

基于 Elman 神经网络集成的诺西肽发酵过程建模

牛大鹏^{1,2}, 王福利^{1,2}, 何大阔², 贾明兴²

(1. 东北大学 流程工业综合自动化教育部重点实验室, 辽宁 沈阳 110004;

2. 东北大学 信息科学与工程学院, 辽宁 沈阳 110004)

摘 要: 针对单神经网络模型外推效果不理想、泛化能力较差的缺点,将神经网络集成用于诺西肽发酵过程的建模.采用 Bagging 技术进行重复取样用于个体神经网络的训练,结论生成时采用加权平均法,各子网络的权重利用差分进化算法来确定.个体神经网络选用典型的动态神经网络 Elman 网络,通过对多个 Elman 神经网络模型的输出进行融合,建立了基于神经网络集成的诺西肽发酵产物浓度模型.最后将所建立的模型与基于单神经网络的模型进行了比较,结果说明该模型具有更高的精度和泛化能力.

关键词: 诺西肽发酵;建模;神经网络集成;差分进化算法;Elman 神经网络

中图分类号: TP 274

文献标识码: A

文章编号: 1005-3026(2009)06-0761-04

Nosiheptide Fermentation Process Modeling Based on Elman Neural Network Ensemble

NIU Da-peng^{1,2}, WANG Fu-li^{1,2}, HE Da-kuo², JIA Ming-xing²

(1. Key Laboratory of Integrated Automation of Process Industry, Ministry of Education, Northeastern University, Shenyang 110004, China; 2. School of Information Science & Engineering, Northeastern University, Shenyang 110004, China. Correspondent: NIU Da-peng, E-mail: niudapeng@ise.neu.edu.cn)

Abstract: In order to improve the poor extrapolation effect and generalizability of the single neural network, the neural network ensemble is used to develop the model of Nosiheptide fermentation process. Each individual network is trained on a bootstrap re-sampling replication of the original training data through the Bagging approach. Then, outputs of the individual neural networks are combined to form an overall output of neural network ensemble through the weighted average method, in which the weight of each individual network is determined by the differential evolution algorithm. The Elman network, a typical dynamic neural network, is applied in each individual network. The model of Nosiheptide fermentation product concentration, based on the neural network ensemble, is thus developed through combination of outputs from multi-Elman neural networks. This model is compared with the single neural network model to illustrate its high accuracy and generalizability.

Key words: Nosiheptide fermentation; modeling; neural network ensemble; differential evolution algorithm; Elman neural network

对诺西肽发酵过程进行优化是提高发酵产量的有效途径,而发酵建模是优化的基础.神经网络凭借其良好的逼近复杂非线性系统的能力,在发酵过程的建模中获得了广泛的应用^[1-3].单一的神经网络虽然可用于发酵过程建模,但由于发酵过程非常复杂,单个神经网络拟合能力有限,现有

的单神经网络模型存在着外推效果不理想、泛化能力较差等缺点.1990年,Hansen和Salamon^[4]提出了神经网络集成方法,通过训练多个神经网络并将其结果进行合成,可以显著地提高神经网络系统的泛化能力.该方法易于使用且效果明显,是一种非常有效的工程化神经网络计算方法.

收稿日期: 2008-08-06

基金项目: 国家自然科学基金资助项目(60774068); 流程工业综合自动化教育部重点实验室开放课题(PAL200509).

作者简介: 牛大鹏(1980-),男,山东淄博人,东北大学助教,东北大学博士研究生;王福利(1957-),男,辽宁辽阳人,东北大学教授,博士生导师.

本文利用神经网络集成对诺西肽(Nosiheptide)发酵过程的产物浓度进行建模.首先,采用 Bagging 技术^[5]由原始训练数据集生成新的训练数据集,经训练生成用于集成的个体网络.然后,对个体网络的输出进行加权合成得到网络集成的输出结果,为了改善集成的效果,本文采用差分进化算法对合成权重进行优化.由于发酵过程是一个动态过程,个体神经网络选用典型的动态神经网络 Elman 网络,它在前馈神经网络基本结构的基础上,通过存储内部状态使其具有映射动态特征的功能,非常适合于发酵过程的建模.本文最后将所建立的 Elman 神经网络集成模型与基于单神经网络的模型进行了比较.

