

基于集成神经网络的水泥生产能耗建模

黄 堃, 杨 文, 丁孝华

(国电南瑞科技股份有限公司, 江苏 南京 211106)

摘要: 为了提高水泥生产过程的能耗建模和预测的精度, 提出了一种基于神经网络与马尔科夫修正的水泥生产集成能耗预测模型: 在数据预处理阶段, 为了减小处理数据的规模, 采用平均影响值法进行数据降维, 筛选敏感变量, 从 12 个变量中选出对能耗输出影响较大的 6 个, 构建一个 6 输入单输出的神经网络, 使能耗建模阶段选用的神经网络模型结构更为简单, 可以有效减少训练神经网络所需的时间。在能耗建模阶段, 为了建立性能更佳的能耗模型, 在以神经网络作为能耗建模元学习器的基础上采用集成学习思想, 组合数个元学习器成一个性能更佳的强学习器, 即将多个神经网络的预测输出值求平均作为集成模型的预测结果, 采用水泥烧成系统的依赖变量和对应的炉窑能耗值作为试验数据进行模型的训练、验证和预测, 结果表明, 集成模型预测结果的决定系数 R^2 值较单个神经网络提高了 0.019, 预测值与真实值的相对残差的均值较单个神经网络也减少了 0.027, 模型性能有所提高。在能耗预测阶段, 为了进一步提高模型的预测精度, 引入马尔科夫残差修正法, 即依据历史预测能耗值与实际能耗值的残差修正当前预测值, 提升集成能耗模型的预测精度。结果表明, 经马尔科夫修正法修正的预测值相对残差从 -0.6% 降至 -0.25%, 能耗预测值更加接近实际值, 预测精度提升, 可更好地挖掘水泥炉窑烧成系统电能耗变化与依赖变量的规律, 实现能耗精确预测, 为水泥生产过程的能耗监管提供了更精确的参考依据。根据水泥生产能耗建模 3 个阶段的描述, 提出一种基于神经网络与马尔科夫修正的水泥生产集成能耗预测模型, 在水泥生产能耗预测上有更佳的预测效果和更高的预测精度。

关键词: 神经网络; 能耗建模; 水泥生产; 集成算法; 马尔科夫修正

中图分类号: TQ53; TK114 文献标志码: A 文章编号: 1006-6772(2020)05-0103-08

Energy consumption modeling of cement production based on integrated neural network

HUANG Kun, YANG Wen, DING Xiaohua

(NARI Technology Co., Ltd., Nanjing 211106, China)

Abstract: In order to improve the accuracy of energy consumption modeling and prediction in the cement production process, an integrated energy consumption prediction model for cement production based on neural network and Markov correction was proposed in this paper. In the data preprocessing stage, in order to reduce the scale of processing data, the average influence value method was used to reduce the data dimension, and the sensitive variables were filtered, six of the 12 variables that have a greater impact on the energy consumption output were selected to construct a 6-input single-output neural network, which made the energy consumption modeling stage selection. The structure of the neural network model is simpler, and could effectively reduce the time required to train the neural network. In the energy consumption modeling stage, in order to establish a better performance energy consumption model, the integrated learning idea was adopted on the basis of the neural network as the energy consumption modeling meta-learner, and several meta-learners were combined into a stronger performance. The average of the predicted output values of multiple neural networks was used as the prediction result of the integrated model. The dependent variables of the cement firing system and the corresponding furnace energy consumption value were used as experimental data for model training, verification and prediction. The experimental results show that the determination coefficient of the

收稿日期: 2020-08-06; 责任编辑: 张晓宁 DOI: 10.13226/j.issn.1006-6772.IF20080609

基金项目: 国家重点研发计划资助项目(2016YFB0601501)

作者简介: 黄堃(1985—), 男, 江苏南京人, 高级工程师, 从事企业能效管理及新能源研究工作。huangkun1@sgepri.sgcc.com.cn。通讯作者: 杨文, 高级工程师, 从事企业能效管理及新能源研究工作。yangwen@sgepri.sgcc.com.cn

引用格式: 黄堃, 杨文, 丁孝华. 基于集成神经网络的水泥生产能耗建模[J]. 洁净煤技术, 2020, 26(5): 103-110.

HUANG Kun, YANG Wen, DING Xiaohua. Energy consumption modeling of cement production based on integrated neural network[J]. Clean Coal Technology, 2020, 26(5): 103-110.



移动阅读

prediction results of the integrated model is improved by 0.019 compared with a single neural network. The mean value of the relative residual size between the predicted value and the true value is also reduced by 0.027 compared with the single neural network. The performance of the model is improved. In the energy consumption prediction stage, in order to further improve the prediction accuracy of the model, the Markov residual correction method is introduced, that is, the current predicted value is corrected based on the residual of the historical predicted energy consumption value and the actual energy consumption value to improve the prediction of the integrated energy consumption model accuracy. The experimental results show that the relative residual error of the predicted value corrected by Markov's correction method is reduced from -0.6% to -0.25% . The energy consumption predicted value is closer to the actual value, the prediction accuracy is significantly improved, and the law of electric energy consumption change and dependent variable of cement furnace firing system can be better excavated. The energy consumption is accurately predicted, which provides a more accurate reference basis for energy consumption supervision in the cement production process. Based on the description of the three stages of cement production energy consumption modeling, a cement production integrated energy consumption prediction model based on neural network and Markov correction is proposed in this paper, which has better prediction effects and higher prediction accuracy on cement production energy consumption.

