

Evaluation of Application College Teachers' Scientific Research Ability Based on GFMG-SVM*

BAO Yongqiang^{1*}, WANG Mulan², LIANG Ruiyu¹,
WANG Qingyun¹, SONG Yufei¹, ZHU Hao¹

(1. School of Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China;
2. Dept of Science, Technology and Industries, Nanjing Institute of Technology, Nanjing 211167, China)

Abstract: The key to improving the level of scientific research is to comprehensively evaluate teachers' scientific research ability. This paper analyzes the problems existing in the evaluation of applied university scientific research, constructs an evaluation model of Applied University teachers' scientific research ability based on expert evaluation, uses support vector machine (SVM) to simulate peer expert evaluation, and proposes a generalized Gaussian mixture kernel function (GFMG) to improve the recognition rate of SVM model. Through the empirical analysis of an applied university in Nanjing, the Mixed Gaussian Kernel Function can improve the correct recognition rate by 95.79%, and the model can effectively evaluate the scientific research ability of applied teachers.

Key words: scientific research management in college; evaluation of teachers' scientific research ability; GFMG-SVM; Gaussian kernel function

EEACC: 0120; 0140

doi: 10.3969/j.issn.1005-9490.2019.05.001

基于 GFMG-SVM 模型的应用型高校教师 科研能力评价研究*

包永强^{1*}, 汪木兰², 梁瑞宇¹, 王青云¹, 宋宇飞¹, 朱昊¹

(1. 南京工程学院通信工程学院, 南京 211167; 2. 南京工程学院科技与产业处, 南京 211167)

摘要: 提升学校科研水平关键在于全面评价教师科研能力。文章分析了应用型高校科研评价存在的问题, 构建了基于专家评价应用型高校教师科研能力评价模型, 采用支持向量机 (SVM) 来模拟同行专家评价; 提出了一种广义混合高斯核函数 (GFMG) 用于提高 SVM 模型识别率。通过南京某应用型高校实证分析, 混合高斯核函数能够提高正确识别率达 95.79%, 模型能够有效评价应用型教师科研能力。

关键词: 高校科研管理; 教师科研能力评价; GFMG-SVM; 高斯核函数

中图分类号: G644.4

文献标识码: A

文章编号: 1005-9490(2019)05-1081-04

教师科研能力评价方法主要有两大类: 同行评价法和成果计量法。同行评价法^[1-4]是国外普遍采用的方法, 能够相对客观地反映教师科研活动, 但存在成本高、周期长等不足。国内高校普遍采用成果计量法。指标体系主要包括论文、专著、专利、科研项目、科技成果获奖等方面^[5-6]。成果计量法具有

操作简单优点, 但权值不合理易导致评价误差大, 同时也会引起教师片面追求成果数量^[7]。

本文从模式识别理论出发, 采用 SVM 模型来模拟专家评价结果, 构建教师科研能力指标体系, 将其作为 SVM 模型特征参数。根据从教师科研成果中提取特征参数, 将专家评价结果对 SVM 模型进行训

练和测试。

1 应用型高校教师科研能力评价模型的构建

1.1 评价模型构建

本文根据国内高校科研成果和现行办法,从科研投入、科研成果、人才培养 3 个方面构建应用型高校教师科研评价模型。具体描述如下:学校每隔若干年邀请来自省内外不同层次高校的同学科专家对各学科教师近三年科研业绩进行主观评价,评价等级分为优秀、良好、中等、及格、基本合格和不合格 6 个等级,评价结果作为后几年科研评价系统的训练数据,科技管理系统采用支持向量机 SVM (Support

Vector Machine) 作为训练和识别模型,特征参数采用表 1 给出的应用型高校教师科研评价指标体系。新一年度教师指标完成情况可作为 SVM 模型测试数据,测试结果等级作为年度科研考核或职称评审的依据,年度考核完成后,本年度的教师科研指标更新到训练数据中,重新训练 SVM 模型,同时淘汰部分旧的训练数据。

1.2 应用型高校教师科研评价三级指标体系

表 1 给出了三级指标体系共 56 个指标向量,表 1 中的其他成果主要包括:科研平台、科研团队获批、标准制定、艺术作品获奖等。除科技成果获奖和其他成果外,其他指标必须为第 1 完成人;教师获得科技成果获奖和其他成果得分为教师排序倒数。

