

基于自适应核参数优化的小波核相关向量机算法

高明哲, 许爱强, 张 伟

(海军航空工程学院 科研部, 山东 烟台 264001)

摘 要: 传统相关向量机算法在处理大规模数据集时训练速度较慢, 并且高斯径向核无法完备表示特征空间。为此, 基于自适应核参数优化, 提出一种小波核相关向量机算法。以小波核作为基函数, 在训练中, 采取增量学习流程实现各个小波核参数的快速自适应优化。将提出算法应用于混沌时间序列预测及 UCI 数据集分类实验, 结果表明, 自适应参数优化小波相关向量机算法在预测精度、训练速度上均优于传统相关向量机算法。

关键词: 相关向量机; 小波核函数; 自适应参数优化; 增量学习; 稀疏度先验

中文引用格式: 高明哲, 许爱强, 张 伟. 基于自适应核参数优化的小波核相关向量机算法[J]. 计算机工程, 2017, 43(9): 245-249.

英文引用格式: GAO Mingzhe, XU Aiqiang, ZHANG Wei. Wavelet Kernel Relevance Vector Machine Algorithm Based on Adaptive Kernel Parameter Optimization[J]. Computer Engineering, 2017, 43(9): 245-249.

Wavelet Kernel Relevance Vector Machine Algorithm Based on Adaptive Kernel Parameter Optimization

GAO Mingzhe, XU Aiqiang, ZHANG Wei

(Department of Scientific Research, Naval Aeronautical and Astronautical University, Yantai, Shandong 264001, China)

[Abstract] The traditional Relevance Vector Machine(RVM) algorithm is slow to train on large scale datasets, and the Gauss radial kernel cannot express the feature space completely. So based on adaptive kernel parameter optimization, this paper proposes a wavelet kernel relevance vector machine algorithm. Regarding the wavelet kernel as the basis function, incremental learning process is used to realize the fast adaptive optimization of each wavelet kernel parameter in training. The proposed algorithm is used on prediction of chaotic time series and classification of UCI data sets. Simulation results show that the adaptive parameter optimization wavelet correlation vector machine algorithm is superior to the traditional correlation vector machine algorithm in forecasting accuracy and training speed.

[Key words] Relevance Vector Machine(RVM); wavelet kernel function; adaptive parameter optimization; incremental learning; sparsity prior

DOI: 10.3969/j.issn.1000-3428.2017.09.043

0 概述

相关向量机(Relevance Vector Machine, RVM)是一种基于稀疏贝叶斯推理的机器学习算法^[1], 与支持向量机(Support Vector Machine, SVM)相比, RVM在保留SVM较高精度的基础上, 还具有参数设置简单、稀疏度更高、概率式输出、基函数不受Mercer条件限制等优点, 目前已在图像分类^[2]、故障检测^[3]及各类预测^[4-6]等应用领域取得了较好的效果。然而传统RVM中广泛采用的高斯径向基函数(Radial Basis Function, RBF)并不具备完备表示解空间的能力^[7], 在样本数据结构比较复杂的情况下, 单一RBF核模型的精度并不理想^[8]。同时, 传统

RVM中核参数的选取大多采用优化算法与交叉验证相结合的方法, 计算量比较大, 尤其是在训练样本量较大的情况下, 传统RVM的训练时间远长于SVM^[9]。

多分辨分析理论催生了小波核(Wavelet Kernel)的诞生。文献[7]指出, 在平方可积空间中, 小波核是比RBF核更完备的一类核函数, 能够在特征空间中更好地描述训练数据的特性。文献[10]将小波核RVM应用于图像压缩, 并采用一维直接搜索法来确定小波核的伸缩因子。文献[11]对小波核RVM进行了深入研究, 指出小波核在回归问题中具有冗余度小(正交或近似正交)、分辨率高等优点, 其在高频数据回归中相比其他核函数具有更好的性

能。然而,RVM 中常规的超参数约束方法并不能很好地控制模型的复杂度,可能会导致小波核 RVM 在低频数据部分产生过拟合,影响其总体应用效果。

自适应核学习相关向量机(Adaptive kernel learning Relevance Vector Machine, ARVM)^[12]是在增量学习方法^[13]和平滑先验控制^[11]基础上提出的一种新型稀疏贝叶斯学习方法。相比传统 RVM,其优点在于:采用了平滑稀疏度先验直接限制模型中有效参数的个数,能够更好地防止过拟合;增量学习流程有效减少了训练过程的计算复杂度,大幅减少了大规模数据训练时的时间;基函数的参数设置更加灵活,各相关向量对应的基函数可选取不同的核参数值。

