

# 以架构演化为中心的负荷预测深度神经网络

Santiago Gomez-Rosero\*, Miriam AM Capretz†, Syed Mir S  
电气和计算机工程系\*†西部大学,伦敦,安大略省,加拿大,  
N6A 5B9 电子邮件: \*sgomezro@uwo.ca, †mcapretz  
@uwo.ca ‡London Hydro,伦敦,安大略省,加拿大,  
N6A 4H6 电子邮件:mirs@londonhydro.com

**摘要** 如今,电力需求预测对电力公司来说至关重要。准确的住宅负荷预测作为集成区域(如邻里负荷消耗)的独立组成部分起着至关重要的作用。短期负荷预测可以帮助电力公司减少浪费,因为电力的存储成本很高。本文提出了一种用于住宅负荷预测的时间序列预测的深度神经网络演化新方法。该方法在演化过程中将其工作重点放在神经网络架构上。然后,使用进化优化技术调整模型权重以自动调整模型性能。包含安大略省伦敦住宅每小时负载消耗的大型数据集的实验结果表明,未经调整的权重架构的性能可与其他最先进的方法相媲美。

过去十年表明,人工神经网络(ANN)是预测建筑物[6]、工业[7]和城市[8]电力消耗的强大且适用的方法。用电量可以用多变量时间序列模型表示,其中输出预测是具有多个输入的函数的结果,例如一天中的时间、星期几、天气和社会因素。基于这个观点,最近使用ANN预测负荷的方法已经从简单的单变量模型发展到复杂的多变量模型。下一个改进是将隐藏层的复杂性从一层增加到多层;这种方法称为多层感知器(MLP)

此外,当调整架构权重时,模型精度超过了最先进的LSTM方法3.0%。

**索引词** 神经网络搜索、深度神经网络、进化算法、短期负荷预测、住宅负荷预测。

## 一、引言

随着可再生能源进入电网系统,电力市场和电控调度变得更加复杂和难以管理。这种现象的结果是需要提前知道电力负载需求。因此,准确的短期负荷预测(STLF)对于电力管理至关重要。此外,准确的STLF模型有助于避免电力过度发电,从而提高收入。最后,良好的STLF模型有助于维持电网系统的稳定和运行。

根据国际能源署[1]的数据,全球家庭能源消耗占所有消耗的27%。在欧盟28国,占29%,特别是在加拿大,占2017年调度电力的33%。因此,预测居民用电量成为短、中、长期研究的重点。和长期预测。最近,提出了多种预测住宅负荷消耗的方法[2]、[3]。其他研究与统计模型相关,例如ARIMA方法[4],还有一些是机器学习对齐方法,例如支持向量机[5]。

人工神经网络模型[9]。在过去十年中,负载预测的最新技术是使用递归神经网络(RNN)[2]、[10]、[11]。这些ANN模型中的绝大多数仅限于固定架构,他们的工作重点是调整模型,以便ANN学习最佳权重集。因此,用于负荷预测的ANN模型更注重权重而不是架构。

在ANN设计过程中,许多优化和规范化技术可用于提高模型的性能。尽管如此,人工神经网络架构[12]可以带来相当多的进步,但它们通常被忽视了。因此,人们对自动化神经架构搜索方法越来越感兴趣。这项研究工作应用进化算法来进化ANN架构,以提高负荷预测模型的性能。

在过去的20年中,随着计算能力的提高,用于时间序列预测的ANN技术的发展一直在上升。在大多数情况下,训练ANN或深度神经网络(DNN)模型的重点一直放在权重上,而忽略了架构的重要性。本文提出了一种开发用于时间序列预测的DNN的方法,该方法增强了神经架构搜索在进化过程中的重要性。本文的贡献可以概括如下:1)一种以架构进化(DNN-CAE)为中心的深度神经网络进化方法。该方法基于DNN的两个主要阶段演化,将演化工作集中在第一阶段的架构开发和第二阶段的权重调整。将演化过程集中在DNN上

体系结构,在第一阶段,每个网络代理的适应度使用共享权重值进行测量。在第二阶段,CMA-ES [13] 技术用于调整 DNN 权重。

- 2) 一种用于负荷预测的神经结构搜索方法,可用于研究学校、建筑物和住宅设施的用电量。出于本文的目的,该方法用于住宅负荷预测。**据作者所知,这是第一个使用神经架构搜索的作品**

**基于进化技术来学习STLF问题的电力消耗行为。**

- 3) 所提出的方法是用真实世界的数据进行评估的,这些数据包括来自安大略省伦敦市一所房子的大约三年的每小时负载住宅数据。

本文的其余部分安排如下。第二部分介绍了用作所提出方法基础的技术背景。第三部分介绍了与住宅数据的 STLF 相关的工作。第四部分描述了 DNN-CAE 模型的方法。第五部分解释了各种实验和获得的结果。最后,第六节提出了未来研究的结论和方向。

二.背景

神经架构搜索 (NAS) [14] 是过去五年开发的一种用于自动化 DNN 架构工程的方法。NAS 方法专门通过遵循性能最大化的搜索策略从所有可能的体系结构中找到最佳体系结构。根据 Liu 等人的说法。[15],NAS 优化方法可分为强化学习、进化算法 (EA)、基于梯度的算法和贝叶斯技术。在这些技术中,使用频率较高的是强化学习和 EA。强化学习技术通常使用策略优化来估计 DNN 参数和结构。钟等。[16] 提出了一种应用强化学习在 RNN 上找到最佳架构的方法。另一方面,EA 技术将 DNN 结构编码到基因组中,这些基因组被进化以找到最佳架构 [17]。优化方法在很大程度上依赖于衡量 DNN 性能的评估方法。用于 NAS 评估的一些专门方法是网络态射 [18]、权重共享 [19] 和超网络 [20]。

