

[文章编号] 1003-4684(2023)02-0001-06

基于多任务学习的非侵入式负荷分解

刘 辉, 江煦成

(湖北工业大学太阳能高效利用湖北省协同创新中心, 湖北 武汉 430068)

[摘 要] 为解决目前非侵入式负荷分解研究中存在的模型数量多及训练时间长等问题, 将多任务学习引入到非侵入式负荷分解研究中, 提出一种基于多门多专家模型的非侵入式负荷分解方法, 首先通过 seq2point 模型将用电设备的功率分解转换为总功率序列与用电设备在序列中点时刻功率值的映射关系, 其次利用 MMoE 模型的门控函数及共用的 Expert 网络组兼顾不同用电设备功率分解任务的独特性和关联性, 最终通过单个 MMoE 模型同时完成多个用电设备的功率分解。在公开数据集上进行验证, 测试算例验证了方法的有效性。

[关键词] 非侵入式负荷分解; 多任务学习; MMoE; seq2point

[中图分类号] TM714 **[文献标识码]** A

目前主流的非侵入式负荷分解(Non-intrusive Load Decomposition, NILD)研究中, 通过低频采样的功率、电流、电压等数据, 使用不同算法得到总负荷数据序列与不同用电设备的负荷数据序列之间的映射关系, 实现对不同用电设备运行情况的识别, 在深度学习首次引入 NILD 研究取得很好效果之后^[1], 越来越多的学者开始使用深度学习模型构建 NILD 模型, 通过深度模型使用回归或者分类的方法实现对不同电器功率的识别, 例如时序卷积网络^[2], 卷积注意力块模型^[3], 以及 seq2seq 模型^[4]。以上研究, 都是通过单任务学习模型, 不同用电设备使用不同结构的模型进行负荷分解。虽然能做到较高的准确率, 但不同用电设备负荷分解任务需要重新构造模型, 操作较为繁琐, 也存在训练单个模型时间较长的问题。

解决上述问题, 可以将用电设备的负荷分解作为不同的子任务导入至多任务学习模型, 通过训练单个多任务学习同时对多个用电设备进行负荷分解。但是一般多任务学习模型训练中, 不同任务之间的相关性对单个任务的最终结果有很大的影响, 在任务之间相关性较弱时, 多任务学习模型的表现并不比单任务模型更好^[5], 主要原因在于基于深度神经网络的多任务学习模型对数据分布差别与任务之间的关系较为敏感^[6], 且真实数据模型比较复杂, 很难明确各个任务之间的区别, 即使不需要明确任务差别的模型, 为适应不同的任务, 其模型参数也较

多, 计算成本较高^[7]。

综上所述, 本文提出一种基于 MMoE 模型的非侵入式负荷监测方法, MMoE 通过建模来描述不同用电设备分解任务之间的相关性, 利用共享表示学习各分解任务特定的函数, 同时自动分配模型参数去捕捉任务共享或者任务独有的信息, 避免参数增多的同时, 也能提高不同用电设备识别任务的准确度。最后在公开数据集上验证本文所提方法的有效性。

1 非侵入式负荷分解模型

1.1 模型流程图

本文的非侵入式负荷分解模型如图 1 所示。

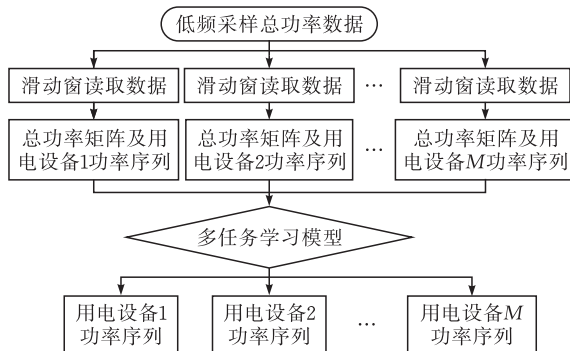


图 1 NILD 模型流程

本文所需的数据为低频采样的有功功率数据, 首先通过滑动窗口顺序读取总功率数据, 同时对应该读取滑动窗口中点时刻不同用电设备的功率值, 之

[收稿日期] 2022-03-10

[基金项目] 国家自然科学基金项目(61903129)

