文章编号：1003-0417(2020)09-25-06

基于深度神经网络构建风电机组性能模型的超参数选择

吴　莎[1*]，汪　健[1]，谢　新[2]，蒋紫虓[2]，邓少平[1]，卢　胜[1]
(1. 湖北省电力勘测设计院有限公司，武汉 430040；2. 美迪欧动力气象技术 (北京) 有限公司，北京 100027)

摘　要：风电行业中基于深度神经网络技术来构建风电机组性能模型受到了业内人员的广泛关注，然而对于深度神经网络自身超参数选取对模型结果的影响仍缺少系统性分析。以深度神经网络隐含层层数、深度神经网络形状及隐含层神经元数量为研究基础，研究了超参数取值对风电机组性能模型构建效果的影响，并通过实际运行数据验证了参考风电机组的最优配置及其推广性。结果表明，深度神经网络隐含层层数、深度神经网络形状及隐含层神经元数量这 3 个超参数的不同配置对最终构建的风电机组性能模型存在一定影响，但从工业应用的角度来看，这些影响基本可以忽略；同时，基于参考风电机组得到的最优超参数配置能够直接推广到同一风电场相同型号的所有风电机组上应用。

关键词：风电机组性能模型；深度神经网络；超参数；后评估
中图分类号：TM513.5　　**文献标志码**：A

0 引言

随着计算机计算能力的不断提高，机器学习这一技术在各个工业领域中逐渐得到应用。与此同时，随着国内风电平价上网进程的推进，风电场后评估工作在风电行业内不断受到关注；而风电机组性能评估作为风电场后评估工作中最重要的内容之一，其与机器学习技术的结合自然受到了行业内的广泛关注[1-2]。

风电机组性能评估依赖于建立风电机组性能模型，即建立风电机组输出功率与风电机组各关键参数 (如风速、桨距角、偏航角等) 之间的函数关系，这属于典型的非线性回归问题；在机器学习领域，深度神经网络是解决该问题的经典手段，因此，利用其构建风电机组性能模型自然得到了行业内研究者的广泛尝试。比如，文献 [3] 尝试将神经网络拟合的风电机组功率曲线与测量的风电机组功率曲线进行对比分析；文献 [4] 验证了采用卷积神经网络构建的风电机组齿轮箱故障识别分类器比一般的支持向量机 (SVM) 分类器具备更好的性能表现；文献 [5] 提出了基于引力搜索算法优化人工神经网络的风电机组传动链故障诊断方法；文献 [6] 采用深度神经网络构建了风电机组齿轮箱状况检测框架，以用于故障预警；文献 [7] 利用线性混合自组织映射神经网络对风力机状态特征进行分类，提出了一种特征变化累积趋势差分法，用于预测风力机的异常和故障。上述研究的关注点主要在于深度神经网络在具体场景中的实际应用，所构建的模型大多仅针对研究中所使用的特定风电机组。

综上，本文尝试从深度神经网络本身出发，探讨其自身超参数的取值对最终风电机组性能模型的准确度的影响，同时将验证以上超参数在参考风电机组训练中得到的最优配置是否可以直接推广应用于同一风电场同一型号的其他风电机组上。针对以上研究内容，本文首先介绍了深度神经网络的基本概念及所选取的超参数，并从某

收稿日期：2019-11-26
通信作者：吴莎 (1988—)，男，硕士，主要从事新能源发电系统方面的研究。651521101@qq.com

2020-09杂志.indd　25　　2020-09-25　14:09:32

实际运行的风电场 50 台 2000 kW 风电机组中选取 1 台作为参考风电机组，基于该机组的 SCADA 数据采用深度神经网络进行训练，通过网格搜索 (Grid Search) 法得到最优的超参数配置方案，并验证该配置方案是否为整个风电场的全局最优配置。

1 深度神经网络的概念及超参数的选取

1.1 深度神经网络的概念

深度神经网络 (Deep Neural Network，DNN) 是具有多个隐含层的人工神经网络 [8]，其由 3 部分构成，分别为输入层、隐含层及输出层。每一层都由一定数量的神经元构成，同时每个神经元都具有 1 个自己的激活函数，当其接收到的信号达到激活函数的阈值时，其将被激活，并向下层连接的神经元输出信号，最终各层的神经元逐一连接在一起构成一个类似生物神经的网状结构。具有 3 个隐含层的深度神经网络结构示意图如图 1 所示。其中，X_i 为选定的进入深度神经网络输入层参与风电机组性能模型构建的、与风电机组输出功率相关性较高的 SCADA 数据类型，如风速等；Y 为风电机组实际输出功率。