1 Elman 神经网络集成

1.1 Elman 神经网络

Elman 网络是一类有局部反馈的神经网络,其结构如图 1 所示. Elman 网络除了输入层、隐层、输出层之外,还有一个特殊的上下文单元,上下文单元用来记忆隐层单元以前时刻的输出值.网络的前馈连接部分可以进行连接权修正,而递归部分是固定的,不进行修正.具体地,上下文单元 k 时刻的输出,等于隐层在 $k-1$ 时刻的输出加上上下文单元 $k-1$ 时刻输出值的 α 倍^[6],即

$$x_{c,l}(k) = \alpha \cdot x_{c,l}(k-1) + x_l(k-1), \quad l = 1, 2, \dots, n. \quad (1)$$

式中: $x_{c,l}(k)$ 和 $x_l(k)$ 分别表示第 l 个上下文单元和第 l 个隐层单元的输出; α 为自连接反馈增益因子.

Elman 神经网络通常采用 BP 算法训练,其学习过程详见文献^[6].

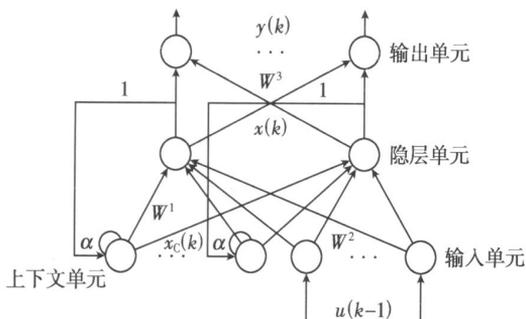


图 1 Elman 神经网络结构

Fig. 1 Architecture of Elman neural network

1.2 神经网络集成

构造神经网络集成包括两个步骤,首先生成个体神经网络,然后将个体神经网络的输出进行结合^[7].在生成个体神经网络方面, Bagging 是最常用的技术之一,它通过重复取样(bootstrap sampling)

增加了神经网络集成的差异度,从而提高了泛化能力,并具有较强的鲁棒性和稳定性^[8].本文采用 Bagging 技术生成训练样本集来训练个体神经网络.

用以生成神经网络模型的原始训练样本集记为 $\{U, Y\}$, 其中 $U \in \mathbf{R}^{n \times p}$ 为输入数据, $Y \in \mathbf{R}^{n \times q}$ 为期望输出, n 为采样个数, p 为输入变量个数, q 为输出变量个数.构造神经网络集成的过程如下^[9].先由原始训练样本集通过重复取样生成 M 个训练样本集 $\{U_1, Y_1\}, \{U_2, Y_2\}, \dots, \{U_M, Y_M\}$, 其中, $U_m \in \mathbf{R}^{n \times p}$, $Y_m \in \mathbf{R}^{n \times q}$, $m = 1, 2, \dots, M$.每个训练样本集分别训练相应的个体神经网络 N_1, N_2, \dots, N_M , 个体神经网络的输出分别记为 $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_M$.网络集成的输出结果由生成的个体神经网络的输出加权决定,

记为 $\hat{Y} = \sum_{m=1}^M m \hat{Y}_m$, 其中, $m > 0, \sum_{m=1}^M m = 1$.

1.3 合成权重优化

有研究表明^[10],在神经网络集成过程中,通过对个体网络的合成权重进行优化,可以改善集成的效果.本文采用差分进化算法来优化个体网络的合成权重.集成神经网络输出的平方和误差(SSE)反映了集成神经网络的精度与泛化能力,作为权重优化的目标函数.

差分进化(DE)算法以遗传算法作为基本框架,针对实数编码的遗传个体设计了差分操作,并利用该操作实现了算法中的杂交和变异^[11].DE 算法速度快,鲁棒性好,在实数域上搜索能力强.下面给出 DE 算法的核心操作.

1) 变异操作. DE 最基本的变异成分是父代的差分向量,每个向量由父代(第 G 代)群体中两个不同的个体 (Z_{r1}^G, Z_{r2}^G) 生成.差分向量定义为

$$D_{1,2} = Z_{r1}^G - Z_{r2}^G. \quad (2)$$

在差分向量基础上进行变异操作,其方程为

$$\hat{Z}_i^{G+1} = Z_i^G + F(D_{1,2}). \quad (3)$$

式中: \hat{Z}_i^{G+1} 为变异个体; Z_i^G 为父代个体; Z_{r1}^G, Z_{r2}^G 为不同于 Z_i^G 且互不相同的个体; F 为缩放因子,表示差分向量对下一代个体的影响大小.

