

# 粒子群和遗传算法优化支持向量机的破产预测

杨钟瑾

YANG Zhongjin

广东财经大学 信息学院, 广州 510320

School of Information Science and Technology, Guangdong University of Finance and Economics, Guangzhou 510320, China

**YANG Zhongjin. Bankruptcy prediction based on Support Vector Machine optimized by Particle Swarm Optimization and Genetic Algorithm. Computer Engineering and Applications, 2013, 49(18):265-270.**

**Abstract:** A method based on Support Vector Machine optimization by Particle Swarm Optimization and Genetic Algorithm is proposed for predicting bankruptcy. The proposed method integrates the merits of Particle Swarm Optimization, Genetic Algorithm and Support Vector Machine, which simultaneously searches optimal regularization parameter and kernel parameter of Support Vector Machine for optimal prediction model. A sample dataset comprised of bankruptcy and non-bankruptcy data derived from the UCI machine learning repository is used. The data are randomly read from the dataset and automatically preprocessed by normalization. A 7-fold cross-validation test is used to objectively evaluate the prediction results. The simulation results indicate that the proposed method can automatically and efficiently construct optimal Support Vector Machine. Compared with other methods, the proposed method has better generalization capability, faster learning speed and better bankruptcy prediction accuracy than the other methods.

**Key words:** Particle Swarm Optimization(PSO) algorithm; Genetic Algorithm(GA); Support Vector Machine(SVM); optimization; parameter; bankruptcy prediction

**摘 要:**介绍了一种基于粒子群算法和遗传算法优化支持向量机预测破产的方法。这种方法融合了粒子群算法、遗传算法和支持向量机诸多优点,并行地搜寻支持向量机最优的正则化参数和核参数,由此构建优化的预测模型。采用源自UCI机器学习数据库的破产和非破产混合样本数据集,随机地读入数据和进行数据预处理,运用7重交叉校验方法客观地评价预测结果。仿真结果显示,这种方法能自动有效地构建优化的支持向量机,与其他方法比较,具有更强的推广能力和更快的学习速度,而且具有更好的破产预测准确率。

**关键词:**粒子群算法;遗传算法;支持向量机;优化;参数;破产预测

**文献标志码:**A **中图分类号:**TP18 **doi:**10.3778/j.issn.1002-8331.1303-0352

## 1 引言

破产预测能让人类做出理性的决策,从而最小化风险造成的损失。破产预测广泛应用于各行各业,尤其在学术界和商界备受关注。但是,现实世界所面对的破产预测问题经常是复杂多变和难于捉摸。由于统计模型方法固有的缺陷,因此难于胜任破产预测。与传统的统计模型方法相比,虽然神经网络<sup>[1]</sup>对破产预测的效果有所改进,但是神经网络具有以下的缺点:(1)难以构建合适的预测模型;(2)常常陷入局部极小值;(3)经验风险最小化原则的运用不能确保良好的推广能力;(4)不容易正确选定训练集的大小。

支持向量机<sup>[2]</sup>是一种新型的破产预测智能技术,不仅备受瞩目,而且日益盛行。支持向量机以统计学习理论为

基础,与神经网络明显不同之处在于运用结构风险最小化原理,克服了神经网络容易陷入局部极小值的问题,提高了推广能力。同时,在运用支持向量机构建预测模型时,在正确选定了核函数<sup>[3]</sup>的基础上,只需选择最优参数。因此,相对神经网络而言,支持向量机更简易快捷。目前,启发式优化技术<sup>[4-5]</sup>得到众多研究者的追捧,其中粒子群优化<sup>[6]</sup>和遗传算法<sup>[7]</sup>等最为流行,它们都具有并行处理的特征。与其他的优化技术相比,粒子群优化有许多优点:减少内存、快速运算、容易与其他优化技术集成和易于实现。同样,遗传算法也具有以下优点:有效地扫描和探测复杂空间、避免陷入局部极小值、出色地平衡效果与效率之间的关系和不乏灵活性。