Key words: neural network; energy consumption modeling; cement production; integrated algorithm; Markov correction

0 引言

目前,水泥生产作为我国国民经济发展不可或缺的支柱产业,同时作为高能耗产业,水泥生产工艺流程的优化与节能环保的需求日益突出。而对水泥生产的能耗预测是减少电能消耗与污染物排放、提高水泥生产过程的品质和效率的前提条件^[1-3]。国内外研究人员对构建水泥炉窑的能耗预测模型进行了研究,一般可分为基于机理建模和基于数据建模2类^[4-5],但水泥炉窑煅烧涉及众多环节与设备,参数具有变量多、关联性强等特点。机理建模无法准确描述烧成系统主要参数与能耗的关系。因此数据建模是当前的主流研究方向,其主要是采用专家系统、自适应回归、模糊系统、人工神经网络等人工智能方法对水泥炉窑的能耗建模进行研究^[6-8]。

马尔科夫过程常被用于船舶交通流量、瓦斯浓度、需水量、股票走势的预测。吕鹏飞等^[9]利用BP神经网络建立船舶交通量的预测模型,预测时结合马尔科夫修正法有效提高了预测精度。韩婷婷等^[10]采用马尔科夫修正法修正灰色神经网络模型预测值,使预测的瓦斯浓度变化趋势更贴近实际瓦斯浓度的变化曲线。景亚平等^[11]结合马尔科夫修正法建立了灰色神经网络的城市需水量预测模型,试验表明其获得了优于单一灰色神经网络预测模型的预测效果。WANG等^[12]建立了基于马尔科夫过程的模糊神经网络预测模型,较精准地预测了股票指数的走势。

相关学者通常采用单个神经网络建立水泥生产窑的能耗预测模型。为提高预测模型的精度,本文在数据预处理时采用平均影响值法筛选能耗敏感变量。建模时以RBF神经网络为元学习器,结合集成算法的思想,建立精度较单个神经网络更高的集成

能耗模型。在基于集成模型预测能耗时采用马尔科夫修正法,即依据历史能耗预测值与实际值的残差修正网络输出的能耗预测值,从而提高水泥生产的能耗预测精度。

1 水泥生产的工艺流程

水泥的生产过程主要包含生料制备、熟料煅烧和水泥磨粉3个阶段^[13-14](简称“两磨一烧”),其主要生产过程如图1所示。

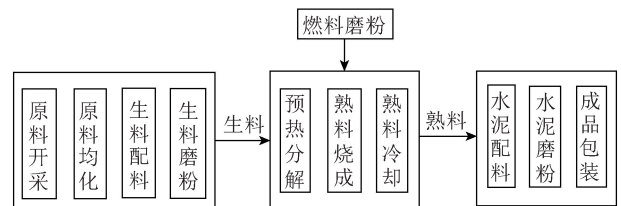


图1 水泥的生产过程

Fig.1 Cement production process

水泥生料制备所用原材料包括石灰石、砂岩、黏土、铁矿石、粉煤灰及煤碳等,其中石灰石是水泥生产的主要原材料,具有颗粒较大、硬度较高等特点,需经破碎机破碎处理,实现原料的均化;再将原料按一定比例混合,实现生料的调配;最后经生料磨冲击磨粉的方式处理成细颗粒。

熟料煅烧是水泥生产中能耗最大的阶段,主要包括预热分解、烧成和熟料冷却3个部分。将生料输送至预热器顶端的进料口,生料在预热器内进行预热和少量的预分解处理,而后在预热器下端的分解炉内热分解大部分生料,经过分解的生料进入回转窑,在烧成反应下生成熟料,最后将熟料导入冷却机冷却。水泥磨粉阶段,在刚出冷却机的熟料内加入适量的石膏和矿渣,经水泥磨磨成细粉状的水泥出厂。

综上,水泥烧成系统能耗可能的依赖变量主要

包括:喂煤量、CO 体积分数、生料流量、冷却剂鼓风管道压力、入冷却机空气温度、熟料流量、预热器的出口压力、预热器废水温度、冷却机出口熟料量、冷却机出口熟料温度、冷却剂烟筒压力、冷却剂烟筒废气温度、分解炉中段表面温度、回转炉中段表面温度和环境温度等。

2 水泥生产能耗预测模型

2.1 数据预处理

从水泥窑分散控制系统(distributed control system, DCS)中按时间序列收集喂煤量、CO 体积分数、生料流量、冷却剂鼓风管道压力、入冷却机空气温度、熟料流量等变量值及对应电能消耗的实际数据,作为试验样本数据集。

