本文在 Keras 平台构建模型;之后,训练以及校验 HS-MoE 模型和 MH-MoE 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;最

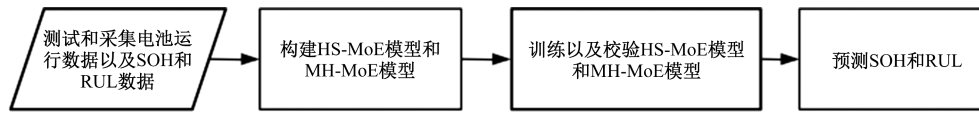


图 1 电池 SOH 和 RUL 预测流程

后,预测电池 SOH 和 RUL,在 HS-MoE 模型和 MH-MoE 训练完毕之后,在模型效果良好的前提下,可以部署训练好的模型,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。

当采集电池电化学阻抗谱(electrochemical impedance spectroscopy, 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  $10^{-2} \sim 10^5$  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 构建模型

#### 2.2.1 构建 HS-MoE 模型

HS-MoE 的模型结构通常包含以下几个关键组成部分。

(1)专家层。由多个专家网络组成,每个专家网络专门负责学习数据的一个子集或特定特征。这些专家可以是简单的神经网络或其他可学习的模型。

(2)门控网络(gating network)。每个层次都有一个门控网络,它的作用是为每个专家分配权重,决定哪些专家在当前输入数据上最为活跃。门控网络通常也是由神经网络构成,其输出是一个概率分布,指示每个专家的激活程度。

(3)稀疏性约束。在门控网络上施加稀疏性约束,确保在任何给定时间内只有少数专家被激活。这可以通过各种稀疏性诱导技术实现,如 L1 正则化或软阈值化。

(4)层次结构。HS-MoE 的层次结构意味着模型由多个层级组成,每个层级都有自己的专家网络和门控网络。较低层级的输出可以作为较高层级的输入,从而形成层次化的处理流程。

基于 HS-MoE 使用电池运行数据预测 SOH 和

RUL 的单目标预测模型结构图如图 2 所示。

输入数据是电池运行数据,包括但不限于电流、电压、温度、电阻和阻抗 EIS 等。

接下来是多层级的专家层和门控网络层,使用 2 层级的专家层和门控网络层;每个层级设置 3 个 CNN-Transformer 专家子模型;门控网络模型使用多层感知机(MLP);每个层级均选取概率最高的 2 个专家子模型参与计算。

首先,输入数据分别输入到第一层级的 3 个专家子模型中,从不同角度提取相应的特征。

第 1 个层级的 3 个专家子模型的输出分别是 output\_1、output\_2 和 output\_3,这 3 个输出均是高维向量,本文设置 32 维,分别表示每个子专家模型对于输入数据的特征表示;同时,输入数据输入到第 1 层级的门控网络模型中。

第 1 个层级的门控网络输出是 1 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。

选取概率最高的 2 个专家的输出,与对应的权重进行加权求和,得到第一层级专家层的综合特征表示。

然后将第 1 层级专家层的综合特征表示,作为输入数据,输入到第 2 层级的专家模型和门控网络模型中。

当预测目标为 SOH 时,第 2 个层级的 3 个专家子模型的输出分别为 SOH\_1、SOH\_2 和 SOH\_3,这 3 个输出均是 1 个连续变量,分别代表各个专家预测 SOH 的结果。

当预测目标为 RUL 时,第 2 个层级的 3 个专家子模型的输出分别是 RUL\_1、RUL\_2 和 RUL\_3,这 3 个输出均是 1 个连续变量,分别代表各个专家预测 RUL 的结果。