Shin等人运用支持向量机预测公司破产,所得预测准

**作者简介:**杨钟瑾,男,工学博士,副教授,主要研究领域为信息分析、预测与决策和人工智能等。E-mail: yzj136@hotmail.com

**收稿日期:**2013-03-25 **修回日期:**2013-06-24 **文章编号:**1002-8331(2013)18-0265-06

**CNKI出版日期:**2013-07-03 <http://www.cnki.net/kcms/detail/11.2127.TP.20130703.1144.020.html>

准确率超过神经网络与其他统计预测方法<sup>[8]</sup>。Min 和 Lee 采用栅格搜索技术寻找支持向量机最优的核函数参数值,并在此基础上构建破产预测模型,实验结果显示支持向量机的预测准确率高于神经网络与其他统计预测方法<sup>[9]</sup>。Min 等人集成遗传算法与支持向量机来预测破产,采用遗传算法优化特征子集和支持向量机的参数,预测结果表明这种方法优于其他没有优化方法<sup>[10]</sup>。杨和蒙采用支持向量机在商业银行构建破产预测模型,并将结果与 BP 神经网络相比,研究结果表明支持向量机优于 BP 神经网络<sup>[11]</sup>。Wu 等人采用真值遗传算法优化支持向量机的参数,同时构建优化的破产预测模型,实验结果显示这种模型的预测准确率强于神经网络和常规支持向量机<sup>[12]</sup>。鲁和徐建立基于支持向量机的企业破产分析预测系统模型,实证表明该模型具有良好的预测能力和较强的实用性<sup>[13]</sup>。Gao 等人采用 K-最近邻与支持向量机对公司破产进行预测,这种预测方法比传统支持向量机有更好的分类性能<sup>[14]</sup>。Zhong 等人运用支持向量机构建破产预测模型,实验结果显示该种模型性能优于传统统计模型<sup>[15]</sup>。Yoon 和 Kwon 运用栅格搜索方法寻找支持向量机的最优参数,并建立模型预测小企业的破产,这种引荐模型的预测性能优于其他模型<sup>[16]</sup>。Chaudhuri 和 De 运用模糊支持向量机预测上市大公司的破产,与其他方法相比,这种方法改进了预测准确率<sup>[17]</sup>。Moradi 等人运用支持向量机预测上市公司的破产,结果显示该方法的平均预测准确率高于辨别分析方法<sup>[18]</sup>。

预测准确率取决于支持向量机的推广能力,支持向量机的推广能力又深受相关参数选择所左右。在很多情况下,为某一个具体问题而选择最优参数是艺术多于科学,而且这其中经常涉及大量的尝试和失败。因此,预测模型的优化经常导致昂贵的过程。本研究将运用粒子群算法和遗传

算法优化支持向量机的参数,构建优化的支持向量机预测模型,采用来自 UCI 机器学习数据库<sup>[19]</sup>的样本数据并进行归一化数据预处理,将这些引荐方法与随意选择相关参数方法进行比较研究,同时与输入未经预处理的样本原值数据方法比较,由此提出有效优化支持向量机的方法,进而加快支持向量机的学习速度和提高支持向量机的推广性能,从而提高破产预测准确率。

## 2 支持向量机

支持向量机和核函数方法日愈流行,二者珠联璧合成为核函数支持向量机<sup>[20]</sup>,同时成功地在预测领域应用。在模式分类与回归分析方面,支持向量机基于统计学习理论和运用 VC 维思想。支持向量机的基本思想是构造一个超平面作为决策面,由此最大化两类样本的间隔。核函数具有平滑和相似性度量的作用,支持向量机主要由核函数确定。在特征空间中,选择不同形式的核函数就会构成不同的支持向量机,因而也构造出不同的超平面。

在预测领域,径向基核函数是一种常用的核函数,该核函数的表达式如下:

$$k(x, x_i) = \exp(-\gamma \|x - x_i\|^2) \tag{1}$$

其中,  $\gamma$  是核参数。

在支持向量机的学习过程中,破产公司和非破产公司的样本数据连续地被输入,这些输入数据通过非线性核函数映射到特征空间。在高维特征空间中寻找破产公司和非破产公司样本之间的最大间隔,由此构造超平面的一个最优分界来增强分类性能。因此,可以运用支持向量机预测破产。

