

doi: 10.7690/bgzdh.2025.05.013

基于迁移学习和 RBF 神经网络的小子样产品性能参数预测方法

毛廷鏊, 赵建印, 杨根庆, 孙伟赫, 崔爽
(海军航空大学岸防兵学院, 山东 烟台 264001)

摘要: 针对小子样产品预测模型不够精确的问题, 提出一种产品性能参数预测方法。在径向基函数 (radial basis function, RBF) 神经网络学习算法的基础上, 加入迁移学习的思想, 将小子样产品自身的历史测试数据和同型号同批次其他产品的测试数据当作源领域知识来充分学习, 弥补当前领域因已标签样本数据少而导致的产品性能参数预测精度差的问题。结果表明, 该方法的预测精度较高。

关键词: 迁移学习; RBF 神经网络; 小子样产品; 参数预测

中图分类号: TP183 **文献标志码:** A

Prediction Method of Performance Parameters of Small Sample Products Based on Transfer Learning and RBF Neural Network

Mao Tingliu, Zhao Jianyin, Yang Genqing, Sun Weihe, Cui Shuang
(College of Coast Guard, Naval Aviation University, Yantai 264001, China)

Abstract: Aiming at the problem that the prediction model of small sample product is not accurate enough, a prediction method of product performance parameters is proposed. Based on the learning algorithm of radial basis function (RBF) neural network, the idea of transfer learning is added, and the historical test data of small sample products and the test data of other products with the same model and batch are fully learned as the source domain knowledge, which makes up for the problem of poor prediction accuracy of product performance parameters caused by the lack of labeled sample data in the current field. The results show that the prediction accuracy of this method is high.

Keywords: transfer learning; RBF neural network; small sample product; parameter prediction

0 引言

对于小子样产品来说, 由于造价高昂和测试的特点, 导致可获取的数据较少, 而仅仅使用少量测试数据对其性能参数进行径向基函数 (RBF) 神经网络学习预测, 得到的结果误差较大。这是因为过少的数据样本量, 不能很好地反映数据的整体分布情况, 从而导致 RBF 神经网络学习更容易发生过拟合, 其泛化性能就会下降^[1-5]。迁移学习可以解决这个问题, 即通过源领域知识的学习, 使得目标领域中的数据样本得到学习训练, 从而丰富数据信息^[6-7]。可以考虑将迁移学习思想引入 RBF 神经网络算法之中, 即在文献[8]提出的 RBF 算法的基础上, 引入 ϵ 不敏感损失函数^[9-11], 同时加入结构风险项^[12], 通过学习源领域选择的径向基函数的各类参数^[13-16], 并学习源领域所建立模型的各类模型参数^[17-20], 帮助目标领域建立模型。该模型可以合理处理待评估产品历史数据和其他同型号同批次产品的数据, 将这些数据当作源领域信息, 将待评估产品的当前数据信息当作目标领域信息, 利用源领域

信息进行学习训练, 解决因目标领域标签数据少而造成预测结果不准确的问题。

1 问题描述

某小子样产品造价高昂, 且具有“长期储存、一次使用”的特点, 同时采取周期性的测试方式进行状态评估; 因此, 获取数据较为困难。在数据量较少的情况下对该产品某些参数进行预测, 显然难以保证精确度。同类型同批次的其他产品与其拥有相似的特点, 其性能参数的特征也具有相似性; 因此, 可以考虑利用迁移学习的方法将这些数据融入预测模型的建立之中, 提高预测模型的精准度。

现有某个小子样产品的性能参数数据如表 1 所示, 同时与该产品同型号同批次的其他产品提供了一批相同性能参数的数据, 利用这些数据对该产品性能参数进行预测。

2 基于迁移学习和 RBF 神经网络建模

2.1 建模思路

因该产品获取的样本数据量比较少, 不能反映

其性能参数的整体特点,而同型号同批次其他产品与该产品存在相似性,故将同型号同批次其他产品的数据用作源领域数据,选择径向基函数,构建 RBF 神经网络模型;接着将迁移学习思想融入 RBF 神经网络模型,建立新的预测模型,该模型通过学习源领域各类参数,将源领域数据的相似信息迁移到目标领域之中;同时将该产品的前 8 组数据用作目标领域数据集,后 2 组数据用作目标领域测试数据集,建立新的目标领域模型,用该模型实现对目标域测试集的预测。