2) 交叉操作.群体中第 i 个个体 Z_i^G 将与 \hat{Z}_i^{G+1} 进行交叉操作,产生试验个体 Z_T .对于 Z_T 中的各个位,利用交叉概率因子 C_R 决定哪位由 \hat{Z}_i^{G+1} 贡献,哪位由 Z_i^G 贡献.交叉操作的方程为

$$z_{jT} = \begin{cases} \hat{z}_{ji}^{G+1}, & \text{rand}() < C_R; \\ z_{ji}^G, & \text{其他}. \end{cases} \quad (4)$$

式中, $\text{rand}()$ 为 $[0, 1]$ 之间的均匀随机数, $j = 1, 2, \dots, n$, 表示第 i 个变量, n 为变量的维数。

3) 选择操作 DE 采用“贪婪”的搜索策略, 经过变异和交叉操作后生成的试验个体 Z_T 与 Z_i^G 进行竞争, 只有当 Z_T 的适应度较 Z_i^G 更优时才被选作子代; 否则, 直接将 Z_i^G 作为子代。对于有约束的优化问题, 本文利用以下准则来选出子代个体^[12]: 若 Z_T 和 Z_i^G 都可行, 有较高适应度的个体胜出; 若只有一个个体可行, 可行的个体胜出; 若两个个体都不可行, 偏离约束条件较少的个体胜出。

2 诺西肽发酵神经网络集成模型的建立

2.1 诺西肽发酵实验

诺西肽是一种典型的硫肽类抗生素, 用作饲料添加剂时, 可以明显促进动物的生长, 是一种理想的新型非吸收性动物饲料添加剂。通过发酵方式获取诺西肽, 发酵实验采用 100 L 搅拌式发酵罐, 按照发酵工艺提供的培养基配方, 进行分批发酵实验。发酵过程状态变量数据采用传感器系统自动获取与存储, 可在线测量的变量有温度、压力、pH 值、溶解氧、搅拌转速、通气流量等。生物量每 3 h 取样一次, 通过离线化验获得, 离线测定的生物量包括菌体、基质 (葡萄糖) 和诺西肽产物质量浓度。发酵初始基质质量浓度为 40 g/L, 初始菌体质量浓度为 1.2 g/L, 初始产物质量浓度为 0。

2.2 模型结构及输入变量选取

发酵过程优化是提高发酵产量的有效途径, 而过程建模是优化的基础。对诺西肽发酵过程而言, 虽然已有发酵产物浓度的机理模型, 但是由于发酵过程机理复杂, 影响因素很多, 机理推导中进行了简化并设置了限制, 致使所建模型精度不高。本文采用神经网络集成对诺西肽发酵产物浓度进行建模。模型结构如图 2 所示。

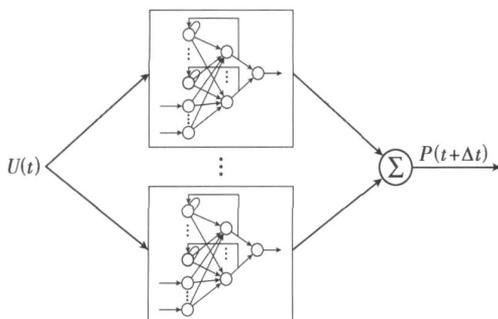


图 2 神经网络集成模型结构

Fig. 2 Model architecture of neural network ensemble

该模型将在线可测变量 $U(t)$ 和发酵产物浓度 $P(t+\Delta t)$ 分别作为输入和输出, 组成训练样本, Δt 为采样间隔。通过重复取样由原始训练样本集生成新的训练样本集对个体 Elman 神经网络进行训练, 直到神经网络收敛, 得到网络的权值和阈值。然后, 将生成的个体神经网络的输出进行加权平均得到集成的输出结果, 合成权重采用差分进化算法进行优化。下面通过机理分析确定模型的输入变量。

诺西肽发酵需在合适温度下进行, 温度对菌体代谢具有调节作用, 并会影响产物的生成速度; pH 值对菌体生长影响很大, 会影响培养基中某些营养物和中间代谢产物的解离, 从而影响菌体对营养物质的吸收、利用和产物的生成; 溶解氧 (DO) 是需氧微生物生长所必需的, 培养液中的溶解氧和细胞内溶解氧之差形成氧传递的推动力, 溶解氧的变化会影响菌体的代谢过程, 在诺西肽发酵中溶解氧的大小由通气量 (Q) 和搅拌转速 (n) 决定。综上, 在建立诺西肽发酵产物浓度模型时, 选取 T , pH 值, Q 和 n 作为输入变量。

2.3 实验与结果

用于模型训练与测试的数据都来自诺西肽分批发酵实验。发酵周期为 96 h, 取 3 h 为采样周期, 其中在线可测变量利用传感器自动获取, 产物浓度通过取样化验分析获得。取 10 批次诺西肽分批发酵数据, 其中 9 批次作为训练数据, 1 批次用于测试。由于不同过程变量数据的原始值相差较大, 为避免小数据被淹没, 在进行模型训练和测试时, 需要对数据进行标准化处理。