运用支持向量机预测破产模型如图 1,预测模型从混合样本数据集中读取数据,接着归一化预处理这些数据,

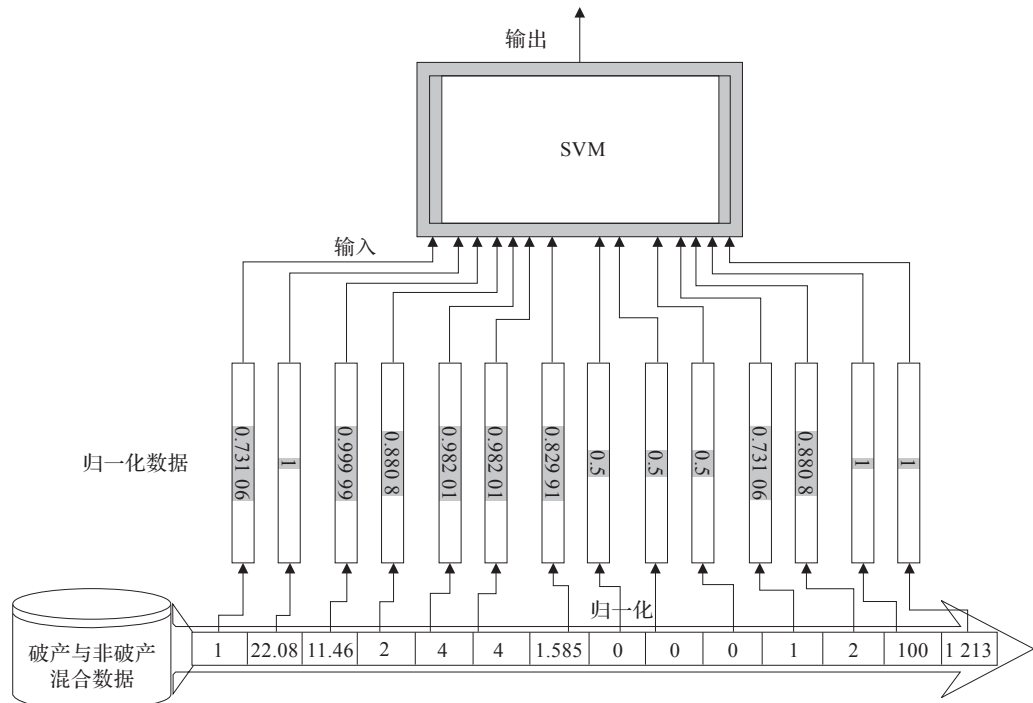


图1 运用支持向量机和归一化输入数据预测破产模型

随后已预处理数据被送进输入空间并转换为输入向量。输入向量经由核函数映射到高维特征空间,在高维特征空间中,通过优化获得一个超平面来分离破产公司和非破产公司的样本,从而实现破产预测功能。

### 3 粒子群算法和遗传算法

#### 3.1 粒子群算法

粒子群算法是一种新颖且有效的优化智能技术,算法思想来自由个体组成的群体与周边环境以及个体之间互动行为的启示。

在求解优化问题时,个体不仅在搜索空间里有确定的位置,而且具有一定的运动速度,这些个体被称为粒子。根据粒子本身具有最佳性能时所处位置与邻近全局最优粒子的位置,粒子动态地调整它的状态。这个迭代过程直到寻获最优解或是超出了计算限值才终止。在搜索空间里,每个粒子都是一个潜在解,而粒子适应度是评价这些解优劣的标准。粒子适应度计算公式如下:

$$f(t) = \frac{1}{2}(y(t) - d(t))^2 \quad (2)$$

在此,  $f(t)$  为第  $t$  回合迭代时当前粒子的适应度,  $y(t)$  表示第  $t$  回合迭代时当前粒子的实际输出,  $d(t)$  代表第  $t$  回合迭代时当前粒子的期望输出。

#### 3.2 遗传算法

遗传算法是一种具有随机全局搜索性能的优化智能技术,它模拟自然生物进化原理——适者生存,优胜劣汰。遗传算法根据这个进化原理,将优化问题的求解表示为个体在种群中优胜劣汰的过程,通过对个体进行选择、交叉和变异等遗传操作,实现种群中个体一代接一代地进化,最终获得问题的最优解或满意解。

选择操作确定了新个体基因的来源,在每一代中,选择操作根据种群中个体适应度的评价水平选出优良个体。交叉操作是新个体产生的主要途径,并且决定遗传算法的全局搜索性能,交叉操作是将来自父母代的不同个体基因进行交换和重组,由此产生下一代的新个体。变异操作增加了种群变化的途径,决定局部搜索性能并促进遗传算法收敛于全局最优。变异操作是依据存在于染色体中某个体基因的变异概率而改变这个基因的值,由此产生新的个体。

### 4 粒子群算法和遗传算法优化支持向量机

正则化参数  $C$  和核参数  $\gamma$  是支持向量机的重要参数,这两个参数对支持向量机的预测性能具有深远的影响。在本研究中,将分别运用粒子群算法和遗传算法优化正则化参数  $C$  和核参数  $\gamma$ ,然后运用优化参数的支持向量机预测破产。