EA [21] [22] 是一类受生物进化启发的基于人口的随机搜索技术。这些类型的算法与遗传学共享相同类型的进化机制:突变、适应度分配、选择和后代。

进化机制和配置因算法而异。但是,它们可以概括为代理  $P(\theta)$  的总体,代表优化问题的潜在解决方案。每个代理  $\theta_i$  的性能通过适应度函数  $F(\theta_i)$  在每一代中进行评估。然后选择一些代理成为父母并产生下一个后代,后代通过定义的进化技术发生变异。进化需要

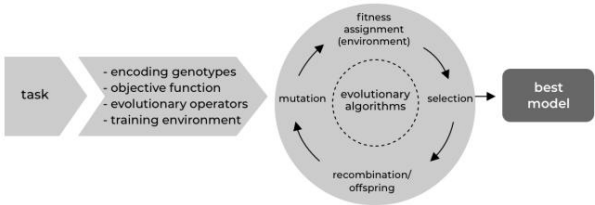


图 1. 进化算法循环。改编自 geneticprogramming.com。

通过 N 个循环重复此过程,或直到满足停止条件 (Back,1996) 。

在将 EA 应用于 DNN 之后,出现了称为神经进化的方法,该方法对 DNN 进行进化,直到找到最适合定义任务的网络代理,如图 1 所示。这些方法从经典学习方法转变为进化方法,其中 DNN 结构、权重等特征,在某些情况下,超参数被编码到基因组中。然后根据性能标准使用 EA 进化基因组。根据 Stanley 等人的说法。[23],使用 EAs 设计 ANNs 在解决方案的搜索空间中呈现出高度并行化和多样性的优势。使用 EA 的另一个优势是,如果在探索过程中将 EA 与质量概念 (例如,最大化奖励) 相结合,群体将分散寻找不同的策略,并迅速指向全局最优。

除了进化策略,另一种用于优化 DNN 的算法称为协方差矩阵自适应进化策略 (CMA-ES) [13]。CMA-ES 是一种基于 EA 的用于非线性和非凸问题的黑盒优化技术。CMA-ES 被认为是进化计算中最先进的技术,并已被用作连续优化问题的标准工具之一。

这种方法创建了一个描述决策变量之间相关关系的协方差矩阵。然后,通过演化机制,最大化矩阵自然生成成功的解决方案。对于 N 维空间,CMA-ES 状态变量由分布均值  $m \in R^n$ 、步长  $\sigma > 0$  和协方差矩阵  $C \in R^{n \times n}$  给出。

CMA-ES 是一种迭代算法,在其每次迭代中,从多元正态分布中采样  $\lambda$  个候选解,对其进行评估,然后调整用于下一次迭代的采样分布 [24]。

三.相关工作

用电量可以通过时间序列模型表示。每个用电量数据集都呈现出与收集数据的位置、天气和社会因素相关的独特属性。因此,由于这种可变性,每个数据集都难以分析和建模。可以从数据中学习模式的统计机器学习的一个重要分支是 DNN 模型。DNN 与高计算能力一起在各种预测任务中显示出出色的性能。郑等。[10] 提出了一种使用 LSTM 对城市智能电网进行负荷预测的方法。马里诺等

阿尔。 [25],介绍了他们使用 DNN 预测建筑能耗的方法。具体来说,对于住宅负荷预测,Zhang 等人。 [26] 提出了使用 SVR 模型预测单个房屋负荷的观点。最近,Wang 等人。 [27] 提出了一种使用应用于 LSTM 的概率方法的方法。提出的大部分改进都与通过改进 ANN 和找出如何有效地学习神经元的权重来增强负荷预测有关。尽管如此,我们还没有找到可以改进用于时间序列预测的 DNN 架构的工作。此外,据作者所知,还没有做任何工作来自动化架构搜索负载

预测。

在 STLF 中,EA 通常用于优化预测模型中的超参数,例如 Zhang 等人提出的工作。 [28] 调整具有差异进化的基于 SVR 的模型。在最近使用 EA 的一项研究中,Zeng 等人。 [10] 将粒子优化技术与一种称为极限学习机的方法相结合。在另一项研究中,Bouktif 等人。 [29] 建议使用 LSTM 模型,结合特征选择和遗传算法进行负荷预测。在应用于负荷预测的神经进化技术方面,Srinivasan [30] 提出了一种进化 ANN 的方法。在这种方法中,架构固定在三层,只有 ANN 的超参数和权重得到改进。

相比之下,到目前为止提出的技术一直专注于调整模型超参数,或者在最佳情况下,以改进 DNN 权重。除了图像优化之外,很少有方法专注于搜索最佳 DNN 架构。Gaier 和 Ha [31] 提出的称为权重神经网络的方法侧重于为强化学习任务演化出一种稳健的体系结构,该体系结构可以用作时间序列预测等其他任务的研究基础。