建立 Elman 神经网络集成模型时, 通过重复取样生成训练样本集, 用于训练个体神经网络。个体 Elman 神经网络的个数取为 10, 采用 BP 算法对 Elman 神经网络进行训练。网络结构采用一个隐含层, 隐含层和输出层的激活函数分别取 S 型函数和线形函数。为提高集成网络的泛化能力, Elman 神经网络集成中的个体网络分别采用不同的结构, 即各个体网络的隐层单元取不同数量^[7], 分别在 10~30 之间随机选择。

为验证 Elman 神经网络集成的性能, 分别用 BP 神经网络、Elman 神经网络和 Elman 神经网络集成建立了诺西肽发酵过程的产物浓度模型。利用验证数据对 3 种方法所建模型进行检验, 验证结果如图 3 所示。可见, Elman 神经网络在诺西肽发酵建模中具有比 BP 神经网络更优异的性能, 而集成的 Elman 神经网络的评价精度和泛化能力要高于单个的神经网络。

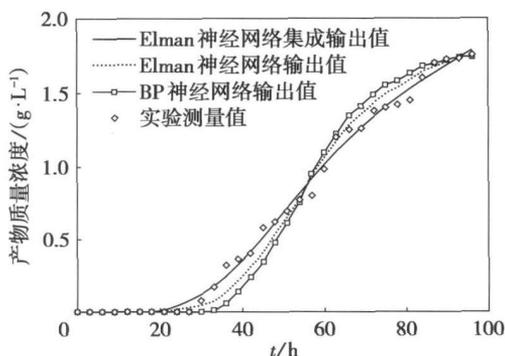


图 3 模型输出值与实验测量值比较图

Fig. 3 Comparison of model output values to measured values

3 结 论

本文建立了诺西肽发酵产物浓度的 Elman 神经网络集成模型,利用 Bagging 技术生成训练样本集来训练单个 Elman 神经网络,为了提高个体网络的差异度从而增强网络集成的泛化能力,个体网络采用了不同的结构和参数设置;然后采用加权平均将个体网络的输出进行结合,结合的权重采用差分进化算法来确定.这种集成策略有效提高了所建模型的泛化能力,增强了模型的实用性.该模型与单神经网络模型相比具有更高的精度,能更加准确地描述发酵过程.

参考文献:

[1] Potocnik P, Grabec I. Empirical modeling of antibiotic fermentation process using neural networks and genetic algorithms[J]. *Mathematics and Computers in Simulation*, 1999,49:363 - 379.

[2] Chen L Z, Nguang S K, Chen X D. Modelling and optimization of fed-batch fermentation processes using dynamic neural networks and genetic algorithms [J]. *Biochemical Engineering Journal*, 2004,22:51 - 61.

[3] Fang B S, Chen H W, Xie X L. Using genetic algorithms coupling neural networks in a study of xylitol production: medium optimization [J]. *Process Biochemistry*, 2005,38: 979 - 985.

[4] Hansen L K, Salamon P. Neural network ensembles [J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 1990,12(10):993 - 1001.

[5] Breiman L. Bagging predictors [J]. *Machine Learning*, 1996,24(2):123 - 140.

[6] 时小虎,梁艳春,徐旭.改进的 Elman 模型与递归反传控制神经网络[J].*软件学报*,2003,14(6):1110 - 1119. (Shi Xiao-hu, Liang Yan-chun, Xu Xu. An improved Elman model and recurrent back-propagation control neural networks [J]. *Journal of Software*, 2003,14(6):1110 - 1119.)

[7] 薛月菊,胡月明,杨敬锋.基于 SFAM 神经网络集成的土地评价[J].*农业工程学报*,2008,24(3):184 - 188. (Xue Yue-ju, Hu Yue-ming, Yang Jing-feng. Land evaluation based on SFAM neural network ensemble [J]. *Transactions of the CSAE*, 2008,24(3):184 - 188.)

[8] 周志华,陈世福.神经网络集成[J].*计算机学报*,2002,25(1):1 - 8. (Zhou Zhi-hua, Chen Shi-fu. Neural network ensemble [J]. *Chinese Journal of Computers*, 2002,25(1):1 - 8.)

[9] Zhang J, Martin E B, Morris A J. Inferential estimation of polymer quality using stacked neural networks [J]. *Computers and Chemical Engineering*, 1997,21:1025 - 1030.

[10] Zhou Z, Wu J, Tang W. Ensembling neural networks: many could be better than all [J]. *Artificial Intelligence*, 2002,137:239 - 263.

[11] Rainer S, Price K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces [J]. *Journal of Global Optimization*, 1997,11(4): 341 - 359.

[12] Deb K. An efficient constraint handling method for genetic algorithms [J]. *Computer Methods in Applied Mechanics and Engineering*, 2000,186(2/3/4):311 - 338.