基于此,本文将小波核函数与 ARVM 算法相结合,构建一种基于自适应核参数优化的小波核相关向量机(Wavelet kernel with Adaptive kernel parameter optimization Relevance Vector Machine, WARVM)算法。在保留小波核分辨率高和 ARVM 核参数设置灵活的优势下,利用超参数的先验约束防止小波核可能出现的过拟合,并通过增量学习流程减少大规模数据训练时的训练时间。最后采用混沌时间序列预测和 UCI 标准数据集分类实验将新算法与传统 RVM 算法的性能进行对比。

1 小波核函数分析

RVM 是一种稀疏核方法,通过预先选择的基核函数将数据从原始输入空间映射到一个高维特征空间,从而实现数据在高维空间的线性化。在平方可积空间中,小波分析能够有效地对数据进行多分辨率的细化分析,基于小波分析理论的小波核自提出以来广泛应用于神经网络^[14]和支持向量机^[15]中,取得了较好的效果。

小波核是一种平移不变核,设置为一个小波母函数,则可构造 d 维空间上的小波核函数^[15]:

$$K(\mathbf{x}, \mathbf{x}'; \theta) = \sum_{i=1}^d \varphi\left(\frac{x_i - x'_i}{\theta}\right) \quad (1)$$

其中, θ 为小波核的尺度因子。常用的小波核函数主要有:

1) Shannon 小波核:

$$K(x, x') = \frac{1}{d} \prod_{i=1}^d \left(2 \operatorname{sinc}\left(2\left(\frac{x_i - x'_i}{\theta}\right)\right) - \operatorname{sinc}\left(\frac{x_i - x'_i}{\theta}\right) \right) \quad (2)$$

2) Morlet 小波核:

$$K(x, x') = \prod_{i=1}^d \cos\left(1.75 \frac{x_i - x'_i}{\theta}\right) \exp\left(-\frac{\|x_i - x'_i\|^2}{2\theta^2}\right) \quad (3)$$

$$K(x, x') = \prod_{i=1}^d \left(\exp\left(-\frac{\|x_i - x'_i\|^2}{2\theta^2}\right) - \frac{1}{2} \exp\left(-\frac{\|x_i - x'_i\|^2}{8\theta^2}\right) \right) \quad (4)$$

除此之外,还有 Mexican hat 小波核^[10]、Symmle 小波核^[11]等。图 1 给出了 Symmle 小波核和 lspline 核用于传统 RVM 算法时,对带噪声的 Sinc 函数和 Bumps 数据的逼近能力。

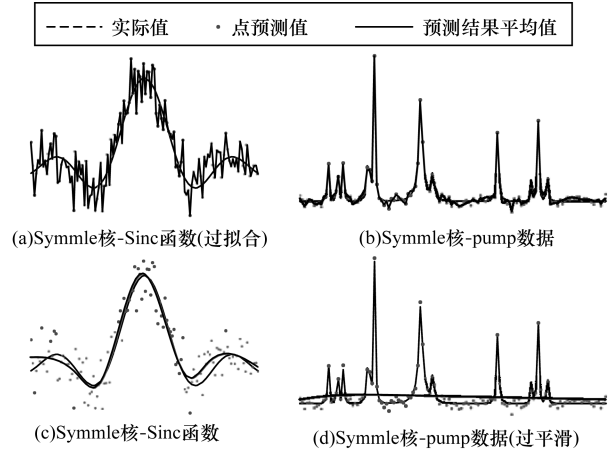


图 1 Symmle 小波核与 lspline 核性能对比

由图 1 可以看出,在传统 RVM 中, Symmle 小波核对于跳变较快的高频数据具有很好的逼近能力,然而在 Sinc 函数中却产生了明显的过拟合现象;而 lspline 核虽然能够较好地还原 Sinc 函数,却无法描述 Bumps 数据。

因此,虽然小波核函数具有很强的分析和表述能力,但由于传统 RVM 中的先验控制方法并不能很好地约束模型的复杂度,致使采用小波核的 RVM 很容易对较为平滑的训练数据产生过拟合。