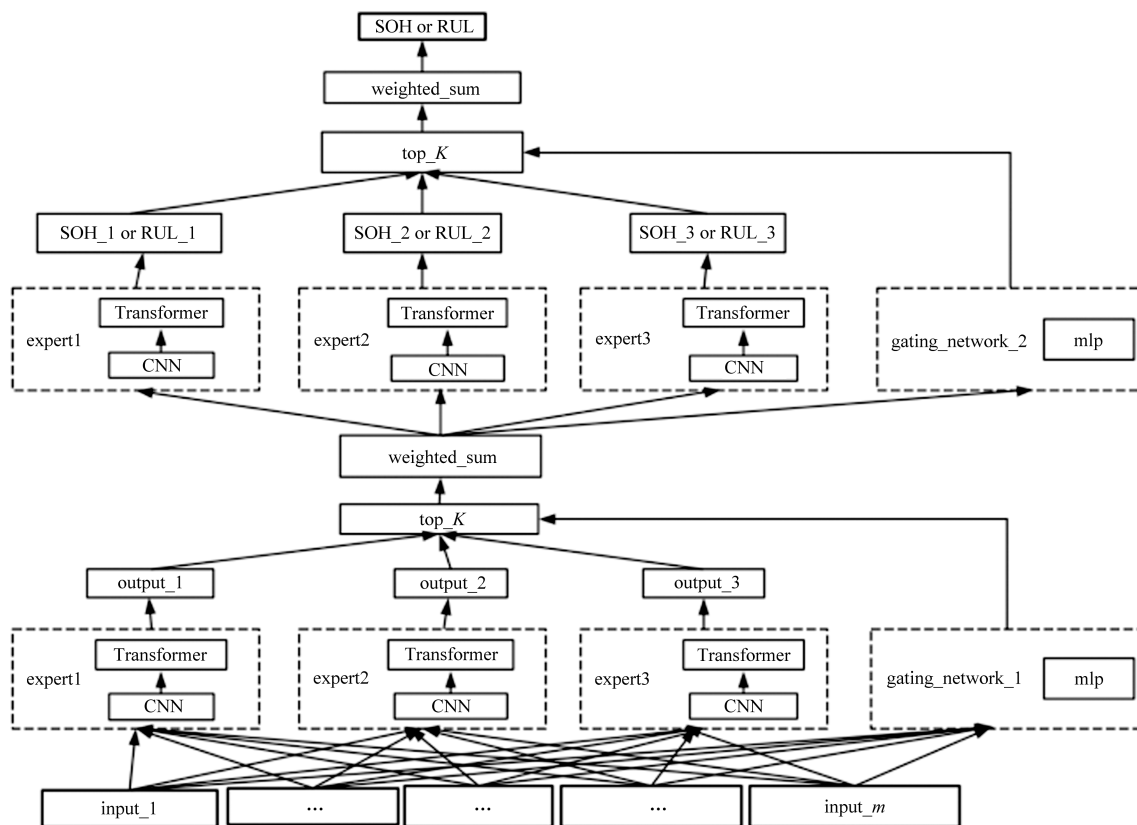
第 2 个层级的门控网络输出是 1 个三维向量,分别对应 3 个专家子模型的权重,3 个权重之和为 1。

选取概率最高的 2 个专家的输出,与对应的权重进行加权求和,得到最终的预测值 SOH 或者 RUL。

#### 2.2.2 构建 MH-MoE 模型

MH-MoE 的模型结构如下。

(1)专家网络。专家网络(experts)包含多个独立的子网络,每个专家网络从不同角度处理输入



input<sub>1</sub>, ..., input<sub>m</sub> 表示输入数据; expert1、expert2 和 expert3 表示专家子模型; gating\_network 表示门控网络模型; mlp 表示多层感知机模型; output<sub>1</sub>、output<sub>2</sub> 和 output<sub>3</sub> 表示专家子模型的输出值; top<sub>K</sub> 表示激活前 K 个专家子模型; weighted\_sum 表示加权求和

图 2 基于 HS-MoE 模型预测 SOH 和 RUL 的模型结构图

数据。

(2) 门控网络。门控网络 (gating network) 的每个门控网络为一个“头”，根据输入数据选择和激活适当的专家网络。多个门控网络 (头) 可以并行处理输入数据。

(3) 多头机制。多头机制 (multi-Head mechanism) 支持多个门控网络同时工作，每个门控网络生成自己的权重分布，选择和激活不同的专家网络。

文本基于 MH-MoE 的使用电池运行数据预测 SOH 和 RUL 的单目标预测模型结构图如图 3 所示。

输入数据是电池运行数据，包括但不限于电流、电压、温度、电阻和阻抗 EIS 等。

接下来是专家层和门控网络层，本发明以 3 个 CNN-Transformer 专家子模型为例，以 3 个基于多层感知机 (MLP) 的门控网络为例。

输入数据分别输入到 3 个专家子模型中和 3 个门控网络模型中。

3 个专家子模型从不同角度提取相应的特征，3 个门控网络模型分别独立地计算每个专家子模型

的权重。

当预测目标是 SOH 时，3 个专家子模型的输出分别是 SOH<sub>1</sub>、SOH<sub>2</sub> 和 SOH<sub>3</sub>，这 3 个输出均是一个连续变量，分别表示每个子专家模型对于 SOH 的预测值。

当预测目标是 RUL 时，3 个专家子模型的输出分别是 RUL<sub>1</sub>、RUL<sub>2</sub> 和 RUL<sub>3</sub>，这 3 个输出均是 1 个连续变量，分别代表各个专家预测 RUL 的结果。