在利用神经网络进行能耗建模前,不能明确各输入变量对输出能耗结果的影响程度,选择的变量较多时,影响较小的变量会使模型结构更复杂,甚至可能会降低能耗模型的精度。因此需通过合适的预处理方法筛选自变量,获得对能耗输出影响较大的变量作为神经网络的输入变量,本文选用平均影响值法选择参数^[15]。其具体实现步骤如下:

若原始训练数据集含有一定数量的样本,每个样本包含 n 个属性,即样本集 $P = \{P_1, P_2, \dots, P_n\}$,输出为一个变量 $Y = [y_1, y_2, \dots, y_m]$ 。

Step1:在原始神经网络训练结束后,将训练样本集 P 中的数据列 $P_j(j = 1, 2, 3, \dots, n)$ 的每个数值在原来基础上分别加和减 10%,形成 2 个新的数据列 P_j^1 和 P_j^2 。

Step2:利用已训练好的模型分别对由 P_j^1 和 P_j^2 组成的新训练样本集进行预测,分别得到 2 组预测结果 A_j^1 和 A_j^2 ,二者求差后的差值表示该变量对输出结果的影响变化值 IV_j 。

Step3:对所有的输出差值 IV_j 求和并取平均值,得到第 j 个输入变量的平均影响值 MIV_j ,本文输出结果仅包含能耗值一项,所以 $IV_j = MIV_j$ 。该值的正负号表示该输入变量与输出变量的相关方向,该输入变量对输出变量的影响程度由其绝对值大小表示。

Step4:对 MIV_j 的绝对值按照降序排列,若排序后的前 k 个平均影响值 MIV 绝对值的累计贡献率满足

$$\eta_k = \frac{\sum_{i=1}^k |MIV_i|}{\sum_{j=1}^n |MIV_j|} > \eta_0, \quad (1)$$

选择对应的 k 个输入变量代表全部输入变量重新构建神经网络建模,其中取 $\eta_0 = 85\%$ 。

2.2 集成能耗模型

在建模阶段,以 RBF 神经网络为元学习器,结合集成算法的思想,集成学习本质是使用多个元学习器并行学习,并按照某种规则将多个学习结果进行组合,获得一个强学习器,建立比单个模型精度更高的集成模型。神经网络的集成学习算法由个体生成和结论组合 2 个步骤完成。集成学习算法用于回归建模问题时,结论组合阶段通常将各神经网络的输出经平均或加权平均处理后作为集成模型的输出。个体学习器的数目和集成的效果无关,且个体分类器数量多会增加计算量,占用更多的计算机资源。因此选择差异度大的个体神经网络作为个体生成阶段的重要内容。本文采用差异化训练样本的方式完成个体的生成,具体运行流程为

Step1:将预处理后的试验样本按时间段随机分为 T 份(时间段间可交叉重叠)。

Step2:对每份样本分别划分训练集和测试集,并分别训练 T 个 RBF 神经网络,训练完成后得到 T 个能够反映不同能耗特征的能耗模型。

Step3:分别用 T 个模型对同一测试集测试得到 T 个网络输出结果,对 T 个结果采用求平均值的方式得出最终的集成输出结果。

集成神经网络建立能耗模型的过程如图 2 所示。

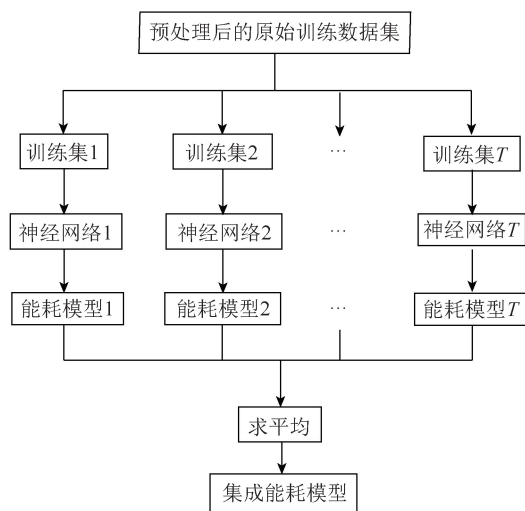


图2 集成神经网络建立能耗模型的过程示意

Fig.2 Schematic diagram of the process of integrating neural networks to build energy consumption models

元学习器(RBF神经网络)结构与学习过程如下:

RBF(Radial basis function)神经网络,又称径向

基神经网络,是一种包括输入层、单隐层、输出层^[16]的3层神经网络。输入空间经非线性变换到隐藏层空间,而隐层空间经线性变换到输出层空间,其结构具体如图3所示。

RBF网络的基本思想是以RBF作为隐单元的基,从而构成隐层空间,将输入矢量直接映射到隐空间,输入层和隐藏层之间不需要对信号进行处理,只起到信号传递的作用,没有通过权连接。这种映射关系在RBF的中心点确定后即可确定。隐层到输出层通过权连接,即网络输出对可调参数权值而言是线性的。网络的权可由梯度下降法修正获得,RBF神经网络学习速度快且无局部极小问题。理论上隐藏层节点数即中心点越多,RBF神经网络的函数逼近性能越好。