2 WARVM 算法

2.1 基本模型

为解决小波核容易过拟合的问题,同时提高 RVM 模型的训练效率与灵活性,提出一种基于自适应核参数优化的小波核相关向量机(WARVM)。WARVM 的基本模型与传统 RVM 相同,其形式为:

$$y(\mathbf{x}; \boldsymbol{\omega}) = \sum_{i=1}^N \omega_i K(\mathbf{x}; \theta_i) \quad (5)$$

其中, $K(\mathbf{x}; \theta_i)$ 为小波核函数; θ_i 为对应的尺度因子; $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_N)^T$ 为权值。学习的目标是求得式(5)中各参数的值。由于训练后 $\boldsymbol{\omega}$ 大部分为 0,因此该模型最终是由少数权值为非 0 的相关向量(Relevance Vector, RV)组成的稀疏模型。

给定训练样本集 $\{\mathbf{x}_n, t_n\}_{n=1}^N$, \mathbf{x}_n 为输入向量, t_n 为其对应的目标值。在回归问题中,目标值 $t_n = y(\mathbf{x}_n) + \varepsilon_n$, 其中, ε_n 为服从高斯分布 $N(0, \beta_n^{-1})$ 的噪声,则目标值 $\mathbf{t} = (t_1, t_2, \dots, t_N)^T$ 的似然函数可写为:

$$p(\mathbf{t} | \boldsymbol{\omega}, \boldsymbol{\beta}) = N(\mathbf{t} | \boldsymbol{\Phi} \boldsymbol{\omega}, \mathbf{B}^{-1}) \quad (6)$$

其中, $\mathbf{B} = \operatorname{diag}(\boldsymbol{\beta})$; $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \dots, \boldsymbol{\varphi}_N]^T$ 为 $N \times (N+1)$ 维基函数矩阵,列向量 $\boldsymbol{\varphi}_i = [1, K(\mathbf{x}_i; \theta_1), K$

$(\mathbf{x}_i; \boldsymbol{\theta}_2), \dots, K(\mathbf{x}_i; \boldsymbol{\theta}_N)]^T$ 。

在分类问题中, 对应的目标值 t_n 为类别标识, 在二分类问题中, $t_n = \{0, 1\}$ 。则根据伯努利分布, 目标值 $\mathbf{t} = (t_1, t_2, \dots, t_N)^T$ 的似然函数可写为:

$$p(\mathbf{t} | \boldsymbol{\omega}) = \prod_{i=1}^N \sigma\{y(\mathbf{x}_i; \boldsymbol{\omega})\}^{t_i} [1 - \sigma\{y(\mathbf{x}_i; \boldsymbol{\omega})\}]^{1-t_i} \quad (7)$$

其中, $\sigma\{y(\mathbf{x}_i; \boldsymbol{\omega})\} = [1 + e^{-y(\mathbf{x}_i; \boldsymbol{\omega})}]^{-1}$ 。

2.2 自适应核参数优化

针对权值 $\boldsymbol{\omega}$, 赋予其对应超参数 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$ 作为先验来控制模型的稀疏度, 两者满足高斯分布:

$$p(\boldsymbol{\omega} | \boldsymbol{\alpha}) = \prod_{i=1}^N N(\omega_i | 0, \alpha_i^{-1}) \quad (8)$$

根据贝叶斯定理可得权值的后验概率满足高斯分布:

$$p(\boldsymbol{\omega} | \mathbf{t}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{p(\mathbf{t} | \boldsymbol{\omega}, \boldsymbol{\beta}) p(\boldsymbol{\omega} | \boldsymbol{\alpha})}{p(\mathbf{t} | \boldsymbol{\alpha}, \boldsymbol{\beta})} = N(\boldsymbol{\omega} | \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (9)$$

其中, $\boldsymbol{\mu} = \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{B} \mathbf{t}$; $\boldsymbol{\Sigma} = (\boldsymbol{\Phi}^T \mathbf{B} \boldsymbol{\Phi} + \mathbf{A})^{-1}$; $\mathbf{A} = \text{diag}(\boldsymbol{\alpha})$ 。