表 1 某型装备性能参数数据

序号	电流/mA				
	1	2	3	4	5
1	38	35	38	36	35
2	37	36	36	36	37
3	40	37	37	37	38
4	35	36	36	35	37
5	36	38	37	38	36
6	35	33	32	40	36
7	43	45	44	49	44
8	35	39	38	39	37
9	39	39	39	41	39
10	47	48	49	49	48

2.2 建模过程

2.2.1 建立 RBF 神经网络模型

将高斯函数定义为径向基函数,输入的样本设为 \mathbf{x} ,将函数的中心表示为 $\mathbf{b}_i \in R^d$,其核宽表示为 σ_i ,该函数表示为:

$$\varphi(\|\mathbf{x} - \mathbf{b}_i\|) = \exp(-\|\mathbf{x} - \mathbf{b}_i\|^2 / \sigma_i). \quad (1)$$

设隐含层、输出层之间的连接权值为 ω_i ,径向基函数为 $\varphi(\bullet)$,输出设为 y ,且 $y \in R$,隐含层节点数为 N ,则 $R^d \rightarrow R$ 的映射可以表示为:

$$y = f(x) = \sum_{i=1}^N \omega_i \varphi(\|\mathbf{x} - \mathbf{b}_i\|); \quad i = 1, 2, \dots, N. \quad (2)$$

样本 \mathbf{x}_j 对第 i 类的模糊隶属度设为 v_{ji} ,设 ρ 为

$$\sum_{i=1}^n |y'_i - y_i|_\varepsilon = \sum_{i=1}^n \max(0, |y'_i - y_i| - \varepsilon) = \sum_{i=1}^n \max(0, |\mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i| - \varepsilon). \quad (6)$$

在式(6)的基础上,引入松弛因子 β_i 和 β_i^* ,使 $|\mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i| < \varepsilon$ 尽可能满足。同时,要使 ε 不敏感损失函数值取得最小值,可以构造出如下的约束条件:

$$\begin{aligned} & \min \sum_{i=1}^n (\beta_i + \beta_i^*); \\ \text{s.t.} & \begin{cases} y_i - \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} < \varepsilon + \beta_i, \\ \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i < \varepsilon + \beta_i^*, \end{cases} \beta_i \geq 0, \beta_i^* \geq 0. \end{aligned} \quad (7)$$

为使算法的结构风险能够降到最低,在式(7)的基础上加入结构化风险项 $\omega_g^T \omega_g / 2$,具体如式(8)

缩放参数且该参数可调,样本训练规模用 n 表示,则通过式(3)可以计算径向基函数的中心 \mathbf{b}_i 和核宽 σ_i :

$$\mathbf{b}_i = \frac{\sum_{j=1}^n v_{ji} \mathbf{x}_j}{\sum_{j=1}^n v_{ji}}, \quad \sigma_i = \rho \frac{\sum_{j=1}^n v_{ji} \|\mathbf{x}_j - \mathbf{b}_i\|^2}{\sum_{j=1}^n v_{ji}}. \quad (3)$$

将 \mathbf{b}_i 和 σ_i 代入式(1),同时设 $\mathbf{h}_g = [\omega_1, \omega_2, \dots, \omega_N]^T$,令 $\tilde{\mathbf{x}}_i = \varphi(\|\mathbf{x} - \mathbf{b}_i\|)$, $i = 1, 2, \dots, N$,有 $\tilde{\mathbf{x}}_g = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_N]^T$ 。得到 RBF 神经网络输出函数为:

$$y = \mathbf{h}_g^T \tilde{\mathbf{x}}_g. \quad (4)$$

2.2.2 迁移学习与 RBF 神经网络模型的结合

设源领域径向基函数的中心向量集为: $\mathbf{b}_i = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N\}$,核宽集合为: $\sigma_i = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$, ρ_s 为核宽的缩放参数且该参数可调, \mathbf{h}_s 为模型参数。

$X_t = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$ 为目标领域数据集,其中, $\mathbf{x}_i \in R^d$, n 为目标域样本数, N 为隐含层节点个数,且 $n > N$ 。

设目标域核宽的缩放参数为 ρ_g 且可调, $\mathbf{b}_{gi} = \{\mathbf{b}_{g1}, \mathbf{b}_{g2}, \dots, \mathbf{b}_{gN}\}$ 为目标域径向基函数的中心向量集, $\sigma_{gi} = \{\sigma_{g1}, \sigma_{g2}, \dots, \sigma_{gN}\}$ 为核宽集合。其中, \mathbf{b}_{gi} 从学习 \mathbf{b}_i 获得, σ_{gi} 从学习 σ_i 获得,其关系如式(5):

$$\mathbf{b}_{gi} = \mathbf{b}_i, \quad \sigma_{gi} = \rho_g \sigma_i / \rho_s, \quad i = 1, 2, \dots, N. \quad (5)$$