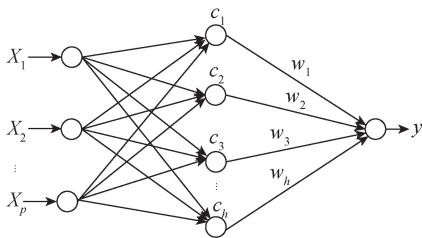


图3 RBF神经网络结构

Fig.3 Diagram of RBF neural network structure

采用高斯函数作为径向基神经网络隐藏层的激活函数,即

$$R(\mathbf{X}-c_j) = \exp\left(-\frac{1}{\beta_j^2} \|\mathbf{X}-c_j\|^2\right), j=1,2,3,\dots,h, \quad (2)$$

其中, \mathbf{X} 为 p 个维度为 h 的输入向量; h 为隐藏层中心点个数; β_j 为隐藏层的中心宽度; c_j 为第 j 个径向基函数的中心点。 $\|\mathbf{X}-c_j\|^2$ 表示向量 $\mathbf{X}-c_j$ 的欧几里德范数平方,即为输入样本点与中心点的距离,当 \mathbf{X} 到 c_j 的距离达到最近时, $\|\mathbf{X}-c_j\|^2$ 最小, $R(\mathbf{X}-c_j)$ 达到极大,此时函数被激活,神经网络的输出为

$$y = b + \sum_{j=1}^h w_j \exp\left(-\frac{\|\mathbf{X}-c_j\|^2}{\beta_j^2}\right), \quad (3)$$

式中, y 为输出单元; b 为隐藏层和输出层之间的偏置; w_j 为隐藏层到输出层的权值。

RBF神经网络的训练包括非监督学习和监督学习2个阶段。在非监督学习阶段使用K-means聚类算法确定RBF的中心点。具体训练过程为

Step1:从输入样本集中随机选取 N 个样本作为初始的中心点,即 $c_1, c_2, c_3, \dots, c_N$ 。

Step2:计算每一个输入样本与这 N 个中心点的

欧式距离,即

$$d_j = \|\mathbf{X}-c_j\|, j=1,2,3,4,\dots,N. \quad (4)$$

Step3:将每个样本点依次划分到与其距离 d_j 最近的中心点所在的簇中。

Step4:计算各个簇中的样本点均值,并将均值作为各簇新的中心点。

Step5:重复Step2~4,直至中心点不再有明显变化,即达到训练要求。

上述训练完成后,各中心点的位置即被确定。

在监督学习阶段使用梯度下降法修正隐藏层与输出层间的权重。具体训练过程为

Step1:初始化权值,并设定的准确率限值 ε 。

Step2:计算神经网络输出值,并求出误差目标函数值 E ,即

$$E = \frac{1}{2} \sum_{i=1}^p e_i^2, \quad (5)$$

其中, p 为输入训练样本的个数; e_i 为第 i 个样本输入后产生的误差量。为使总误差函数 E 达到最小值,权重的修正量应与其负梯度值成正比。

Step3:若 $E < \varepsilon$,训练结束;否则计算权值修正量 Δw 为

$$\Delta w_j = -\eta \frac{\partial E}{\partial w_j} = -\eta \sum_{i=1}^p e_i G(\|\mathbf{X}_i - c_j\|), \quad (6)$$

其中, G 为高斯函数。更新权值 $w_{k+1} = w_k + \Delta w$,然后转向Step2。

上述训练完成后,权值即可确定。RBF神经网络的训练过程如图4所示。

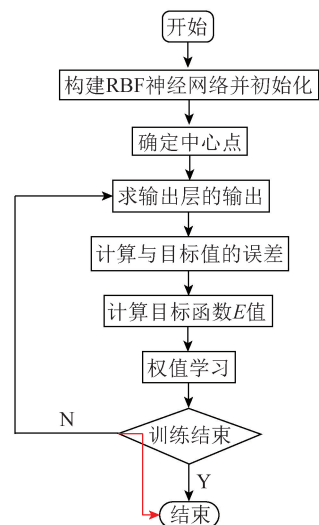


图4 RBF神经网络的训练过程流程

Fig.4 Flow chart of the training process of RBF neural network

2.3 预测值的马尔科夫修正

马尔科夫链预测过程描述的是一个随机时间序列的动态变化过程,该过程指在已知 t_0 时刻状态的

情况下,系统或者过程在 $t (t > t_0)$ 时刻所处状态的条件分布仅由时刻 t_0 的状态决定,而与 t_0 之前的状态无关^[9-10]。马尔科夫过程的处理对象是随机波动性大的离散事件数据,可用数学形式表示为

$$P\{X_{k+1} = i_k | X_1 = i_1, X_2 = i_2, \dots, X_k = i_k\} = P\{X_{k+1} = i_{k+1} | X_k = i_k\}, \quad (7)$$

式中, P 为条件概率; X_k 为子事件; i_k 为 X_k 对应的状态。

水泥生产过程的能耗值经采样后得到的是一组按时间序列的离散数据,本文求出测试样本集中的各样本能耗实际值 Y_1 与网络输出预测值 Y_2 的相对残差值,用马尔科夫过程建立残差修正模型,对集成模型输出的能耗预测值修正,使当前预测值更接近真实值。其具体过程为

