# 机器学习在污水处理领域的应用

王博洋1，张文宇1[[1]](#footnote-1)\*，杜权1，钟声2

（1.辽宁科技大学计算机与软件科学学院，辽宁 鞍山 114000

2.辽宁科技大学化学与工程学院，辽宁 鞍山 114000）

【摘要】

本文介绍了神经网络，支持向量机（SVM）这两类典型的算法以及各种改进型算法在污水处理领域的研究现状，同时指出了机器学习技术应用于污水处理领域所面临的问题和改进措施。

【关键词】污水处理 机器学习 神经网络 支持向量机

引言：

 污水处理是一个多学科交叉融合的复杂过程，我国在该领域长期落后于欧，美，日等发达国家 [[[2]](#endnote-1)]。[在污水处理系统中管控工艺流程，优化参数设置](javascript:;)可以提高污水处理效率也成了研究热点[[[3]](#endnote-2)-][[[4]](#endnote-3)][[[5]](#endnote-4)][[[6]](#endnote-5)]。近年来，[机器学习方法在污水处理中表现优异，国内外专家学者对其做了大量研讨，并且成功应用于污水处理企业](javascript:;) [[[7]](#endnote-6)]。在已有的知识框架下，机器学习算法经过训练，可以分析归纳出新的知识模型来对未知的事物做出预测[[[8]](#endnote-7)]。支持向量机（support vector machine，SVM）、神经网络（artificial neural network，ANN）是众多机器学习方法中应用于污水处理领域较为成熟的两类方法，本文将对这2类较为典型的方法在污水处理中的应用展开综述。

## 1.污水处理工艺及机器学习方法综述

1.1污水处理工艺流程概述

在一般的污水处理过程中，污水池中的污水先被直流泵打入反应器，加入药剂后由搅拌机进行搅拌混合经充分反应后混凝沉淀。沉淀淤泥会自动排放，清水则经过处理达到排放标准后排出。活性污泥法是我国污水处理厂当下使用的主要处理方法，它将有机污染物经过微生物处理转变成无害的气态物质 (如),液态水和固态物质(生物污泥)。预处理、初沉、曝气、二沉、污泥回流是活性污泥处理的5个主要工艺流程[[[9]](#endnote-8)]。

1.2神经网络概述

神经网络模仿人类大脑的构造和功能，具有强大的联想记忆功能和较高的容错性，它自适应性高，自主学能力强，且具备高度并行处理的特点，是污水处理领域运用最多的数据驱动模型[[[10]](#endnote-9)－][[[11]](#endnote-10)][[[12]](#endnote-11)][[[13]](#endnote-12)][[[14]](#endnote-13)]。具有多输入，单输出的非线性阀值器称为神经元，它是构成神经网络的最小单位，常见的多层前馈神经网络由输入层（Input layer），输出层（Output layer），隐含层（Hidden layer）三部分组成，其中隐含层的节点（神经元）数目不定，通常情况下非线性强，鲁棒（robustness）性好的神经网络具有较多的隐含层[[[15]](#endnote-14)]。

1.3 支持向量机概述

支持向量机(Support Vector Machine，SVM)最早提出于上世纪60年代，并在90年代后快速发展，支持向量机（SVM）是一种有监督学习算法，通常情况下，为优化参数而采用改进的粒子群优化算法，它在非线性、高维度的模式识别以及分类回归中应用广泛，与神经网络的启发式学习方法相比，SVM具有严密的数学理论基础。