3 个门控网络模型独立地计算每个专家子模型的权重，每个门控网络的输出都是一个三维向量，分别对应 3 个专家子模型的权重，3 个权重之和为 1。

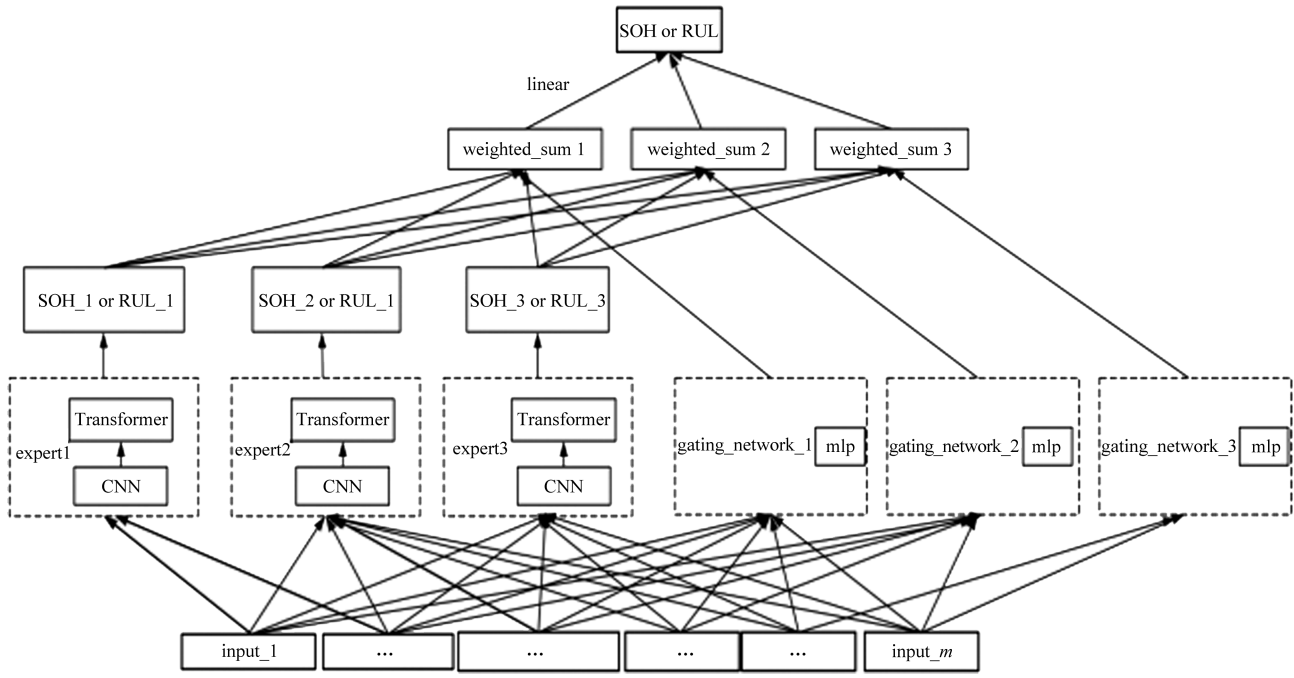
然后，将 3 个专家子模型的 SOH 或 RUL 预测值分别与 3 个门控网络输出的专家权重进行加权求和，得到 3 个 SOH 或 RUL 的加权求和值。

最后，将 3 个 SOH 或 RUL 的加权求和值进行线性回归预测，得到最终的预测值 SOH 或 RUL。

### 2.2.3 专家子模型和门控网络模型

#### 1) 专家子模型

每个子模型均为 CNN-Transformer 模型。



input<sub>1</sub>, ..., input<sub>m</sub> 表示输入数据; expert1、expert2 和 expert3 表示专家子模型; 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; weighted\_sum 表示加权求和

图 3 基于 MH-MoE 模型预测 SOH 和 RUL 的模型结构图

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。

本文使用的 CNN 结构包含 1 个卷积层和一个最大池化层, 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。

卷积计算示意图如图 4 所示。图 4 的卷积层中, 有一个 1×6×6 的矩形作为输入数据, 称为输入特征图, 它经过一个 3×3 的滤波器(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 1×4×4 的输出特征图, 即 Result。

池化层对输入进行非线性下采样, 通过取局部邻域内的最大值(最大池化 Max Pooling)或平均值(平均池化 Avg Pooling)作为输出。

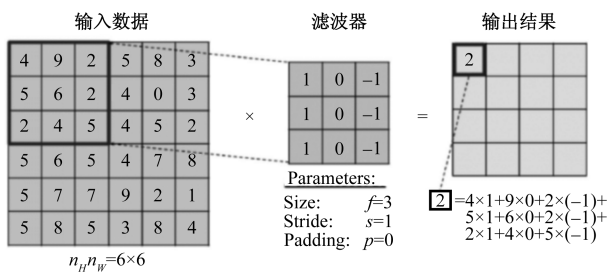


图 4 卷积计算示意图

池化计算示意图如图 5 所示。其中, 池化输入特征图大小为 4×4, 最大池化和平均池化卷积核大小均为 2×2, 池化输出特征图大小为 2×2。

每个卷积层可以提取输入数据的高层特征, 池化层实现下采样缩小数据维度。本文采用最大池化方法。Transformer 模型结构图如图 6 所示。

Transformer 模型主要结构和作用如下。

(1) 位置编码(positional encoding)。为输入序列提供顺序信息, 帮助 Attention 学习序列顺序特征。

(2) 多头注意力(multi-head attention)机制。该模块用于学习输入序列不同位置之间的相关性, 得到输入的特征表示。

(3) 前馈全连接网络(feed forward network)。

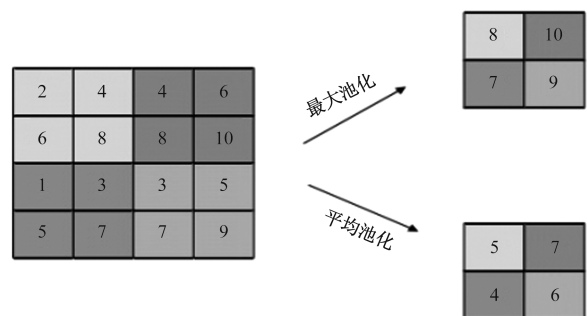
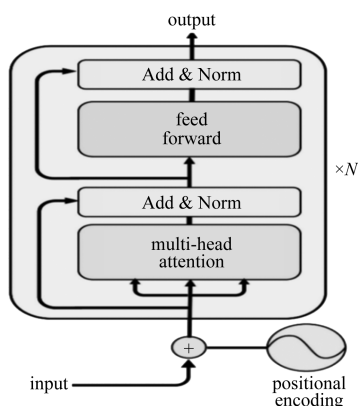