[第一作者] 刘 辉(1962-), 男, 湖北武汉人, 湖北工业大学教授, 研究方向为电网控制与智能制造

[通信作者] 江煦成(1996-), 男, 湖北黄冈人, 湖北工业大学硕士研究生, 研究方向为非侵入式负荷分解

后拼接多个总功率序列得到总功率矩阵,以及不同用电设备功率序列,将多个电器的负荷分解作为多个任务,将总功率矩阵及不同用电设备功率序列输入到多任务学习模型中进行训练,最后输入待识别的总功率矩阵至训练好的多任务学习模型,输出得到各用电设备功率序列。

1.2 seq2point 模型

本文采用 seq2point 模型来实现负荷分解,具体原理如下式:

$$x_{\tau} = F(Y_{t:t+\omega-1}) \quad (1)$$

式中: $Y_{t:t+\omega-1} = [y_t, y_{t+1}, \dots, y_{t+\omega-1}]$ 是一段总功率序列,其中总功率为不同用电器功率总和, ω 为滑动窗口长度, y_t 为 t 时刻对应的总功率值, x_{τ} 为窗口中心时刻不同用电设备的有功功率值,其中 $\tau = t + \omega/2$, F 表示 x_{τ} 与 $Y_{t:t+\omega-1}$ 的映射关系。

本文通过训练多任务模型,找出如式(1)的映射关系,通过输入总功率序列数据至网络,输出得到对应序列中点时刻不同用电设备的功率值。

1.3 MMoE 模型

本文使用的多任务学习模型为 MMoE^[7]。为了更好地说明 MMoE 的改进原理,先介绍传统的共享底层网络模型(Shared-Bottom Model, SBM)^[8],如图 2 所示

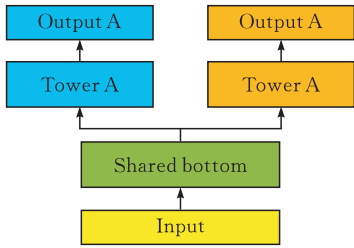


图 2 共享底层模型

图 2 表明当有两个任务的时候,传统的多任务学习模型结构,模型的输入层之后接一个共享底层网络,之后与两个塔层网络相连,塔层网络对共享底层输出的数据处理之后,得到不同任务的结果。

当有 k 个任务时,共享底层网络与 k 个塔层网络连接,假设输入网络的数据为 x ,则第 k 个任务的输出

$$y_k = h^k(f(x)) \quad (2)$$

式中: h^k 为第 k 个塔层网络的函数表示, f 为共享底层的函数表示。

基于共享底层的网络结构可以减少模型过拟合的风险,但是由于所有的任务共同使用共享底层的参数集,所以任务之间的差异容易导致模型优化冲突,虽然已有学者通过加设约束条件等做法来改善这种情况^[9],但是随着任务数目增多,模型中的参数也大大增加,进一步带来了需要加大数据集等问题,

且实验证明,在任务之间相关性不高时,传统的共享底层网络表现不佳。

为解决上述问题,有学者提出将混合专家网络(Mixture-of-Experts, MoE)层替换传统网络中的共享底层,同时对每一个任务都对应配对一个门控网络,得到 MMoE 网络,其具体结构如图 3 所示,MMoE 网络将传统网络中,共享底层网络被替换为一组专家网络,与输入层以及塔层网络连接,每个专家网络为前向网络。当有 k 个任务时,输入层也同时与 k 个门控网络进行连接,输入数据 x 至 MMoE 模型,得到第 k 个任务的输出可以表示为:

$$y_k = h^k\left(\sum_{i=1}^n g^k(x)_i f_i(x)\right) \quad (3)$$

$$\sum_{i=1}^n g^k(x)_i = 1 \quad (4)$$

式(3)中: f_i 为第 i 个专家网络的函数表示, g_i^k 表示第 k 个门控网络在第 i 个专家网络上的概率分布,且满足式(4), h^k 为第 k 个塔层网络的函数表示,门控网络为一层 softmax 变换器,对输入数据进行线性变换:

$$g^k(x) = \text{softmax}(W_{gk}x) \quad (5)$$

式中: $W_{gk} \in R^{n \times d}$ 为待训练矩阵, n 为专家网络数量, d 为特征维度。

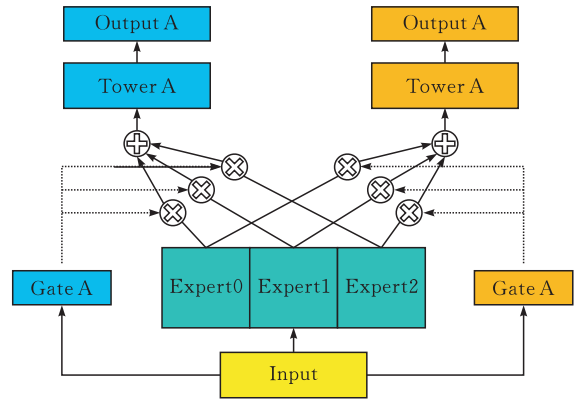


图 3 MMoE 模型

通过替换 MoE 层与对应增加门控网络之后,MMoE 结构可以根据不同任务的实际数据情况,根据式(5)调整各自门控网络矩阵参数,得到不同的概率分布并对应转换为加权数,最后每个任务的输出为不同专家网络的加权和表示,这种做法能够充分考虑单独任务的特殊性。同时由于多个任务共用一组专家网络,在一定程度上也能兼顾任务之间的相关性。故 MMoE 网络相比于传统的多任务学习模型鲁棒性更好,对于关联性较差的任务的训练效果也更好。

2 非侵入式负荷分解

由于实际生活中,不同用电设备的运行并没有

很大的关联性,故使用 MMoE 模型可以很好的解决 NILM 问题。

2.1 数据读取

假设通过对总功率序列使用滑动窗进行数据读取,将每次读取的短序列总功率数据以及用电设备在滑动窗中点时刻对应的功率数据进行存储,最终得到总功率数据矩阵 X 及用电设备 m 的功率序列 Y :

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_w \\ x_2 & x_3 & \cdots & x_{w+1} \\ \vdots & \vdots & \cdots & \vdots \\ x_n & x_{n+1} & \cdots & x_{n+w-1} \end{bmatrix}$$

(6)

$$Y_m = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

(7)

式(6)中: x_i 为*i*时刻总功率值, w 为滑动窗口长度, n 为样本数量。式(7)中: m 为用电设备编号, y_j 为*j*个总功率短序列样本对应滑动窗中点时刻用电设备*m*的功率值。之后将 X 与 Y 数据标准化,导入到 MMoE 模型中进行训练及测试。

2.2 基于 MMoE 的非侵入式负荷分解

为解决不同用电设备如式(1)的非侵入式负荷分解任务,将标准化后的总功率数据矩阵 X 以及不同用电器的功率序列 Y 输入到 MMoE 模型进行训练及测试,其具体流程见图 4。

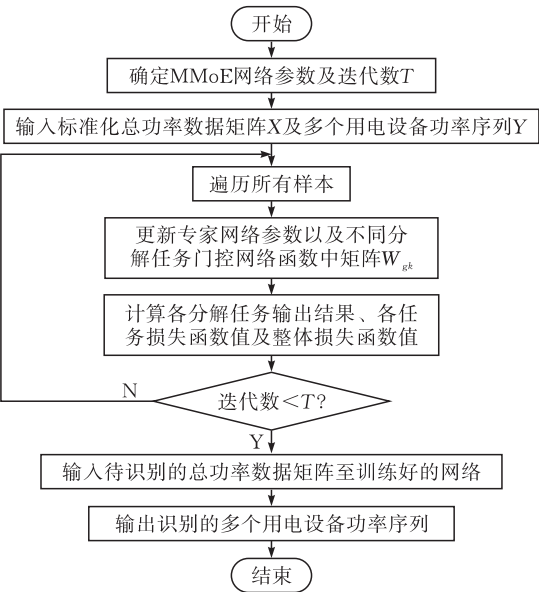


图 4 MMoE 模型训练流程

首先设定好 MMoE 模型的相关参数和训练迭代数 T 之后,将总功率数据矩阵 X 以及多个用电设备的功率序列 Y 至网络。网络通过遍历每个分解任务的训练集样本数据,更新各分解任务中式(5)所示的门控网络数据矩阵 W_{gk} 以及专家网络参数,