#### 4.1 支持向量机参数

参数选择对支持向量机的推广能力至关重要,因此错误选定参数将极大减弱支持向量机的推广能力。优化支持向量机模型是指为增强推广能力而选定最优参数的过程,决定支持向量机模型性能的参数主要是正则化参数  $C$

和核参数  $\gamma$ 。正则化参数  $C$  协调最小学习误差与最小模型复杂度之间的关系,如果参数  $C$  赋值过小,那么就会因为模型过于简单而引起实践风险过高;如果参数  $C$  赋值过大,那么就会因为模型过于复杂而过度学习。核参数  $\gamma$  隐式定义了从输入空间到高维特征空间的非线性映射,通常也将这个高水平参数称为超参数,如果参数  $\gamma$  取值过小,那么就会对上界域产生过大影响。虽然支持向量机目前发展得很好,但其目前尚缺乏结构化方法来实现参数的最优选择<sup>[21]</sup>。在实际应用中,这些最优参数的选择缺乏坚实的理论指导,往往经过多次实验和错误后才能选定。在此,分别通过粒子群算法和遗传算法优化这些参数。

#### 4.2 粒子群算法优化支持向量机过程

粒子群算法依据以下公式调整每个粒子的速度和位置:

$$v_{id}(t+1) = w \times v_{id}(t) + c_1 \times r_1 \times (p_{ibd}(t) - x_{id}(t)) + c_2 \times r_2 \times (p_{gbd}(t) - x_{id}(t)) \quad (3)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (4)$$

其中,  $i = 1, 2, \dots, m$ ,  $d = 1, 2, \dots, D$ ;  $w$  为惯性因子;  $c_1$ ,  $c_2$  是取值为正的学习因子;  $r_1$ ,  $r_2$  是取值范围为  $[0, 1]$  的随机数;  $v_{id}(t+1)$  和  $v_{id}(t)$  分别代表在第  $(t+1)$  回合和第  $t$  回合迭代时,第  $i$  个粒子在第  $d$  维度的速度;  $x_{id}(t+1)$  和  $x_{id}(t)$  分别代表在第  $(t+1)$  回合和第  $t$  回合迭代时,第  $i$  个粒子在第  $d$  维度的位置;  $p_{ibd}(t)$  表示第  $t$  回合迭代时,第  $i$  个粒子在第  $d$  维度的最佳位置;  $p_{gbd}(t)$  表示第  $t$  回合迭代时,整个粒子群在第  $d$  维度的最佳位置。

运用粒子群算法优化支持向量机参数过程的概要如图2。

#### 4.3 遗传算法优化支持向量机过程

图3描绘了运用遗传算法优化支持向量机参数过程的概要。

#### 4.4 性能评价

采用预测准确率作为评价支持向量机预测模型性能的指标,同时运用7重交叉校验方法计算和评价模型的预测准确率,这种校验方法确保样本数据集中的每个样本数据都做并且只做一次检验,因此最低程度地减少了因特别选择训练子集和检验子集所带来的结果偏差。在此,  $PCR$  表示预测准确率,  $TQ$  是指检验样本总数,  $CQ$  代表被正确预测的检验样本个数。预测准确率的计算公式如下:

$$PCR = \frac{CQ}{TQ} \times 100\% \quad (5)$$

### 5 结果和讨论

利用 Matlab 编写程序和进行仿真,相关方法在运行 Windows XP 的个人计算机(AMD Athlon 64 3200+ 2 GHz processor; 4 GB RAM)上实施与测试。

实验仿真数据源自 UCI 机器学习数据库的澳大利亚信用数据集,实验仿真数据集由 383 个破产样本数据与 307 个非破产样本数据所构成,这个混合数据集共包含 690 个样本数据。在支持向量机的 7 重交叉校验仿真实验中,分