通过上式的学习方式,可以得到新空间的数据集 $\tilde{X}_g = \{(\tilde{\mathbf{x}}_{gi}, y_i)\}_{i=1}^n$,其中, $\tilde{\mathbf{x}}_{gi} \in R^N$ 。

接下来引入 ε 不敏感损失函数。根据 ε 不敏感损失函数 $L^\varepsilon(x, y, f)$ 的定义^[21],设网络输出为 y'_i ,真实输出为 y_i ,则式(4)对应的 ε 不敏感损失函数可表示为:

所示:

$$\begin{aligned} & \min_{\mathbf{h}, \beta_i, \beta_i^*} \frac{1}{2} \|\mathbf{h}_g\|^2 + S \sum_{i=1}^n (\beta_i + \beta_i^*); \\ \text{s.t.} & \begin{cases} y_i - \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} < \varepsilon + \beta_i, \\ \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i < \varepsilon + \beta_i^*, \end{cases} \beta_i \geq 0, \beta_i^* \geq 0. \end{aligned} \quad (8)$$

在完成前面一系列的学习过程后,为对源领域模型参数 \mathbf{h}_s 进行学习,将 $\|\mathbf{h}_s - \mathbf{h}_g\|^2$ 项加入式(8),所需解决的问题转化为:

$$\min_{\mathbf{h}, \beta_i, \beta_i^*} \frac{1}{2} \|\mathbf{h}_g\|^2 + S \sum_{i=1}^n (\beta_i + \beta_i^*) + \frac{\eta}{2} \|\mathbf{h}_s - \mathbf{h}_g\|^2;$$

$$\text{s.t.} \begin{cases} y_i - \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} < \varepsilon + \beta_i \\ \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i < \varepsilon + \beta_i^* \end{cases}, \beta_i \geq 0, \beta_i^* \geq 0. \quad (9)$$

式中 S 、 η 、 ε 均为可设置参数。

基于式(9)，利用拉格朗日乘子构造拉格朗日函数如下：

$$C = \frac{1}{2} \|\mathbf{h}_g\|^2 + S \sum_{i=1}^n (\beta_i + \beta_i^*) + \frac{\eta}{2} \|\mathbf{h}_s - \mathbf{h}_g\|^2 + \sum_{i=1}^n \delta_i (y_i - \mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - \varepsilon - \beta_i) + \sum_{i=1}^n \delta_i^* (\mathbf{h}_g^T \tilde{\mathbf{x}}_{gi} - y_i - \varepsilon - \beta_i^*). \quad (10)$$

将式(10)转化为相应的对偶问题：

$$\begin{aligned} \max_{\delta} &= \frac{\sum_{i=1}^n \sum_{j=1}^n (\delta_i - \delta_j^*) (\delta_j - \delta_j^*) \tilde{\mathbf{x}}_{gi}^T \tilde{\mathbf{x}}_{gj}}{2(1+\eta)} + \sum_{i=1}^n \delta_i (y_i - \varepsilon - \frac{\eta}{1+\eta} \mathbf{h}_s^T \tilde{\mathbf{x}}_{gi}) + \sum_{i=1}^n \delta_i^* (-y_i - \varepsilon + \frac{\eta}{1+\eta} \mathbf{h}_s^T \tilde{\mathbf{x}}_{gi}); \\ \text{s.t.} & \quad 0 \leq \delta_i^* \leq S. \end{aligned} \quad (11)$$

利用式(11)得到最优解：

$$\mathbf{h}_g = \left(\sum_{i=1}^n (\delta_i - \delta_i^*) \tilde{\mathbf{x}}_{gi} + \eta \mathbf{h}_s \right) / (1 + \eta). \quad (12)$$

将式(12)求得的经过迁移学习方式得到的 \mathbf{h}_g 值代入式(4)，即可输出最终的预测值。

2.3 实验仿真

为验证本文中方法的可行性，在仿真中加入另外 2 种方法：1) 仅采用目标领域数据进行预测；2) 利用源领域数据和目标数据，但不引入迁移学习思想进行预测。用折线图将上述 2 种方法的预测结果与本文中方法的预测结果进行对比，如图 1 和 2 所示。表 2 和 3 为 3 种方法预测结果的误差对比。

