

# 基于文本的抑郁情感倾向识别模型<sup>①</sup>

施志伟, 高俊波, 胡雯雯, 刘志远

(上海海事大学, 上海 201306)

**摘要:** 针对学生在新浪微博文本中所表现出来的抑郁情感倾向, 提出了一种识别抑郁情感倾向的模型. 通过在本校广泛发动学生在线填写抑郁自评量表, 获得学生的量表得分. 采集学生的微博文本, 并请本校心理学老师对微博进行人工标注. 在预处理阶段, 利用抑郁情感词典重新组合在分词阶段被拆分的抑郁情感词, 以提高识别正确率. 然后基于支持向量机构建一个情感分类器对微博数据进行训练, 经过不断的学习反馈, 获得较好的分类效果; 最后, 定义了抑郁指数来衡量个体在一段时间内的抑郁倾向程度. 实验结果表明, 抑郁指数衡量的抑郁程度大致与量表结果吻合, 该方法识别准确率达到 82.35%.

**关键词:** 抑郁倾向识别; 抑郁自评量表; 抑郁情感词典; 支持向量机; 抑郁指数; 新浪微博

引用格式: 施志伟, 高俊波, 胡雯雯, 刘志远. 基于文本的抑郁情感倾向识别模型. 计算机系统应用, 2017, 26(12): 155-159. <http://www.c-s-a.org.cn/1003-3254/6088.html>

## Depression Tendency Identification Model Based on Text Content Analysis

SHI Zhi-Wei, GAO Jun-Bo, HU Wen-Wen, LIU Zhi-Yuan

(Shanghai Maritime University, Shanghai 201306, China)

**Abstract:** In order to solve the problem of identifying depression tendency among students on sina microblog platform, this paper proposes a depression tendency identification model. By inviting students widely to fill in the self-rating depression scale online on campus we can get the students' score. We collect students' microblog text and ask the psychology teacher to annotate the microblog artificially. In the pretreatment stage, we use the depression emotional dictionary to reassemble the depressed emotion words that are split at the segmentation stage so as to improve the recognition accuracy rate. And then we build a classifier based on the support vector machine to train the data. Through continuous learning and feedback, we get a better classification result. Finally, this paper defines the depression index and uses it to measure the degree of depression for a period of time. The experimental results indicate that the degree of depression measured by depression index is approximately consistent with the results of the scale, the accuracy of the method being 82.35%.

**Key words:** depression tendency identification; self-rating depression scale; depression emotional dictionary; support vector machine (SVM); depression index; sina micro-blog

微博作为一种开放化的互联网社交服务, 由于文本格式自由、使用方便, 越来越多的用户通过微博发布自己的所见所闻, 所感所思. 在海量微博文本中, 有很多蕴藏着个人情感, 可以利用这些文本进行情感分析研究<sup>[1]</sup>. 当前, 情感分析研究主要是针对某个具体事

物的评论<sup>[2]</sup>, 如影评, 产品评论等, 而针对文本中所蕴藏的抑郁情感的研究还比较少.

在世界范围内, 抑郁症是人们面临的最为普遍的心理疾病之一. 随着微博等社交网络平台的发展壮大, 许多研究者借助用户网络特征行为来判别其心理抑郁

<sup>①</sup> 基金项目: 上海海事大学研究生创新基金 (2016ycx036)

收稿时间: 2017-03-09; 修改时间: 2017-03-27; 采用时间: 2017-03-31

情况<sup>[3]</sup>. 文献[4]基于产妇在产后社交网络中对话频度、语言风格等进行分析,建立产妇抑郁统计模型. Wang<sup>[5]</sup>等将抑郁患者视为一个节点,并以此为中心构建一个图网络,根据网络中相邻节点的属性及连接权重,给出模型来计算抑郁状况. 文献[6]采用脑成像方法研究静息态下不同性别的抑郁症患者的脑功能差异. 文献[7]从用户发微博的时间、粉丝数及关注数等方面来分析用户的抑郁情况.

本文将对学生微博文本中所蕴含的抑郁情感倾向展开研究,并定义抑郁指数来衡量个体在一段时间内的抑郁倾向程度,为高校的心理工作者及医院医护人员识别抑郁患者提供辅助手段.