Step1:按时间序列将测试样本集中的各样本能耗实际值 Y_1 与网络输出预测值 Y_2 比较,求出两者的相对残差 Z 为

$$Z = \frac{Y_2 - Y_1}{Y_1} \times 100\%, \quad (8)$$

其中网络输出预测值 Y_2 为 4 个能耗模型输出的均值。将相对残差值归一化为

$$Z^* = \frac{Z - Z_{\min}}{Z_{\max} - Z_{\min}}, \quad (9)$$

式中, Z_{\min} 为序列中相对残差的最小值; Z_{\max} 为序列中相对残差的最大值; Z^* 为归一化结果。

Step2: 求出所有 Z^* 的均值 \bar{Z}^* , 根据黄金分割法

$$\lambda_\gamma = \Delta \bar{Z}^*, |s| < n, \gamma = 1, 2, 3, \dots, n, \quad (10)$$

其中, Δ 为黄金分割率 0.618; s 为任意整数; n 为所划分的区间个数。将相对残差值按大小划分出 n 个状态 $E_1, E_2, E_3, \dots, E_n$ 。 n 个状态对应的残差区间为

$$Q_i \in (a, b), i = 1, 2, \dots, n. \quad (11)$$

Step3: 求出状态 E_i 只经 1 步转移到特定状态 E_j 的概率,即

$$P_{ij} = \frac{m_{ij}}{\sum_{j=1}^n m_{ij}}, i = 1, 2, \dots, n, \quad (12)$$

其中, m_{ij} 为序列中状态 E_i 转移到状态 E_j 的次数。1 步状态转移概率矩阵由 P_{ij} 组合形成,即

$$A_1 = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}, \quad (13)$$

根据 $C-K$ 方程求出 k 步状态转移概率矩阵为

$$A_k = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1n} \\ P_{21} & P_{22} & \dots & P_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nn} \end{bmatrix}^k = A^k, \quad (14)$$

Step4: 建立马尔科夫链预测模型为

$$p_{k+1} = p_0 A^k, \quad (15)$$

其中, p_0 为初始时刻的概率分布; p_{k+1} 为 $k+1$ 时刻的概率分布。由 $k+1$ 时刻的概率分布可得该时刻对应的状态和残差区间 $Q \in (Q_1, Q_2)$, 并根据式 (16) 修正模型预测值。

$$X_1 = \frac{\hat{X}}{1 - Q_1}, X_2 = \frac{\hat{X}}{1 - Q_2}, \quad (16)$$

其中, \hat{X} 为神经网络预测值; Q_1, Q_2 分别为对应的残差区间上限、下限。最后求 X_1 与 X_2 的平均值,即为经马尔科夫修正后的神经网络预测值。

马尔科夫修正法修正能耗预测值的过程如图 5 所示。

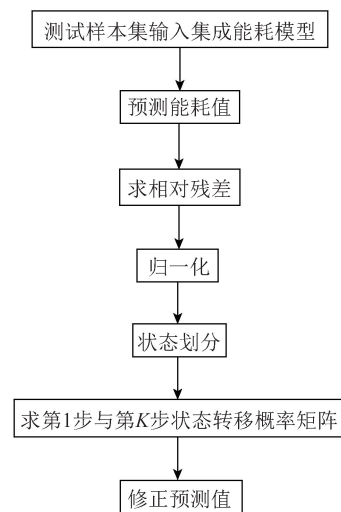


图 5 马尔科夫修正法修正能耗预测值的过程示意

Fig.5 Schematic diagram of the process of Markov correction method to correct energy consumption forecast

3 实例验证及数据分析

3.1 数据预处理

为了建立水泥生产的集成能耗模型,需要大量的水泥生产过程中的样本数据用于模型训练,因此根据要求在某水泥厂生产现场以 5 h 作为每组数据的采样间隔,从水泥窑分散控制系统 (distributed control system, DCS) 中随机收集 4 个时间段的 (时间段可交叉重叠) 喂煤量、CO 体积分数、生料流量、冷却剂鼓风管压力、入冷却机空气温度、熟料

流量等变量值及对应电能消耗的实际数据,共采集2 100个实测数据组,作为试验样本数据集,按2.1

节平均影响值算法步骤进行变量筛选,各变量参数及其MIV值见表1。

表1 各变量参数及其MIV值

Table 1 Variable parameters and their MIV values

| 参数 | 生料流量/t | 分解炉喂煤量/t | 窑头喂煤量/t | 高温风机转速/(r·s ⁻¹) | EP风机转速/(r·s ⁻¹) | 预热器出口压/Pa | 解炉中段炉内温度/℃ | 环境温度/℃ | CO体积分数 | 冷却剂鼓风机管道压/Pa | 入冷却机空气温度/℃ | 熟料流量/t |
|-----|---------|----------|---------|-----------------------------|-----------------------------|-----------|------------|---------|---------|--------------|------------|---------|
| MIV | 0.093 9 | 0.089 2 | 0.074 2 | 0.060 4 | 0.060 1 | 0.043 | 0.038 3 | 0.036 4 | 0.016 7 | 0.013 5 | 0.009 7 | 0.009 7 |