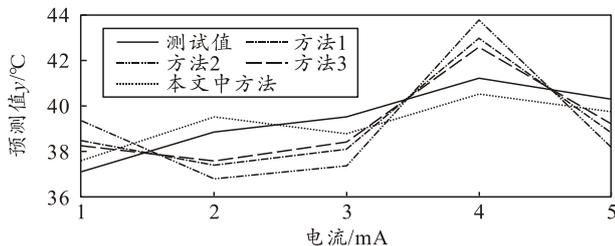


图 1 3 种方法对第 9 组数据预测结果对比

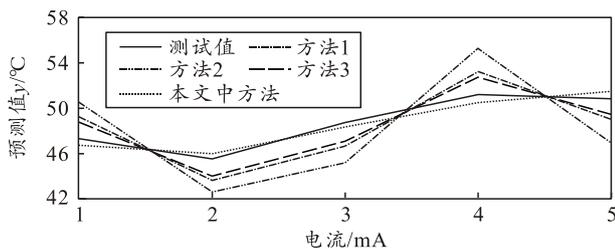


图 2 3 种方法对第 10 组数据预测结果对比

表 2 3 种方法对第 9 组数据的预测误差比较

项目	电流/mA				
	1	2	3	4	5
测试值	37.12	38.87	39.53	41.24	40.31
方法 1 误差	2.23	-2.07	-2.12	2.57	-2.17
方法 2 误差	1.35	-1.47	-1.41	1.76	-1.52
本文中方法误差	0.53	-0.67	0.72	-0.79	-0.94

表 3 3 种方法对第 10 组数据的预测误差比较

项目	电流/mA				
	1	2	3	4	5
测试值	47.00	48.00	49.00	49.00	48.00
方法 1 误差	3.27	-2.91	-3.53	4.07	-3.89
方法 2 误差	2.24	-2.21	-2.31	2.73	-2.59
本文中方法误差	-1.14	1.47	-1.37	-1.03	1.12

由图 1—2 和表 2—3 可以看出：仅仅利用小样本数据或者在没有加入迁移学习的情况下拓展利用数据，在此基础上得到的预测结果精确度并不是很高。在迁移学习的思想下，利用相似数据对目标域数据进行扩充学习，能够较好地提升预测结果的精确度。

3 结论

笔者针对已有的 RBF 神经网络算法在小样本情况下，不能反映数据整体分布特点而导致预测性能降低的情况；同时，解决相似数据利用中不能合理融合数据导致数据利用度不高的问题。利用迁移学习的思想，将能够利用的相似数据用作源领域知识，使目标域的小样本数据在充分学习源领域知识的情况下，将自身不够丰富或缺失的数据进行补充，从而帮助目标域建立更加完善的模型，使预测精准度得以提高。通过仿真实验验证了该方法的可行性。

参考文献：

- [1] MOODY J, DARKEN C J. Fast learning in networks of locally-tuned processing units[J]. Neural computation, 1989, 1(2): 281-294.
- [2] STASINAKIS C, SERMPINIS G, THEOFILATOS K, et al. Forecasting us unemployment with radial basis neural networks, Kalman filters and support vector regressions[J]. Computational economics, 2016, 47(4): 569-587.
- [3] 韩敏, 穆云峰. 一种改进的 RAN 网络结构优化算法[J]. 控制与决策, 2007, 22(10): 1177-1180.
- [4] SARIMVEIS H, DOGANIS P, ALEXANDRIDIS A. A classification technique based on radial basis function neural networks[J]. Advances in engineering software, 2006, 37(4): 218-221.
- [5] RAITOHARJU J, KIRANYAZ S, GABBOUJ M. Training radial basis function neural networks for classification

via class-specific clustering[J]. IEEE transactions on neural networks and learning systems, 2016, 27(12): 2458-2471.

[6] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.

[7] 张雅俊, 高陈强, 李佩, 等. 基于卷积神经网络的人流量统计[J]. 重庆邮电大学学报: 自然科学版, 2017, 29(2): 265-271.

[8] 桑庆兵, 邓赵红, 王士同, 等. 基于 ϵ 不敏感准则和结构风险的鲁棒径向基函数神经网络学习[J]. 电子与信息学报, 2012, 34(6): 1414-1419.