## 2.神经网络在污水处理领域的研究现状

神经网络在污水处理领域应用广泛，取得了大量研究成果。S . Azimi[[[16]](#endnote-15)]在其所著文章中，运用人工神经网络以及改进模糊聚类方法，建立了GIS环境下潜在的饮用水水质预测模型，在使用数据充分性检验以及干旱指数标准化方法对模型进行验证的过程中得到了较好的结果。Antwi等[[[17]](#endnote-16)]使用上流式厌氧污泥床（UASB）小试反应器所得检测数据构建神经网络，希望通过反应器中废水的COD浓度、pH、总凯氏氮浓度、挥发性脂肪酸（VFA）以及水力停留时（HRT）等指标预测厌氧产沼这一过程中产生甲烷的量。所构建的神经网络在甲烷产量的预测准确度上高于传统非线性回归模型。在预测UASB反应器中厌氧生物的除硫和脱氮性能研究中Cai等[[[18]](#endnote-17)]建立了前馈神经网络模型，为了分析参数对反应器性能的影响进而研究了神经元的连接权重，并得出水力停留时间会显著影响反应器同步厌氧生物脱氮除硫性能的结论。Mirbagjier等[[[19]](#endnote-18)]在比较BP神经网络与径向基函数（RBF）神经网络在模拟膜生物反应器跨膜压差（TMP）的过程中，研究了上流和下流两种不同方式在膜污染方面的影响。与此类似，为了预测好氧-缺氧膜生物反应器的TMP，Schmitt等[[[20]](#endnote-19)]建立了BP神经网络预测模型。通过分析优化神经网络中不同神经元的权值，得出COD、pH、MLSS同溶解氧在预测TMP时低于废水中的氮、磷浓度，得出采用定期检测废水中氮、磷浓度能够判断膜生物反应器的膜污染情况。

近些年，国内学者在神经网络及其改进算法运用于污水处理做了大量研究，余颖等[[[21]](#endnote-20)]在研究活性污泥处理过程中先是采用主元分析（PCA）方法来降低输入变量维度并去除相关性，又分别使用BP和BRF神经网络建立模型,分析两种模型在污水处理过程中的适应性。结果显示，在逼近速度和收敛能力上RBP神经网络模型优于BF神经网络模型，准确反映了实际情况,且适应性良好。为了模拟污水处理厂的脱氮工艺，林佳敏等[[[22]](#endnote-21)]建立了BP 神经网络模型，引入自回归移动平均模型（ARIMA）来预测污水处理厂未来短期出水 TN 浓度。结果显示BP 神经网络模型在训练集上达到了15.9%的相对误差，在测试集上为 16.5%，模型平稳性不高，然而ARIMA 模型在出水 TN 浓度时序预测中的误差只有4.41%，为了对污水处理厂更加高效地实时监测，文中得出可以将2个模型相结的结论。在污水处理过程中出水总磷往往难以精确控制，张璐等[[[23]](#endnote-22)]在研究出水总磷控制时提出了一个基于模糊神经网络(fuzzy neural network，FNN)的方法。先是分析污水处理过程的生化除磷机理，确定操作变量是溶解氧(dissolved oxygen, DO)传递系数和生化反应池第五分区外部碳源(external carbon, EC)，又设计出一种基于FNN 的出水总磷控制器，并使用梯度下降法更新参数进而将基于FNN的出水总磷控制器用在污水处理过程的基准仿真平台(benchmark simulation model No.1，BSM1)，结果显示，所提出的模型可以保证出水总磷达标排放，且控制效果良好。孟彦京等[[[24]](#endnote-23)]采用三层网络结构预测出水 COD，先是建立样本集，进而将学习来的样本集数据送入神经网络得到网络输出值，计算出差值后，再通过梯度下降算法来调整网络权矩阵；重复这一过程，当网络经过 6313 次训练后收敛，最终结果显示，模型计算出的平均相对误差为 1.9%具有较高的精度且稳定性良好。传统神经网络模型多是黑箱模型,虽能模拟系统的输入–输出关系, 却不能对系统的行为机理做出必要解释。污水处理过程具有串级结构特点，对于这类机理过程明显的对象，将神经网络同机理知识相融合构造出的灰箱模型可以更好地描述该过程。丛秋梅等[[[25]](#endnote-24)]针对这一特点,将建模理论同串级过程相结合[[[26]](#endnote-25),[[27]](#endnote-26)], 把各反应器子系统与递阶多层神经网络结合在一起，提出了基于活性污泥过程机理的建模方案, 并通过输入–状态稳定性(ISS, input-to-state stability)方法研究了不同子过程建模误差之间的关系。为了保证算法稳定性，使用了带有时变学习率的误差反传类学习方法。结果表明，递阶神经网络建模方法对处理串级工业过程建模这一类复杂问题具有重要意义。

