

DOI:10.11992/tis.201609020

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20170317.1937.006.html>

基于高维 k -近邻互信息的特征选择方法

周红标^{1,2,3}, 乔俊飞^{1,2}

(1. 北京工业大学 信息学部, 北京 100124; 2. 计算智能和智能系统北京市重点实验室, 北京 100124; 3. 淮阴工学院 自动化学院, 江苏 淮安 223003)

摘要:针对多元序列预测建模过程中特征选择问题,提出了一种基于数据驱动型高维 k -近邻互信息的特征选择方法。该方法首先将数据驱动型 k -近邻法扩展用于高维特征变量之间互信息的估计,然后采用前向累加策略给出全部特征最优排序,根据预设无关特征个数剔除无关特征,再利用后向交叉策略找出并剔除冗余特征,最终得到最优强相关特征子集。以 Friedman 数据、Housing 数据和实际污水处理出水总磷预测数据为例,采用多层感知器神经网络预测模型进行仿真实验,验证了所提方法的有效性。

关键词:特征选择;互信息; k -近邻;高维互信息;多层感知器

中图分类号:TP183 **文献标志码:**A **文章编号:**1673-4785(2017)05-0595-06

中文引用格式:周红标,乔俊飞.基于高维 k -近邻互信息的特征选择方法[J].智能系统学报,2017,12(5):595-600.

英文引用格式:ZHOU Hongbiao, QIAO Junfei. Feature selection method based on high dimensional k -nearest neighbors mutual information[J]. CAAI transactions on intelligent systems, 2017, 12(5): 595-600.

Feature selection method based on high dimensional k -nearest neighbors mutual information

ZHOU Hongbiao^{1,2,3}, QIAO Junfei^{1,2}

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China; 3. Faculty of Automation, Huaiyin Institute of Technology, Huai'an 223003, China)

Abstract: Feature selection plays an important role in the modeling and forecast of multivariate series. In this paper, we propose a feature selection method based on data-driven high-dimensional k -nearest neighbor mutual information. First, this method extends the k -nearest neighbor method to estimate the amount of mutual information among high-dimensional feature variables. Next, optimal sorting of all these features is achieved by adopting a forward accumulation strategy in which irrelevant features are eliminated according to a preset number. Then, redundant features are located and removed using a backward cross strategy. Lastly, this method obtains optimal subsets that feature a strong correlation. Using Friedman data, housing data, and actual effluent total-phosphorus forecast data from wastewater treatment plant as examples, we performed a simulation experiment by adopting a neural network forecast model with multilayer perception. The simulation results demonstrate the feasibility of the proposed method.

Keywords: feature selection; mutual information; k -nearest neighbor; high-dimensional mutual information; multilayer perceptron

通过特征选择实现数据降维,构建结构精简的辨识模型,能够有效避免输入特征太多造成“维数灾难”以及给学习模型带来“过拟合”等问题^[1-3]。

目前,特征选择的方法主要有偏最小二乘回归(partial least squares regression, PLSR)^[4]、灰色关联分析(grey relational analysis, GRA)^[5]和互信息(mutual information, MI)^[6-7]等。MI对样本的分布类型无特别要求,可有效捕捉特征间的非线性关系,特别适合多元序列特征选择问题。

收稿日期:2016-09-21. 网络出版日期:2017-03-17.

基金项目:国家自然科学基金重点项目(61533002).

通信作者:乔俊飞. E-mail:hyitzhb@163.com.

互信息是一种衡量两个随机变量之间相互依赖强弱程度的准则,其来源于信息论中熵的概念^[8]。采用互信息进行特征选择主要有两个难点:1)特征评价策略;2)互信息的准确估计。针对评价策略,Battiti等^[9]提出了MIFS(mutual information feature selection)方法,通过平衡相关特征和冗余特征的选择,取得了较好的效果,此后,Peng等^[10]、Fleuret^[11]、Yang等^[12]、Brown^[13]和韩等^[14]都分别提出了各自的评价策略。上述方法主要是采用一维互信息来衡量特征好坏,当存在冗余特征时,并不能保证得到最优特征子集。互信息的估计一般通过非参数方式^[15-16]求解概率密度函数来实现,主要有直方图法和核函数法。近几年发展的 k -近邻法(k -nearest neighbors, kNN)^[17-18]构建的数据驱动型特征选择框架,无需假设概率密度函数形式,避免了对概率密度函数的估计,其基础理论完备,非常适合具有非线性不规则分布特点的实际数据,但是高维 k -近邻互信息的估计存在一定困难。

