

基于SVM的高新技术企业财务危机预警研究

朱发根, 刘拓, 傅毓维

(哈尔滨工程大学 经济管理学院, 黑龙江 哈尔滨 150001)

摘 要:研究高新技术企业的财务危机预警,不仅有助于经营者预防和监控财务危机,也有利于投资者规避或降低财务风险。应用新型机器学习方法——支持向量机(SVM)构建了高新技术企业财务预警模型。实证分析表明,该模型简单易行,准确可靠。

关键词:SVM;高新技术企业;财务危机;预警

中图分类号:F276.44

文献标识码:A

文章编号:1001-7348(2009)11-0073-03

0 引言

高新技术企业是高新技术产业的载体,是国家创新体系的重要组成部分。与一般企业相比,高新技术企业具有“高投入,高风险,高收益”的特征,使得企业主要利益相关者对其财务状况更为敏感,对其进行财务危机预警的要求更为迫切。对于经营者而言,高新技术企业的财务状况和潜在的财务风险决定了它们在市场上的信誉和融资能力,如何提高经营者的风险防范和危机识别能力,直接关系到高新技术企业的生存与发展。对于投资者而言,高新技术企业往往是高成长性、高回报率和高风险率的混合体,如何有效对其进行危机预警、分析其财务走势,对于投资决策能否获得预期收益至关重要。

根据Dimitras^[1]、Balcaen^[2]、龚小君^[3]和何荣华^[4]等人的研究成果,常见的企业财务危机预警模型有:一元判定模型(Univariate),其简单适用,但总体判断精度不高;多元线性判断模型(Z Score),该模型精度较高,但现实样本往往难以满足其假设前提;对数比例模型(Logit),其克服了统计假设的局限性,但精度受到影响;概率单位回归模型(Probit),其精度较高但假设复杂,且计算复杂;神经网络模型(ANN),该模型具有较好的学习能力和预测能力,但依赖大样本,实用性较差。为克服以上模型在假设前提、样本容量等方面的缺陷,同时满足精度要求,本文应用具有扎实理论基础的新型机器学习方法——支持向量机构建了财务危机预警模型,并进行了实证分析。

1 高新技术企业财务危机预警指标体系

高新技术企业财务危机预警,即以财务信息为基础,

通过设置并观察一些敏感性预警指标的变化,对高新技术企业可能要面临的财务危机所实施的实时监控和预测警报^[5]。显然,财务危机预警的关键是构建一套科学合理的指标体系。从已有的相关研究成果来看,姜毅^[6]构建的财务预警指标体系包括现金状况、债务状况、经营状况、周转状况和盈利状况5类指标;Huang等^[7]用于企业破产预测的财务指标有财务结构、偿债能力、营运能力、获利能力、现金能力;刘开瑞^[8]认为财务预警分析指标包括偿债能力、资产运营状况、获利能力、发展能力和现金流量。从国内证券网站采用的财务预警指标来看,中国证券网(www.cn-stock.com)用于财务评价的指标有营利能力、资产运营能力、现金获取能力、偿债能力;证券之星(www.stockstar.com)采用的财务评分指标包括偿债能力、盈利能力、成长能力、现金能力和资产负债管理能力。

综上所述,偿债能力、营运能力、盈利能力和现金能力是用于财务预警分析的普遍性指标。与其它企业相比,高新技术企业具有高失败率(高风险),也意味着高成长性(高收益),其成长能力更能揭示企业收益状况和财务风险,因而对其进行财务预警分析还应考虑成长能力。在上述分析的基础上,遵循指标设计的系统性和科学性,本文设计的高新技术企业财务预警指标体系如下:

(1) 盈利能力:净资产收益率(X_1)、每股收益(X_2)、主营业务利润率(X_3);

(2) 偿债能力:流动负债与经营活动净现金流比(X_4)、资产负债比率(X_5)、速动比率(X_6);

(3) 现金能力:每股营业现金流量(X_7)、现金股利支付率(X_8)、现金自给率(X_9);

(4) 营运能力:总资产周转率(X_{10})、应收账款周转率(X_{11})、存货周转率(X_{12});