计算各用电设备分解任务的输出结果、各任务的损失函数值及整体损失函数值。迭代次数到达最大迭代数 T 之后,导出训练好的网络,输入待识别的总功率数据矩阵,网络输出识别的各用电设备的功率序列,完成负荷分解。

本文能够设置的网络参数为专家网络的数量以及单元数以及训练迭代次数。通过多次实验,本文设置专家网络数量为 270,每个专家网络中含有的神经单元数为 350,迭代数为 150,优化器选用 Adam,学习率设为 10^{-5} 。

3 算例分析

本文选用 REDD^[11] 数据集来验证本文所提模型的有效性,REDD 数据集包含了 6 个家庭的家用电器功率能耗数据,功率采样频率为 1/6 Hz。本文选用数据集中冰箱、洗碗机、浴室电器以及微波炉的有功功率数据。

3.1 数据预处理

选取的用电设备其运行情况各有特点,冰箱属于长时间运行变状态设备,洗碗机运行具有多状态及变状态组合的特点,浴室电器和微波炉皆为启停状态设备,但是浴室电器的高峰大小并不一致,启动运行功率不稳定,微波炉则相对稳定。

为最大限度利用数据,本文取用各电器450 000个采样点数据进行重采样处理,每个采样点功率数据计算平均值,得到的平均功率作为新的采样点数据,将每个设备的原始功率序列转换为长度为 30 000 的功率序列,并将几个用电器的功率序列求和,得到长度为30 000的总和功率序列。

本文通过多次实验,设置滑动窗口长度为 5,同时在总功率序列首端和末端补零,通过滑动窗取值,得到一共30 000条总功率短序列,通过拼接短序列,得到大小为(30 000,5)输入数据矩阵,同时得到 4 个用电设备长度为30 000的功率序列。将数据集的 80% 作为训练集,10% 作为验证集,10% 作为测试集。各设备功率数据信息如表 1 所示。

表 1 用电设备功率信息				W
	最小功率	最大功率	平均值	标准差
总功率	—	—	118.42	258.44
冰箱	0	437	55.63	85.37
洗碗机	0	1144	22.58	140.14
浴室用电	0	1664	7.08	93.21
微波炉	0	1784	21.08	143.43

3.2 负荷分解评价指标

本文选用平均绝对误差 MAE,归一化信号误差 SAE 以及标准化分解误差 NDE 来评价实验结果,如下:

$$MAE = \frac{1}{M} \sum_{i=1}^m |(\hat{y}_i - y_i)| \tag{8}$$

$$NDE = \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i y_i^2} \tag{9}$$

$$SAE = \frac{|\sum_i \hat{y}_i - \sum_i y_i|}{\sum_i y_i} \tag{10}$$

式中： y_i 表示用电设备在 i 时刻的真实功率值， \hat{y}_i 表示用电设备在 i 时刻的预测功率值。

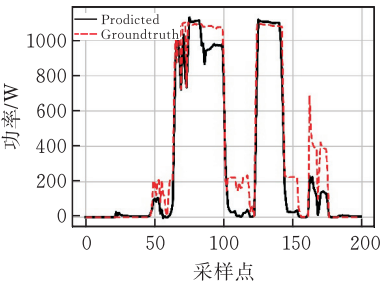
3.3 实验结果对比

本文选用 One-gate Mixture-of-Experts (OMoE) 模型^[7]、传统多任务学习模型 (SBM)^[8] 来做多任务学习模型的实验结果对比,常规 seq2seq 预测模型,LSTM 预测模型做单任务学习模型的实验结果进行对比。得到的 MAE 结果见表 2。