本文提出了一种寻找以架构演化 (DNN-CAE) 为中心的神经网络以进行短期负荷预测的方法。该方法侧重于使用 EA 技术的 NAS。DNN-CAE 搜索用于寻找最佳架构,在训练模型时撇开权重的重要性不谈。最后,在找到最佳架构后,使用 CMA-ES 算法调整权重,从而提高负载预测性能。

四、DNN-CAE方法论

本文提出的方法已分三个阶段实施,如图 2 所示,即:工程化数据集、神经网络架构的演化和权重调整。接下来将详细解释这些阶段中的每一个阶段:

A. 第 1 阶段 :设计数据集

在此阶段,在将数据集分为训练集和测试集之前执行三个步骤。清洗数据集、特征工程和归一化和分裂的步骤

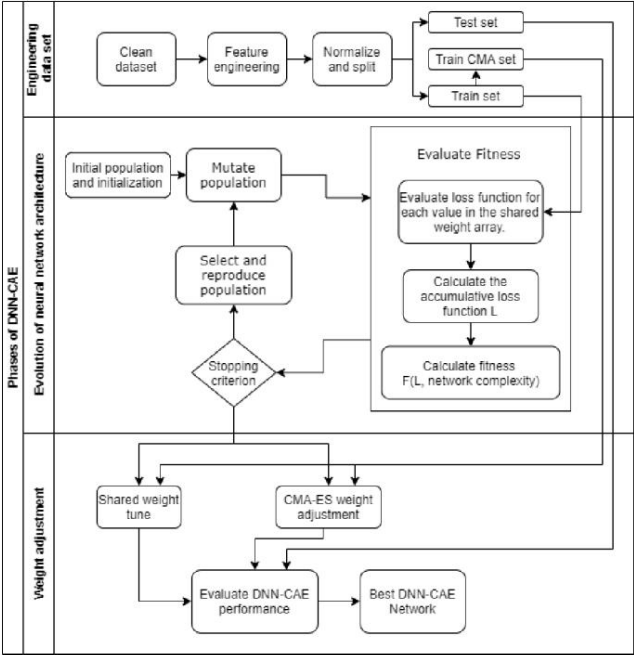


图 2. DNN-CAE 模型的框架。

数据集对于算法的充分训练和性能至关重要。

1)清理数据集 :通过删除所有 NaN 和重复值来清理数据集。此外,还会检测并删除异常,例如非顺序数据、单位错误的值和填充缺失值。在这一步之后,收集到的整个数据变得一致。

2)特征工程 :在这一步中,增加了数据集中的特征数量。天气属性相关

与添加数据相同的日期和时间,例如温度和天气状况作为分类值。新特征也被添加到数据集中,例如星期几、工作日、周末、假期和一年中的季节。最后,通过正弦函数将月、日、小时、星期几等循环特征转化为循环值。还添加了来自与新功能相同的集合的过去值,例如从最后一小时到最后 48 小时的先前目标值。前 24 小时和 48 小时的平均值,以及上周同一时间、上个月和去年的值。

3)归一化和分割数据集 :在这一步中,整个数据集通过应用 (1)中给出的最小-最大归一化来归一化。然后,将数据集按照80%-20%的比例分成训练集和测试集。在 DNN 的演化过程中,模型使用训练集进行训练。然后 CMA-ES 算法使用最后 20% 的训练集来调整权重。采用这种方法来改进 DNN 的权重,使其具有与测试集类似的行为,但使用训练集数据。执行此过程还可以提高权重调整阶段的模型准确性。最后,测试集用于模型性能测量。

表一

进化过程中使用的参数说明。
参数 说明Ps网络代理的人口规模。
Swb Shared 权重负边界和正边界。
Sws共享权重列表大小。
Gmax最大世代数。
Facebook健身阈值。
Mnode插入节点的概率。
Mconn添加连接的概率。
Mact改变激活函数的概率。
Afl可用激活函数列表。
Ts锦标赛规模。

$$x = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1}$$

B. 阶段 2:神经网络架构的演变

在此阶段,执行具有 EA 优化方法的 NAS。这种方法通过降低权重的重要  
性来关注架构的演变。一般来说,在进化阶段,架构搜索通过使用共享权重值  
对不同的适应度测量进行采样来避免权重训练和调整。每个网络代理都在一  
组共享权重值上进行评估,并记录累积损失函数。最后,将创建后代的父母是  
通过比赛选择来选择的。

按照进化机制重复此过程,直到找到具有最佳未调整权重架构的模型。表 1  
描述了搜索架构演变过程中使用的各种参数。

算法 1 描述了 DNN-CAE 方法的神经网络架构的演变,如图 2 所示。在以  
下小节中,将描述如图 2 所示的神经网络架构演变的每个组件;描述参考算  
法 1 中的相应行。

- 1) 初始种群和初始化:创建具有最小神经网络拓扑结构的初始种群。

每个网络代理 $\theta_i$ 从一个输入层和一个输出层开始。在这一步中,输入层和输出  
层之间没有连接。根据Ps 设置网络代理种群  $P(\theta)$  的大小。此外,在此步骤中,  
共享权重 $w_j \in asw$ 的各种值由 (2) 定义。共享权重数组 $asw$ 、当前代  $g$  和当  
前代的最佳适应度 $fg$ 也被初始化。如算法 1 第 3 行所示,该步骤在演化过程  
中仅执行一次。

$$w_j = Swb - 1 + \frac{2j}{瑞士}, \forall j = 0, 1, \dots, Sws \tag{2}$$

2) 停止标准:这里算法检查是否已达到Gmax或Fb。如果达到其中一个标  
准,则进化交互停止,最佳代理 $\theta_{best}$ 被保存,并且程序前进到下一阶段。如果  
两个条件都不满足,则迭代演化过程继续。参数Gmax和Fb都在第 1 行中设  
置