因此,本文在 k -近邻互信息基础上将其扩展用于高维互信息的估计,然后采用前向累加后向交叉(forward accumulation and backward cross, FABC)的最优特征子集选择策略,构建kNN-FABC的特征选择方法,最后以Friedman数据、Housing数据和实际污水处理出水总磷预测数据为例进行仿真实验,并结合多层感知器(multilayer perceptron, MLP)预测模型来验证所提特征选择方法的有效性。

1 互信息理论

互信息是一种衡量两个随机变量之间相互依赖强弱程度的准则,其来源于信息论中熵的概念。两个随机变量之间的互信息越大,则两者之间的相关性就越强^[17]。

对随机变量 X 和 Y 来说,设其联合概率密度函数为 $\rho_{X,Y}(x,y)$,则其边缘概率密度函数为

$$\rho_X(x) = \int \rho_{X,Y}(x,y) dy \quad (1)$$

$$\rho_Y(y) = \int \rho_{X,Y}(x,y) dx \quad (2)$$

根据信息论的有关理论^[17],随机变量 X 和 Y 之间的互信息定义为

$$I(X;Y) = \iint_S \rho_{X,Y}(x,y) \log \frac{\rho_{X,Y}(x,y)}{\rho_X(x)\rho_Y(y)} dx dy \quad (3)$$

式中 S 为 X 和 Y 的定义域。

根据Shannon熵的定义^[17], X 、 Y 的熵和联合熵分别为

$$H(X) = - \int \rho_X(x) \log \rho_X(x) dx \quad (4)$$

$$H(Y) = - \int \rho_Y(y) \log \rho_Y(y) dy \quad (5)$$

$$H(X,Y) = - \iint_S \rho_{X,Y}(x,y) \log \rho_{X,Y}(x,y) dx dy \quad (6)$$

根据式(4)、(5)和(6),式(3)还可以表示为

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \quad (7)$$

2 高维 k -近邻互信息理论

2.1 高维 k -近邻互信息的估计

Kraskov等^[17-18]提出的 k -近邻法通过数据直接近似估计特征的互信息,避免了对概率密度分布的近似估计。其基本思想是,在由随机变量 X 和 Y 构成的空间中对于给定的 N 个样本首先找出它的 k 个近邻,再找出其他样本分别在 X 和 Y 方向到当前样本的距离小于当前样本到 k 个近邻距离的最大值的数目,通过统计数目从而进行互信息的估计。

现将其扩展到高维互信息的估计,高维互信息计算的具体思路如下^[18]:

给定 N 个样本 $\mathbf{z}^{(i)} = [(\mathbf{x}^{(i)})^T (\mathbf{y}^{(i)})^T]^T$ ($i=1, 2, \dots, N$),其中 $\mathbf{x}^{(i)} \in \mathbb{R}^d$ 且 $\mathbf{y}^{(i)} \in \mathbb{R}^p$ 。如果 \mathbf{z} 和 \mathbf{z}' 为数据集中两个不同的向量,则

$$\|\mathbf{z} - \mathbf{z}'\| = \max(\|\mathbf{x} - \mathbf{x}'\|, \|\mathbf{y} - \mathbf{y}'\|) \quad (8)$$

式中 $\|\cdot\|$ 表示取最大范数。

若 $\mathbf{z}^{(i)}$ 的 k -近邻为 $\mathbf{z}^{(k(i))} = [(\mathbf{x}^{(k(i))})^T (\mathbf{y}^{(k(i))})^T]^T$,则 $\mathbf{z}^{(i)}$ 与 $\mathbf{z}^{(k(i))}$ 之间的Euclidean距离为

$$\varepsilon^{(i)} = \|\mathbf{z}^{(i)} - \mathbf{z}^{(k(i))}\| = \max(\|\mathbf{x}^{(i)} - \mathbf{x}^{(k(i))}\|, \|\mathbf{y}^{(i)} - \mathbf{y}^{(k(i))}\|) \quad (9)$$