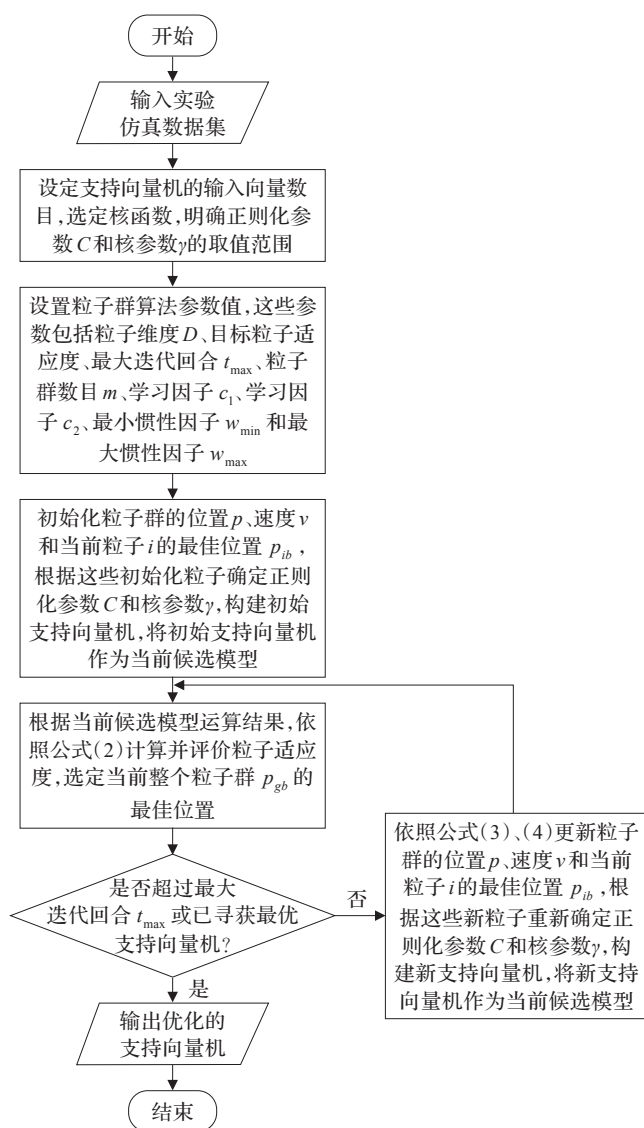


图2 粒子群优化支持向量机参数算法流程图

别从混合数据集中选取 370 个破产样本数据和 295 个非破产样本数据作为测试样本数据, 这 665 个样本数据首先随机地分成各自含有 95 个样本数据的 7 个子集。每一重交叉校验仿真实验都选取 7 个子集中的一个子集作为检验集, 剩下的 6 个子集共同构成训练集。最后所得预测准确率是 7 重交叉校验预测准确率的平均值。

在此, 粒子群算法和遗传算法的相关参数值是根据预测对象与所获得实验结果而选择设置。在运用粒子群算法优化支持向量机参数过程中, 粒子维度设置为 2, 目标粒子适应度设置为 0.001, 最大迭代回合设置为 3, 粒子群数目