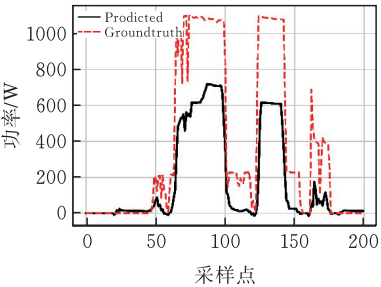
表 2 各模型 MAE 结果对比

	冰箱	洗碗机	浴室用电	微波炉
OMoE	14.06	19.36	13.17	21.22
SBM	30.53	36.32	29.24	28.45
seq2seq	21.12	30.13	27.34	26.31
LSTM	19.81	22.89	15.21	25.31
MMoE	9.69	13.76	7.23	15.10

由表 2 可知,MMoE 模型大部分电器识别结果的 MAE 值都较低,说明了 MMoE 模型识别结果拟合较好,准确度更高。相对于其他模型,OMoE 模型的 MAE 结果与 MMoE 较为接近,即 OMoE 识别的准确性较高,模型识别的拟合程度较好,故本文重点比较两个模型的功率分解结果,洗碗机功率分解结果对比图 5,其余电器对比结果见图 6~图 8。

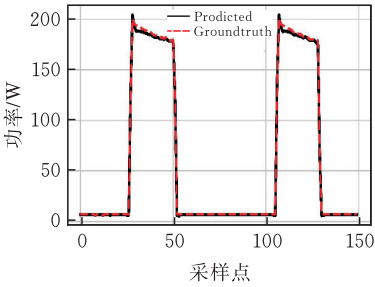


(a)MMoE 分解结果

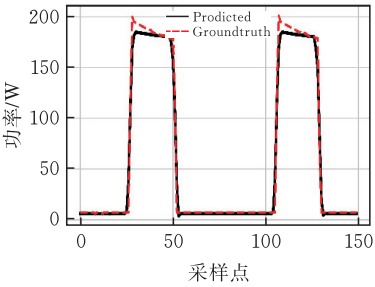


(b)OMoE 分解结果

图 5 洗碗机分解结果对比

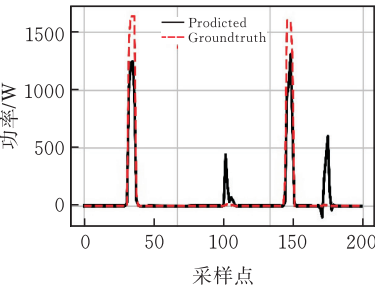


(a)MMoE 分解结果

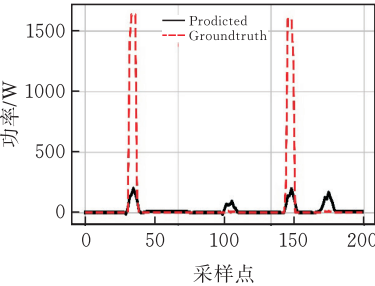


(b)OMoE 分解结果

图 6 冰箱分解效果对比

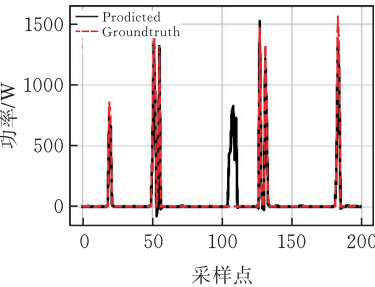


(a)MMoE 分解结果

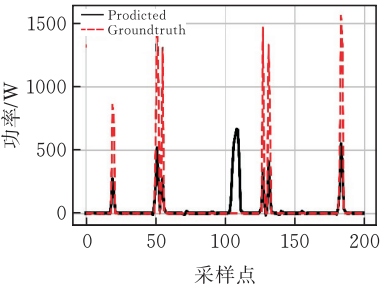


(b)OMoE 分解结果

图 7 微波炉分解效果对比



(a)MMoE 分解结果



(b)OMoE 分解结果

图 8 浴室用电分解效果对比

由图 5~图 8 知,MMoE 模型相比于 OMoE 模型能更准确识别电器的状态变化,如图 5 中洗碗机的运行状态识别,MMoE 模型能够更加准确判断电器的运行情况,同时分解功率值更接近于电器真实运行时的功率值。MMoE 模型与各模型功率分解结果的 SAE 及 NDE 指标数据见表 3 及表 4。