图 5 池化计算示意图



input 表示输入层;positional encoding 表示位置编码;multi-head attention 表示多头注意力;feed forward 表示前馈全连接网络; output 表示输出层;Add & Norm 表示残差连接和 Layer Norm 层

图 6 Transformer 模型结构图

该模块对注意力模块输出再进行非线性转换,以学习更复杂的特征表示。

(4) 残差连接和 Layer Norm 层 (Add & Norm)。残差连接 Add 用来避免深度网络中的梯度消失问题;Layer Norm 用来正则化模型,加快训练速度,增加模型稳定性。

本文采用 1 层 Transformer 结构,参数  $d_{model}$  为 32,num\_heads 为 4,dff 为 1。 $d_{model}$  表示特征嵌入(embedding)的维度;num\_heads 表示多头注意力机制的头的个数;dff 表示残差连接的步长。

### 2) 门控网络模型

门控网络模型采用多层感知机模型,模型结构图如图 7 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,输出层是一个三维向量,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。

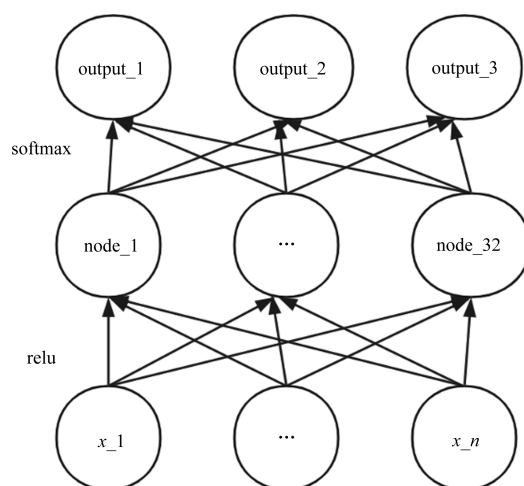
### 2.3 训练和校验模型

本文在 Keras 平台实现 HS-MoE 模型和 MH-MoE 模型代码并训练,模型训练的优化器 optimizer 为 adam 算法,损失函数均方误差(mean square error, 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。

### 2.4 预测电池 SOH 和 RUL

在 HS-MoE 模型和 MH-MoE 模型训练完毕之后,在模型效果良好的前提下,可以部署训练好的



$x_1, \dots, x_n$  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 7 门控网络模型结构图

模型,并对电池的 SOH 和 RUL 进行预测。

## 3 预测效果展示

### 3.1 电池数据展示

#### 3.1.1 NASA 电池运行数据展示

本文的测试数据来源有 2 个,一是 NASA(美国国家航空和宇宙航行局)电池数据集合中的 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 电池数据集合 1 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。其中,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	°C
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	A·h

表2 电池运行数据和 SOH 样例

Voltage_ measured	Current_ measured	Temperature_ measured	Current_ charge	Voltage_ charge	Capacity
4.188 108 651	0.000 130 667	23.819 520 25	0.000 6	0	1.855 004 521
4.188 195 943	0.001 459 081	23.828 807 16	0.000 6	4.203	1.855 004 521
...	...	...	...	...	...
3.477 276 739	-0.001 940 069	34.581 659 54	0.000 6	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

### 3.1.2 EIS 电池数据展示

本文在公开数据集上测试效果,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,特征序号依次是 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 数据集样例见表 5。

表 5 中 f1~f120 表示 EIS 频率特征,这里的 Capacity 表示电池最大容量即 SOH。

电池 EIS 和 RUL 数据集样例见表 6,其中 RUL 表示电池剩余循环次数。

## 3.2 HS-MoE 模型预测结果展示

### 3.2.1 SOH 预测结果展示

#### 1) NASA 数据集 SOH 预测效果展示

电池容量的测试集合的均值为 1.584 1,标准方差为 0.157 1,预测集合的均值为 1.567 0,标准方差为 0.128 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 的预测集合和测试集合的均方根误差只有 0.087 77,说明模型的预测效果良好。预测结果见表 7。预测结果样例如图 8 所示。从图 8 中可以看出,预测效果良好。

#### 2) EIS 数据集 SOH 预测效果展示

电池容量的测试集合的均值为 30.624 9,标准方

表5 电池 EIS 和 SOH 数据集样例(部分)

f1	f2	...	f120	Capacity
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6 电池 EIS 和 RUL 数据集样例(部分)

f1	f2	...	f120	RUL
0.384 70	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