对于 $\mathbf{z}^{(i)}$ 中的分量 $\mathbf{x}^{(i)}$ 和 $\mathbf{y}^{(i)}$ 有

$$\varepsilon_x^{(i)} = \|\mathbf{x}^{(i)} - \mathbf{x}^{(k(i))}\| = \max_{s=1,2,\dots,d} |\mathbf{x}_s^{(i)} - \mathbf{x}_s^{(k(i))}| \quad (10)$$

$$\varepsilon_y^{(i)} = \|\mathbf{y}^{(i)} - \mathbf{y}^{(k(i))}\| = \max_{t=1,2,\dots,p} |\mathbf{y}_t^{(i)} - \mathbf{y}_t^{(k(i))}| \quad (11)$$

将与 $\mathbf{x}_s^{(i)}$ 和 $\mathbf{y}_t^{(i)}$ 之间距离小于 $\varepsilon^{(i)}$ 的样本点个数分别记为 $N_{x_s}^{(i)}$ 和 $N_{y_t}^{(i)}$ 。因此,变量 X 和 Y 之间互信息的估计量为

$$\hat{I}(X;Y) = \Psi(k) - \frac{d+p-1}{k} + (d+p-1)\Psi(N) - \frac{1}{N} \sum_{p=1}^N \left(\sum_{s=1}^d \Psi(N_{x_s}^{(i)}) + \sum_{t=1}^p \Psi(N_{y_t}^{(i)}) \right) \quad (12)$$

式中: $\Psi(\cdot)$ 为digamma函数, $\Psi(t+1) = \Psi(t) + \frac{1}{t}$, $\Psi(1) \approx -0.577\ 215\ 6$; k 一般取2~6。通过式(12)可以发现, $\hat{I}(X;Y)$ 不仅与 k 和 N 有关,而且与变量 X 和 Y

的维数有关,能够实现高维互信息的估计。

2.2 高维 k -近邻互信息特征选择流程

由于输入特征之间并非互相独立,存在非线性耦合,同时根据信息论知识,增加相关特征,能更好表征输出特征。基于高维 k -近邻互信息特征选择的基本思想为,首先利用前向累加搜索策略找出由强相关特征和弱相关特征组成的相关特征候选子集,然后采用后向交叉搜索策略剔除候选子集中的冗余特征,最终形成最优强相关特征子集。具体的流程如下。

1) 参数初始化,设置 k 值和无关特征个数。

2) 根据式(12)计算输入特征 X 的每一维分量与输出特征 Y 之间的一维互信息,提取出互信息最大值对应的特征,作为第1个相关特征。

3) 在2)得到的特征基础上,利用式(12)计算其他输入特征与输出特征 Y 之间的高维互信息,找出第2个相关特征。

4) 重复2)~3),直至所有特征处理完毕,得到全部特征变量的排序,根据预设无关特征个数,剔除无关特征,得到相关特征子集。

5) 根据相关特征子集,计算两两之间的互信息,找出互信息最大值对应的特征组。

6) 结合2)的一维互信息值,剔除冗余特征,最终得到最优强相关特征子集。

3 仿真实验

为了验证本文特征选择方法的有效性,本文分别采用标准 Friedman 函数数据和实际污水处理出水总磷预测数据进行特征选择的实验分析,并与 PLSR、MRMR 和 JMI 等方法进行比较,最后利用 MLP 预测模型^[19]进行误差分析。实验时, k 设置为 5,MLP 预测模型的隐含层节点取 20 个,迭代次数为 20 000 次,学习率 η 取为 0.001。由于神经网络建模受初值影响具有随机性,本文实验结果为运行 30 次取平均值和标准差。

3.1 Friedman 数据

Friedman 由式(13)描述:

$$Y = 10\sin(\pi x_1 x_2) + 20(x_3 - \frac{1}{2})^2 + 10x_4 + 5x_5 + 0 \times \sum_{i=6}^{10} x_i + x_{11} + x_{12} + \varepsilon \quad (13)$$

式中: $x_1 \sim x_{10}$ 服从 $[0,1]$ 的均匀分布, $x_{11} = 0.5x_1 + \varepsilon$, $x_{12} = 0.5x_2 + \varepsilon$ 。随机产生 500 个含高斯噪声数据作样本,且高斯噪声 ε 满足 $N(0,0.1)$ 。由式(13)可知, Y 只与 $x_1 \sim x_5$ 有关, $x_6 \sim x_{10}$ 是无关特征, x_{11} 和 x_{12} 属于冗余特征。