提取8个对生产能耗影响较大的关键特征参数,包括生料流量、分解炉喂煤量、窑头喂煤量、高温风机转速、EP风机转速、预热器出口压、解炉中段炉

内温度、环境温度,将其作为神经网络输入的自变量,电能消耗作为输出的因变量,其中部分数据见表2。

表2 水泥生产过程的部分样本数据

Table 2 Partial sample data of cement production process

| 生料流量/t | 分解炉喂煤量/t | 窑头喂煤量/t | 高温风机转速/(r·s ⁻¹) | EP风机转速/(r·s ⁻¹) | 预热器出口压力/Pa | 解炉中段炉内温度/℃ | 环境温度/℃ | 电能消耗/kWh |
|--------|----------|---------|-----------------------------|-----------------------------|------------|------------|--------|----------|
| 235 | 7.5 | 9.7 | 40 | 35 | 6.1 | 877 | 17 | 15 302 |
| 225 | 6.8 | 9.8 | 39 | 35 | 6.2 | 882 | 11 | 12 513 |
| 235 | 7.4 | 9.2 | 41 | 39 | 6.8 | 883 | 19 | 15 256 |
| 234 | 8.1 | 9.2 | 40 | 35 | 6.5 | 896 | 17 | 15 789 |
| 235 | 7.3 | 9.7 | 40 | 35 | 6.5 | 877 | 17 | 15 125 |
| 213 | 7.5 | 9.2 | 40 | 28 | 6.2 | 875 | 18 | 14 921 |

3.2 集成能耗模型

以Matlab作为试验平台,将2 100个预处理后的数据使用randperm函数打乱顺序,并从中随机抽取100个作为测试集T-train,剩余2 000个数据建立4个训练集T-test,每个包含500个数据样本。RBF神经网络输入层神经元的个数设置为8,输出层神经元个数为1。按2.2节RBF非监督学习阶段的K-means聚类算法的步骤,经多次试验后最终确定中心的个数为50个较合适,即可构成8-50-1的神经网络结构。使用newrbe函数创建并训练RBF神经网络,训练完成后得到4个能够反映不同能耗特征的能耗模型,再组合为集成模型。使用sim函数仿真测试,以误差值与决定系数R²作为评价指标。

$$R^2 = \frac{SSR}{SST}, \quad (17)$$

其中, $SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$, \hat{y}_i 为测试集中第*i*个实际能耗值, \bar{y} 为测试集中所有实际能耗值的均值;

$SST = \sum_{i=1}^n (\hat{y}_i - f_i)^2$, f_i 为第*i*个预测能耗值。 R^2 的范围为[0,1],该值越趋近于1,拟合优度越大,即输入变量对输出量的解释程度越高,模型精度越高。