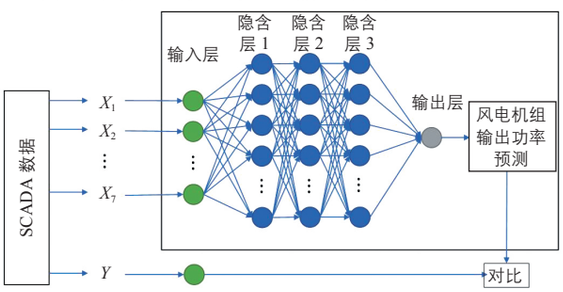


图 1 具有 3 个隐含层的深度神经网络结构示意图

Fig. 1 Structure schematic diagram of deep neural network with three hidden layers

1.2 超参数的选取

深度神经网络的结构和特性是通过超参数来决定的，超参数是在深度神经网络训练过程开始之前需要设置数值的参数，而不是通过训练得到的参数数值。通常情况下，需要对超参数进行优化，以寻找到一组最优的超参数，提高训练效果。

深度神经网络的超参数很多，包括但不局限于：1) 隐含层的层数；2) 深度神经网络的形状；3) 每层隐含层神经元的数量；4) 神经元的激活函数；5) 损失函数；6) 优化器的选择；7) 隐含层传递权重初始化方法。

这些超参数对于深度神经网络的影响不尽相同，其中有些会影响最终的训练效果，有些会影响训练过程的速度和收敛。由于本研究仅针对风电机组性能模型构建中影响深度神经网络最终训练效果的超参数，因此基于参考风电机组的训练过程，选取深度神经网络隐含层层数、深度神经网络形状及隐含层神经元数量这 3 个超参数作为主要研究对象，并寻找其最优配置方案 [9]。

1.2.1 深度神经网络隐含层层数

深度神经网络模型中，输入层与输出层之间隐含层的层数决定了模型的复杂程度。一般情况下，隐含层的层数越多，模型的复杂程度就越高，也就越能更好地模拟复杂问题；但相应地，模型训练所需要的时间也会相应增大，并且可能会伴随出现“过拟合”的问题。针对本研究中风电机组发电性能的问题，输入层的神经元数量为 7，输出层的神经元数量为 1，属于机器学习中相对简单的案例，因此，所采用的隐含层层数不需要过多。

1.2.2 深度神经网络形状

当深度神经网络的隐含层层数确定后，每一层隐含层的神经元数量决定了深度神经网络的形状，实际应用中较为常见的形状为漏斗型和均衡型。

1.2.3 隐含层神经元数量

本研究中的隐含层神经元数量是指输入层之后的第一层隐含层的神经元数量。对于漏斗型深度神经网络形状来说，输入层之后的每一层隐含层的神经元数量与前一层相比，均按一定比例递减 (本研究中采用 0.5 倍递减)；对于均衡型深度神经网络形状而言，输入层之后的每一层隐含层神经元数量均与第一层相同。图 1 所示的即为 1

2020-09杂志.indd 26　　2020-09-25 14:09:32

个典型的均衡型 3 层隐含层深度神经网络形状。与深度神经网络的层数类似，每一层隐含层的神经元数量越多，表明该模型越复杂，增加隐含层神经元数量可以更好地模拟复杂问题，但机器学习的时间成本也相应更高，并且随着神经元数量的增长，也会出现“过拟合”问题。

2 超参数寻优

本研究选取了某个实际运行的风电场进行方案验证。该风电场的装机容量为 100 MW，由 50 台 2000 kW 风电机组构成，风电机组采用 SCADA 系统，共可采集 90 种数据类型，采用 10 min 时间步长，共记录了 9 个月的数据。

本研究采用基于互信息的关键数据提取方法 [10]，从中选取了 7 个与风电机组输出功率“相关性”较高的数据类型作为深度神经网络训练输入层的入口数据。这 7 个数据类型的名称及对应单位如表 1 所示。

表 1 深度神经网络输入层的数据类型

Table 1 Data types of input layer of deep neural network

SCADA 数据类型名称	单位
发电机后轴承温度	℃
发电机前轴承温度	℃
网侧三相电压	V
变流器无功功率	kW
网侧三相电流	A
轮毂高度平均风速	m/s
发电机平均转速	rps