[9] 许敏, 史荧中, 葛洪伟, 等. 一种具有迁移学习能力的 RBF-NN 算法及其应用[J]. 智能系统学报, 2018, 13(6): 959-966.

[10] RYU D, LIANG F, MALLICK B K. Sea surface temperature modeling using radial basis function networks with a dynamically weighted particle filter[J]. Journal of the American statistical association, 2013, 108(501): 111-123.

[11] PRATHIBA R, BALASINGHMOSES M, DEVARAJ D, et al. Multiple output radial basis function neural network with reduced input features for on-line estimation of available transfer capability[J]. Control engineering and applied informatics, 2016, 18(1): 95-106.

[12] 韩建立, 高山, 陆巍巍, 等. 装备综合评价理论与方法[M]. 北京: 电子工业出版社, 2021: 184-190.

[13] 樊劲辉, 贾松敏, 李秀智. 基于 RBF 神经网络的全向

(上接第 46 页)

[8] LI Z, JIN H, ZOU D, et al. Exploring new opportunities to defeat low-rate DDoS attack in container-based cloud environment[J]. IEEE Transactions on Parallel and Distributed Systems, 2019, 31(3): 695-706.

[9] ELVENY M, WINATA A, SIREGAR B, et al. A tutorial: Load balancers in a container technology system using docker swarms on a single board computer cluster[J]. Elementary Education Online, 2020, 19(4): 744-751.

[10] 左旭彤, 王莫为, 崔勇. 低时延网络: 架构, 关键场景与研究展望[J]. 通信学报, 2019, 40(8): 22-35.

[11] MARINKOVIĆ D, KOJIĆ V, AVRAMOVIĆ Z Ž. Software Application Development Using Container Technology[J]. JITA-JOURNAL OF INFORMATION TECHNOLOGY AND APLICATIONS, 2021, 21(1): 54-60.

智能轮椅自适应控制[J]. 华中科技大学学报: 自然科学版, 2014, 42(2): 111-115.

[14] ALEXANDRIDIS A, CHONDRODIMA E, SARIMEIS H. Radial basis function network training using a nonsymmetric partition of the input space and particle swarm optimization[J]. IEEE transactions on neural networks and learning systems, 2013, 24(2): 219-230.

[15] ALI S H A, OZAWA S, NAKAZATO J, et al. An online malicious spam email detection system using resource allocating network with locality sensitive hashing[J]. Journal of intelligent learning systems and applications, 2015, 7(2): 55866.

[16] 李方伟, 郑波, 朱江, 等. 一种基于 AC-RBF 神经网络的网络安全态势预测方法[J]. 重庆邮电大学学报: 自然科学版, 2014, 26(5): 576-581.

[17] 蒋亦樟, 邓赵红, 王士同. ML 型迁移学习模糊系统[J]. 自动化学报, 2012, 38(9): 1393-1409.

[18] 林杨, 高思煜, 刘同舜, 等. 基于深度学习的高速铣削刀具磨损状态预测方法[J]. 机械与电子, 2017, 35(7): 12-17.

[19] 沈美杰, 赵龙章, 周兵, 等. 基于 PSO 优化的 RBF 网络液压泵故障诊断研究[J]. 液压与气动, 2016(5): 87-92.

[20] 孙文珺, 邵思羽, 严如强. 基于稀疏自动编码深度神经网络的感应电动机故障诊断[J]. 机械工程学报, 2016, 52(9): 65-71.

[21] 邓乃扬, 田英杰. 支持向量机: 理论与拓展[M]. 北京: 科学出版社, 2009: 63-80.

[12] 高会生, 于文思, 韩东升. 基于时间敏感网络的风电场通信网时延特性分析[J]. 现代电力, 2021, 38(4): 465-472.

[13] DANG L M, PIRAN M, HAN D, et al. A survey on internet of things and cloud computing for healthcare[J]. Electronics, 2019, 8(7): 768.

[14] 卢莎莎, 肖海力, 王小宁. 容器技术在高性能计算环境中的应用[J]. 数据与计算发展前沿, 2022, 3(6): 118-126.

[15] RASHID A, CHATURVEDI A. Cloud computing characteristics and services: a brief review[J]. International Journal of Computer Sciences and Engineering, 2019, 7(2): 421-426.

[16] 张海波, 李虎, 陈善学, 等. 超密集网络中基于移动边缘计算的任务卸载和资源优化[J]. 电子与信息学报, 2019, 41(5): 1194-1201.