表 3 NDE 实验结果对比

	冰箱	洗碗机	浴室用电	微波炉
MMoE	0.13	0.12	0.32	0.48
OMoE	0.17	0.28	0.85	0.89
LSTM	0.23	0.33	0.91	0.95
seq2seq	0.56	0.67	0.96	0.98

由表 3 可知,MMoE 模型功率分解结果的 NDE 指标相对于其他模型均较低,即 MMoE 功率分解曲线的离散程度较小,更符合实际的功率运行情况。

表 4 SAE 实验结果对比

	冰箱	洗碗机	浴室用电	微波炉
MMoE	0.02	0.14	0.20	0.32
OMoE	0.12	0.26	0.11	0.47
LSTM	0.21	0.32	0.22	0.53
seq2seq	0.34	0.45	0.39	0.61

由表 4 的 SAE 指标结果可知,MMoE 模型对浴室用电的分解结果精度较 OMoE 模型分解结果有所下降,但是其他电器的分解结果精度则相较 OMoE 分解结果较好且整体也均优于单任务学习模型,即 MMoE 模型的分解精度结果更好。综上,MMoE 在曲线拟合程度上更好,分解精度也较高,负荷分解的效果更好。

由于深度学习存在模型训练时间长短的差别,本文将 MMoE 模型训练时间与对照模型的训练时间进行对比见表 5。

表 5 模型训练时间对比

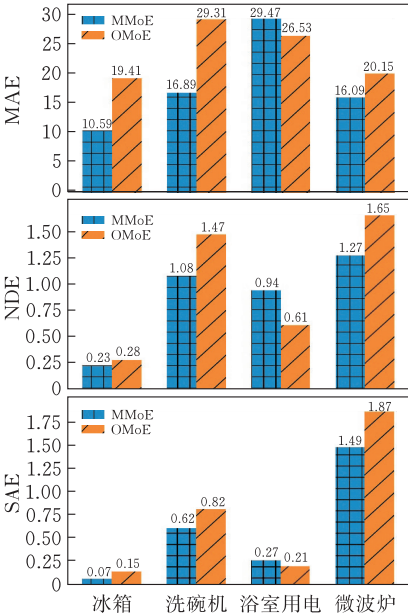
	seq2seq	LSTM	MMoE
训练时间	59.2min	48.3min	13.5min

训练时间选择的对照模型为单任务学习模型,不同用电设备的分解选用参数不同的模型进行训

练,训练时间为训练不同参数模型所用时间的总和。由表 5 可知,相对于单任务学习模型,MMoE 模型可以同时分解多个用电设备的功率且用时最短,效率更高。且由表 2、表 3 及表 4 的实验结果,MMoE 模型的功率分解结果拟合效果更好,离散度更小,精度更高,模型更具优越性。

3.4 泛化性验证

为验证本文模型的泛化性,通过提取其他房间相同电器的功率数据及对应的总功率数据,进行相同的数据处理操作,一共得到 20000 条功率数据,以及大小为(20000,5)的总功率数据矩阵,输入之前实验训练完成的 MMoE 模型,得到 MMoE 模型对其他房间用电设备识别结果,并计算相关指标,且由不同实验指标结果中,OMoE 模型的实验结果与 MMoE 实验结果较为接近,故本文重点对两模型的泛化性实验结果进行对比,具体结果见图 9。