收稿日期:2008-01-09

基金项目:黑龙江省科技计划项目(GC06D209);黑龙江省教育科学“十一五”规划重点课题(HZG036)

作者简介:朱发根(1983-),男,湖南株洲人,哈尔滨工程大学博士研究生,研究方向为财务风险管理;傅毓维(1940-),男,辽宁盖县人,哈尔滨工程大学经济管理学院教授,博士生导师,研究方向为决策优化。

(5)成长能力:主营业务收入增长率(X_{13})、净利润增长率(X_{14})、固定资产投资扩张率(X_{15})。

2 基于SVM的高新技术企业财务危机预警模型

2.1 SVM(支持向量机)的基本思想

支持向量机(Support vector Machine, SVM)是Vapnik等人^[9]提出的一种新型机器学习方法,它遵循结构风险最小原则和有限样本假设,克服了传统机器学习(如神经网络)过学习、久学习、局部收敛、高维灾难等问题,具有较好的学习能力和推广能力。

支持向量机的基本思想就是构造一个最优分类超平面,以此作为决策曲面将样本准确划分为两类,即假定 n 个样本总体 $D=\{x_i, y_i \mid i=1, 2, \dots, n\}$ ($x \in R^p, y \in R^q$),能被一个超平面 $H: wx+b=0$ 正确分开,并且分类间隔最大。对于线性可分样本 $(x_i, y_i), i=1, 2, \dots, n$,构造其最优分类超平面可用如下凸二次规划描述:

$$\min_{w,b} \frac{1}{2} \|W\|^2 = \min_{w,b} \frac{1}{2} W^T W$$

$$\text{s.t. } y_i[(w \cdot x_i) + b] - 1 \geq 1 \quad i=1, 2, \dots, n$$

引入Lagrange系数 α_i ,上述问题的对偶形式为:

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$$

$$\text{s.t. } \sum_{i=1}^n y_i \alpha_i = 0, \alpha_i > 0, i=1, 2, \dots, n$$

求解得到决策函数:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n \alpha_i y_i (x \cdot x_i) + b \right\}$$

其中,sgn(\cdot)为符号函数, α_i 为Lagrange系数, b 为分类域值。

对于非线性问题,可通过非线性变换转化为在高维特征空间求取最优分类面。通常采用满足Mercer条件的核函数 $K(x, x_i)$ 来实现这一非线性变换,引入非负松弛变量 ξ_i (≥ 0)和规则化常数 $C(>0)$,相应的决策函数为:

$$f(x) = \text{sgn} \left\{ \sum \alpha_i y_i K(x \cdot x_i) + b \right\}$$

2.2 SVM在高新技术企业财务危机预警模型中的应用

对于每一个样本企业,假设反映样本企业财务状况的预警指标有 p 个,即支持向量机的输入为 p 维;判别企业处于危机状态与否的结论(预警结果)为 q 维,即支持向量机的输出为 q 维。则可选取 n 个高新技术企业构成训练样本集 $D=\{x_i, y_i \mid i=1, 2, \dots, n\}$ ($x \in R^p, y \in R^q$)。

实际经济活动中,反映高新技术企业财务状况的预警指标(输入)和预警结果(输出)是非线性关系,因此引入满足Mercer条件的核函数 $K(x, x_i)$ 来实现这一非线性变换。常用的核函数有以下4种^[10]:

- (1)Dot核函数, $K(x, x_i) = x \cdot x_i$
- (2)Polynomial核函数, $K(x, x_i) = [(x \cdot x_i) + 1]^d$;
- (3)Neural核函数, $K(x, x_i) = \tanh(ax \cdot x_i + b)$;

(4)Radial核函数, $K(x, x_i) = \exp(-\gamma \|x - x_i\|^2)$ 。

通过训练,选择核函数和其它参数,得到高新技术企业财务危机预警模型的决策函数:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n \alpha_i y_i K(x \cdot x_i) + b \right\}$$