## 3.支持向量机在污水处理领域的研究现状

支持向量机（SVM）是应用于污水处理领域的较为成熟的技术，国内外学者对H . Haghiabil[[[28]](#endnote-27)]研究了人工祌经网络（ANN）、数据处理分组方法（GMDH）和SVM等智能化技术，以伊朗太雷河为对象对其水质成分进行预测，然后结合相应的误差指标对这几个模型景区进行评价，得出SVM模型具有最佳精确度。M . Zh . Huang[[[29]](#endnote-28)]在研究水质预测问题时提出了一种包括神经网络（ANN）、模糊逻辑（FL）、小波变换（WT）和遗传算法（GA）的模糊小波神经网络(FWNN)预测模型，并将模型运用到实际工作中。M . Hameed[[[30]](#endnote-29)]在研究热带环境（马来西亚）水质数据时，发现径向基函数神经网络（RBFNN）和反向传播神经网络模型是两种各不同的神精网络算法，并在此基础上建立了精度更高的水质预测模型。

国内学者张永利等[15]在研究UCI 数据库中污水处理厂日常监控数据时，对其中527组38维数据进行处理，得到完整记录375个，使用其中的245个样本进行训练，剩下的130个样本做测试验证分类效果。实验结果显示SVM算法在污水处理监测数据的模式识别中达到了92.15%的正确率，展现出巨大的优越性。面对具体环境具体问题，各种基于支持向量机的衍生或改进模型也显示出了良好的适用性。李晓东等[[[31]](#endnote-30)]选用长沙市第二污水净化中心数据做实验，当直接使用SVM方法进行训练时640个数据项产生了高达198个的预测错误。使用基于RLOO(α)的GA改进多类SVM方法，为了得到最优的参数，使用智能算法进行了120轮迭代最终达到88%的正确率。曝气池SVI是活性污泥法工艺中难以实时测量的一种可以反应出水水质的重要参数，宋贤民等[5]对样本数据采用粗糙集技术化简最小属性约简集, 分别用 BPNN 、PCA-BPNN 、RS -BPNN三种方法建模，预测效果并不是很理想, 转而采用泛化能力较强的支持向量机来建立模型，实验表明在小样本的情况下,基于支持向量机的模型相对于神经网络模型,拟合精度和预测精度更高,而且运算快,实时性较好。COD是反映水质的重要指标，辽宁省地方排放标准是COD≤50 mg/L[[[32]](#endnote-31)]。近年来使用三维电极来处理难降解污水的技术发展较快，周文君等[[[33]](#endnote-32)]选择曝气生物滤池作为研究对象，先是使用 SVM的改进型 LSSVM 建立模型，为解决模型准确度不高的问题，利用改进的粒子群优化算法寻找最优模型参数使得改进后的模型准确度大幅度提高，又使用模糊隶属度解决了LSSVM方法不能很好应对诸如噪声、缺失项等异常数据的不足。实验显示，模糊 LSSVM 模型相比于传统SVM无论是在模型的适用性、稳定性、还是准确度上都有明显提升，能够较好预测出水COD浓度。

## 4.总结与反思

在污水处理过程中基于传统数学模型建立的系统结构复杂难以解决实际问题，基于神经网络和支持向量机建立的污水处理模型也面临着集约程度不够、自主学习能力不强、数值计算能力差、经验推理能力不足等缺点，在处理高维度数据上面临着较大困难，系统的可靠性和应用效率不高。

在未来的研究中应注重：1）加强前期的数据收集和整理工作，结合领域专家的建议对海量数据进行预处理。2）依照不同污水处理厂之间的差异，选取适当的模型及参数。3）不断完善和修改现有知识库，并根据事物的发展增加新知识以此提高系统的准确性，提升效率。随着计算机性能的不断提升以及数据信息的不断丰富，应用机器学习技术必将提高污水处理效率，促进污水处理事业发展，这些研究对我国生态文明建设以及经济事业发展具有重大意义。