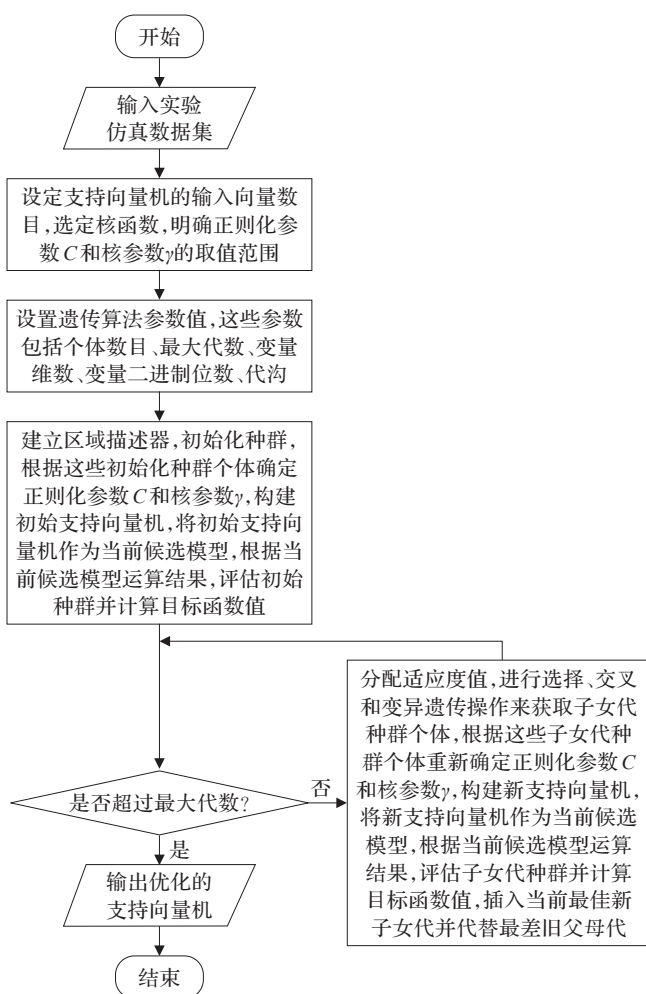


图3 遗传算法优化支持向量机参数算法流程图

设置为 15, 学习因子  $c_1$  与学习因子  $c_2$  都设置为 1.48, 最小惯性因子与最大惯性因子分别设置为 0.4 和 0.9。在运用遗传算法优化支持向量机参数过程中, 通过 Genetic Algorithm (GA) Toolbox<sup>[22]</sup> 应用软件解决优化问题。采用二进制字符串编码形式, 遗传算法的主要参数分别设置为: 个体数目等于 20, 最大代数等于 10, 变量维数等于 2, 变量二进制位数等于 20, 代沟等于 0.9。采用随机全体采样选择操作方法, 同时采用单点交叉操作方法, 并将交叉概率设置为 0.7。

在破产预测过程中, 预测模型运用二进制支持向量机并通过 STPRtool<sup>[23]</sup> 应用软件实现, 这些支持向量机由 SVMlight<sup>[24]</sup> 软件包 (5.00 版本)<sup>[25]</sup> 所构建, SVMlight 软件包是处理复杂问题的强大优化工具<sup>[26]</sup>。预测模型的支持向量机采用径向基核函数。表 1 给出了每一个预测模型的模型

表1 性能比较

MN	IDT	IVN	C	$\gamma$	OVN	PCR/(%)	TIME/s
CSOODSVM	原值数据	14	128	16	1	69.172 9	—
CSOPDSVM	归一化数据	14	4	8	1	79.548 9	—
PSOODSVM	原值数据	14	32 768	1 024	1	84.360 9	7 991.406 3
PSOPDSVM	归一化数据	14	128	1	1	86.616 5	1 253.896 0
GAOODSVM	原值数据	14	1 024	64	1	74.887 2	11 852.156 0
GAOPDSVM	归一化数据	14	8	0.5	1	86.315 8	3 026.719 0

名称(MN)、输入数据类型(IDT)、输入向量个数(IVN)、正则化参数( $C$ )、核参数( $\gamma$ )和输出向量个数(OVN),同时该表也分别列出了各个预测模型的预测准确率(PCR)与平均每一回合模型优化时间(TIME)。

从表1可以得到以下观点:

(1)采用经归一化预处理数据时,运用粒子群算法优化支持向量机参数所获得的预测模型PSOPDSVM的预测准确率是86.616 5%,平均每一回合模型优化时间为1 253.896 s;运用遗传算法优化支持向量机参数所获得的预测模型GAOPDSVM的预测准确率是86.315 8%,平均每一回合模型优化时间为3 026.719 s;而随意设置支持向量机参数所获得的预测模型CSOPDSVM的预测准确率仅是79.548 9%。

(2)采用原值数据时,运用粒子群算法优化支持向量机参数所获得的预测模型PSOODSVM的预测准确率是84.360 9%,平均每一回合模型优化时间为7 991.406 3 s;运用遗传算法优化支持向量机参数所获得的预测模型GAOODSVM的预测准确率是74.887 2%,平均每一回合模型优化时间为11 852.156 s;而随意设置支持向量机参数所获得的预测模型CSOODSVM的预测准确率只是69.172 9%。

(3)当采用经归一化预处理数据时,模型PSOPDSVM的预测准确率比模型GAOPDSVM和模型CSOPDSVM的预测准确率分别高出0.300 7%和7.067 6%,模型GAOPDSVM比模型CSOPDSVM的预测准确率高出6.766 9%。模型PSOPDSVM的平均每一回合模型优化时间比模型GAOPDSVM的平均每一回合模型优化时间约少1.4倍。

(4)当采用原值数据时,模型PSOODSVM的预测准确率比模型GAOODSVM和模型CSOODSVM的预测准确率分别高出9.473 7%和15.188%,模型GAOODSVM比模型CSOODSVM的预测准确率高出5.714 3%。模型PSOODSVM的平均每一回合模型优化时间比模型GAOODSVM的平均每一回合模型优化时间约少0.5倍。

(5)预测模型PSOPDSVM的预测准确率比预测模型PSOODSVM的预测准确率高2.255 6%,预测模型GAOPDSVM的预测准确率比预测模型GAOODSVM的预测准确率高11.428 6%,预测模型CSOPDSVM的预测准确率比预测模型CSOODSVM的预测准确率高10.376%。模型PSOODSVM的平均每一回合模型优化时间比模型PSOPDSVM的平均每一回合模型优化时间约多5.4倍,模型GAOODSVM的平均每一回合模型优化时间比模型GAOPDSVM的平均每一回合模型优化时间约多2.9倍。