集成模型的能耗预测效果如图6所示,单个RBF神

经网络的能耗预测效果如图7所示,2种方法预测结果比较如图8所示。

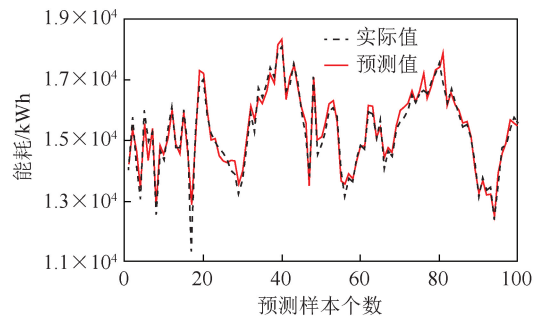


图6 集成模型的能耗预测结果

Fig.6 Energy consumption prediction results of the integrated model

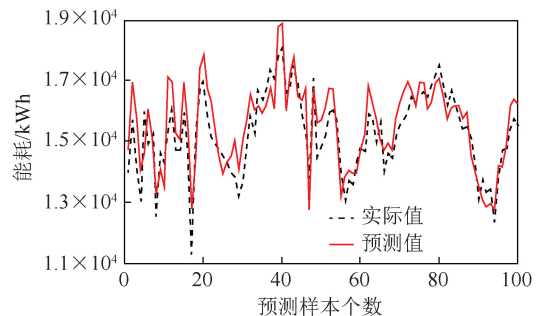


图7 单个RBF神经网络的能耗预测结果

Fig.7 Energy consumption prediction results of a single RBF neural network

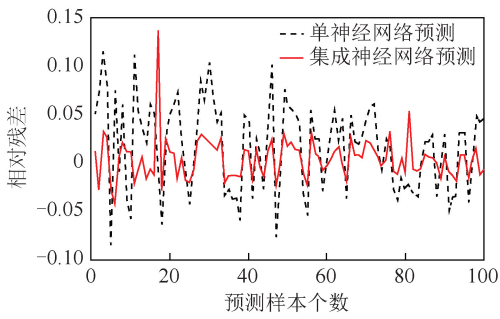


图8 2种方法预测的相对残差对比

Fig.8 Comparison of relative residuals predicted by the two methods

仿真结果表明,集成模型的 $R^2 = 0.924\ 93$,单个神经网络的 R^2 为0.906,集成模型有更高的 R^2 值,说明模型解释性更好。

根据图7两种方法能耗预测与实际值相对残差的比较,定义相对残差均值为

$$E = \frac{1}{100} \sum_{i=1}^{100} |Z_i|, \quad (18)$$

表3 集成模型的能耗预测结果

Table 3 Energy consumption prediction results of the integrated model

| 样本 | 实际值/kWh | 预测值/kWh | 相对残差/% | 相对残差归一化 | 状态 |
|----|---------|---------------|--------|---------|-------|
| 1 | 14 033 | 14 220.799 03 | 1.3 | 0.73 | E_3 |
| 2 | 15 784 | 15 371.366 10 | -2.6 | 0.20 | E_1 |
| 3 | 14 189 | 14 656.310 36 | 3.3 | 1.00 | E_3 |
| 4 | 13 077 | 13 432.284 19 | 2.7 | 0.92 | E_3 |
| 5 | 16 003 | 15 569.137 32 | -2.7 | 0.19 | E_1 |
| 6 | 14 981 | 14 369.311 54 | -4.1 | 0 | E_1 |
| 7 | 15 273 | 15 394.603 02 | 0.7 | 0.65 | E_2 |
| 8 | 12 549 | 12 822.296 76 | 2.1 | 0.84 | E_3 |
| 9 | 14 667 | 14 854.968 56 | 1.2 | 0.72 | E_2 |
| 10 | 14 367 | 14 545.776 58 | 1.2 | 0.72 | E_2 |
| 11 | 15 398 | 15 084.425 57 | -2.0 | 0.29 | E_1 |
| 12 | 16 112 | 16 009.663 94 | -0.6 | 0.47 | E_2 |
| 13 | 14 783 | 14 888.319 29 | 0.7 | 0.65 | E_2 |
| 14 | 14 769 | 14 547.065 94 | -1.5 | 0.35 | E_1 |
| 15 | 16 011 | 15 922.499 27 | -0.6 | 0.47 | E_2 |

第13个能耗值经2步转移到第15个能耗值的2步状态转移概率矩阵为

$$A_2 = \begin{bmatrix} 0.38 & 0.41 & 0.21 \\ 0.34 & 0.45 & 0.21 \\ 0.325 & 0.462\ 5 & 0.212\ 5 \end{bmatrix},$$

同理可求出 k 步状态转移概率矩阵。

由表3可知,第14个能耗值处于状态 E_1 ,其初始向量可表示为 $P_0 = [1\ 0\ 0]$ 经1步转移后,第15个能耗值的概率向量为

$$P_1 = P_0 \cdot A_1 = [1\ 0\ 0] \cdot \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.5 & 0.25 & 0.25 \end{bmatrix} =$$

其中, Z_i 为第 i 个样本输入后产生的相对残差值, $i=1,2,\dots,100$ 。计算得到单神经网络预测 $E = 0.043$,集成模型预测 $E = 0.016$,即集成模型的预测误差的损失函数值更小,模型预测精度更高。

3.3 预测值的马尔科夫修正

为提高预测精度,按时间序列求出测试样本集中的各样本能耗实际值与网络输出预测值的相对残差值(表3),并归一化至 $[0,1.00]$ (表2的第5列),求出归一后的相对残差的平均值为0.55。根据黄金分割法(取 $s=1$)的规则,将能耗值划分为3个状态 E_1 、 E_2 、 E_3 , E_1 区间为 $[0,0.47)$, E_2 区间为 $[0.47,0.72]$, E_3 的区间为 $(0.72,1.00]$ 。根据马尔科夫修正过程的描述,第14个能耗值经1步转移到第15个能耗值的1步状态转移概率矩阵为

$$A_1 = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}。$$

$$[0.2\ 0.6\ 0.2],$$

则第15个能耗值处于状态 E_2 的概率较大,并对集成能耗模型的预测值进行马尔科夫修正。

$$X_1 = \frac{\hat{X}}{1 - Q_1} = \frac{15\ 922.5}{1 + 0.006} = 15\ 827.5,$$

$$X_2 = \frac{\hat{X}}{1 - Q_2} = \frac{15\ 922.5}{1 - 0.012} = 16\ 115.9,$$

其中, \hat{X} 为第15个样本的能耗预测值,为15 922.5; Q_1 、 Q_2 分别为第15个样本能耗值对应的残差区间下限和上限,为-0.6%、1.2%。最后求出 X_1 与 X_2 的平均值为15 971.7,即为修正后的预测值,与实际值

16 011 更接近,此时两者的相对残差由原来的-0.6%降至-0.25%。

4 结 论

1)在利用神经网络进行能耗建模前,采用平均影响值算法明确各输入变量对输出能耗结果的影响程度,并从12个变量中剔除对网络输出影响较小的6个,使神经网络模型的结构更为简单。

2)在建模中引入集成算法的思想后决定系数 R^2 提高了0.019,预测值与真实值的相对残差均值减少了0.027,说明能耗模型具有了更好的解释性与更高的预测精度。

