

基于特征内相关和互信息的加权 SVM 算法

彭晓冰^{1,2} 朱玉全¹

(江苏大学计算机科学与通信工程学院 江苏 镇江 212013)¹ (江苏大学信息化中心 江苏 镇江 212013)²

摘要 特征加权支持向量机没有考虑特征间的相关性,因此产生的冗余会形成干扰并对最后的分类结果产生负面影响。为解决这个问题,提出了一种基于特征内相关和互信息的特征加权算法,并将其应用于支持向量机。该算法引入了特征间相关系数作为衡量冗余度的一个指标,以此计算出惩罚因子,在特征加权向量机的基础上对权值进行处理,尽可能真实地体现出特征对分类的贡献度。经过多个数据集以及几种不同算法的实验比较,提出的新算法具有更好的鲁棒性和泛化能力。

关键词 特征加权,互信息,相关系数,惩罚因子,支持向量机

中图分类号 TP309 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.12.029

Weighted Support Vector Machine Algorithm Based on Inner-correlations and Mutual Information of Features

PENG Xiao-bing^{1,2} ZHU Yu-quan¹

(School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, Jiangsu 212013, China)¹

(Information Center, Jiangsu University, Zhenjiang, Jiangsu 212013, China)²

Abstract The feature-weighted support vector machine (FWSVM) does not take into account the correlation among the features, thus the redundancy and the interference caused by it will have a negative impact on the final classification result. A feature weighting algorithm based on inner-feature correlation and mutual information was proposed and applied in support vector machines. The algorithm introduces the inter-feature coefficient as an index to measure the redundancy, and then calculates the penalty factor to deal with the weight on the basis of the feature weighting vector machine. Thus it realizes the contribution of the feature to the classification as much as possible. The comparison of experiments on multiple data sets with several different algorithms shows that the proposed new algorithm has better robustness and generalization ability.

Keywords Feature weighting, Mutual information, Correlation coefficient, Penalty factor, Support vector machine

1 前言

分类是数据挖掘中的一个重要分支,大多数的分类方法都需要用到一个特征的集合来达到分类的目的,可见特征提取和特征选择是分类的基础,很大程度上决定了分类的准确度。特征提取总是尽可能地将对象的特征全面且完整地提取出来,而特征选择是根据不同的分类目的,选择出对目标有用的特征组成训练集来作为分类模型的输入,这也避免了不相关特征对分类性能的负面影响。然而,在这些选择出的特征中,不同特征对于分类的贡献是不同的。显而易见,若将每个特征对分类的重要程度视为相同,则得到的结果将不够准确。特征加权算法就是尽可能地估计出特征对分类的贡献程度,量化后作为权值加载在对应的特征上^[1],以此作为分类模型的输入,从而生成最优分类器,提高分类器的泛化能力^[2-3]。

汪廷华等^[4]和 Xing 等^[5]基于信息论^[6-7]提出了 FWSVM

(Feature-Weighted Support Vector Machine)方法,以特征和分类的互信息作为权重应用在 SVM(Support Vector Machine)分类方法上,获得了更高的分类准确度。Giveki 等^[8]结合 MCS^[9](Modified Cuckoo Search)改进了 FWSVM 并将其具体应用于糖尿病诊断。有的研究人员则用信息增益作为特征权重,分别改进了朴素贝叶斯算法和 KNN(k-Nearest Neighbor)算法,并将其应用于文本分析以及股市分析,取得了良好的效果^[10-11]。相关系数反映两个变量之间的相互关系,可用于衡量特征对于分类的重要程度,文献[12-13]以特征与分类间的相关系数为权重,用神经网络、KNN 和 DGC(Data Gravitation Classification)等算法做测试,证明了这种特征权重算法提高了分类准确率。此外,还有研究人员基于粗糙集的分辨矩阵和训练数据的分布距离,提出了各种特征加权算法^[14-17],有效地降低了不相关特征对分类的负面影响。

收稿日期:2017-11-28 返修日期:2018-01-24 本文受江苏省自然科学基金(BK20150531),江苏省高校研究生科研创新资助项目(1291170028)资助。

彭晓冰(1977-),男,博士生,实验师,主要研究方向为数据挖掘、隐私保护和计算机视觉等,E-mail:pxb@ujs.edu.cn;朱玉全(1965-),男,博士,教授,博士生导师,主要研究方向为数据挖掘、模式识别、人工智能等,E-mail:yqzhu@ujs.edu.cn(通信作者)。

然而,上述各种方法仅考虑了特征与分类之间的关系,未考虑到特征之间的关联也会对分类结果造成影响,这些特征内的冗余信息会导致过拟合。本文提出了一种基于互信息和相关系数特征加权的支撑向量机(Mutual Information and Correlation Based Feature Weighted SVM, MCB-FWSVM)。MCB-FWSVM 首先用互信息评估特征空间中各个特征对分类的贡献度,然后计算特征空间的相关系数矩阵,据此重新衡量各个特征的权重,最后将获得的权重值应用到 SVM 核函数的计算中,从而避免了生成的分类器被不相关特征以及冗余特征所支配。