(6)在预测准确率方面,运用粒子群算法优化支持向量机参数所获得的预测模型比运用遗传算法优化支持向量机参数与随意设置支持向量机参数所获得的预测模型高,运用遗传算法优化支持向量机参数所获得的预测模型比随意设置支持向量机参数所获得的预测模型高,采用经归一化预处理数据比采用原值数据高。

(7)在模型优化时间方面,运用粒子群算法优化支持向量机参数比运用遗传算法优化支持向量机参数少,采用

经归一化预处理数据比采用原值数据少。

## 6 结论和未来研究

本文介绍了运用粒子群算法和遗传算法优化支持向量机参数的研究。此外,针对破产预测构建了二进制支持向量机,采用源自UCI机器学习数据库的破产和非破产混合样本数据集,分别输入未经预处理的原值数据和经归一化预处理的数据进行仿真实验,运用7重交叉校验方法计算和客观地评价模型的预测准确率。仿真实验证明,运用粒子群算法和遗传算法优化支持向量机参数有效地加速学习和提高推广性能,从而进一步提高破产预测准确率。同时,采用经归一化预处理数据可以进一步提高破产预测准确率。实验结果比较分析表明,在支持向量机预测模型优化方面,无论是预测准确率或模型优化时间,粒子群算法都强于遗传算法。

支持向量机由于不同核函数的构造和选择而各异,核函数的合理构造与正确选择是支持向量机执行学习过程的重要一步,并且对支持向量机性能有至深的影响。为特定问题而构造和选择核函数是一个耗时与耗力的过程,因此,如何构造与如何选择最优核函数是亟需有待进一步研究的重要问题。

## 参考文献:

- [1] Haykin S. Neural networks: a comprehensive foundation[M]. Upper Saddle River, NJ, USA: Prentice Hall, 1998.
- [2] Vapnik V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.
- [3] Schoelkopf B, Smola A. Learning with kernels[M]. Cambridge, Massachusetts: The MIT Press, 2002.
- [4] Jiao Licheng, Li Yangyang, Gong Maoguo, et al. Quantum-inspired immune clonal algorithm for global optimization[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2008, 38(5): 1234-1253.
- [5] Karaboga D, Basturk B. On the performance of artificial bee colony algorithm[J]. Applied Soft Computing, 2008, 8(1): 687-697.
- [6] Kennedy J, Eberhart R C. Particle swarm optimization[C]//Proceedings of IEEE International Conference on Neural Networks. Piscataway: IEEE Press, 1995: 1942-1948.
- [7] Goldberg D E. Genetic algorithms in search, optimization and machine learning[M]. New York: Addison-Wesley, 1989.
- [8] Shin K, Lee T S, Kim H. An application of support vector machines in bankruptcy prediction model[J]. Expert Systems with Applications, 2005, 28(1): 127-135.
- [9] Min J H, Lee Y. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters[J]. Expert Systems with Applications, 2005, 28(4): 603-614.
- [10] Min S, Lee J, Han I. Hybrid genetic algorithms and support vector machines for bankruptcy prediction[J]. Expert Systems with Applications, 2006, 31(3): 652-660.