(b)OMoE 分解结果

图 9 泛化性实验结果对比

由图 9 所示,虽然相对于原房间的测试结果各指标都有所下降,但是也仍然维持在较高水平,且也基本优于 OMoE 模型,证明本文所提的模型具有良好的泛化性。

4 结论

本文提出一种基于 MMoE 模型的非侵入式负荷分解方法,首先将低频采样的功率数据进行预处理,通过滑动窗读取用电设备及总功率数据之后,将不同用电设备的分解作为不同的任务,通过训练 MMoE 模型,输入总功率数据矩阵,同时得到多个用电设备的功率序列,实现负荷分解。通过在 REDD 上进行实验,可以得到本文所提模型相对于

其他多任务学习模型具有较高的分解精度以及准确性,同时相对于单任务学习模型,多任务学习模型具有更短的训练时间,更具有实用性。但大功率用电设备运行时,对其他用电设备的分解影响,未来的研究将进一步解决这个问题,也将进一步研究提高模型泛化性的方法。

[参 考 文 献]

[1] KELLY J, KNOTTENBELT W J. Neural NILM: deep neural networks applied to energy disaggregation[C].// Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments .Seoul, Korea: ACM, 2015, 55-64.

[2] 刘仲民,侯坤福,高敬更.基于时间卷积神经网络的非侵入式居民用电负荷分解方法[J].电力建设, 2021, 42 (03): 97-106.

[3] 徐晓会,赵书涛,崔克彬.基于卷积块注意力模型的非侵入式负荷分解算法[J/OL].电网技术: 1-8[2021-09-06]. <https://doi.org/10.13335/j.1000-3673.pst.2020.1487.>].

[4] 王轲,钟海旺,余南鹏,等.基于 seq2seq 和 Attention 机制的居民用户非侵入式负荷分解[J].中国电机工程学报, 2019, 39(01): 75-83.

[5] YANG Y, HOSPEDALES T. Deep multi-task representation learning: a tensor factorisation approach[J/

OL]. (2022-03-09). <https://arxiv.org/abs/1605.06391>.

[6] MISRA I, SHRIVASTAVA A, GUPTA A, et al. Cross-stitch networks for multi-task learning[C].// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.

[7] MA J, ZHAO Z, YI X, et al. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts[C]. ACM, 2018.

[8] Rich Caruana. 1998. Multitask learning. In Learning to learn. Springer, 95-133.

[8] CARUANA R. Multitask learning: a knowledge-based source of inductive bias[C].// Machine Learning, Proceedings of the Tenth International Conference, University of Massachusetts, Amherst, MA, USA, June 27-29, 1993. DBLP, 1993.

[9] EIGEN D, RANZATO M, SUTSKEVER I. Learning factored representations in a deep mixture of experts [J/OL]. (2022-03-09). <https://arxiv.org/abs/1312.4314>.

[10] JACOBS R, JORDAN M, NOWLAN S, et al. Adaptive mixtures of local experts[J]. Neural Computation, 2014, 3(01): 79-87.

[11] KOLTER J Z, Johnson M J. REDD: A public data set for energydisaggregation research[C].// Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA. 2011, 25(Citeseer): 59-62.

Non-intrusive Load Decomposition Based on Multi-task Learning

LIU Hui, JIANG Xucheng

(Hubei Collaborative Innovation Center for Efficient Use of Solar Energy, Hubei Univ. of Tech., Wuhan 430068, China)

Abstract: Non-intrusive load decomposition is one of the key technologies for obtaining user electricity information in the construction of the power grid. In order to solve the problems of the large number of models and long training time in the current non-intrusive load decomposition research, this paper introduces multi-task learning In the study of non-intrusive load decomposition, a non-intrusive load decomposition method based on the Multi-gate Mixture of-Experts (MMoE) model is proposed. First, the power of the electrical equipment is decomposed and converted through the seq2point model. It is the mapping relationship between the total power sequence and the power value of the electrical equipment at the midpoint of the sequence. Secondly, the gating function of the MMoE model and the shared Expert network group are used to take into account the uniqueness and relevance of the power decomposition tasks of different electrical equipment, and finally pass A single MMoE model completes the power decomposition of multiple electrical equipment at the same time. This article verifies the method proposed in this article on the public data set, and the test case verifies the effectiveness of the method in this article.

Keywords: non-intrusive load decomposition; multi-task learning; mmoe; seq2point