3)集成能耗模型预测的能耗值经马尔科夫修正后的相对残差从-0.6%降至-0.25%,说明经马尔科夫修正后的水泥生产的能耗预测值更接近实际能耗值。

参考文献 (References):

- [1] 石建屏,李新,吕淑珍,等.中国水泥工业CO₂排放现状及减排对策[J].环境科学学报,2012,32(8):2028-2033.
SHI Jianping, LI Xin, LYU Shuzhen, et al. China's cement industry CO₂ emissions status and emission reduction countermeasures[J]. Journal of Environmental Science, 2012, 32(8): 2028-2033.
- [2] 顾阿伦,史霄鸣,汪澜,等.中国水泥行业节能减排的潜力与成本分析[J].中国人口资源与环境,2012,22(8):16-21.
GU Alun, SHI Xiaoming, WANG Lan, et al. Analysis of the potential and cost of energy saving and emission reduction in China's cement industry[J]. China Population, Resources and Environment, 2012, 22(8): 16-21.
- [3] 王俊杰,魏丽颖,汪澜.水泥行业节能减排技术评价体系研究[J].硅酸盐通报,2014,33(5):1268-1274.
WANG Junjie, WEI Liying, WANG Lan. Research on the evaluation system of energy conservation and emission reduction technology in cement industry[J]. The Silicate Bulletin, 2014, 33(5): 1268-1274.
- [4] 赵洁.水泥回转窑系统的建模与控制研究[D].郑州:郑州大学,2014.
ZHAO Jie. Research on modeling and control of cement rotary kiln system [D]. Zhengzhou: Zhengzhou University, 2014.
- [5] 谢胜.通过热平衡计算分析水泥窑煤耗低的原因[J].水泥,2015(1):18-19.
XIE Sheng. Analyze the reasons for low coal consumption in cement kilns through heat balance calculations [J]. Cement, 2015(1): 18-19.
- [6] 覃新颖,余乾坤,彭奎,等.基于神经网络的回转窑建模及其优化控制设计[J].计算机仿真,2012,29(1):160-163.
QIN Xinyin, SHE Qianzhong, PENG Kui, et al. Modeling and optimal control design of rotary kiln based on neural network [J].

- Computer Simulation, 2012, 29(1): 160-163.
- [7] 王盛慧,赵二卫.基于多元自适应回归样条的水泥能耗评估[J].化工自动化及仪表,2017,44(12):1130-1134.
WANG Shenghui, ZHAO Erwei. Cement energy consumption assessment based on multiple adaptive regression splines [J]. Chemical Industry Automation and Instrumentation, 2017, 44(12): 1130-1134.
- [8] 郭峰.基于模糊ARX模型的水泥回转窑预测控制算法研究[D].秦皇岛:燕山大学,2012.
GUO Feng. Research on predictive control algorithm of cement rotary kiln based on fuzzy ARX model [D]. Qinhuangdao: Yanshan University, 2012.
- [9] 吕鹏飞,庄元,李洋,等.船舶交通量的BP神经网络-马尔科夫预测模型[J].上海海事大学学报,2017,38(2):18-28.
LYU Pengfei, ZHUANG Yuan, LI Yang, et al. BP neural network-Markov forecast model for ship traffic volume [J]. Journal of Shanghai Maritime University, 2017, 38(2): 18-28.
- [10] 韩婷婷,吴世跃,王鹏军.基于马尔科夫残差修正的瓦斯浓度预测[J].工矿自动化,2014,40(3):28-31.
HAN Tingting, WU Shiyue, WANG Pengjun. Gas concentration prediction based on Markov residual correction [J]. Industry and Mine Automation, 2014, 40(3): 28-31.
- [11] 景亚平,张鑫,罗艳.基于灰色神经网络与马尔科夫链的城市需水量组合预测[J].西北农林科技大学学报(自然科学版),2011,39(7):229-234.
JING Yaping, ZHANG Xin, LUO Yan. Combination forecast of urban water demand based on grey neural network and Markov chain [J]. Journal of Northwest Sci-Tech University of Agriculture and Forestry (Natural Science Edition), 2011, 39(7): 229-234.
- [12] WANG Yifan, CHENG Shimin, SHU Meihua. Incorporating the Markov chain concept into fuzzy stochastic prediction of stock indexes [J]. Applied Soft Computing, 2010, 10(2): 613-617.
- [13] 赵正平.水泥熟料烧成系统热工建模的研究[D].长春:长春工业大学,2018.
ZHAO Zhengping. Research on thermal modeling of cement clinker burning system [D]. Changchun: Changchun University of Technology, 2018.
- [14] 胡静洋.水泥回转窑热平衡分析及系统优化研究[D].青岛:山东大学,2014.
HU Jingyang. Thermal balance analysis and system optimization research of cement rotary kiln [D]. Qingdao: Shandong University, 2014.
- [15] 高倩.基于大数据的能耗特性分析方法研究[D].保定:华北电力大学,2017.
GAO Qian. Research on analysis method of energy consumption characteristics based on big data [D]. Baoding: North China Electric Power University, 2017.
- [16] POWELL M J D. Radial basis function approximations to polynomials [C] // Proc. of the Numerical Analysis Conference. 1987.