算法 1:体系结构演化1参数: Ps、
Swb、Sws、Gmax、b、 Mconn、Mnode、 Afl、 Ts。
2输出:具有最佳适应度的 $\theta_{best}$ 。
3初始化:大小为Ps的 $P(\theta)$ , $asw$ ,其中 (2), $g = 0$ , $fg = -\infty$ 。
4 while $g < Gmax \wedge fg < Fb$ do /* Mutate population for $i = 1, \dots, Ps$ $\theta_i \in P(\theta_i)$ do - 插入一个概率为
Mnode的节点, - 添加一个概率为 Mconn 的连接, - 改变一个将现有的激
活函数替换为Afl 中列出的另一个。 */
5
6
7 结束
/* 评估 $i = 1, \dots, Ps$ $\theta_i \in P(\theta_i)$ do for
8 $j = 0, 1, \dots, Sws$ $w_j \in asw$ do (3)计算 $l_j(\theta_i, w_j)$ ,
9 $w_j$ ) , 计算 $L(\theta_i) = l_j(\theta_i, w_j)$ 。
10
11
12 用 (4)计算 $F(\theta_i)$ 。
13 结束
/* 选择并复制initialize: offspring. for $i =$
14 $1, \dots, Ps$ do从 $P(\theta_i)$ 中随机选择Ts个候选者。 */
15
16
17 从锦标赛中选择具有最佳适应度的 $\theta_i$ 并将其复制给后代。
18
19 结束 $P(\theta_i) \leftarrow$ 后代。
20 $fg \leftarrow$ 最好的 $F(\theta_i) \in P(\theta_i)$ 。
21 $g = g + 1$ 。 22结束
23 $\theta_{best} \leftarrow \theta(fg)$

算法 1,第 4 行中的 while 循环运行,直到满足一个或两个条件。

3) 变异种群:为了避免局部最优,必须在每一代中对每个智能体  $\theta_i$ 实施变  
异。根据 Gaier 和 Ha [31] 的工作,每个代理 $\theta_i$ 通过直接实施三种可能的突  
变之一来提高其性能,这三种突变是 1) 插入节点,2) 添加连接,3) 改变激活  
函数。选择前两个突变是因为它们增强了 $\theta_i$  的行为和性能。请注意,这些突变  
增加了架构的复杂性。

为了插入一个节点,将两个节点之间的现有连接拆分为两个连接,然后使  
用随机激活函数添加一个新节点。要添加一个连接,随机选择两个不属于同  
一层的未连接节点,然后创建一个新连接。最后,激活函数的突变将一个函数  
更改为另一个随机选择的函数。

算法 1 的第一个 for 循环 (第 5 至 6 行)执行变异步骤,插入节点、添加  
连接和更改激活函数的概率分别由参数Mnode、 Mconn和Mact设置。  
Afl中给出了允许的激活函数列表。

4) 评估适应度:该步骤分为三个动作,如图 2 和算法 1 的第 7 至 13 行所  
示,即 1) 评估损失函数,2) 计算累积损失函数,以及 3) 计算每个 $\theta_i$  的适应度  
值。

在第一个动作中,所有的 $\theta_i$ 权重被设置为一个单一的

共享价值 $w_j$ 。然后评估网络代理在批训练集上的性能,并使用算法 1 的 (3) (第 8 至 10 行) 计算损失函数 $l_j(\theta_i, w_j)$ 。

$$l_j(\theta_i, w_j) = - \sum_{k=0}^{批} \text{abs} (^yk(\theta_i, w_j) - y_k) \tag{3}$$

其中 $\theta_i$ 是要评估的网络代理,  $w_j$ 是共享权重值,  $y_k(\theta_i, w_j)$  是使用 $w_j$ 评估的网络的预测,  $y_k$ 是训练集中的目标值。损失函数是负的,因为算法最小化每一代的损失函数。

第二个动作计算累积损失函数  $L(\theta_i)$  (第 11 行)。这里,将为每个 $w_j$ 计算的损失函数 $l_j$ 相加。最后,第三个动作计算算法 1 中每个网络代理的适应度  $F(\theta_i)$  (第 12 行)。适应度函数如 (4)所示:

$$F(\theta_i) = -L(\theta_i) - N_c(\theta_i) - N_n(\theta_i) \tag{4}$$

其中  $L(\theta_i)$ 是计算出的累积损失函数,  $N_c$ 是网络中活动连接的总数,  $N_n$ 是网络中的节点数。使用 (4)的优点是当两个代理**呈现相同的累积损失函数时;具有更简单架构的智能体获得更高的适应度分数。**

5)选择和重现:在可能的父母选择技术中,选择锦标赛选择是因为这种方法在 NAS 方法中取得了良好的效果,例如 Real 等人提出的方法。[17] 和刘等人。[15]。

这一特征避免了进化过程因最适合个体的支配而停滞,并允许在父母选择过程中产生多样性以创造下一代。

通过使用算法 1 的锦标赛选择 (第 14 至 18 行),根据锦标赛规模 $T_s$  随机选择个人。从比赛中选出的最佳个体是下一代后代的一部分。重复此过程,直到达到 $P_s$ 。