参考文献：

1. 作者简介：王博洋（1993—），男，硕士研究生，

   研究方向：机器学习，

   联系电话：15281066101

   E-mail：461284299@qq.com。

   \* 通信作者：张文宇（1973—），男，博士，教授，

   研究方向：信息管理系统，计算机控制，人工智能算法，

   联系电话：13188008518

   E-mail：zhangwenyu8518@126.com [↑](#footnote-ref-1)
2. [] 周腾腾，戚永洁，戴建军，基于大数据思维的环境工程发展趋势分析，广东化工，1007-1865(2019)08-0130-01 [↑](#endnote-ref-1)
3. [] 刘杰，李佟，李军，基于改进支持向量回归机的污水处理厂出水总氮预测模型，环境工程学报，1673-9108( 2018) 01-0119-08 [↑](#endnote-ref-2)
4. [] NAZARIZ，YUANZS，SANTOPOD，et.al．Low-temperature thermal pretreatment of municipal wastewater sludge: Process optimization and effects on solubilization and anaerobic degradation［J］．Water Research，2017，113:111-123 [↑](#endnote-ref-3)
5. [] HVALAN，VRECKOD，LEVSTEK M，et al． The use of dynamic mathematical models for improving the designs of upgraded wastewater treatment plants［J］．Journal of Sustainable Development of Energy，Water and Environment Systems，2017， 5( 1) :1531 [↑](#endnote-ref-4)
6. [] 宋贤民, 基于支持向量机的曝气池 SVI 软测量研究, 科学技术与工程, 1671-1815（2011）4-0841-03 [↑](#endnote-ref-5)
7. [] 包星星，赵璨，饶家声，机器学习在 MRI 图像脑肿瘤分割中的研究进展，医疗卫生装备，2019，40（11）：90-96 [↑](#endnote-ref-6)
8. [] 邓乃扬,等.数据挖掘中的新方法─支持向量机[M].北京:科学出版社,2004. [↑](#endnote-ref-7)
9. [] 黄道平,邱禹,刘乙奇,李艳, 面向污水处理的数据驱动故障诊断及预测方法综述, 华南 理工大学学报 (自然科学版) ,2015.03.017 [↑](#endnote-ref-8)
10. [] HAMED MM，KHALAFALLAH M G，HASSANIEN E A． Prediction of wastewater treatment plant performance using artificial neural networks［J］．Environmental Modelling ＆ Software，2004，19( 10) :919928 [↑](#endnote-ref-9)
11. [] HONG Y S T，ＲOSEN M Ｒ，BHAMIDIMAＲＲI Ｒ．Analysis of a municipal wastewater