## 1 相关工作

### 1.1 抑郁症

抑郁症是一种普遍的心理疾病,成因非常复杂,研究者对抑郁症的发病机理提出了许多理论假说<sup>[8]</sup>. 心理及医学研究者还提出了各种抑郁症诊断量表,为相关实践提供了重要的实验依据. Zung<sup>[9]</sup>提出的抑郁自评量表具有高度可操作性及适应性,很多医疗机构也使用此量表来度量患者的抑郁程度. 它根据得分将抑郁情绪分为四类, [20, 41]表示正常, [42, 49]表示轻度抑郁, [50, 57]表示中度抑郁, [58, 80]表示重度抑郁. 本文采用 SDS 抑郁自评量表做对比实验.

### 1.2 构建抑郁情感词典

目前,开源情感词典众多<sup>[10]</sup>,但还没有一部完整的抑郁情感词典. 本文在已有情感词典基础上,构建一个基础的抑郁情感词典,共统计常见抑郁词 1041 个. 考虑到微博用语的特征及时代性,采集抑郁情感网络流行词语 54 个;从新浪微博抽取 26 个常用抑郁情感符号加入抑郁情感词典. 抑郁情感词典中部分词如表 1 所示.

表 1 抑郁情感词典

|         |   |
|---------|---|
| 词典抑郁情感词 | 煎熬, 压抑, 郁闷, 寂寞, 无眠, 厌恶, 鄙视, 悲伤, 悲哀, 大吼        |
| 网络抑郁情感词 | 傻逼, 草泥马, 奇葩, 然并卵, 累成狗, 废柴                     |
| 表情抑郁情感词 | [泪], [躁狂症], [崩溃], [抓狂], [怒], [别烦我], [衰], [生病] |

## 2 抑郁情感倾向识别模型

本文建立的抑郁情感倾向识别模型主要包括以下几个部分,如图 1 所示.

采集数据之前,在上海海事大学校内广泛发动学

生在线(网页或者微信)填写抑郁自评量表 (<https://sojump.com/jq/9743549.aspx>),获得每个学生的抑郁倾向得分. 在学生知情并授权的情况下采集微博语料,并请本校的心理咨询师对学生的微博文本进行逐条标注. 用预处理后的训练样本训练分类器,得到可靠的抑郁情感识别器,并对抑郁指数和量表得分进行皮尔逊相关性分析,最后给出抑郁指数和抑郁程度之间的关系式.

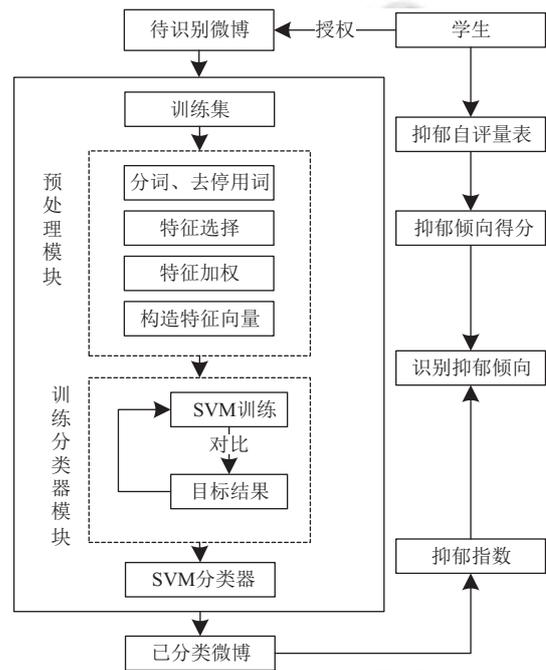


图 1 学生抑郁情感倾向识别模型

### 2.1 预处理

#### 2.1.1 去停用词

对微博文本构造特征向量前,先进行分词和去停用词处理,本文使用中科院分词系统 ICTCLAS. 在分词后,许多抑郁情感词被拆分,如“草泥马”被分割成“草/泥/马”,为了更好保留抑郁情感词,参照已经构建的抑郁情感词典,将这类词恢复原状. 去停用词是利用构建好的停用词表过滤掉无用的字词. 鉴于微博文本特点及多次去停用词的结果,本文在哈工大去停用词表的基础上,加入了一些特殊的人名、地名及无关的字词,如“上海”、“TFBOY”等. 经过反复试验,此方法有更好的实验效果.