C. Phase 3: Weight adjustment 此阶段

接收来自阶段 2 的具有最佳未调整权重架构的模型,并通过权重调整,增强 DNN 模型性能。因此,在此阶段,分析权重以创建 DNN-CAE。

1) 共享权重调整:在这一步中,分析了共享权重对最佳未调整权重架构模型的影响。该模型在参数 $Swb$  指定的值范围内进行测试。通过减小采样步长,找到最大化模型精度的共享权重值。然后这个最小化错误的共享值被拾取并发送到 DNN-CAE 性能评估步骤,如图 2 所示。

2) CMA-ES 权重调整:在这一步中,CMA-ES 技术用于演化最佳未调整权重架构模型的权重并调整 DNN。此步骤使用可被视为基线版本的版本,具有非精英  $(\mu, \lambda)$  选择功能。所有调整常数都设置为其默认值

值,由 Hansen [24] 规定。在此步骤中,CMA-ES 算法与模型交互,直到它找到权重的最佳个体值集并将这些值传递到下一步。

3) DNN-CAE 性能评估:最后,在这一步中,测试了具有**优化共享权重和 CMA-ES 权重值的最佳未调整权重架构模型。**

为此,评估了测试集上的模型性能,首先使用共享权重的值,然后使用 CMA-ES 值。最后,选择最能提高性能的一组权重。因此,DNN CAE 模型是通过以架构为中心的进化方法找到的最佳架构和最佳权重集。最后,计算 RMSE (5) 和 MAPE (6) 以衡量模型在测试集上的性能。

$$均方根误差 = \sqrt{\frac{1}{n} \sum_{k=0}^n (y_k - \hat{y}_k)^2} \tag{5}$$

$$地图 E = \frac{1}{n} \sum_{k=0}^n \frac{y_k - \hat{y}_k}{y_k} \times 100 \tag{6}$$

其中 $y_k$ 表示预测用电量,  $\hat{y}_k$ 表示家庭实际用电量, $n$  是观察次数。

五、实验

DNN-CAE 作为 DNN 模型用于短期负荷预测,该模型访问当前时间步长的输入和之前的用电量,以预测下一小时的用电量。所有模型和实验都在配备 24 个英特尔(R) 至强(R) E5-2630 v2 处理器的 Linux 服务器上运行。

A. 数据

住宅用电模式呈现出复杂的非线性关系,这使它们成为一个具有挑战性的问题。电力消耗的特征取决于工作日和一天中的时间,如图 3 所示。一个家庭的住宅负荷消耗数据由伦敦水电公用事业公司提供。来自智能电表的原始数据集包含在一小时时间尺度上以千瓦时 (kW-h) 测量的用电量数据。历史每小时天气和温度数据来自加拿大政府官方网站[32]。

该数据集包含从 2014 年 1 月 1 日到 2016 年 12 月 31 日的每小时数据。

B.共享权重范围评估和参数选择为了充分运行模型,有必要评估各种共享权重值的操作范围。

表 II 显示了共享权重性能测试的结果,每个测试都使用不同范围的共享权重和不同的采样步长。在每个性能测试中,测量中值适应度和最佳个体适应度。



表 I 共享权重		
性能测试。		
共享权重值	Best agent Range Sampling	最佳共享重量
fitness between 0.5 and 1.0	0.5 1.0 0.5 1.0 0.5	
[-10, 10]	3000	0
	6130	0
[-5, 5]	1020	0
	1550	0
[-2, 2]	663	0.1
	419	0.1
[-1.5, 1.5]	96	0.4
	85	0.4
* Gmax = 256 和 Ps = 128 的性能测试。		

表三	
实验1中使用的参数	
范围	值128
附言	1024
Gmax	0.10
b	0.25
麦肯	0.25
节点	0.50
马夫	线性、
AFL	阶跃、正弦、高斯、tanh、S 型、
Ts	逆、绝对、ReLU 32

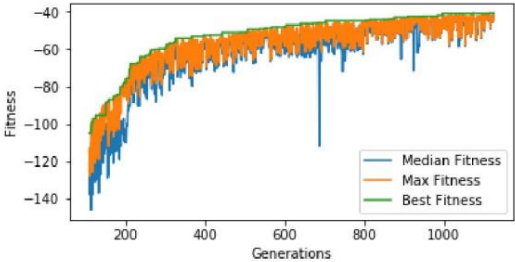


图 4. 适应度图显示了每一代的适应度演变。为了更好地可视化,删除了前 30 代的值。

然后选择最佳共享权重值,并调整 DNN CAE 权重以衡量健身表现。

表 II 显示最佳适应度值随着评价范围变大。因此,失去了对最佳选择权重的收敛性。相反,如果所选范围变窄,则最佳适应度值会降低并开始收敛到一个不同于零的值。根据表 II 的结果,Swb 的限值设置为 [-1, 1],Sws = 0.1。

C. 实验 1:用于住宅负荷预测的 DNN-CAE 为运行实验 1 选择的参数如表 III 所示。这些参数是在几次模型运行后根据经验选择的。图 4 显示了超过 1024 代的模型演化。模型演化耗时 11 小时 46 分钟,深度为 644 个节点。