3 实证分析

3.1 样本数据的选取

本文采用的是被认定为高新技术企业或所在行业为高新技术产业的上市公司财务数据(数据来源于证券之星网站),视被特别处理(ST)的上市公司为陷入财务危机的企业。共选择50家企业作为训练和验证样本,包括2006年新增的23家ST公司(视为财务异常),其余27家非ST公司视为财务正常;取其中30家企业构成训练样本,另外20家企业构成验证样本。对所有企业,输入2年之前(2004年)的财务数据(经过标准化处理),输出为-1(ST公司)和1(非ST公司)。运用这样的样本数据得到的预警模型,可以在高新技术企业发生财务危机两年之前提供较准确的预警信号,避免因时间间隔过短而高估模型预测能力,从而真正起到预警作用,而非事后分析。

3.2 模型训练——确定核函数及其参数

应用Winsvm软件对训练样本进行1 000次优化训练,根据最小均方差(MSE)准则,选用Radial核函数($\gamma=0.6$),并得到规则化常数 $C(1.0 \times 10^5)$ 、非负松弛变量 ξ_i (均取 1.0×10^{-5})、Langerange乘数 α_i (限于篇幅不予列出)和阈值 $b(-9.529 \times 10^{-3})$ 。运用Radial核函数和上述参数对训练样本进行测试,训练结果见表1。

表1 模型训练结果

类型	训练样本		训练精度(%)
	实际值(家)	预测值(家)	
ST公司	13	13	100
非ST公司	17	17	100
训练样本	30	30	100

3.3 验证分析

利用训练得到的核函数和各项参数,对验证样本进行测试,验证结果和详细数据如表2和表3所示。

表2 验证结果

类型	验证样本		验证精度(%)
	实际值(家)	预测值(家)	
ST公司	10	10	100
非ST公司	10	8	80
训练样本	20	18	90

从表2可以看出,该模型对验证样本的预测精度为90%,其中对ST企业的预测精度为100%,对非ST企业的预测精度为80%。换言之,模型将ST企业误判为非ST企业的概率为0,将非ST企业误判为ST企业的概率为20%。究其原因,可能是某些濒临财务危机的企业为了避免被特别处理(ST)而粉饰公开的财务报表,导致出现模型判断误差。

表3 验证样本数据和预测值

企业	X1*	X2*	X3*	X4*	X5*	X6*	X7*	X8*	X9*	X10*	X11*	X12*	X13*	X14*	X15*	实际值	预测值
st 方大	0.84	0.64	0.93	0.61	0.66	0.04	0.60	0.02	0.36	0.11	0.05	0.02	0.58	0.92	0.30	-1	-1
st 创智	0.00	0.48	0.00	0.90	0.64	0.03	0.61	0.00	0.00	0.00	0.00	0.01	0.10	0.95	0.00	-1	-1
st 数码	0.55	0.55	0.93	0.98	0.73	0.02	0.71	0.01	0.52	0.42	0.20	0.07	0.39	0.85	0.29	-1	-1
st 大通	0.87	0.73	0.94	0.57	0.69	0.01	0.64	0.03	0.43	0.06	0.01	0.00	0.00	0.93	0.28	-1	-1
st 上风	0.82	0.62	0.88	0.04	0.76	0.25	0.65	0.00	0.48	0.10	0.01	0.05	0.53	0.91	0.23	-1	-1
st 实达	0.80	0.68	0.95	0.92	0.59	0.02	0.48	0.02	0.43	0.31	0.09	0.06	0.32	0.92	0.24	-1	-1
st 天颐	0.73	0.60	0.88	0.86	0.78	0.03	0.83	0.00	0.51	0.07	0.05	0.02	0.11	0.93	0.33	-1	-1
st 长城	0.85	0.66	0.91	0.17	0.76	0.14	0.67	0.01	0.48	0.09	0.02	0.02	0.43	0.79	0.29	-1	-1
st 科苑	0.41	0.00	0.76	0.99	0.63	0.03	0.51	0.00	0.47	0.12	0.11	0.08	0.54	0.00	0.20	-1	-1
st 华光	0.61	0.32	0.67	0.75	0.63	0.00	0.54	0.00	0.01	0.02	0.01	0.07	0.80	0.72	0.35	-1	-1
用友	0.93	0.86	0.99	0.00	1.00	1.00	0.79	0.02	1.00	0.19	0.18	0.23	0.60	0.94	0.38	1	-1
信雅达	0.93	0.81	0.98	0.48	0.82	0.17	0.83	0.03	0.59	0.23	0.04	0.10	0.78	0.94	0.33	1	1
中兴通讯	0.97	1.00	0.97	0.60	0.76	0.19	1.00	0.01	0.59	0.50	0.11	0.04	0.72	0.94	0.34	1	1
长园新材	0.97	0.86	1.00	0.19	0.87	0.08	0.74	0.01	0.69	0.25	0.03	0.05	0.62	0.94	0.33	1	1
思达高科	0.93	0.77	0.97	0.75	0.71	0.09	0.67	0.06	0.60	0.21	0.03	0.04	0.94	0.94	0.31	1	-1
华立科技	0.93	0.77	0.97	0.74	0.73	0.05	0.68	0.04	0.73	0.23	0.06	0.05	0.61	0.94	0.28	1	1
长城电工	0.90	0.76	0.97	0.55	0.73	0.04	0.71	0.05	0.51	0.16	0.02	0.02	0.58	0.94	0.28	1	1
复旦复华	0.92	0.76	0.98	0.58	0.70	0.04	0.65	0.04	0.53	0.08	0.03	0.01	0.59	0.95	0.29	1	1
工大高新	0.91	0.76	0.97	0.28	0.58	0.14	0.58	0.02	0.50	0.10	0.25	0.01	0.62	0.94	0.26	1	1
宝信软件	0.97	0.80	0.97	0.37	0.82	0.14	0.69	0.00	0.53	0.85	0.09	0.14	0.69	0.94	0.32	1	1