#### 2.1.2 特征选择

特征空间的高维性和稀疏性是文本分类中面临的重大问题. 将特征项全部保留,这将导致维度太大,降低分类效率和准确率,因此必须进行特征降维. 本文采

用 CHI 进行特征选择<sup>[11]</sup>, 其主要思想是假设特征  $t$  和类别  $c_i$  之间符合 CHI 分布, CHI 统计值越大, 特征与类别之间的相关性越强, 对类别的贡献度越大. 计算公式如式 (1) 所示.

$$\text{CHI}(t, c_i) = \frac{N \times (AD - BC)^2}{(A + C) \times (B + D) \times (A + B) \times (C + D)} \quad (1)$$

最后, 取特征  $t$  的最大值作为其全局 CHI 统计量, 如式 (2).

$$\text{CHI}_{\max}(t) = \max_{i=1}^{|c|} \{\text{CHI}(t, c_i)\} \quad (2)$$

### 2.1.3 特征加权

文本特征的权重综合反映了该特征区分文本类别的能力, 特征权重的计算对文本分类有着极其重要的作用. 本文采用文本处理领域经典的 TF-IDF 权重计算方法<sup>[12]</sup>. 特征权重  $W_{ik}$  计算如公式 (3):

$$W_{ik} = tf_{ik} \cdot idf_{ik} \quad (3)$$

其中,  $tf_{ik}$  表示特征词  $t_i$  在文本  $d_k$  中出现的次数,  $idf_{ik}$  表示特征词  $t_i$  的反文档频率.  $idf_{ik}$  的计算一般会设置一个常数以保证平滑, 考虑到某个特征词在某类文档中可能并没有出现, 给出  $idf_{ik}$  计算公式 (4), 其中  $D_k$  表示文本集中所有文本数,  $d_{ik}$  表示文本集中包含特征  $t_i$  的文本数:

$$idf_{ik} = \log_2 \left( \frac{D_k}{d_{ik} + 0.5} + 0.01 \right) \quad (4)$$

则特征词  $t_i$  在文档  $d_k$  中的权重为:

$$W_{ik} = tf_{ik} \cdot \log_2 \left( \frac{D_k}{d_{ik} + 0.5} + 0.01 \right) \quad (5)$$

### 2.1.4 构造特征向量

本文采用向量空间模型对微博文本进行刻画. 在向量空间模型中, 文本被视为由特征权重构成的特征矩阵, 关于文本的所有处理都在向量空间上进行. 根据特征选择和特征加权的結果, 微博文本将被表示成如下的向量空间模型.

$$\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1m} \\ w_{21} & w_{22} & \cdots & w_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix}$$

每一行表示一条微博,  $w_{ij}$  表示第  $i$  条微博的第  $j$  个特征词的权重;  $n$  表示文本集中微博总数,  $m$  表示每条微博的特征维数.

微博长短不一, 按向量空间模型构造的矩阵是一个有大量特征值为零的稀疏矩阵. 为了避免浪费存储

空间, 提高运算效率, 本文采用 (L T:W) 形式来表示每条微博, 其中 L 表示每条微博的标签, T 表示特征项, W 为特征项的权重. 如: “我真的爱你, 闭上眼, 以为我能忘记, 但流下的眼泪, 却没有骗到自己”, 这句话特征选择后得到 5 个特征词, 分别为“骗”、“爱”、“眼泪”、“闭上眼”、“流下”. 故而这句话可以表示成“1.0 28:0.4528 39:0.2295 49:0.3215 862:0.5811 1832:0.54878”, 其中 1.0 表示标签, 28 是特征词“骗”的索引号, 0.4528 是特征词的权重.