1) DNN-CAE 共享权重适应度评估:图 5 显示了测试集上共享权重 DNN-CAE 适应度性能的结果。在此评估期间,模型在 [-1, 1] 范围内进行了测试。图 5 显示 DNN CAE 共享权重的最佳性能出现在 0.6 和 0.8 之间。最大适应度值为 0.71,适应度为 -45.44。

2)CMA-ES调整性能评估:运行CMA-ES算法单独调整权重后,评估DNN-CAE模型在测试集上的性能

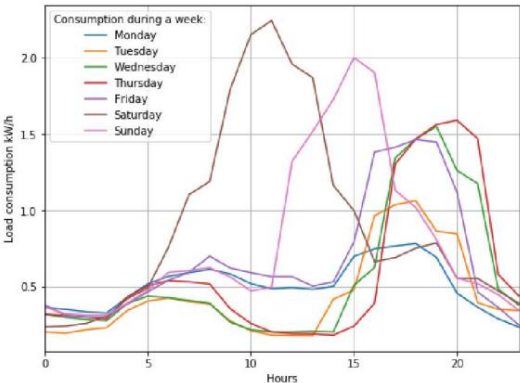


图 3. 2015 年随机一周的住宅日用电量。

吃了。CMA-ES 权重演化运行了 500 多代,历时 2 小时 46 分钟。

3) DNN-CAE 调整后的性能:在用每种调整技术的最佳值调整 DNN CAE 模型后,进行评估。每个实验都进行了三个实验:最佳适应度、RMSE 和 MAPE;计算这些值的平均值并列于表 IV 中。对于 RMSE 和 MAPE,值越低,模型越准确。图 6 显示了模型如何在这两种技术的测试集的第一周内拟合测试集。

D. 实验二:基于激活函数选择的进化行为在这个实验中,进化了三个模型来分析它们的进化行为,并限制可用的

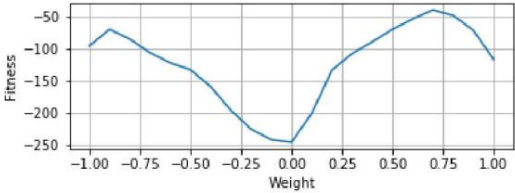


图 5. [-1, 1] 和 Sws = 0.01 范围内共享权重的 DNN-CAE 适应度性能。

表IV			
实验1性能调优技术比较			
调整技术	最佳适应度	RMSE (kW-h)	MAPE (%)
共享权重	-61.41	0.2623	22.18
CMA-ES	-34.75	0.1885	9.51

表五

实验2期间修改的参数			
范围	WALF-3AF	WALF-5AF	WALD-9AF线
AFL	逆,abs,ReLU	tanh,sigmoid,逆函数,abs,ReLU	性,阶跃,正弦,高斯,tanh,sigmoid,逆,abs,ReLU
Gmax	500		

表六

限制激活函数的比较结果				
模型	进化时间 (分)	CMA 时间 (分钟)	马佩 (%)	
			共享权重	CMA-ES
WALF-3AF	93 122	35 57	18.80	8.16
WALF-5AF		71	31.32	7.74
WALF-9AF				9.49

激活函数。第一个模型,CAE-3AF,只设置了逆激活函数、绝对激活函数和 ReLu 激活函数;第二个模型 CAE-5AF 具有最常用的可用激活函数,它们是: S 型函数、双曲正切函数、逆函数、绝对函数和 ReLu。最后,第三个模型 CAE-9AF 具有所有可用的激活函数。这三个模型是

与实验 1 中的参数相同,仅对Ps和Afl 进行了修改。表 V 显示了为运行实验 2 而修改的参数值。

对于这个实验,错误阈值被移除,模型不断进化,直到达到最大世代数。还记录了进化每个模型所需的时间。表 VI 显示了该实验的结果。

显然,CAE-9AF 的误差高于其他两个模型,这可能是在所有激活函数中进行广泛搜索的结果。另一方面,CAE-3AF 和 CAE-5AF 在 500 代后显示出最接近的误差值。尽管如此,具有五个激活函数的模型比具有三个激活函数的模型快 29%。到目前为止,这些实验已经确定 CAE-5AF 模型表现良好,并且进化它的时间是最佳的。

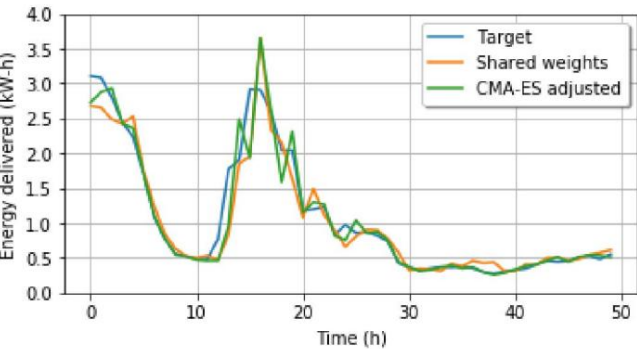


图 6. 共享权重和 CMA-ES 调整的 DNN-CAE 性能。出于可视化目的,仅显示了从测试集的第一周开始的 48 小时。

表七

基准模型参数			
方法	# 特征	# 隐藏层 2 2	# 隐藏节点/层 5 20 12 20
香草 LSTM	1	2 2	5 20 12 20
多层板			
长短期记忆特征	33 6 33	2 2	5 20 12 20
LSTM 一次性			