验证结果同时表明,就某个企业而言,若预警结果为“-1”,则有83.33%(10/12)的概率表明其两年后会被特别处理;若预警结果为“1”,则几乎可以肯定该企业两年之后财务状况正常。这说明运用支持向量机对高新技术企业进行财务危机预警的精度较高,根据预警结果可以对企业两年后的财务状况作出较为准确的判断。

4 结语

高新技术企业“高投入,高收益,高风险”的特点使得对其进行财务危机预警具有重要意义。本文运用具有扎实理论基础的新型机器学习方法——支持向量机,构建了高新技术企业的财务危机预警模型。训练结果表明,该模型具有100%的训练精度,学习能力强;验证结果表明,该模型具有90%的预警精度,具有较好的推广能力。综上所述,对于小样本、非线性、高维(输入)的机器学习问题——高新技术企业财务危机预警,应用支持向量机建模能够获得比较精确、可靠的结果。

参考文献:

[1] A I DIMITRAS.A survey of business failures with an emphasis on prediction methods and industrial applicationl[J].European

Journal of Operational Research, 1996(90):487-513.

[2] SOFIE BALCAEN, HUBERT OOGHE.35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems [J].The British Accounting Review, 2006(38):63-93.

[3] 龚小君,王光华.我国财务预警研究的回顾与展望[J].重庆邮电学院学报(社会科学版),2005(2):186-188.

[4] 何荣华.财务预警模型—分析、比较、评价[J].会计之友, 2006(8):54-55.

[5] 张鸣,张艳,程涛.企业财务预警研究前沿[M].北京:中国财政经济出版社,2004:3-4.

[6] 姜毅.企业财务预警指标体系构建[J].商业会计,2006(8): 18-19.

[7] S M HUANG.A hybrid financial analysis model for business failure prediction [J].Expert Systems with Applications, 2007 (10):1-7.

[8] 刘开瑞.财务预警分析指标[J].生产力研究,2007(4):138-141.

[9] 李应红,尉询楷,刘建勋.支持向量机的工程应用[M].北京:兵器工业出版社,2004:2-14.

[10] S R GUNN.Support vector machines for classification and regression[D].University of Southampton,1997.

(责任编辑:王尚勇)