利用本文方法进行特征选择,第1步得到的互信

息如图1所示,可见仅依靠一维互信息会误选 x_{11} ,误剔除 x_3 和 x_5 ;第2步得到的高维互信息如图2所示,设定无关变量个数为 5,则剔除 $x_6 \sim x_{10}$ 得到最优相关特征子集为 $x_1 \sim x_5, x_{11}$ 和 x_{12} ;第3步得到的互信息如图3所示,确定 x_1, x_{11} 和 x_2, x_{12} 中存在冗余特征,再结合一维互信息值,可以判定 x_{11} 和 x_{12} 为冗余特征,剔除后得到最优强相关特征子集为 $x_1 \sim x_5$,完全与式(13)特征属性相吻合。

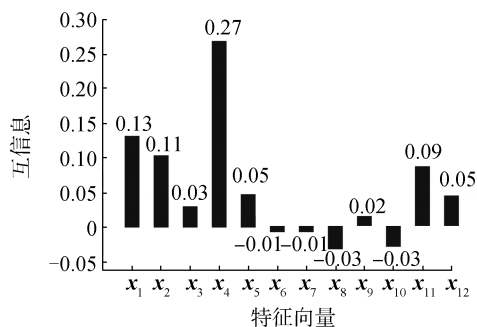
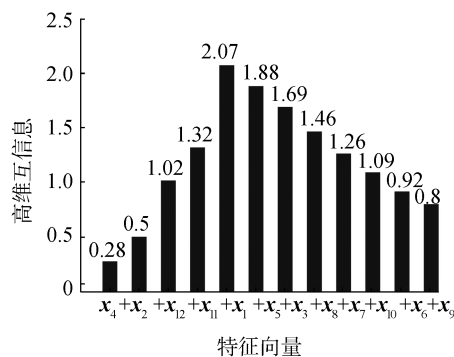


图1 一维互信息图

Fig.1 One-dimensional mutual information



注: $+x_2$ 表示求取 $I(x_4, x_2; Y)$; $+x_{12}$ 表示求取 $I(x_4, x_2, x_{12}; Y)$; 其余类似。

图2 FA 策略高维互信息图

Fig.2 High-dimensional mutual information under FA

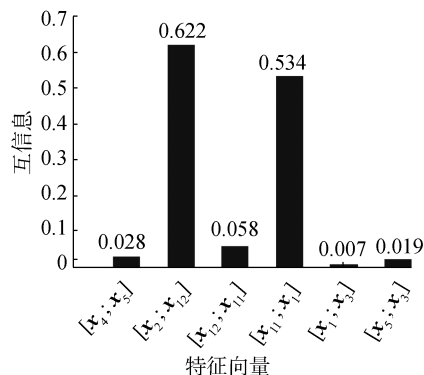


图3 BC 策略互信息图

Fig.3 Mutual information under BC

利用本文特征选择方法,并利用 MLP 网络对 Friedman 数据进行测试,随机产生 240 个带噪声数据

作训练样本,1 000 个不含噪声数据作测试样本,变量选择及网络测试结果如表 1 所示。

表 1 Friedman 数据特征选择及测试结果

Table 1 Feature selection and test results for Friedman

方法	变量选择结果	RMSE
ALL	$x_1 \sim x_{12}$	3.101±1.274
PLSR	$x_4, x_2, x_1, x_5, x_{11}, x_{12}$	3.543±1.440
MIFS ^[9]	x_4, x_2, x_1, x_5, x_6	2.002±0.095
MRMR ^[10]	x_4, x_2, x_1, x_5, x_6	2.002±0.095
CMIM ^[11]	x_4, x_2, x_1, x_5, x_6	2.002±0.095
JMI ^[12]	$x_4, x_2, x_1, x_5, x_{12}$	3.854±1.611
MMI ^[13]	$x_4, x_2, x_1, x_5, x_{12}$	3.854±1.611
STEP ^[14]	x_4, x_2, x_1, x_5, x_3	1.143±0.070
kNN-FABC	x_4, x_2, x_1, x_5, x_3	1.143±0.070