## 2.2 SVM 分类器的构建

支持向量机是一种实用高效的机器学习方法<sup>[13,14]</sup>, 和其他方法相比, 它在处理非线性及高维分类问题中表现出其特有的优势. 处理非线性的文本分类问题, 首先构造非线性映射函数将数据  $x$  变换到高维空间 F, 然后在高维空间 F 使用线性分类器分类. 公式 (6) 实现把数据从低维空间转换到高维空间 F.

$$K(x_i \cdot x_j) = \varphi(x_i) \varphi(x_j) \quad (6)$$

在空间变换后, 采用高斯核函数实现非线性问题线性化. 高斯核函数如下<sup>[15]</sup>:

$$K(x \cdot x_i) = \exp \frac{-\|x - x_i\|^2}{2\sigma^2} \quad (7)$$

在高维空间的数据样本线性可分时, SVM 通过计算最小化  $\|w\|^2/2$  来获得最优超平面. 同时, 考虑到部分数据点可能偏离超平面, 故在已有的约束条件基础上引入松弛变量  $\xi_i$ , 约束条件如公式 (8) 所示.

$$y_i(w^T x_i + b) \geq 1 - \xi_i \quad i = 1, 2, \dots, n \quad (8)$$

此时若对  $\xi_i$  取任意值, 则任意的超平面都是符合要求的. 因此, 需要在原最小化函数  $\|w\|^2/2$  后加上一项, 使得所有  $\xi_i$  的和也是最小的, 新目标函数如公式 (9).

$$\min \left\{ \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right\} \quad 0 \leq \alpha_i \leq C \quad (9)$$

其中  $C$  为参数, 用于控制上述目标函数中两项之间的权重. 此外, 将约束条件加入到目标函数中, 构造拉格朗日函数, 解得  $0 \leq \alpha_i \leq C$ ,  $\alpha_i$  表示约束条件所对应的拉格朗日乘子. 相应的分类函数如公式 (10).

$$f(x) = \sum_{i=1}^n \alpha_i y_i K(x \cdot x_i) + b \quad (10)$$

其中,  $x$  表示待分类微博文本,  $x_i$  表示对文本的支持向量,  $y_i$  为  $x_i$  对应的分类, 当  $f(x) \geq 0$ , 则  $x$  有抑郁倾向,  $f(x) < 0$  则正常.

## 2.3 抑郁指数

对微博文本分类后, 只能看出单条微博的抑郁情况, 而不能刻画个体在一段时间内整体抑郁倾向状况.

通过观察、比较大量学生的抑郁微博数,在上述抑郁情感识别模型基础上,本文提出抑郁指数来衡量个体在一段时间内的抑郁倾向程度,计算公式如(11).其中 $N_d$ 表示分类结果中抑郁微博条数, $N_t$ 表示总的微博数,DI的值越大表示抑郁情感倾向越严重.

$$DI = \frac{N_d}{N_t} \quad (11)$$

鉴于本文仅从微博文本这一特征来分析个体抑郁倾向情况,不涉及微博关注数、粉丝数等其他结构特征,故而只取抑郁微博数与总微博数的比值作为抑郁指数.

### 3 实验及结果分析

#### 3.1 语料收集与标注

从2016年9月22到2016年12月23号,共381名学生在线作答,剔除无效量表,实际有效271人.从得分结果来看,有抑郁倾向80人,正常191人.按照3:1的比例,分别从两类学生中随机抽取训练用例和测试用例,训练集和测试集中人数如表2所示.按照每人采集500条,不足采集全部的原则,共获得7321条微博文本.对训练集中微博语料进行人工标注后,共获得抑郁微博1512条,正常微博3786条;测试集共2023条微博.

表2 训练集和测试集人数及微博数

|         | 训练集  | 测试集  |
|---------|------|------|
| 有抑郁情感倾向 | 60   | 20   |
| 正常情感    | 143  | 48   |
| 总微博数    | 5298 | 2023 |

#### 3.2 评价标准

本文采用分类正确率来评价分类效果,公式如下:

$$Ac = \frac{a+c}{a+b+c+d} \quad (12)$$

其中, $a$ 表示被正确判断为抑郁的微博数, $b$ 表示被误判为抑郁的微博数, $c$ 表示被正确判断为正常的微博数, $d$ 表示被误判为正常的微博数.

#### 3.3 结果分析

实验1.对测试集中2023条微博,分别在采用抑郁情感词典和不采用抑郁情感词典的情况下进行实验,结果如表3所示.

从表3可以看出,采用人工构建的抑郁情感词典后,对单条微博的识别正确率显著提高.

实验2.利用分类器对测试集中每个学生的微博文本进行分类,得到68个测试用例的抑郁指数和正确率,部分数据如表4所示.