表4 电池运行数据和 RUL 样例(部分)

Sense_ current_real	Sense_ current_imag	Battery_ current_real	Battery_ current_imag	Current_ ratio_real	Current_ ratio_imag	
832.749 572 8	-27.377 193 45	327.527 618 4	-90.784 202 58	2.382 645 118	0.576 834 847	
826.062 866 2	-41.281 600 95	331.814 27	-78.227 035 52	2.386 236 081	0.438 156 484	
...	...	...	...	...	...	
803.516 357 4	443.114 929 2	98.716 178 89	158.175 430 3	4.297 792 306	-2.397 684 157	
Battery_ impedance_real	Battery_ impedance_imag	Rectified_ impedance_real	Rectified_ impedance_imag	Re	Rct	RUL
0.165 409 535	-0.129 054 344	0.095 314 82	-0.001 691 894	0.065 158 152	0.095 553 696	1
0.165 409 535	-0.129 054 344	0.095 314 82	-0.001 691 894	0.065 158 152	0.095 553 696	1
...	...	...	...	...	...	...
0.164 978 108	-0.126 231 668	0.091 522 433	-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。

表 7 基于 HS-MoE 模型的电池 SOH 预测结果样例

SOH 真实值	SOH 预测值	SOH 误差
1.351 9	1.483 3	-0.131 4
1.438 1	1.512 8	-0.074 7
1.681 9	1.649 6	0.032 3
1.389 4	1.447 6	-0.058 2
1.480 7	1.460 3	0.020 4

差为 4.130 9,预测集合的均值为 30.559 1,标准方差为 4.151 6。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 0.296,说明模型的预测效果良好。预测结果见表 8。预测结果样例如图 9 所示。从图 9 中可以看出,预测效果良好。

### 3.2.2 RUL 预测结果展示

#### 1)NASA 数据集合 RUL 预测效果展示

电池循环次数的测试集合的均值为 139.783 9,标准方差为 96.904 6,预测集合的均值为 139.278 1,标准方差为 96.033 4,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有 4.649,说明模型在预测 RUL 时,预测效果良好。预测结果见表 9。预测结果样例如图 10 所示。

表 8 基于 HS-MoE 模型和 EIS 数据的电池 SOH 预测结果

SOH 真实值	SOH 预测值	SOH 误差
32.491 0	32.377 1	0.11
39.290 3	39.322 6	-0.03
32.402 4	32.176 6	0.23
26.718 1	26.589 7	0.13
30.047 2	30.046 8	0.00

表 9 基于 HS-MoE 模型的电池 RUL 预测结果

RUL 真实值	RUL 预测值	RUL 误差
3	1	2
267	262	5
220	219	1
185	184	1
39	43	-4

从图 10 中可以看出,预测效果良好。

#### 2)EIS 数据集合 RUL 预测效果展示

电池循环次数的测试集合的均值为 149.295 2,标准方差为 105.195 3,预测集合的均值为 149.838 1,标准方差为 104.339,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有 3.900 5,说明模型在预测 RUL 时,预测效果良好。预测结果见表 10。预测结果样例如图 11 所示。