通过边缘化权值 $\boldsymbol{\omega}$, 训练问题简化为求解后验概率 $p(\boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{t}) \propto p(\mathbf{t} | \boldsymbol{\alpha}, \boldsymbol{\beta}) p(\boldsymbol{\alpha} | \boldsymbol{\beta}) p(\boldsymbol{\beta})$ 关于 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的最大值问题。超参数先验 $p(\mathbf{t} | \boldsymbol{\alpha}, \boldsymbol{\beta})$ 的对数为:

$$L = \log_a p(\mathbf{t} | \boldsymbol{\alpha}, \boldsymbol{\beta}) = -1/2 [N \log_a 2\pi + \log_a |\mathbf{C}| + \mathbf{t}^T \mathbf{C}^{-1} \mathbf{t}] \quad (10)$$

其中, $\mathbf{C} = \mathbf{B} + \boldsymbol{\Phi} \mathbf{A}^{-1} \boldsymbol{\Phi}^T$ 。

不同于传统 RVM 采用 Gamma 先验分布来约束 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ ^[1], 为更好地防止过拟合, 对 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 引入稀疏度先验来直接控制模型有效参数的个数^[11]:

$$p(\boldsymbol{\alpha} | \boldsymbol{\beta}) = \propto \exp(-c \text{trace}(\mathbf{S})) \quad (11)$$

其中, $\mathbf{S} = \boldsymbol{\Phi} \boldsymbol{\Sigma} \boldsymbol{\Phi}^T \mathbf{B}$ 为平滑矩阵; 矩阵 \mathbf{S} 的迹 $\text{trace}(\mathbf{S})$ 称为 \mathbf{S} 的自由度, 用以表示模型参数的有效个数; c 为稀疏度参数, 根据贝叶斯信息准则 (Bayesian Information Criterion, BIC), 在此可以取 $c = \log_a(N)/2$ 。

结合式 (10)、式 (11), 可得对数似然函数为:

$$L^S = L - c(M - \sum_{i=1}^M \alpha_i \Sigma_{ii}) \quad (12)$$

其中, $\Sigma_{ii} = 1/(\alpha_i + s_i)$; $s_i = \boldsymbol{\varphi}_i^T \mathbf{C}_{-i}^{-1} \boldsymbol{\varphi}_i$; $\mathbf{C}_{-i} = \mathbf{B} + \sum_{j \neq i} \alpha_j \boldsymbol{\varphi}_j \boldsymbol{\varphi}_j^T$; M 为当前模型中参与训练的基函数个数。

训练的目标是求得使式 (12) 最大化的超参数最优值 $\boldsymbol{\alpha}_{\text{MP}}$ 和 $\boldsymbol{\beta}_{\text{MP}}$ 以及小波尺度因子 $\boldsymbol{\theta}$ 。将式 (12) 分解得到^[9]:

$$l_i^S = l_i - c \left(1 - \frac{\alpha_i}{\alpha_i + s_i} \right) \quad (13)$$

其中, l_i^S 为模型中只包含 α_i 所在基函数时的后验概

率大小, 其数值只与单个超参数 α_i 和对应的小波核尺度因子 θ_i 有关。

WARVM 通过增量学习方法^[13] 求得各参数的最优估计值。在训练初始假设所有 $\alpha_i \rightarrow \infty$, 之后采用迭代方式向模型中逐个添加满足要求的基函数, 并更新模型中所有参数的值, 随后剔除更新后模型中不满足要求的基函数, 直至收敛。每一步迭代的运算量取决于模型当前包含的基函数数量。迭代中超参数 α_i 及 θ_i 的更新值由 l_i^S 对两者偏导数为 0 的解得到, θ_i 采用拟牛顿法来计算最优值。同样, 超参数向量 $\boldsymbol{\beta}$ 的值则取后验概率 L^S 对 $\boldsymbol{\beta}$ 偏导数为 0 的数值解来更新。

2.3 预测结果

模型训练完成后, 在回归问题中, 对于新的输入 \mathbf{x}_* , 其输出 t_* 满足高斯分布:

$$p(t_* | \mathbf{t}, \boldsymbol{\alpha}_{\text{MP}}, \boldsymbol{\beta}_{\text{MP}}, \boldsymbol{\theta}) = N(t_* | y_*, \sigma_*^2) \quad (14)$$

其中:

$$y_* = \boldsymbol{\mu}^T K(\mathbf{x}_*; \boldsymbol{\theta}) \quad (15)$$

$$\sigma_*^2 = \boldsymbol{\beta}_{\text{MP}} + K(\mathbf{x}_*; \boldsymbol{\theta})^T \boldsymbol{\Sigma} K(\mathbf{x}_*; \boldsymbol{\theta}) \quad (16)$$