    treatment plant using a neural network based pattern analysis［J］ ． Water Research，2003， 37( 7) :16081618 [↑](#endnote-ref-10)
12. [] MJALLI F S，AL-ASHEH S，ALFADALA H E．Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performanc［J］． Journalof Environmental Management， 2007， 83( 3) :329338 [↑](#endnote-ref-11)
13. [] ALVANI V，NABIZADEH Ｒ，ANSAＲIZADEH M，et al． Predicting TOC removal efficiency in hybrid biological aerated filter using artificial neural network［J］．Desalination and Water Treatment，2016，57( 43) :2028320291 [↑](#endnote-ref-12)
14. [] ZHAO Y，GUO L，LIANG J B，et al．Seasonal artificial neural network model for water quality prediction via a clustering analysis method in a wastewater treatment plant of China［J］．Desalination and Water Treatment，2016，57( 8) :34523465 [↑](#endnote-ref-13)
15. [] 张永利,朱艳伟, 基于支持向量机的污水状态识别, 唐山师范学院学报, 1009-9115(2011)02-0030-03 [↑](#endnote-ref-14)
16. [] S. Azimi, M. Azhdary Moghaddam, S.A. Hashemi Monfared. Prediction of annual drinking water quality reduction based on Groundwater Resource Index using the artificial neural network and fuzzy clustering[J]. Journal of Contaminant Hydrology,2018. [↑](#endnote-ref-15)
17. [] Antwi P, Li J, Boadi P O, et al, Estimation of biogas and methane yields in an UASB treating Potato starch processing wastewater with backpropagation artificial neural network[J], Bioresource Technology,2017,228:106-115. [↑](#endnote-ref-16)
18. [] Cai J, Zhen P, Qaisar M, et al. Prediction and quantifying parameter importance simultaneous anaerobic sulfide and nitrate removal process using artificial neural network[J], Environmental Science and Pollution Research Inrernational,2015,22(11):8272-8279. [↑](#endnote-ref-17)
19. [] Mirbagheri S A, Bagheri M, Bagheri Z, et al. Evaluation and prediction of membrane fouling in a submerged membrane bioreactor with simultaneous upward and downward aeration using artificial neural network-genetic algorithm[J]. Process Safety and Environmental Protection,2015,96:111-124. [↑](#endnote-ref-18)
20. [] Schmitt F, Banu R, Yeom I, et al. Development of artificial neural network to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater[J]. Biochemical Engineering Journal,2018,133:47-58. [↑](#endnote-ref-19)
21. [] 余颖，乔俊飞，叶旭东，基于神经网络的污水处理过程建模的研究，第五届全球智能控制与自动化大会，2004，6，15-19。 [↑](#endnote-ref-20)
22. [] 林佳敏，陈金良，林晶晶，李宣辑，马聪，张志强，沈亮，BP神经网络和ARIMA 模型对污水处理厂出水总氮浓度的模拟预测，厦门大学，1674-991X(2019)05-0-0 [↑](#endnote-ref-21)
23. [] 张璐，张嘉成，韩红桂，乔俊飞，基于模糊神经网络的污水处理生化除磷过程控制，化工学报，：0438-1157（2019）00-0000-00 [↑](#endnote-ref-22)
24. [] 孟彦京，高希明，神经网络在污水处理过程建模中的应用，仿真技术，1008-0570(2009)09-1-0165-02 [↑](#endnote-ref-23)
25. [] 丛秋梅, 柴天佑, 余文，污水处理过程的递阶神经网络建模，控制理论与应用，1000-8152（2009）01-0008-07 [↑](#endnote-ref-24)
26. [] DUAN J C, CHUNG F L. Cascaded fuzzy neural network model based on syllogistic fuzzy reasoning[J]. IEEE Transactions on Fuzzy Systems,2001, 9(2): 293 – 306. [↑](#endnote-ref-25)
27. [] RICARDO J G, CAMPELLO B. Hierarchical fuzzy relational models: Linguistic interpretation and universal approximation[J]. IEEE Transactions on Fuzzy Systems, 2006, 14(3): 446 – 453. [↑](#endnote-ref-26)
28. [] Amir Hamzeh Haghiabi, Ali heidar Nasrolahi, Abbas Parsaie. Water quality prediction using machine learning methods[J]. Water Quality Research Journal of Canada,2018,53(1). [↑](#endnote-ref-27)
29. [] Mingzhi Huang, Di Tian, Hongbin Liu, Chao Zhang, Xiaohui Yi, Jiannan Cai, Jujun Ruan, Tao Zhang, Shaofei Kong, Guangguo Ying, Hassan Zargarzadeh. A Hybrid Fuzzy Wavelet Neural Network Mode with Self-Adapted Fuzzy c-Means Clustering and Genetic Algorithm for Water Quality Prediction in Rivers[J]. Complexity,2018. [↑](#endnote-ref-28)
30. [] Mohammed Hameed, Saadi Shartooh Sharqi, Zaher Mundher Yaseem, Haitham Abdulmohsin Afan, Aini Hussain, Ahmed Elshafie, Application of artificial intelligence (AI) techniques in water quality index prediction: a case study in tropical region, Malaysia[J], Neural computing and Applications,2017,28(1). [↑](#endnote-ref-29)
31. [] 李晓东 ,曾光明,蒋茹,李峰,石林,梁婕,韦安磊,黄国和，改进支持向量机对污水处理厂运行状况的故障诊断，湖南大学学报(自然科学版)，1000-2472(2007)12-0068-04 [↑](#endnote-ref-30)
32. [] 大连理工大学. DB 21/1627—2008 辽宁省污水综合排 放标准［S］. 沈阳：辽宁科学技术出版社，2009 [↑](#endnote-ref-31)
33. [] 周文君，基于改进的最小二乘支持向量机污水软测量建模研究，安徽工业大学，2019 [↑](#endnote-ref-32)