- [11] 杨毓, 蒙肖莲. 用支持向量机(SVM)构建企业破产预测模型[J]. 金融研究, 2006(10): 65-75.
- [12] Wu C, Tzeng G, Goo Y, et al. A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy[J]. Expert Systems with Applications, 2007, 32(2): 397-408.
- [13] 鲁海波, 徐云. 基于支持向量机的企业破产分析[J]. 数学的实践与认识, 2007, 37(12): 13-18.
- [14] Gao Zhong, Cui Meng, Po Laiman. Enterprise bankruptcy prediction using noisy-tolerant Support Vector Machine[C]//2008 International Seminar on Future Information Technology and Management Engineering. Leicestershire: IEEE Press, 2008: 153-156.
- [15] Zhong Ping, Xi Bin, Cen Yong. Support Vector Machines modeling for prediction[J]. International Journal of Systems and Control, 2008, 1: 42-47.
- [16] Yoon J S, Kwon Y S. A practical approach to bankruptcy prediction for small businesses: substituting the unavailable financial data for credit card sales information[J]. Expert Systems with Applications, 2010, 37(5): 3624-3629.
- [17] Chaudhuri A, De K. Fuzzy Support Vector Machine for bankruptcy prediction[J]. Applied Soft Computing, 2011, 11(2): 2472-2486.
- [18] Moradi M, Sardasht M S, Ebrahimpou M. An application of Support Vector Machines in bankruptcy prediction evidence from Iran[J]. World Applied Sciences Journal, 2012, 17(6): 710-717.
- [19] Asuncion A, Newman D J. UCI machine learning repository [EB/OL]. [2012-03-14]. <http://archive.ics.uci.edu/ml/datasets.html>.
- [20] 杨钟瑾. 核函数支持向量机[J]. 计算机工程与应用, 2008, 44(33): 1-6.
- [21] 顾亚祥, 丁世飞. 支持向量机研究进展[J]. 计算机科学, 2011, 38(2): 14-17.
- [22] Andrew C, Peter F, Hartmut P, et al. Genetic Algorithm toolbox user's guide[R]. Sheffield: University of Sheffield, 1994.
- [23] Franc V, Hlava V. STPRtool[EB/OL]. (2004-12-14). [http://cmp.felk.cvut.cz/cmp/cmp\\_software.html](http://cmp.felk.cvut.cz/cmp/cmp_software.html).
- [24] Joachims T. Making large-scale SVM learning practical[C]//Advances in Kernel Methods—Support Vector Learning. Cambridge, MA: MIT Press, 1999: 169-184.
- [25] Joachims T. SVMlight software[EB/OL]. (2002-03-14). <http://svmlight.joachims.org/>.
- [26] Franc V, Hlava V. Statistical pattern recognition toolbox for Matlab user's guide[EB/OL]. (2004-06-24). <ftp://cmp.felk.cvut.cz/pub/cmp/articles/>.

(上接211页)

变化特征。针对雷电电磁辐射的时域特征,在信号个数远小于数据块长度时,分析了能量检测和特征检测的互补特性,首次将特征检测与能量块检测相结合,基于雷达自动筛选思想和E-CFAR检测技术,提出了自适应联合检测算法,并给出实现原理和具体步骤。利用采集的290 MHz雷电数据进行检测分析与比较,可知自适应联合检测能够明显提高检测概率,表明提出的自适应联合检测的有效性和实用性。且硬件实现时,其复杂度并不增加太大,同时该算法也可用于其他数据块检测。

### 参考文献:

- [1] Du Haiming, Ma Hong, Zhang Yufeng. A novel algorithm for real-time adaptive signal detection in lightning location system[C]//ICISE 2010, Hangzhou, China, 2010.
- [2] 陈成品, 郗秀书, 张泉, 等. 地闪放电过程的甚高频辐射特征分析[J]. 中国电机工程学报, 2005, 25(19): 122-126.
- [3] 张泉, 郗秀书, 张广庶. 短基线时间差闪电辐射源探测系统和初步定位结果[J]. 高原气象, 2003, 22(3): 226-234.
- [4] Urkowitz H. Energy detection of unknown deterministic signals[J]. Proceedings of the IEEE, 1967, 55(4): 523-531.
- [5] 隋丹, 葛临东, 屈丹. 一种新的基于能量检测的突发信号存在性检测算法[J]. 信号处理, 2008, 24(4): 614-617.
- [6] 张霆廷, 张钦宇, 张乃通. 基于能量检测的脉冲超宽带接收机[J]. 吉林大学学报, 2010, 40(1): 281-286.
- [7] 胡啸, 马洪. 归一化峰度及其在弱非线性系统盲辨识中的应用[J]. 信号处理, 2010, 26(9): 1389-1396.
- [8] 赵锡凯, 张贤达. 基于最大峰度准则的非因果AR系统盲辨识[J]. 电子学报, 1999, 27(12): 126-128.
- [9] Smith M E, Varshney P K. Intelligent CFAR processor based on data variability[J]. IEEE Trans on AES, 2000, 36(3): 837-847.
- [10] 何友, 关键, 彭应宁, 等. 雷达自动检测与恒虚警处理[M]. 北京: 清华大学出版社, 1999.
- [11] Chen Yunfei. Improved energy detector for random signals in Gaussian noise[J]. IEEE Trans on Wireless Communications, 2010, 9(2): 558-563.
- [12] 陈渭民. 雷电学原理[M]. 北京: 气象出版社, 2006.
- [13] 姚天任, 孙洪. 现代数字信号处理[M]. 武汉: 华中科技大学出版社, 1999: 184-190.
- [14] 杜海明, 马洪, 苏红超, 等. 115 MHz频段雷电波形特征分析[J]. 电网技术, 2010, 34(10): 60-64.
- [15] 苏红超. 甚高频雷电探测站数字处理系统设计与实现[D]. 武汉: 华中科技大学, 2011.