均值 y_* 可作为预测的最大后验估计值, 方差 σ_*^2 可用来计算预测的置信区间。在二分类问题中, \mathbf{x}_* 对应的类别 t_* 服从分布:

$$p(t_* | \mathbf{t}, \boldsymbol{\alpha}_{\text{MP}}, \boldsymbol{\beta}_{\text{MP}}, \boldsymbol{\theta}) = \sigma\{\boldsymbol{\mu}^T K(\mathbf{x}_*; \boldsymbol{\theta})\} \quad (17)$$

3 仿真实验与分析

3.1 实验设计

将本文提出的 WARVM 算法用于 Mackey-Glass 混沌时间序列预测和 UCI 数据集分类的仿真实验, 并将实验结果与 ARVM^[12] 和传统 RVM^[1] 的实验结果进行对比。ARVM 与传统 RVM 均采用 RBF 核函数, 其中, 传统 RVM 采用 5 折交叉验证与粒子群优化 (Particle Swarm Optimization, PSO) 算法来确定核参数值。仿真实验均在 PC 机 (Intel Xeon E3-1230v3, 16 GB 内存) 上的 Matlab R2015a 中运行。

3.2 混沌时间序列预测实验

Mackey-Glass 混沌时间序列是评价预测模型性能的经典基准数据。实验采用文献 [16] 中非平稳 Mackey-Glass 时间序列的前 1 100 个数据点作为实验样本。设置嵌入维数为 10, 对实验样本进行相空间重构^[17], 将重构后的前 50 个样本作为训练样本, 其余作为测试样本。WARVM, ARVM 和传统 RVM 的预测结果如图 2 所示。

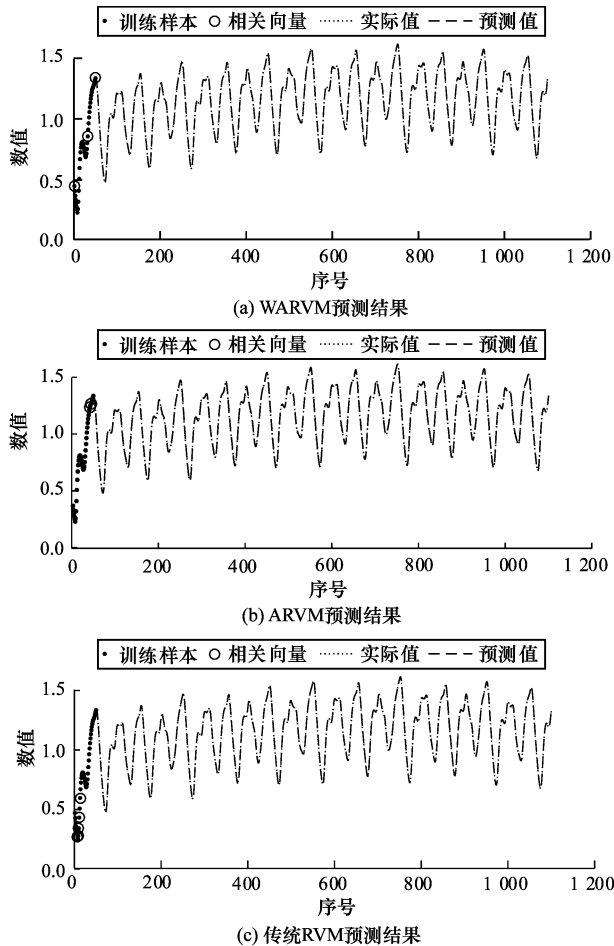


图2 Mackey-Glass 时间序列预测结果对比

表1给出了在Mackey-Glass混沌时间序列预测中,WARVM,ARVM以及传统RVM的各项性能对比。可以看出,WARVM拥有最高的精度,ARVM次之,均高于传统RVM,说明小波核相比高斯核具有更强的函数逼近能力,对于目标函数的预测能力更强;WARVM的训练时间略多于ARVM,但明显少于传统RVM,这是由于传统RVM的核参数自动求解需要所有训练数据多次交叉验证和迭代搜索过

程,而WARVM和ARVM只需对当前模型包含的少量样本点的参数进行更新;稀疏度先验控制使得WARVM和ARVM的相关向量个数少于传统RVM。因此在回归问题上,WARVM具有最好的综合性能。