表 1 应用型高校教师科研评价三级指标体系

一级指标	二级指标	三级指标
科研投入	科研申报	国家级项目申报 (A01)、省部级项目申报 (A02)、市厅级项目申报 (A03)、科技成果奖申报 (A04)、国家发明专利申请 (A05)
	科研项目获批	国家级项目 (B01)、省部级项目 (B02)、市厅级项目 (B03)、校级项目 (B04)、其他纵向项目 (B05)、纵向科研项目经费 (B06)、横向项目到账经费 (B07)
科研产出	论文论著	学术性专著 (C01)、学术性编著 (C02)、译著 (C03)、国家级专业手册或画册 (C04)、Science/Nature 等国际顶级期刊论文 (C05)、SCI 一区论文 (C06)、SCI 二区论文 (C07)、SCI 三区及以下论文 (C08)、EI 检索期刊论文 (C09)、EI/CPCL-S 会议论文 (C10)、SSCI/A&HCI 期刊论文 (C11)、CSSCI 来源刊论文 (C12)、新华文摘等全文转载 (C13)、新华文摘等部分转载 (C14)、北大核心期刊论文 (C15)、普通期刊论文 (C16)、论文获省部级以上奖励 (C17)、论文获市厅级奖励 (C18)
	知识产权	PCT 专利授权 (D01)、发明专利授权 (D02)、实用新型专利授权 (D04)、软件著作权授权 (D05)、知识产权转让金额 (万元) (D06)
	科技成果获奖	国家级科技成果奖 (E01-E02)、省部级科技成果奖 (E03-05)、国家直管行业级/一级学会科技成果奖 (E06-08)、市厅级科技成果奖 (E09-11)
	其他成果	国家级 (F01)、省部级 (F02)、市厅级 (F03)、校级 (F04)
人才培养	指导学生科创	省级以上大学生科创项目 (G01)、校级大学生科创项目 (G02)、省级以上科技竞赛奖 (G03)、校级科技竞赛奖 (G04)、省级毕业论文奖 (G05)、校级优秀毕业论文奖 (G06)

1.3 应用型高校教师科研能力 SVM 识别模型

支持向量机利用核函数将低维空间线性不可分问题转化到高维空间,建立最优分类超平面使得各群体分离边缘最大化。

$$\begin{cases} \min \frac{\|w\|^2}{2} + C \sum_{i=1}^N \xi_i \\ \text{约束条件: } y_i [(w \cdot x_i) + b] \geq 1 \end{cases} \quad (1)$$

x, w, b_i 分别为样本、权重和偏置。对于线性不可分情况,引入松弛变量 ξ_i, C 为惩罚系数。SVM 一般采用径向基函数作为核函数:

$$K(x_i, x_j) = \exp(-g \|x_i - x_j\|^2) \quad (2)$$

研究表明^[5],SVM 的性能优劣和泛化推广能力受到惩罚系数 C 和核函数系数 g 的影响^[8],可以采用遗传算法 GA (Genetic Algorithm) 来对这两个参数进行优化,如图 1 所示。

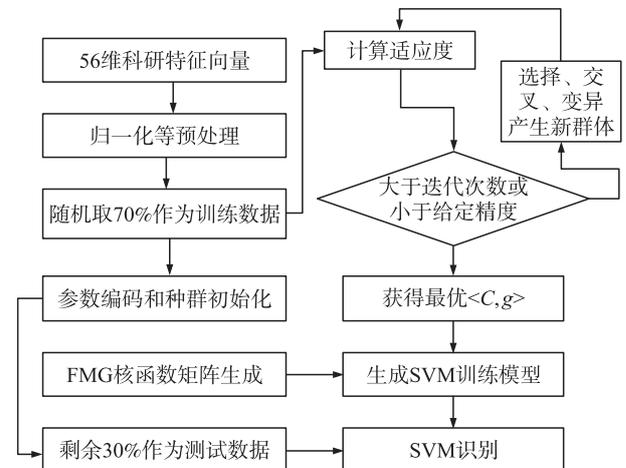


图 1 基于 FMG 核函数的 GA-SVM 模型

1.4 分数混合高斯核函数的提出

表 1 提出的 56 维科研指标向量可作为 SVM 输

人的特征向量,对于 SVM 而言,核函数的选择对识别率影响非常大,在实际过程中一般采用高斯核函数,但目前国内外尚未有对教师科研业绩的分布模型进行研究,高斯核函数不一定是最佳模型。因此,本文将高斯核函数推广至分数域上,寻找符合教师科研业绩分布的更优的核函数。

SVM 核函数必须满足 Mercer's 理论,该理论是 1909 年 Mercer J 提出^[9]的,是积分方程理论中重要的结论,是目前公认的判断是否为核函数的唯一充分条件。目前常用的核函数有线性核函数、多项式核函数、高斯核函数和 sigmoid 核函数。

定理 1 对于所有的平方可积函数 $g(x)$, 存在实函数 $K(x, y)$, 使得下式成立:

$$\iint g(x)K(x, y)g(y) dx dy \geq 0 \quad (3)$$

则 $K(x, y)$ 为核函数。核函数可表示为映射函数 $\Phi(x)$ 之积:

$$K(x, y) = \Phi(x)\Phi(y) \quad (4)$$

SVM 通过核函数将低维线性不可分情况映射到高维空间线性可分,从而达到分类的目的。核函数的选择对 SVM 性能非常重要。由于需要分类的数据统计特性无法预知,采用目前现有的核函数可能达不到最佳分类的目的。高斯核函数是 SVM 中应用最广泛的核函数,本文基于高斯核函数构建新的核函数用于教师科研能力分类。