注: ALL 表示选取全部特征。

由表 1 可知, kNN-FABC 方法和文献[17]的 STEP 方法都成功筛选出全部仅有的 5 个相关特征, MRMR、CMIM 和 MIFS 误剔相关特征 x_3 、误选无关特征 x_6 , JMI 和 MMI 误剔相关特征 x_3 、误选冗余特征 x_{12} , PLSR 误选冗余特征 x_{11} 和 x_{12} 。同时, kNN-FABC 方法的 RMSE 均值和标准差都要远小于其他方法, 表明其极大地提高了网络泛化能力。

3.2 Housing 数据

Housing 数据由 13 个房屋属性的输入特征和 1 个房屋价格的输出特征组成, 共有 506 组样本。随机选择 338 组作训练样本, 剩余 168 组作测试样本。利用本文方法进行特征选择, 并与其他文献特征选择结果进行了比较, 特征选择及 MLP 网络测试结果如表 2 所示。

表 2 Housing 数据特征选择及测试结果

Table 2 Feature selection and test results for Housing

方法	特征选择结果	RMSE
ALL	$x_1 \sim x_{13}$	4.350±0.515
PLSR	x_4, x_5, x_6, x_{11}	5.173±0.643
MIFS ^[9]	x_{13}, x_{11}, x_4, x_2	5.125±0.482
MRMR ^[10]	$x_{13}, x_{11}, x_6, x_{10}$	3.880±0.476
CMIM ^[11]	$x_{13}, x_6, x_{11}, x_{10}$	3.880±0.476
JMI ^[12]	x_{13}, x_6, x_{11}, x_7	4.187±0.598
MMI ^[13]	x_{13}, x_6, x_{11}, x_7	4.187±0.598
STEP ^[14]	x_{13}, x_6, x_8, x_1	4.187±0.488
kNN-FABC	x_{13}, x_6, x_5, x_{11}	3.710±0.406

注: ALL 表示选取全部特征。

由表 2 可知, 在归一化方法、MLP 模型参数选取一致的情况下, kNN-FABC 的 RMSE 均值和标准差均表明了采用本文方法所选特征建立的预测模型精度更高。

3.3 污水处理出水总磷预测数据

近年来, 总磷(total phosphorus, TP)污染问题开始凸显。研究表明, 水体富营养化程度与总磷等指标密切相关。因此加强对 TP 的检测有助于解决水体富营养化问题^[20]。目前, 污水处理厂普遍采用的生化方法成本高、耗时长, 在线检测仪表对检测环境要求高并且维护成本昂贵, 而基于数据驱动的软测量技术能够在线、快速、准确预测, 受到研究者青睐。但是污水处理生化反应过程中可测变量较多, 在建立 TP 软测量模型时, 若选取全部可用特征, 则会产生冗余, 降低网络泛化能力, 同时也大幅增加检测成本, 故对 TP 预测模型进行输入特征选择具有重要意义。污水处理生化反应过程中可测变量及其意义如表 3 所示。本文采用 kNN-FABC 方法实现对 TP 预测特征变量的选择。

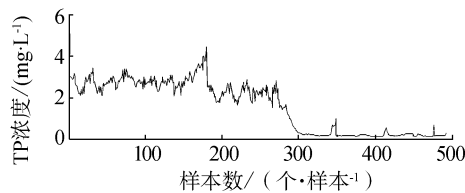
表 3 可测变量的意义

Table 3 Significance of measurable variables

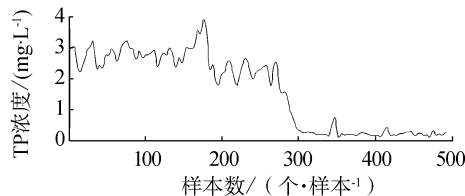
标记	参数	意义	单位	取样点
x_1	T	水温	℃	厌氧池
x_2	ORP ₁	氧化还原电位	mV	厌氧池
x_3	DO ₁	溶解氧浓度	mg/L	好氧前端
x_4	DO ₂	溶解氧浓度	mg/L	好氧末端
x_5	TSS	总固体悬浮物浓度	g/L	好氧末端
x_6	PH	酸碱度	—	出水端
x_7	ORP ₂	氧化还原电位	mV	出水端
x_8	NH ₄ -N	铵态氮浓度	mg/L	出水端
x_9	NO ₃ -N	硝态氮浓度	mg/L	出水端