表1 WARVM,ARVM及传统RVM的回归性能对比

算法	均方根误差	训练时间/ms	相关向量个数
本文算法	0.005 6	56.21	3
ARVM算法 ^[12]	0.006 1	50.57	3
RVM算法 ^[1]	0.007 7	314.74	5

3.3 UCI数据集分类实验

UCI数据库是加州大学欧文分校(University of California Irvine)提出的用于机器学习的数据库。本节实验选取UCI数据库中常用的4个二分类数据集来测试算法的分类性能,如表2所示。

表2 UCI数据集及其属性

数据集	属性数	样本总数	训练样本数		测试样本数	
			类别1	类别2	类别1	类别2
vote	16	455	50	50	118	217
ionosphere	34	351	50	50	175	76
pima	8	768	50	50	218	450
german	24	1 000	50	50	250	650

表3给出了WARVM,ARVM与传统RVM在各数据集分类实验中的性能对比。从中可以看出:除了第1组中WRVM和传统RVM的分类准确率相同,略低于ARVM外,其他3组中WARVM均拥有最高的分类准确率,并且在这3组中,ARVM的分类准确率高于传统RVM;4组实验中,WARVM的训练时间略多于ARVM,但大幅少于传统RVM,并且WARVM的模型相比其他两者更稀疏。因此在分类问题上,WRVM具有最佳的综合性能。

表3 UCI数据集及其属性

数据集	分类正确率/%			训练时间/s			相关向量个数		
	本文算法	ARVM算法	RVM算法	本文算法	ARVM算法	RVM算法	本文算法	ARVM算法	RVM算法
vote	92.836	93.432	92.836	0.826	0.779	2.531	9	9	14
ionosphere	87.251	86.056	84.462	0.920	0.852	2.805	8	9	12
pima	75.599	74.551	73.802	0.765	0.737	2.372	5	6	9
german	70.556	69.889	66.444	0.861	0.853	2.664	7	7	11

4 结束语

本文针对传统相关向量机在大规模数据集上训练速度较慢以及RBF核无法完备表示特征空间的问题,提出一种基于自适应核参数优化的小波核相

关向量机算法。相比已有方法,所提方法采用了增量学习方法进行训练,减少了训练时的计算复杂度,且各参数的求解无需交叉验证,相比传统RVM有效缩短了训练时间。小波核的高分辨率与模型的稀疏

度控制相结合,既提高了模型的数据分析能力,又防止了可能的过拟合,使得算法在提高稀疏度的同时预测精度也有所提高。同时,训练后模型中相关向量对应的小波尺度因子为自适应选择的最优值,各尺度因子可选取不同的值,使得模型能够更好地描述样本数据的结构特点。