表八

负荷预测评估总结			
方法	均方根误差 (kW-h)	MAPE (%)	最好的时代
Vanilla LSTM 0.2074	1500 16.84	1475 15.69	
0.1744 MLP-T/2 小时后退 0.2117	500 15.41	1200 15.06	
6 小时后退 0.2157	16.74	1000 14.94	
0.2188 MLP-D/1 天后退 0.2130	900 16.54	1100 13.69	
2 天后退 0.2153	13.69	1500 11.40	
0.2211 LSTM/2 小时回溯 0.1938	10.72	700 14.04	
6 小时回溯 0.1610	= 1000 7.72	gmax =	
0.1565 DNN-CAE SW 0.2198	1000 7.72	gmaxgmax	
0.1707	= 1000 gmax	gmaxgmax	

E. 实验 3:基准比较 在实验 3 中,对所提出的 DNN-

CAE 和传统的 ANN 方法进行了基准比较。为了进行比较,Kong 等人提出的基准。[33] 被转载。选择用于比较的方法是 MLP, vanilla LSTM 和 LSTM One-shot (LSTMOS)。Vanilla LSTM 是最简单的方法,住宅负荷时间序列只有一个输入。然而,LSTM-features 模型具有额外的训练特征,例如时间、工作日和假期。LSTMOS [33] 是住宅负荷预测的最新技术。表 VII 显示了用于配置 ANN 的参数,如 [33] 中所示。

MLP 和 LSTMOS 使用不同的反向时间步长进行训练。遵循 [33] 中使用的约定 -T 用于向后的小时步长, -D 用于向后的天步长。如表 VIII 所示,具有共享权重的 DNN-CAE 的结果与 MLP 方法的结果接近。调整 DNN-CAE 权重后,模型精度超过最佳 LSTM 一次性模型 3.0%。因此,DNN-CAE 是列出的住宅负荷预测 STLF 任务算法中的有力竞争者。

六.结论

本文提出了一种基于 EA 的 NAS 方法,在演化阶段首先关注架构,从而创建了用于短期住宅电力负荷预测的神经架构。后来,通过调整权重来调整模型,从而提高了模型性能。进化从简单的神经架构开始,随着模型的进化而变得复杂。

如图所示,DNN-CAE 已经过优化,可以与

在一系列共享权重值上的单个共享权重参数。因此,该模型很容易调整,以通过共享权重实现可接受的性能。此外,使用 CMA-ES 技术可以单独调整权重,从而提高模型的准确性。此外,在运行了三个不同的实验后,DNN-CAE 显示出 90% 的准确率。这项工作中提出的模型可以在一般意义上用于其他类似的时间序列预测领域,例如能源公用事业、天气和金融。

DNN-CAE 共享权重呈现的结果与 MLP 的结果相似。此外,权重调整后的 DNN-CAE 优于 LSTM 单次方法,这是住宅负荷预测的最新技术。这项工作中提出的结果是首次应用于短期负荷预测的此类结果。作为未来的工作,作者计划研究如何提高模型的准确性。本文还仅评估了一个住宅数据集的模型性能。在未来的工作中,DNN-CAE 将用于为同一地区的不同房屋创建通用架构,然后通过调整权重,该模型将能够将其预测从一个房屋调整到另一个房屋。最后,将对 DNN-CAE 进行敏感性分析,以确定每个参数对模型的重要性。

参考

[1] “国际能源署,世界人均用电量”,  
<https://www.iea.org>,访问时间:2020-03-30。

[2] W. Kong,ZY Dong,DJ Hill,F. Luo 和 Y. Xu,“基于居民行为学习的短期住宅负荷预测”,  
IEEE 电力系统交易,卷。 33,没有。 1,第 1087–1088 页,2018 年 1 期。

[3] M. Imani 和 H. Ghassemlian,“使用小波和协作表示变换的住宅负荷预测”,Applied Energy, 卷。 253,页。 113505,2019 年 11 月。

[4] D. Alberg 和 M. Last,“使用基于滑动窗口的 ARIMA 算法在智能电表中进行短期负荷预测”,越南计算机科学杂志,卷。 5,没有。 3-4,第 241–249 页,2018 年第 9 页。

[5] Y. Chen,P. Xu,Y. Chu,W. Li,Y. Wu,L. Ni,Y. Bao 和 K. Wang,“使用支持向量回归进行短期电力负荷预测 (SVR)模型来计算办公楼的需求响应基线”,Applied Energy,卷。 195,第 659–670 页,2017 年第 6 页。

[6] MQ Raza 和 A. Khosravi,“基于人工智能的智能电网和建筑负荷需求预测技术综述”,  
可再生和可持续能源评论,卷。 50,第 1352–1372 页,2015 年第 10 页。

[7] YT Chae,R. Horesh,Y. Hwang 和 YM Lee,“预测商业建筑每小时用电量的人工神经网络模型”,能源与建筑,卷。 111,第 184–194 页,2016 年第 1 页。

[8] A. Ghani,M. Baqar,JE Khuda,K. Raza 和 S. Yasin,“卡拉奇市基于人工神经网络的电力负荷预测 横断面研究”,2019 年第 13 届国际数学会议,精算学、计算机科学和统计学 (MACS)。 IEEE,2019 年 12 月,第 1-5 页。

[9] S. Fallah,R. Deo,M. Shojafar,M. Conti 和 S. Shamshirband,“智能能源管理电网中能源负荷预测的计算智能方法:最新技术、未来挑战和研究方向”,能源,卷。 11,没有。 3,页。 596, 3 2018。