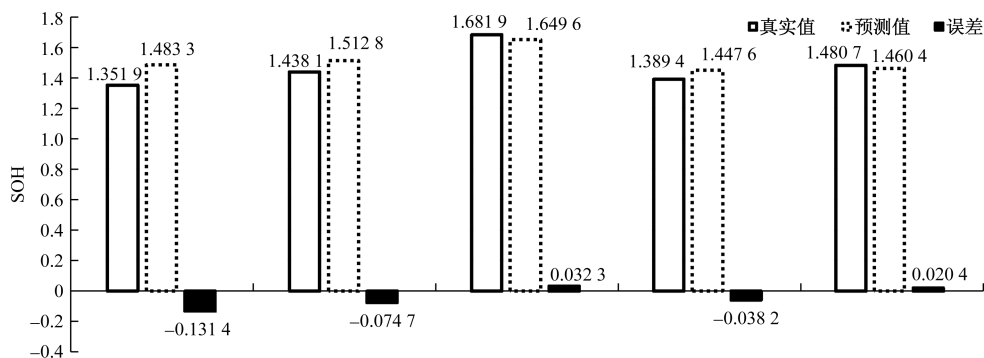


图 8 基于 HS-MoE 模型的电池 SOH 预测结果样例

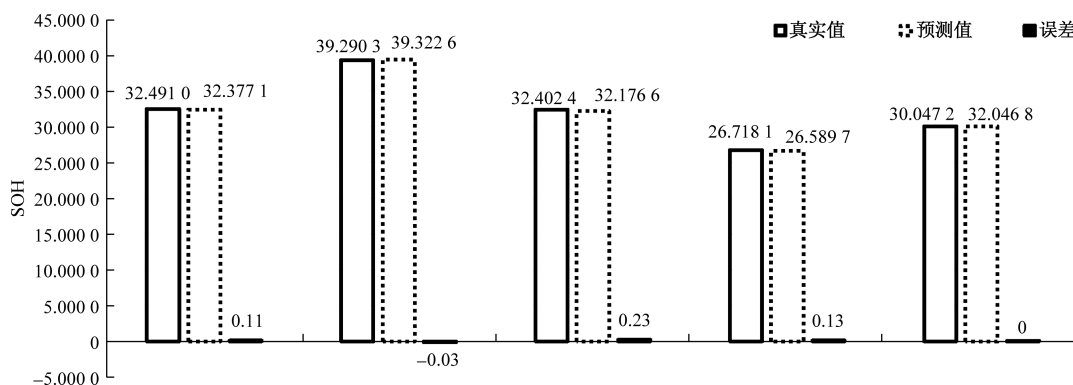


图 9 基于 HS-MoE 模型和 EIS 数据的电池 SOH 预测结果样例图

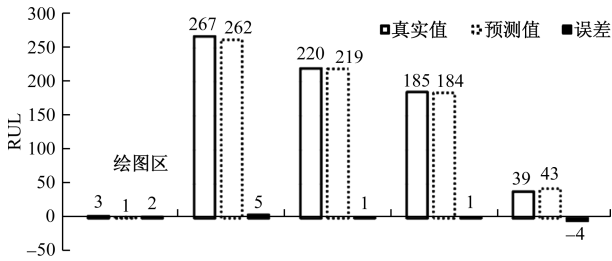


图 10 基于 HS-MoE 模型的电池 RUL 预测结果样例

表 10 基于 HS-MoE 模型和 EIS 数据的电池 RUL 预测结果

RUL 真实值	RUL 预测值	RUL 误差
12	14	-2
268	267	1
222	222	0
30	33	-3
78	76	2

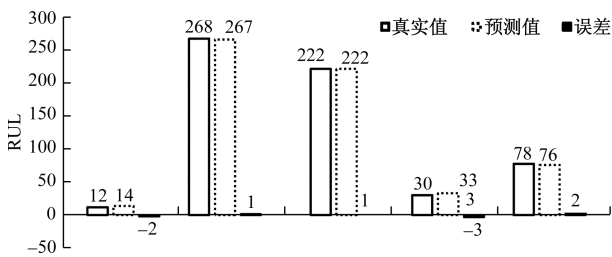


图 11 基于 HS-MoE 模型和 EIS 数据的电池 RUL 预测结果样例

从图 11 中可以看出,预测效果良好。

### 3.3 MH-MoE 模型预测结果展示

#### 3.3.1 SOH 预测结果展示

##### 1) NASA 数据集合 SOH 预测效果展示

电池容量的测试集合的均值为 1.584 1,标准方差为 0.157 1,预测集合的均值为 1.596 2,标准方差为 0.125 8,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 0.086 44,说明模型的预测效果良好。预测结果见表 11。预测结果样例如图 12 所示。从图 12 中可以看出,预测效果良好。

##### 2) EIS 数据集合 SOH 预测效果展示

电池容量的测试集合的均值为 30.624 9,标准方差为 4.130 9,预测集合的均值为 30.539 5,标准方差为 4.141 1 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 0.266 3,说明模型的预测效果良好。预测结果见表 12。预测结果样例如图 13 所示。从图 13 中可以看出,预测效果良好。

表 11 基于 MH-MoE 模型的电池 SOH 预测结果

SOH 真实值	SOH 预测值	SOH 误差
1.351 9	1.519 1	-0.167 2
1.438 1	1.581 9	-0.143 8
1.681 9	1.668 4	0.013 5
1.389 4	1.491 5	-0.102 1
1.480 7	1.474 2	0.006 5

表 12 基于 MH-MoE 模型和 EIS 数据的电池 SOH 预测结果样例

SOH 真实值	SOH 预测值	SOH 误差
32.491 0	32.331 9	0.16
39.290 3	39.164 7	0.13
32.402 4	32.410 8	-0.01
26.718 1	26.646 3	0.07
30.047 2	29.950 5	0.10

#### 3.3.2 RUL 预测结果展示

##### 1) NASA 数据集合 RUL 预测效果展示

电池循环次数的测试集合的均值为 139.783 9,标准方差为 96.904 6,预测集合的均值为 138.581 7,标准方差为 95.283 5,预测结果的数据分布与真实数据的数据分布基本一致。