本文第 2 节介绍了基于互信息的特征支撑向量机;第 3 节提出了一种新的特征加权支撑向量机,介绍了惩罚因子的计算方法;第 4 节通过对 UCI 的 13 个数据集的实验,对新算法、FWSVM 以及几种传统分类算法进行了比较;最后总结全文并展望未来。

2 基于互信息的特征加权向量机

2.1 互信息

互信息(Mutual Information)是信息论中一种有用的信息度量,它可以看作是一个随机变量中包含的关于另一个随机变量的信息量,或者说是一个随机变量由于已知另一个随机变量而减少的不肯定性。

定义 1 设两个随机变量 (x, y) 的联合分布为 $p(x, y)$, 边际分布分别为 $p(x)$ 和 $p(y)$, 互信息 $MI(X, Y)$ 是联合分布 $p(x, y)$ 与乘积分布 $p(x)p(y)$ 的相对熵,即:

$$MI(X, Y) = \sum_{x \in X, y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (1)$$

设 F 为特征空间 F 中特征值的集合, C 为类标值的集合,由定义 1 可知各个特征和分类的互信息为:

$$MI(C, F) = \sum_{f \in F, c \in C} p(f, c) \log_2 \frac{p(f, c)}{p(f)p(c)} \quad (2)$$

由式(2)可知,一个特征和分类之间共享的信息量越多,该特征的权重就越大,如果一个特征独立于分类,那么它与分类的互信息就为 0,即该特征的权重为 0。

2.2 特征加权支撑向量机——FWSVM

支撑向量机(Support Vector Machine, SVM)是在高维特征空间中使用线性函数假设空间的学习系统,它由一个来自最优化理论的学习算法训练,该算法实现了一个由统计学习理论导出的学习偏置。

设样本集为 (x_i, y_i) , $x_i \in R^d$, $y_i \in \{0, 1\}$ 是类别标签, $i = 1, 2, \dots, n$ 。SVM 方法就是在 d 维空间中寻找一个形为 $\omega x + b = 0$ 的最优分类面,其中, ω 为权重向量, $b \in R$ 为偏置。

其决策函数为:

$$f(x) = \text{sgn} \sum_{i=1}^n \alpha_i y_i K(x, x_i) + b \quad (3)$$

特征加权支撑向量机以特征空间中各个特征与分类的互信息值作为权重,即 $\omega_i = MI(C, f_i)$, 由此获得特征加权矩阵:

$$P = \begin{bmatrix} \omega_1 & & & \\ & \omega_2 & & \\ & & \dots & \\ & & & \omega_d \end{bmatrix}$$

其中, P 为 d 阶对角矩阵。 $(P_{ii}) = \omega_i (1 \leq i \leq d)$ 代表第 i 个特征的权值。由式(3)可得特征加权向量机的决策函数为:

$$f(x) = \text{sgn} \sum_{i=1}^n \alpha_i y_i K(x^T P, x_i^T P) + b \quad (4)$$

3 基于特征内相关和互信息的特征加权支撑向量机——MCB-FWSVM

本文采用皮尔逊相关系数作为反映特征之间相关关系密切程度的统计指标。特征内的相关系数 r_{ij} 为:

$$r_{ij} = \frac{\text{cov}(f_i, f_j)}{\delta_{f_i} \delta_{f_j}} \quad (5)$$

其中, $\text{cov}(f_i, f_j)$ 为 f_i 与 f_j 的协方差, δ_{f_i} 为 f_i 的标准方差, δ_{f_j} 为 f_j 的标准方差。特征间的相关系数越大,存在的冗余就越多。皮尔逊相关系数的取值为 $[-1, 1]$, 负值代表负相关,正值代表正相关,而对于本文来说,只考虑特征间的相关程度,因此取其绝对值 $|r_{ij}|$ 作为指标。特征 f_i 与特征空间 F 中其余特征的相关系数之和代表了 f_i 的冗余程度,我们采用其平均值作为衡量该特征的冗余量,即:

$$\bar{r}_{ij} = \frac{1}{d-1} \sum_{j \in F'} |r_{ij}| \quad (6)$$

其中, $F' = F / \{f_i\}$ 。当 $\bar{r}_{ij} = 0$ 时,特征 f_i 与其余特征完全独立,这时 $MI(C, f_i)$ 代表了 f_i 对分类的真实贡献度;当 $\bar{r}_{ij} \neq 0$ 时,值越大,该特征相对于其他特征对分类的真实贡献就越小。还存在一种特殊情况,当 $|r_{ij}| = 1$ 时,说明特征 f_i 与 f_j 完全相关,这时 f_i 对分类的贡献是多余的。综上,使用惩罚因子 β_i 来表示特征 f_i 具有的冗余量对分类造成的负面影响, β_i 如式(7)所示:

$$\beta_i = \begin{cases} 0, & \exists r_{ij} = 1 \\ \frac{1}{1 + (d-1)r_{ij}}, & \text{otherwise} \end{cases} \quad (7)$$