首先选取了编号为 T_1 的风电机组作为参考风电机组进行训练，然后以训练后模型预测的风电机组输出功率 P_{DNN} 与 SCADA 数据中实际记录的风电机组输出功率 P_{SCADA} 间的均方误差 *MSE* 作为损失函数来进行深度神经网络 3 个超参数的寻优。均方误差 *MSE* 的计算公式为：

$$MSE=\frac{1}{N}\sum_{i=1}^{N}(P_{SCADA}-P_{DNN})^2 \tag{1}$$

式中，N 为 SCADA 系统中的总数据量。

针对深度神经网络隐含层层数、深度神经网络形状及隐含层神经元数量这 3 个超参数，本研究设置的可选配置如表 2 所示。

对于前文提到的其他超参数，本研究中则采用了机器学习领域中的经典配置，具体的配置方案如表 3 所示。

表 2 待寻优超参数可选配置方案

Table 2 Optional configuration scheme of hyperparameters to be optimized

待寻优超参数名称	可选配置
深度神经网络隐含层层数	2、3、4、5
深度神经网络形状	均衡型、漏斗型
隐含层神经元数量	64、128、256、512

表 3 固定超参数配置方案

Table 3 Configuration scheme of fixed hyperparameters

固定超参数名称	配置方案
神经元激活函数	ELU 函数
隐含层传递权重初始化方法	Xavier & He 初始化

本研究通过 Grid Search 的方法对超参数进行寻优，即将 3 个超参数所有配置排列组合后的全部方案均进行模型训练，最终选取结果最优的方案。每个不同超参数组合配置的训练采用 2000 批次 (epoch)，每批次 300 批大小 (batch size)，最终参考风电机组 T_1 在不同超参数配置下的风电机组性能模型训练效果如表 4 所示。

从表 4 的训练结果可以看出，在参考风电机组 T_1 上，最优的超参数配置为：隐含层层数取 5、深度神经网络形状为漏斗型、隐含层神经元数量为 64；该配置下达到的测试数据集误差为 6.95 kW，相对于风电机组的额定功率 2000 kW 而言，该值所占比例仅为 0.35%。而表现最差的配置方案为：隐含层层数取 4、深度神经网络形状为均衡型、隐含层神经元数量为 256；该配置下达到的测试数据集误差为 16.77 kW，相对于风电机组的额定功率 2000 kW 而言，该值所占比例为 0.84%。由此可以得到以下结论：

1) 在不同的超参数配置方案下，通过深度

2020-09杂志.indd　27　　2020-09-25　14:09:32

表 4 在不同超参数配置下，参考风电机组 T_1 的风电机组性能模型训练结果

Table 4 Wind turbine performance model training results of reference wind turbine T_1 under different hyperparameters configuration

隐含层层数	深度神经网络形状	隐含层神经元数量	测试数据集误差 /kW	隐含层层数	深度神经网络形状	隐含层神经元数量	测试数据集误差 /kW
2	漏斗型	64	9.23	4	漏斗型	64	7.75
2	漏斗型	128	8.78	4	漏斗型	128	7.88
2	漏斗型	256	9.84	4	漏斗型	256	13.16
2	漏斗型	512	8.87	4	漏斗型	512	16.56
2	均衡型	64	11.02	4	均衡型	64	7.33
2	均衡型	128	10.82	4	均衡型	128	10.66
2	均衡型	256	7.47	4	均衡型	256	16.77
2	均衡型	512	9.24	4	均衡型	512	10.20
3	漏斗型	64	8.38	5	漏斗型	64	6.95
3	漏斗型	128	8.05	5	漏斗型	128	9.43
3	漏斗型	256	13.72	5	漏斗型	256	10.05
3	漏斗型	512	13.05	5	漏斗型	512	11.70
3	均衡型	64	8.56	5	均衡型	64	10.98
3	均衡型	128	7.12	5	均衡型	128	12.54
3	均衡型	256	14.13	5	均衡型	256	13.88
3	均衡型	512	13.95	5	均衡型	512	9.89

注：测试数据集误差是指模型最终预测输出功率与实际输出功率的均方误差

神经网络训练得到的各个风电机组性能模型应用于参考风电机组后的预测准确度之间存在一定差异。

2) 从工业角度来看，在不同的超参数配置方案下，深度神经网络训练得到的风电机组性能模型应用于参考风电机组后的准确度误差基本可以忽略不计。因此，若只是针对参考风电机组进行模型构建，可以直接将超参数选择为训练速度最快的配置方案，即隐含层层数最少、神经元数量最小。