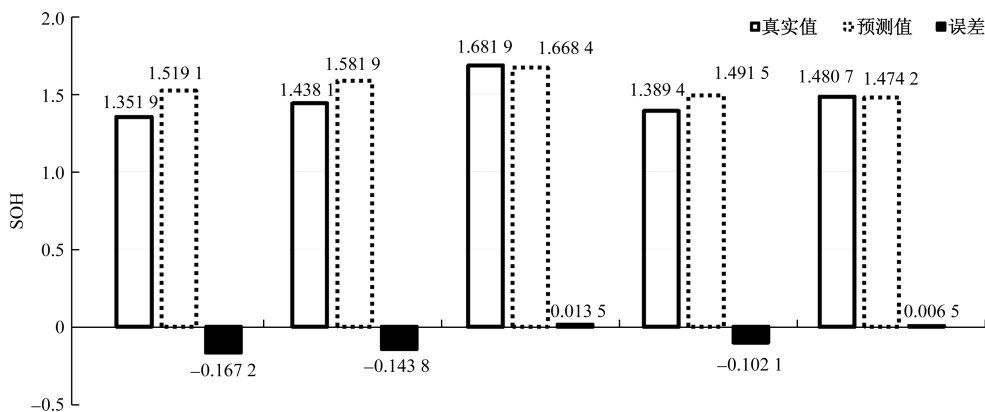


图 12 基于 MH-MoE 模型的电池 SOH 预测结果样例

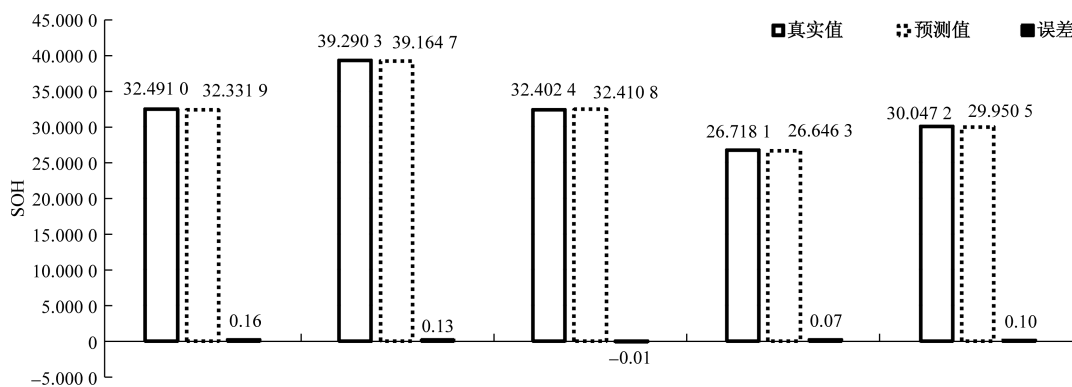


图 13 基于 MH-MoE 模型和 EIS 数据的电池 SOH 预测结果样例

电池 RUL 的预测集合和测试集合的均方根误差为 4.2007,说明模型在预测 RUL 时,预测效果良好。预测结果如表 13 所示。预测结果样例如图 14 所示。从图 14 中可以看出,预测效果良好。

2)EIS 数据集 RUL 预测效果展示

电池循环次数的测试集合均值为 149.2952,标准方差为 105.1953,预测集合的均值为 148.9436,标准方差为 105.1521,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有 3.0188,说明模型在预测 RUL 时,预测效果良好。预测结果见表 14。预测结果样例如图 15 所示。从图 15 中可以看出,预测效果良好。

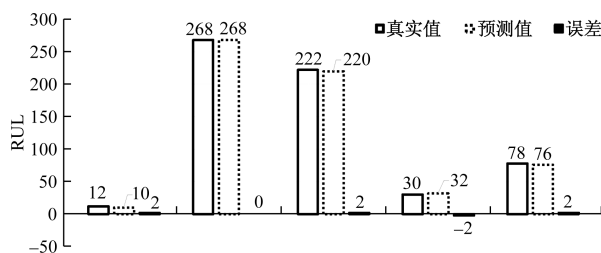


图 15 基于 MH-MoE 模型和 EIS 数据的电池 RUL 预测结果样例

表 13 基于 MH-MoE 模型的电池 RUL 预测结果

RUL 真实值	RUL 预测值	RUL 误差
3	0	3
267	260	7
220	219	1
185	182	3
39	45	-6

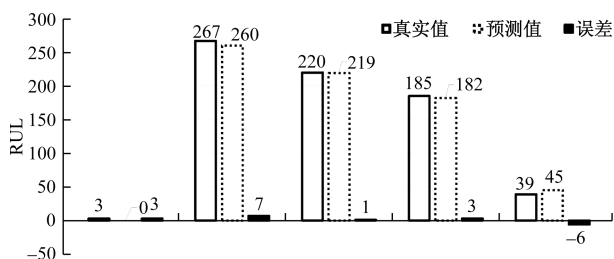


图 14 基于 MH-MoE 模型的电池 RUL 预测结果样例

表 14 基于 MH-MoE 模型和 EIS 数据的电池 RUL 预测结果

RUL 真实值	RUL 预测值	RUL 误差
12	10	2.00
268	268	0.00
222	220	2.00
30	32	-2.00
78	76	2.00

3.4 图预测结果对比分析

将 HS-MoE 模型和 MH-MoE 模型在 NASA 数据集和 EIS 数据集上的测试结果进行汇总,见表 15。

HS-MoE 模型和 MH-MoE 模型在 NASA 数据集和 EIS 数据集的 SOH 预测效果对比如图 16 所示。

CHS-MoE 模型和 MH-MoE 模型在 NASA 数据集和 EIS 数据集的 RUL 预测效果对比如图 17 所示。

表 15 HS-MoE 模型和 MH-MoE 模型测试对比分析

模型	数据集	SOH 预测效果	RUL 预测效果
HS-MoE	NASA	0.08777	4.6490
MH-MoE	NASA	0.08644	4.2007
HS-MoE	EIS	0.2960	3.9005
MH-MoE	EIS	0.2663	3.0188