表3 采用抑郁情感词典的实验准确率

| 训练样本 | 未采用词典(%) | 采用词典(%) |
|------|----------|---------|
| 2000 | 76.62    | 80.58   |
| 3000 | 77.83    | 85.68   |
| 4000 | 78.85    | 87.36   |
| 5298 | 79.45    | 90.04   |

对实验结果中抑郁指数DI和量表得分Score进行分析,发现它们之间有一定的联系,画出两者之间的散点图,如图2,其中Index值为1表示有抑郁倾向,Index值为-1表示正常.对两者之间的相关性进行皮尔逊检测,发现DI和Score在0.01水平(双侧)上显著相关,且 $r=0.544$ ,说明两者有较强的相关性.根据相关性检测结果以及量表得分与抑郁程度的关系,给出抑郁指数与抑郁程度之间的关系式为公式(13).

表4 不同学生抑郁指数及抑郁量表得分情况

|    | 微博昵称      | 抑郁指数   | 量表得分 | 正确率    |
|----|-----------|--------|------|--------|
| 正常 | 盛夏SQH     | 0.10   | 34   | 0.96   |
|    | 谭华要做江南小清新 | 0.0402 | 33   | 0.9821 |
|    | 小疯子白日梦想家  | 0.0566 | 34   | 0.9811 |
|    | 夜夜夜先森     | 0.1096 | 31   | 0.9629 |
|    | 真菌菌凹      | 0.2353 | 36   | 0.8118 |
| 抑郁 | 我是一把宝剑    | 0.1111 | 45   | 0.9630 |
|    | 胖胖是笑声     | 0.2524 | 49   | 0.7865 |
|    | Arioster  | 0.1724 | 42   | 0.8793 |
|    | 图拉拉树袋熊    | 0.2619 | 49   | 0.9048 |
|    | 小桑meng    | 0.1395 | 45   | 0.9767 |

$$E(DI) = \begin{cases} normal & DI \in [0, 0.1] \\ mild & DI \in (0.1, 0.2] \\ medium & DI \in (0.2, 0.3] \\ severe & DI \in (0.3, 1] \end{cases} \quad (13)$$

抑郁指数在 $[0, 0.1]$ 时,情绪状态正常;抑郁指数在 $(0.1, 0.2]$ 时,开始有抑郁倾向但处于轻度状态;抑郁指数在 $(0.2, 0.3]$ 时,处于中度抑郁状态;当抑郁指数在 $(0.3, 1]$ 时,抑郁情况已经比较严重了,必须采取有效的治疗手段.其中,当DI在 $[0.08, 0.15]$ 时,正常与抑郁的交叉现象很明显,这与部分学生量表得分在42分左右有明显的关系.

表5给出了68个测试用例的识别结果,正常情绪的识别准确率比抑郁倾向的识别准确率低,可能是由于学生在正常状态下所发的微博中包含了相关抑郁词.

实验3.从2016年3月到2017年2月,采集3位学生在此期间所有微博数据,以月为单位分别求出抑郁指数并画出走势图,结果如图3.

从图3可以看出,抑郁倾向呈现出波动状态,但在一定时间范围内,有抑郁倾向的学生的抑郁指数普遍

比正常人要高,而且情绪波动更剧烈。

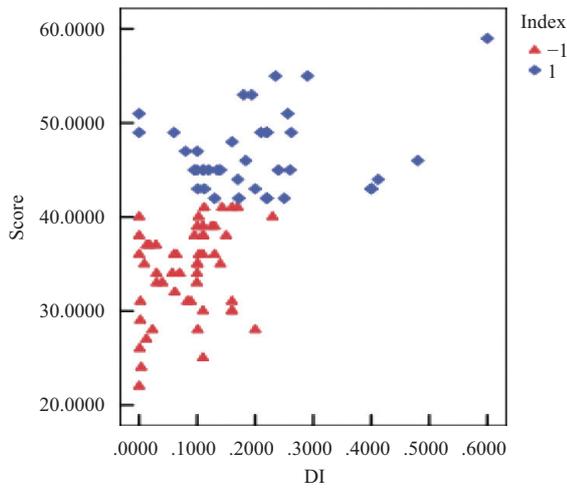


图2 DI与Score的关系散点图

表5 学生测试用例的分类结果

| 抑郁倾向识别成正常 | 正常识别成抑郁倾向 | 正确率    |
|-----------|-----------|--------|
| 3         | 9         | 82.35% |