下文对表 2、表 3 中的超参数设置在参考风电机组上的最优配置与同一风电场中同一型号其他所有风电机组上的最优配置间的相对关系进行探讨。理想情况下希望实现以下 2 点假设：1) 参考风电机组的最优配置同时也是同一风电场中同一型号其他风电机组的最优配置；2) 基于参考风电机组得到的性能模型，应用于同一风电场同一型号其他所有风电机组上时准确度也能满足需求。

假设 1) 和假设 2) 若能同时被验证成功，则说明同一个风电场中同一型号风电机组的最优超参数的配置及基于参考风电机组得到的性能模型具有推广性。这也意味着在实际应用中，不需要针对每台风电机组单独进行训练，只需要选择 1 台参考风电机组进行训练，再推广到整个风电场即可，从而可大幅节省时间，更能满足工业领域的需求。

为了进行以上 2 个假设的验证，本研究将参考风电机组 T_1 训练得到的不同超参数配置下的深度神经网络模型分别在风电机组 $T_2 \sim T_{50}$ 上进行复用，以在风电机组 $T_2 \sim T_{50}$ 上模型的预测输出功率与实际输出功率的均方误差的平均值作为评价标准，结果如表 5 所示。

通过分析 $T_2 \sim T_{50}$ 这 49 台风电机组应用参考风电机组 T_1 经深度神经网络训练得到的模型后

2020-09杂志.indd 28　　2020-09-25 14:09:32

表 5 在不同超参数配置下，风电机组性能模型的推广验证

Table 5 Promotion verification of wind turbine performance model under different hyperparameters configuration

隐含层层数	深度神经网络形状	隐含层神经元数量	测试数据集误差平均值 /kW	隐含层层数	深度神经网络形状	隐含层神经元数量	测试数据集误差平均值 /kW
2	漏斗型	64	26.36	4	漏斗型	64	42.14
2	漏斗型	128	15.26	4	漏斗型	128	16.39
2	漏斗型	256	14.76	4	漏斗型	256	26.06
2	漏斗型	512	115.21	4	漏斗型	512	28.23
2	均衡型	64	18.73	4	均衡型	64	12.59
2	均衡型	128	32.81	4	均衡型	128	20.30
2	均衡型	256	32.15	4	均衡型	256	19.43
2	均衡型	512	55.96	4	均衡型	512	25.29
3	漏斗型	64	18.82	5	漏斗型	64	22.41
3	漏斗型	128	20.46	5	漏斗型	128	24.68
3	漏斗型	256	15.25	5	漏斗型	256	15.78
3	漏斗型	512	24.18	5	漏斗型	512	20.89
3	均衡型	64	15.55	5	均衡型	64	17.65
3	均衡型	128	15.75	5	均衡型	128	26.02
3	均衡型	256	67.96	5	均衡型	256	16.45
3	均衡型	512	33.38	5	均衡型	512	17.37

的最终结果可以发现：

1) 在参考风电机组 T_1 上验证的最优超参数配置方案，在 T_2~T_{50} 风电机组上的表现并非最优。

2) 在 T_2～T_{50} 风电机组上表现最优的超参数配置方案 (隐含层层数为 4、深度神经网络形状为均衡型、隐含层神经元数量为 64)，在参考风电机组 T_1 上的表现也较好。

3) 基于不同超参数配置的深度神经网络训练得到的模型，训练结果在风电机组 T_2～T_{50} 上的表现总体较好，各模型的平均的均方误差平均值为 27.32 kW，最优的均方误差平均值为 12.59 kW，最不理想的均方误差平均值为 115.21 kW，相对于风电机组的额定功率 2000 kW 而言，均方误差平均值占比分别为1.37%、0.63% 和 5.76%。

综上所述，虽然前文中的假设并未完全得到验证，但由于基于参考风电机组得到的性能排名靠前的超参数配置方案应用于本风电场中同一型号其他风电机组时，虽然性能不是最优但基本也较为靠前，同时准确度也在可接受范围内，因此，在工业应用中基本可以认为该方式具备推广性。即实际应用中，出于时间成本考虑，可以将基于参考风电机组得到的超参数配置方案及模型直接应用于本风电场中同型号的其他风电机组。