定义分数混合高斯函数 FMG (Fractional Mixed Gauss function) 如式(5)所示:

$$K(x, y) = x^a y^a \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \quad (5)$$

式中: a 为分数因子。当 $a=0$ 时,该函数为高斯核函数。下面来证明本文定义的分数混合高斯核函数满足 Mercer's 条件。将式(5)代入式(4)可得:

$$\iint g(x)x^a y^a \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right) \exp\left(-\frac{\|y\|^2}{2\sigma^2}\right) \times \exp\left(-\frac{x^T y}{2\sigma^2}\right) g(y) dx dy$$

对式中指数项进行级数展开:

$$\exp\left(-\frac{x^T y}{2\sigma^2}\right) = \sum_{k=0}^{+\infty} \frac{1}{2\sigma^{2k} k!} (x^T y)^k$$

代入式(3)可得:

$$\begin{aligned} & \Rightarrow \iint g(x)K(x, y)g(y) dx dy \\ & = \iint \sum_{k=0}^{+\infty} \frac{1}{\sqrt{2\sigma^{2k} k!}} (x^T)^k g(x) x^a \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right) \times \\ & \quad \sqrt{\frac{1}{2\sigma^{2k} k!}} y^k g(y) y^a \exp\left(-\frac{\|y\|^2}{2\sigma^2}\right) dx dy \end{aligned}$$

$$= \sum_{k=0}^{+\infty} \iint r(x)r(y) dx dy = \sum_{k=0}^{+\infty} \left(\int r(x) dx\right)^2 \geq 0$$

命题得证,显示式(5)为满足 Mercer's 条件的核函数,本文将之作为 SVM 的核函数。

将式(5)进一步推广,可得广义分数混合高斯函数 GFMG (Generalized Fractional Mixed Gauss Function):

$$K(x, y) = x^a y^a \exp\left(-\frac{\|x-y\|^b}{2\sigma^2}\right) \quad (6)$$

特别的,当 $a=0, b=2$ 时,式(6)为普通高斯核函数。

2 实证分析

2.1 实验数据来源

本文实验数据来自南京某应用型高校教学、科研和人事管理部门 2015~2017 年度的师资、科研项目、论文、知识产权、科技成果获奖、大学生科技创新项目、科技创新获奖和毕业设计等 12 000 多条数据。由于原始数据来自多个部门,缺乏完整性和关联性,需要进行关联和核实,数据提取采用 C# 语言编写实现,数据库采用 MSSQL2008,教师科研成果绑定程序编译环境采用 Visual Studio 2010。统计生成 56 维教师科研成果向量,对应于表 2 的 56 维二级指标。

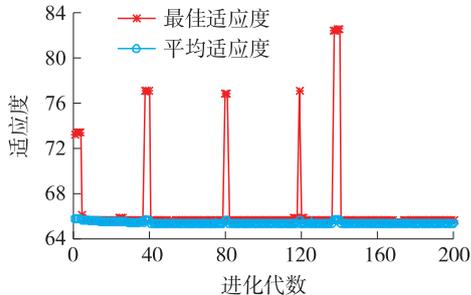
科研管理部门组织专家组针对南京某应用型高校各二级教育教学单位共 1 309 位教师近三年全部科研业绩按照优、良、中等、及格、基本合格、不合格等 6 个等级进行年度科研评价,给出评价等级。随机选择 56 维教师科研数据的 70% 作为训练数据,另外 30% 作为测试数据。教师科研能力评价系统将这 56 维教师科研指标作为特征参数,采用 GA-SVM 神经网络进行训练和识别。SVM 核函数采用本文提出的分数混合高斯核函数,需要利用 SVM 自定义核函数模式。最佳因子 C, g 采用遗传算法进行搜索,最大迭代数为 200,种群最大数量为 50,交叉概率为 0.4,参数 C 的变化范围为 [0.1 1 000], 参数 g 的变化范围为 [0.01 100]。

2.2 基于分数混合高斯核函数的 GA-SVM 性能分析

图 2、图 3 给出了基于高斯核函数和基于混合高斯核函数 (FMG, 因子 $a=0.4$) 的 GA-SVM 训练模型,图 2 中 $C=903.782, g=0.235 7$, 训练精度达 87%; 图 3 中 $C=454.265 8, g=0.113 72$, 训练精度达 91.7%。中可以看出,高斯核函数的训练精度要小于混合高斯核函数的精度。