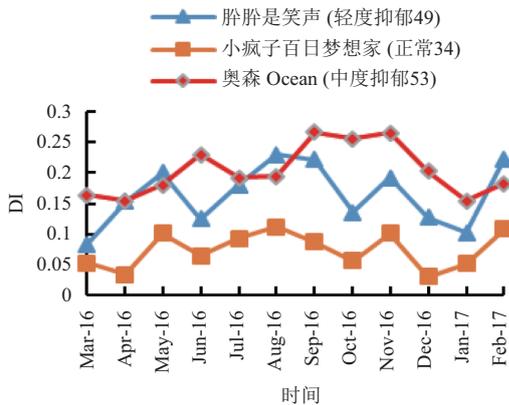


图3 不同学生的抑郁倾向走势图

#### 4 结语

抑郁情感倾向分析试图从微博文本中识别出个体在某段时间内的抑郁倾向状况。在学生知情且授权的情况下,获取学生的微博文本;在学校心理咨询师的帮助下对微博文本进行情感极性标注,并使用机器学习算法训练分类器,实现面向微博文本的学生抑郁情感倾向识别。实验结果表明,该方法对学生的抑郁情感倾向有较好的识别效果。

鉴于微博文本语言极具特色,不少学生的微博都是通网络图片或表情来传达情感,这对深入研究学生的抑郁情感倾向提出了非常高的要求。此外,仅仅通过

微博文本来分析抑郁倾向情况,而未考虑微博周边其他一些因素,如关注数、粉丝数、微博发表时间等,这些都是需要继续研究改进的地方。

#### 参考文献

- 刘楠. 面向微博短文本的情感分析研究[博士学位论文]. 武汉: 武汉大学, 2013: 24-91.
- Pang B, Lee L, Vaithyanathan S. Thumbs up? Sentiment classification using machine learning techniques. Proc. of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10. Stroudsburg, PA, USA. 2002. 79-86.
- Youn SJ, Trinh NH, Shyu I, et al. Using online social media, Facebook, in screening for major depressive disorder among college students. International Journal of Clinical and Health Psychology, 2013, 13(1): 74-80. [doi: 10.1016/S1697-2600(13)70010-3]
- De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. Proc. of the SIGCHI Conference on Human Factors in Computing Systems. New York, USA. 2013. 3267-3276.
- Wang XY, Zhang CH, Ji Y, et al. A depression detection model based on sentiment analysis in micro-blog social network. Proc. of the PAKDD 2013 International Workshops on Trends and Applications in Knowledge Discovery and Data Mining-Volume 7867. New York, USA. 2013. 201-213.
- 王丽, 姚志剑, 滕皋军, 等. 静息态下不同性别抑郁症患者脑功能及其差异的fMRI研究. 中国心理卫生杂志, 2008, 22(4): 271-275.
- 李鹏宇. 微博社交网络中的学生用户抑郁症识别方法研究[硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2014. 17-60.
- 郭珊, 郭克锋. 抑郁症的研究进展. 中国临床康复, 2005, 9(4): 131-133.
- Zung WWK, Richards CB, Short MJ. Self-rating depression scale in an outpatient clinic: Further validation of the SDS. Archives of General Psychiatry, 1965, 13(6): 508-515. [doi: 10.1001/archpsyc.1965.01730060026004]
- 徐琳宏, 林鸿飞, 潘宇, 等. 情感词汇本体的构造. 情报学报, 2008, 27(2): 180-185.
- 谭松波. 高性能文本分类算法研究[博士学位论文]. 北京: 中国科学院计算技术研究所, 2005.
- Aizawa A. An Information-theoretic perspective of tf-idf measures. Information Processing & Management, 2003, 39(1): 45-65.
- 张学工. 关于统计学习理论与支持向量机. 自动化学报, 2000, 26(1): 32-42.
- Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer, 1995.
- 郭丽娟, 孙世宇, 段修生. 支持向量机及核函数研究. 科学技术与工程, 2008, 8(2): 487-490.