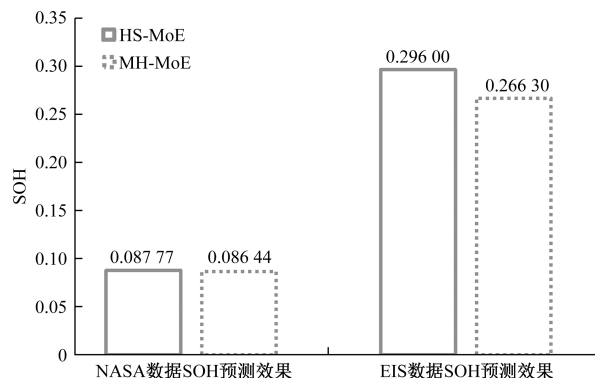


图 16 HS-MoE 模型和 MH-MoE 模型 SOH 预测效果对比

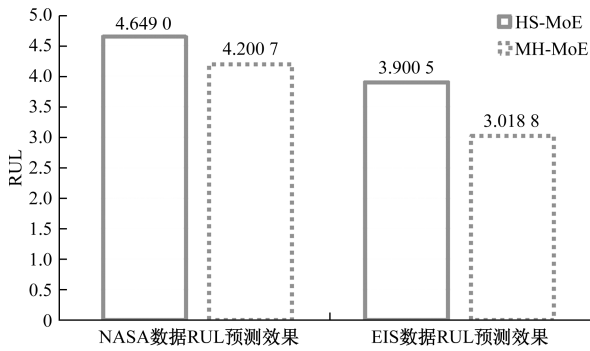


图 17 HS-MoE 模型和 MH-MoE 模型 RUL 预测效果对比

通过观测 HS-MoE 模型和 MH-MoE 模型在 NASA 数据集合和 EIS 数据集合上的测试结果可以发现,在两种数据集合上,MH-MoE 模型的 SOH 和 RUL 的预测效果都要优于 HS-MoE 模型。

#### 4 结论

HS-MoE 和 MH-MoE 模型分别增加了原始 MoE 模型的深度和宽度。本文利用 HS-MoE 和 MH-MoE 模型分别使用电池运行数据和 EIS 数据预测了电池 SOH 和 RUL。通过在 NASA 和 EIS 公开数据集上的实验验证,两种模型都能较好地预测 SOH 和 RUL,并且 MH-MoE 模型的 SOH 和 RUL 的预测效果都要优于 HS-MoE 模型,说明增加 MoE 模型的宽度比增加 MoE 模型的深度更能提升模型的效果。本文的结论,具有一定的参考价值。

#### 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] 常伟,胡志超,潘多昭. 基于神经网络和电化学阻抗谱的锂电池重要频率分析及电池健康状态预测[J]. 科技与产业, 2023, 23(20): 218-224.

[3] 王志福,杨忠义,罗崑,等. 基于数据驱动的锂离子动力电池剩余使用寿命预测方法综述[J]. 科学技术与工程, 2023, 23(15): 6279-6289.

[4] 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.

[5] 陈万,蔡艳平,苏延召,等. 锂离子电池的多状态模型剩余寿命预测方法[J]. 科学技术与工程, 2021, 21(10): 4078-4083.

[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] 杨庆新. 基于深度学习的锂离子电池剩余使用寿命预测方法研究[D]. 烟台:烟台大学, 2024.

[8] 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.

[9] HAO K Q, LÜ Z G, DI R H, et al. Prediction of remaining life of lithium battery based on Whale algorithm optimization of short-duration memory neural network [J]. Science Technology and Engineering, 2019, 22 (29): 12900-12908.

[10] 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.

[11] 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.

[12] 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.

[13] CAI W L, JIANG J Y, WANG F, et al. A survey on mixture of experts[EB/OL]. (2024-06-26) [2024-08-09]. <https://doi.org/10.48550/arXiv.2407.06204>.

### Comparative Study of HS-MoE and MH-MoE Models in Predicting SOH and RUL of Batteries

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 core indicators for evaluating battery performance degradation and potential lifespan. Accurately predicting the SOH and RUL of batteries is crucial in practical applications. To capture changes in battery performance and make predictions, operational data of the battery is typically relied on to train machine learning algorithms, such as neural networks or deep learning methods. However, traditional machine learning models often adopt a single architecture to adapt to the entire dataset, which is insufficient when dealing with complex and highly heterogeneous big data. Such models generally have the risk of insufficient generalization ability and overfitting, and are inefficient in big data processing. Therefore, Hierarchical sparse mixture of experts(HS-MoE) and multi head mixture of experts(MH-MoE) models were used to construct predictive models for battery State of Health(SOH) and Remaining Useful Life(RUL), respectively. Comparative experiments were conducted on publicly available datasets from NASA and EIS, and the results showed that the MH-MoE model outperformed the HS-MoE model in predicting SOH and RUL on both datasets.

**Keywords:** mixture of experts(MoE); hierarchical sparse mixture of experts(HS-MoE); multi head mixture of experts(MH-MoE); state of health(SOH); remaining useful life(RUL)