从北京市某污水处理厂获取了 2015 年 6~8 月共 492 组数据, 首先剔除明显异常值, 然后采用小波包(sym8 小波、2 层分解、软阈值方式)进行降噪处理, 处理效果见图 4。

从原始数据集中每隔 3 个样本选取 1 个样本共 123 个组成测试集, 其余 369 个作为训练集。TP 数据特征选择及 MLP 网络测试结果如表 4 所示。相比 PLSR, kNN-FABC 获得了较少的特征、较低的 RMSE 及其标准差, 寻找到了最有效的特征子集, 构建了更稳定的软测量模型, 验证了 kNN-FABC 方法在特征选择上的有效性。



(a)原始 TP 数据



(b)小波包去噪 TP 数据

图 4 出水 TP 数据小波包去噪

Fig.4 Wavelet packet denoising for TP

表 4 TP 数据特征选择及测试结果

Table 4 Feature selection and test results for TP

方法	特征选择结果	RMSE
ALL	$x_1 \sim x_9$	0.428 ± 0.077
ERR	x_3, x_4, x_8, x_9	0.519 ± 0.071
PLSR	x_3, x_5, x_2, x_1, x_6	0.464 ± 0.033
kNN-FABC	x_7, x_1, x_2, x_6	0.159 ± 0.004

注: ERR 为随机选取特征,作对比。

图 5 和图 6 分别给出了 MLP 模型对 TP 数据预测结果和预测误差。可见,通过 kNN-FABC 方法选出的特征能够很好地表征原始数据,建模误差主要集中在 $-0.3 \sim 0.3$ mg/L 之间,能够较好地满足污水处理厂对总磷检测精度的要求。

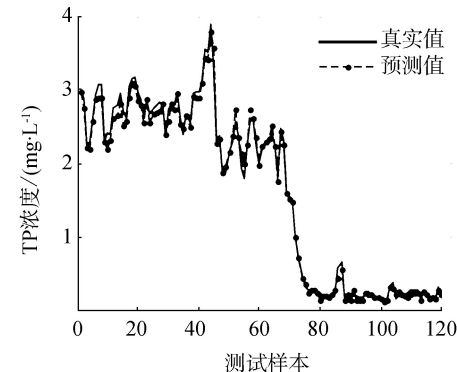


图 5 TP 数据 MLP 预测结果

Fig.5 Predicted output using MLP for TP

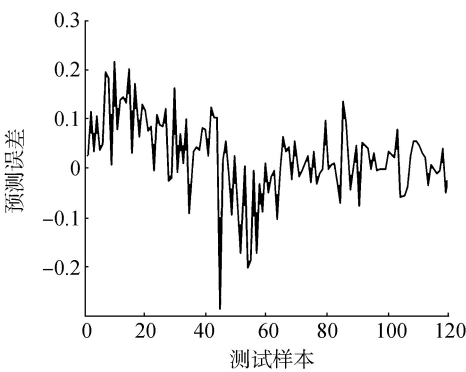


图 6 TP 数据 MLP 预测误差

Fig.6 Predicted error using MLP for TP

4 结束语

通过将数据驱动型 k -近邻方法扩展用于多维特征之间相关性计算,结合前向累加策略,能够较为准确地获得全部特征的最优排序,进而剔除无关特征,再结合后向交叉策略,能够定位并删除冗余特征,最终得到最优强相关特征子集。Friedman、Housing 和实际污水处理出水 TP 预测实验证实了本文方法的有效性。下一步的研究工作是实现无关特征的自动判定,避免人为设置带来的无关特征误剔的可能。

参考文献:

[1] OSELEDETS I V, TYRTYSHNIKOV E E. Breaking the curse of dimensionality, or how to use SVD in many dimensions[J]. SIAM journal on scientific computing, 2009, 31(5): 3744-3759.

[2] GHAMISI P, BENEDIKTSSON J A. Feature selection based on hybridization of genetic algorithm and particle swarm optimization [J]. IEEE geoscience and remote sensing letters, 2015, 12(2): 309-313.