3 结论与建议

本文以某实际运行风电场的 50 台 2000 kW 风电机组的 SCADA 数据为基础，对采用深度神经网络技术建立风电机组性能模型时的隐含层层数、深度神经网络形状及隐含层神经元数量这 3 个超参数的取值对于该模型构建效果的影响进行了分析。通过 Grid Search 方法在参考风电机组上发现这 3 个超参数配置不同时建立的模型的最终性能存在一定差异，但从工业应用角度考虑，此差异基本可以忽略不计。同时分析还发现，参考风电机组采用最优超参数配置方案时训练得到的模型应用于同一风电场中同一型号的其他风电机

2020-09杂志.indd　29　　2020-09-25　14:09:32

组时得到的并非最佳方案，但由于其在同一风电场中同一型号的其他风电机组上的预测准确度在所有配置方案中仍然较优，因此出于时间成本考虑，在工业应用中可以认为同一风电场中同型号的风电机组间可以应用此方案，采用深度神经网络在参考风电机组上得到的最优模型具备推广性。

基于本研究的方法与结论，在未来工作中可以进一步针对不同型号的风电机组或不同风电场的风电机组进行超参数配置方案的寻优及推广性验证。若基于参考风电机组训练得到的风电机组性能模型在不同机型乃至不同风电场间可以直接推广复用，则可进一步为机器学习技术在风电行业后评估工作中的应用确立更为普遍的事实依据。

[参考文献]

[1] 李辉，胡姚刚，唐显虎，等．并网风电机组在线运行状态评估方法 [J]. 中国电机工程学报，2010, 30(33): 103 － 109.
[2] LAPIRA E, BRISSET D, ARDAKANI H D, et al. Wind turbine performance assessment using multi-regime modeling approach[J]. Renewable energy, 2012, 45: 86 － 95.
[3] LI S, WUNSCH D C, HAIR E O', et al. Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation[J]. Solar energy engineering, 2001, 123(4): 327 － 332.
[4] CHEN Z Q, LI C, SANCHEZ R V. Gearbox fault identification and classification with convolutional neural networks[J]. Shock & vibration, 2015(2):1 － 10.
[5] 徐强．风电机组传动链状态诊断方法研究 [D]. 北京：华北电力大学，2015.
[6] WANG L, ZHANG Z J, LONG H, et al. Wind turbine gearbox failureidentification with deep neural networks[J]. IEEE transactions on industrial informatics, 2017, 13 (3):1360 － 1368.
[7] 王双园．风力机健康状态监测及评估关键技术研究 [D]. 上海：上海交通大学，2017.
[8] AURÉLIEN G. Hands-on machine learning with scikit-learn & TensorFlow[M]. America: O'Reilly, 2017.
[9] 元尼东珠，杨浩，房红征．基于卷积神经网络的发动机故障预测方法 [J]. 计算机测量与控制，2019, 27(10): 74 － 78.
[10] DU M, YI J, MAZIDI P, et al. A parameter selection method for wind turbine health management through SCADA data[J]. Energy and power engineering, 2017(4): 36 － 45.

SELECTION OF HYPERPARAMETERS FOR WIND TURBINE PERFORMANCE MODEL BASED ON DEEP NEURAL NETWORK

Wu Sha[1]，Wang Jian[1]，Xie Xin[2]，Jiang Zixiao[2]，Deng Shaoping[1]，Lu Sheng[1]

(1. *POWERCHINA Hubei Electric Engineering Co.，Ltd.，Wuhan* 430040，*China*；
2. *Meteodyn Beijing, Beijing* 100027, *China*)

Abstract：The construction of wind turbine performance models based on deep neural network technology has attracted widely attention in the wind power industry. However，there is still a lack of systematic and in-depth research on the influence of hyperparameter selection in neural networks on the performance model. This paper attempted to study the effect of the hyperparameters values on the wind turbine performance model from three aspects of the number of hidden layers，the shape，and the number of hidden neurons of the deep neural network，to verify the optimal configuration and generalizability of the reference wind turbine through actual operating data. The results show that the three deep neural network hyperparameters have a certain impact on the final turbine performance model under different configurations. However，from the perspective of industrial applications，they can be basically ignored，and the optimal hyperparameter configuration based on the reference wind turbine is suitable to be promoted to all the same type wind turbines in the same wind farm.

Keywords：wind turbine performance model；deep neural network；hyperparameters；post-evaluation

2020-09杂志.indd　30　　2020-09-25　14:09:32