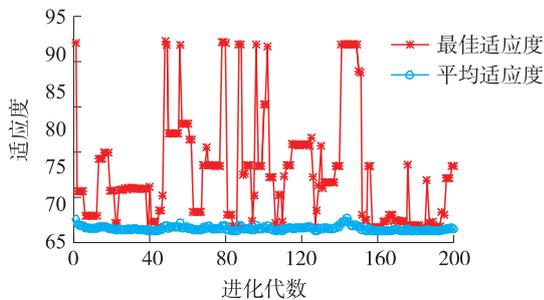
图 4、图 5 给出了在最佳 $\langle C, g \rangle$ 参数 (通过遗传算法寻优获得) 下不同因子 a, b 对应的教师科研能力评

价模型的识别率,在图 4 中 $b=2$,可以看出当 $a=0.6$ 时,可以获得最佳的识别率 88.67%;图 5 中 $a=0.6$,当 $b=3.5$ 时可以获得最佳的识别率 95.79%,要高于普通高斯核函数($a=0, b=2$)的 86.41%。



适应度曲线 Accuracy [GAMethod]; 终止代数 = 200, 种群数量 pop = 50; Best $c = 903.782$ $g = 0.2357$ CV Accuracy = 87%

图 2 基于高斯核函数的 GA-SVM 训练模型



适应度曲线 Accuracy [GAMethod]; 终止代数 = 200, 种群数量 pop = 50; Best $c = 454.2658$ $g = 0.11372$ CV Accuracy = 91.7%

图 3 基于 GFMG 核函数的 GA-SVM 训练模型

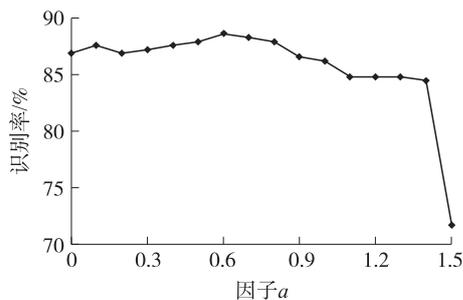


图 4 不同因子 a 对应的 SVM 识别率

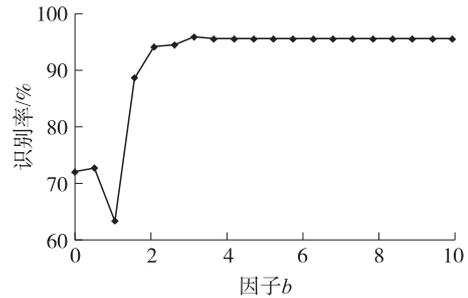


图 5 不同因子 b 对应的 SVM 识别率

3 结论

同行评价是高校科研评价的趋势,采用本文提出的 GFMG+SVM 结合应用型高校教师科研能力评价模型融合了同行评价和成果计量法的特点,既考虑了同行评价模式,又大大降低了运行成本和周期,通过实证研究发现能够有效提高识别率达 95.79%,比采用普通高斯核函数的 SVM 模型提高了 10.9%。

参考文献:

- [1] 王学松. 试析美国高校教师科研绩效评价机制[J]. 合作经济与科技, 2013, 23(20): 40-41.
- [2] 许春东. 美国大学教师评价制度的特点及对我国的启示[J]. 兵团教育学院学报, 2012, 22(6): 38-43.
- [3] 刘兴凯, 梁珣. 英国高校科研水平评估及其对我国的启示[J]. 当代教育科学, 2014, 26(23): 51-55.
- [4] 韩小娇, 高军. 日本大学教师学术评价的理念、内容及方法[J]. 高教发展与评估, 2013, 29(2): 64-71, 107.
- [5] 史万兵, 杨慧. 高等学校教师科研绩效评价方法研究[J]. 高教探索, 2014, 6(6): 112-117.
- [6] 郭涛, 林盛, 刘金培. 高校教师科研绩效评价: 一种多准则决策分析模型[J]. 统计与决策, 2012, 9(9): 66-69.
- [7] 包永强, 梁瑞宇, 宋宇飞. 应用型高校教师科研评价现状与对策[J]. 南京工程学院学报(社会科学版), 2017, 17(3): 67-70.
- [8] 谢国民, 单敏柱, 刘明. 煤与瓦斯突出强度的 FOA-SVM 预测模型与应用[J]. 传感技术学报, 2016, 29(12): 1941-1946.
- [9] Mercer J. Functions of Positive and Negative Type and Their Connection with the Theory of Integral Equations[J]. Philosophical Transactions of the Royal Society A, 1909, 209(441-458): 415-446.



包永强(1973-),男,汉族,江苏江阴人,现为南京工程学院信息与通信工程学院院长,教授,博士,硕士生导师,主要研究方向为音频取证、高校科研管理等, jybyq@163.com。