[3] RAUBER T W, ASSIS BOLDT de F, VAREJÃO F M. Heterogeneous feature models and feature selection applied to bearing fault diagnosis[J]. IEEE transactions on industrial electronics, 2015, 62(1): 637-646.

[4] WOLD S, SJÖSTRÖM M, ERIKSSON L. PLS-regression: a basic tool of chemometrics[J]. Chemometrics and intelligent laboratory systems, 2001, 58(2): 109-130.

[5] SONG Q, SHEPPERD M. Predicting software project effort: a grey relational analysis based method[J]. Expert systems with applications, 2011, 38(6): 7302-7316.

[6] FENG J, JIAO L, LIU F, et al. Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low redundancy [J]. IEEE transactions on geoscience and remote sensing, 2015, 53(5): 2956-2969.

[7] BENNASAR M, HICKS Y, SETCHI R. Feature selection using joint mutual information maximisation [J]. Expert

- systems with applications, 2015, 42(22): 8520–8532.
- [8] SHANNON C E. A mathematical theory of communication [J]. ACM sigmobile mobile computing and communications review, 2001, 5(1): 3–55.
- [9] BATTITI R. Using mutual information for selecting features in supervised neural net learning[J]. IEEE transactions on neural networks, 1994, 5(4): 537–550.
- [10] PENG H, LONG F, DING C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE transactions on pattern analysis and machine intelligence, 2005, 27(8): 1226–1238.
- [11] FLEURET F. Fast binary feature selection with conditional mutual information[J]. Journal of machine learning research, 2004, 5: 1531–1555.
- [12] YANG H H, MOODY J E. Data visualization and feature selection; new algorithms for nongaussian data[C]//Advances in Neural Information Processing Systems. Cambridge, Britain, 1999: 687–693.
- [13] BROWN G. A new perspective for information theoretic feature selection[C]//Proceedings of the 12th International Conference on Artificial Intelligence and Statistics. Florida, USA, 2009: 49–56.
- [14] 韩敏, 刘晓欣. 基于互信息的分步式输入变量选择多元序列预测研究[J]. 自动化学报, 2012, 38(6): 999–1005.
HAN Min, LIU Xiaoxin. Stepwise input variable selection based on mutual information for multivariate forecasting [J]. ACTA automatica sinica, 2012, 38(6): 999–1005.
- [15] BEIRLANT J, DUDEWICZ E J, GYÖRFI L, et al. Nonparametric entropy estimation; an overview[J]. International journal of mathematical and statistical sciences, 1997, 6(1): 17–39.
- [16] MOON Y I, RAJAGOPALAN B, LALL U. Estimation of mutual information using kernel density estimators [J]. Physical review E, 1995, 52(3): 2318–2321.
- [17] KRASKOV A, STÖGBAUER H, GRASSBERGER P. Estimating mutual information[J]. Physical review E, 2004, 69(6): 1–16.
- [18] STÖGBAUER H, KRASKOV A, ASTAKHOV S A, et al. Least-pendent-component analysis based on mutual information[J]. Physical review E, 2004, 70(6): 1–17.
- [19] 霍军周, 王亚杰, 欧阳湘宇, 等. 基于 BP 神经网络的 TBM 主轴承载荷谱预测[J]. 哈尔滨工程大学学报, 2015, 36(7): 965–969.
HUO Junzhou, WANG Yajie, OUYANG Xiangyu, et al. Load spectrum prediction of the main drive bearing of a tunnel boring machine based on BP neural networks[J]. Journal of Harbin Engineering University, 2015, 36(7): 965–969.
- [20] 乔俊飞, 周红标. 基于自组织模糊神经网络的出水总磷预测[J]. 控制理论与应用, 2017, 34(2): 224–232.
QIAO Junfei, ZHOU Hongbiao. Prediction of effluent total phosphorus based on self-organizing fuzzy neural network[J]. Control theory and applications, 2017, 34(2): 224–232.

作者简介:



周红标,男,1980年生,讲师,博士研究生,主要研究方向为神经网络分析与设计。发表论文十余篇,其中被EI检索6篇。



乔俊飞,男,1968年生,教授,博士生导师,主要研究方向为污水处理过程智能优化控制。获教育部科技进步奖一等奖和北京市科学进步奖三等奖各1项,发表论文近100篇,其中被SC

I收录18篇,EI收录60篇,获发明专利20项。