参考文献

- [1] TIPPING M E. Sparse Bayesian Learning and the Relevance Vector Machine [J]. Journal of Machine Learning Research, 2001, 1(3): 211-244.
- [2] 赵春晖,齐滨,张熾. 基于改进型相关向量机的高光谱图像分类[J]. 光学学报, 2012, 32(8): 256-261.
- [3] WIDODO A, KIM E Y, SON J D, et al. Fault Diagnosis of Low Speed Bearing Based on Relevance Vector Machine and Support Vector Machine [J]. Expert Systems with Applications, 2009, 36(3): 7252-7261.
- [4] CAESARENDRA W, WIDODO A, HONG T P, et al. Machine Degradation Prognostic Based on RVM and ARMA/GARCH Model for Bearing Fault Simulated Data [C]//Proceedings of Prognostics and Health Management Conference. Washington D. C., USA: IEEE Press, 2010: 1-6.
- [5] ZHANG Lei. A Multivariate Relevance Vector Machine Based Algorithm for On-line Fault Prognostic Application with Multiple Fault Features [C]//Proceedings of International Conference on Intelligent Computation Technology & Automation. Washington D. C., USA: IEEE Press, 2012: 26-32.
- [6] 胡昌华,王兆强,周志杰,等. 一种 RVM 模糊模型辨识方法及在故障预报中的应用[J]. 自动化学报, 2011, 37(4): 503-512.
- [7] 丁二锐,曾平,丁阳,等. 一种新的回归型约简多分辨率相关向量机[J]. 控制与决策, 2008, 23(1): 65-69.
- [8] 汪洪桥,孙富春,蔡艳宁,等. 多核学习方法[J]. 自动化学报, 2010, 36(8): 1037-1050.
- [9] FAUL A, AVENUSE J J T. Fast Marginal Likelihood Maximisation for Sparse Bayesian Models [C]//Proceedings of the 9th International Workshop on Artificial Intelligence & Statistics. Berlin, Germany: Springer, 2003: 3-6.
- [10] TOLAMBIYA A, KALRA P K. Relevance Vector Machine with Adaptive Wavelet Kernels for Efficient Image Coding [J]. Neurocomputing, 2010, 73(7-9): 1417-1424.
- [11] SCHMOLCK A, EVERSON R. Smooth Relevance Vector Machine: A Smoothness Prior Extension of the RVM[J]. Machine Learning, 2007, 68(2): 107-135.
- [12] TZIKAS D G, LIKAS A C, GALATSANOS N P. Sparse Bayesian Modeling with Adaptive Kernel Learning[J]. IEEE Transactions on Neural Networks, 2009, 20(6): 926-937.
- [13] TZIKAS D, LIKAS A, GALATSANOS N. Incremental Relevance Vector Machine with Kernel Learning [C]//Proceedings of Conference on Artificial Intelligence. Berlin, Germany: Springer, 2008: 301-312.
- [14] CHEN Yuehui, YANG Bo, DONG Jiwen. Time-series Prediction Using a Local Linear Wavelet Neural Network[J]. Neurocomputing, 2006, 69(4): 449-465.
- [15] ZHANG Li, ZHOU Weida, JIAO Licheng. Wavelet Support Vector Machine [J]. IEEE Transactions on Systems Man & Cybernetics Part B, 2004, 34(1): 34-39.
- [16] WANG Xinying, HAN Min. Online Sequential Extreme Learning Machine with Kernels for Nonstationary Time Series Prediction [J]. Neurocomputing, 2014, 145: 90-97.
- [17] TAKENS F. Detecting Strange Attractors in Turbulence[M]//RAND D, YOUNG Lai-sang. Dynamical Systems and Turbulence. Berlin, Germany: Springer, 1981: 366-381.
- [3] 李素建. 基于语义计算的语句相关度研究[J]. 计算机工程与应用, 2002, 38(7): 75-76.
- [4] 刘群,李素建. 基于《知网》的词汇语义相似度计算[D]. 北京:中国科学院计算技术研究所, 2002.
- [5] 江敏,肖诗斌,王弘蔚,等. 一种改进的基于《知网》的词语语义相似度计算[J]. 中文信息学报, 2008, 22(5): 84-89.
- [6] 朱征宇,孙俊华. 改进的基于《知网》的词汇语义相似度计算[J]. 计算机应用, 2013, 33(8): 2276-2279, 2288.
- [7] 李家南. IT领域问答系统的研究与实现[D]. 广州: 华南理工大学, 2016.
- [8] 李彬,刘挺,秦兵,等. 基于语义依存的汉语句子相似度计算[J]. 计算机应用研究, 2003, 20(12): 15-17.
- [9] 张华平. NLPir 简介[EB/OL]. (2014-12-12). <http://ictclas.nlpir.org/docs>.
- [10] 王利局. 基于语义分析树核的句子相似度计算[D]. 大连:大连理工大学, 2008.
- [11] 蓝雁玲,陈建超. 基于词性及词性依存的句子结构相似度计算[J]. 计算机工程, 2011, 37(10): 47-50.
- [12] 腾少冬,王志良,王莉,等. 基于马尔可夫链的情感计算建模方法[J]. 计算机工程, 2005, 31(5): 17-19.
- [13] 王志良,解仑,董平. 情感计算数学模型的研究初探[J]. 计算机工程, 2004, 30(21): 33-34.
- [14] 微软研究院释库[EB/OL]. (2015-07-18). <http://www.datatang.com/data/14263>.
- [15] 李佳媛. 汉语句子相似度计算技术及其应用[D]. 北京:北京信息科技大学, 2013.
- [16] 庄成龙,钱龙华,周国栋. 基于树核函数的实体语义关系抽取方法研究[J]. 中文信息学报, 2009, 23(1): 3-8.
- [17] 刘宏哲. 一种基于本体的句子相似度计算方法[J]. 计算机科学, 2013, 40(1): 251-256.

编辑 顾逸斐

编辑 顾逸斐