[10] Jian Zheng,Cencen Xu,Ziang Zhang 和 Xiaohua Li,“使用基于长期记忆的递归神经网络的智能电网中的电力负荷预测”,2017 年第 51 届信息科学与系统年会 (CISS)。 IEEE,2017 年 3 月,第 1-6 页。

[11] B. Zhang,JL Wu 和 PC Chang,“用于短期负荷预测的基于多时间序列的递归神经网络”,Soft Computing,卷。 22,没有。 12,第 4099–4112 页,2018 年第 6 期。

[12] T. Elsken,JH Metzen 和 F. Hutter,“神经架构搜索:A 调查”,arXiv 预印本 arXiv:1808.05377,2018 年。

[13] N. Hansen 和 A. Ostermeier,“进化策略中完全去随机化的自我适应”,进化计算,卷。 9,没有。 2, pp. 159–195, 6 2001。

[14] B. Zoph,V. Vasudevan,J. Shlens 和 QV Le,“学习可扩展图像识别的可迁移架构”,CVPR,第 8697–8710 页,2018 年。

[15] C. Liu,B. Zoph,M. Neumann,J. Shlens,W. Hua,L.-J. Li, L. Fei-Fei,A. Yuille,J. Huang 和 K. Murphy,“渐进式神经结构搜索”,计算机科学讲义 (包括子系列人工智能讲义和生物信息学讲义), 2018,卷。 11205 LNCS,第 19-35 页。

[16] Z. Zhong,Z. Yang,B. Deng,J. Yan,W. Wu,J. Shao 和 C.-L. Liu,“BlockQNN:高效的块式神经网络架构生成”,IEEE 模式分析和机器智能汇刊,第 1-1 页,2020 年 8 期。

[17] E. Real,S. Moore,A. Selle,S. Saxena,YL Suematsu,J. Tan,QV Le 和 A. Kurakin,“图像分类器的大规模演化”,第 34 届国际机器学习会议,ICML 2017,卷。 6, 2017, 第 4429–4446 页。

[18] T. Wei,C. Wang,Y. Rui 和 CW Chen,“网络态射”,第 33 届国际机器学习会议,ICML 2016,卷。 2, 2016, 第 842–850 页。

[19] R. Luo,F. Tian,T. Qin,E. Chen 和 TY Liu,“神经结构优化”,神经信息处理系统进展,卷。 2018-12 月,2018 年,第 7816–7827 页。

[20] A. Brock,T. Lim,JM Ritchie 和 N. Weston,“SmaSH:通过超网络进行一次性模型架构搜索”,第 6 届国际学习表示会议,ICLR 2018 - 会议论文集,2018 年第 8 期。

[21] J. Holland,“人工和自然系统的适应”,安娜堡: 密歇根大学出版社,p。 232,1975。

[22] I. Rechenberg,“Evolutionsstrategien”,载于 Medizin und Biologie 中的模拟方法。 斯普林格,1978 年,第 83-114 页。

[23] KO Stanley,J. Clune,J. Lehman 和 R. Miikkulainen,“通过神经进化设计神经网络”,《自然机器智能》,卷。 1,没有。 1,第 24-35 页,2019 年 1 期。

[24] N. Hansen,“CMA 进化策略:教程”,arXiv 预印本 arXiv:1604.00772,2016 年。

[25] DL Marino,K. Amarasinghe 和 M. Manic,“使用深度神经网络构建能源负荷预测”,发表于 IECON 2016 - IEEE 工业电子学会第 42 届年会。 IEEE,2016 年 10 月,第 7046–7051 页。

[26] XM Zhang,K. Grolinger,MA Capretz 和 L. Seewald,“预测住宅能源消耗:单户视角”,2018 年第 17 届 IEEE 机器学习与应用国际会议 (ICMLA)。 IEEE,2018 年,第 110-117 页。

[27] Y. Wang,D. Gan,M. Sun,N. Zhang,Z. Lu 和 C. Kang,“使用 pinball 损失引导 LSTM 进行概率个体负荷预测”,Applied Energy,卷。 235,第 10–20 页,2019 年。

[28] F. Zhang,C. Deb,SE Lee,J. Yang 和 KW Shah,“使用加权支持向量回归和差分进化优化技术预测建筑能耗的时间序列”,  
能源和建筑,卷。 126,第 94–103 页,2016 年。

[29] S. Bouktif,A. Fiaz,A. Ouni 和 M. Serhani,“使用特征选择和遗传算法进行电力负荷预测的最佳深度学习 LSTM 模型:与机器学习方法的比较”,  
能源,卷。 11,没有。第 7 页1636 年,2018 年 6 月。

[30] D. Srinivasan,“用于短期负荷预测的进化人工神经网络”,Neurocomputing,卷。 23, 没有。 1-3,第 265–276 页,1998 年第 12 页。

[31] A. Gaier 和 D. Ha,“重量不可知神经网络”,在 Advances in 神经信息处理系统,2019 年,第 5364–5378 页。

[32] “加拿大政府,历史 <https://climate.weather.gc.ca>,访问: 气候 数据,” 2020-03-28。

[33] W. Kong,ZY Dong,Y. Jia,DJ Hill,Y. Xu 和 Y. Zhang,“基于 lstm 递归神经网络的短期住宅负荷预测”, IEEE 智能电网交易,卷。 10,没有。 1,第 841–851 页,2019 年。