改进后的权重 $\omega_i^* = \beta_i MI(C, f_i)$ 。特征加权矩阵为:

$$P^* = \begin{bmatrix} \omega_1^* & & & \\ & \omega_2^* & & \\ & & \dots & \\ & & & \omega_d^* \end{bmatrix}$$

由式(4)可得 MCB-FWSVM 的决策函数为:

$$f(x) = \text{sgn} \sum_{i=1}^n \alpha_i y_i K(x^T P^*, x_i^T P^*) + b \quad (8)$$

综上所述,MCB-FWSVM 算法如算法 1 所示。

算法 1 MCB-FWSVM

输入: 训练样本集

输出: 特征加权矩阵 P^*

1. 初始化各参数值;

2. for $i=1$ to d do begin

3. $\omega_i = \sum_{f_i \in F, c \in C} p(f_i, c) \log_2 \frac{p(f_i, c)}{p(f_i)p(c)}$

4. for $j=1$ to $d-1$ do begin

5. $r_{ij} = \frac{\text{cov}(f_i, f_j)}{\delta_{f_i} \delta_{f_j}}$

6. $\bar{r}_{ij} = \frac{1}{d-1} \sum_{j \in F'} |r_{ij}|$

7. $\beta_i = \begin{cases} 0, & \exists r_{ij} = 1 \\ \frac{1}{1 + (d-1)r_{ij}}, & \text{otherwise} \end{cases}$

- 8. $w_i^* = \beta_i w_i$
- 9. End for
- 10. End for

4 实验结果和分析

4.1 参数设置和数据集

本节将用 MCB-FWSVM、FWSVM 以及几种传统分类算法进行实验,以比较各算法的分类性能,其中 SVM 的核函数采用径向基函数。为了使实验更具说服力,随机选取 5 组 SVM 参数 (C, γ) ($0 < C < 100, 0.01 < \gamma < 1$) 值进行比较,如表 1 所列。MCB-FWSVM 与 FWSVM 的参数与传统 SVM 的参数保持一致;KNN 中的 K 值通过交叉验证的方法取最优赋值,Nai. Bay. 方法的先验概率函数采用高斯分布。

表 1 5 种参数组合

Table 1 Five sets of parameters

组号	1	2	3	4	5
C	1	56	25	70	93
γ	0.01	0.8	0.2	0.05	0.08

数据集采用 UCI 中的 13 个数据集,包括了文献[6]中使

用的两个数据集,如表 2 所列。

表 2 数据集属性

Table 2 Characteristics of data sets

数据集	样本数	特征数	类别数	样本数 × 特征数
Wisconsin Breast Cancer	683	9	2	6147
Letter-recognition	20000	16	26	320000
ionosphere	351	34	2	11934
Spambase	4601	57	2	262257
Pima	768	8	2	6144
Bupa	345	6	2	2070
german	10000	24	2	240000
wine	178	13	3	2314
magicGammaTelescope	19020	10	2	190200
waveform	5000	40	3	200000
glass	214	9	7	1926
segment	2310	19	7	43890
iris	150	4	3	600

4.2 实验结果和比较

采用 python2.7 版本作为实验工具,PC 配置为 Intel Core2 Duo CPU E7500@2.93 GHz,3.21 GB 内存。实验对所有算法采用 8 折交叉验证,取所有结果的平均值作为最后的结果,表 3—表 7 分别比较了 5 组参数下各算法的结果。

表 3 几种算法的分类准确率比较(组合 1)

Table 3 Comparison of classification accuracy of several algorithms(set 1)

(单位:%)

数据集	SVM	FWSVM	KNN	MCB-FWSVM	Nai. Bay.	Best result
Wisconsin Breast Cancer	97.07	97.21	94.79	97.76	95.35	MCB-FWSVM
Letter-recognition	94.28	90.61	95.20	95.27	64.42	MCB-FWSVM
ionosphere	88.60	87.15	87.76	88.11	88.09	MCB-FWSVM
Spambase	84.11	81.79	83.65	85.31	81.93	MCB-FWSVM
Pima	65.89	73.05	65.52	75.00	75.19	Nai. Bay.
Bupa	67.55	68.42	60.41	68.44	57.07	MCB-FWSVM
german	73.00	70.70	71.42	73.70	71.97	MCB-FWSVM
magicGammaTelescope	67.87	69.35	79.76	82.28	72.53	MCB-FWSVM
waveform	86.98	87.14	78.43	87.64	79.17	MCB-FWSVM
glass	38.30	50.57	71.93	49.98	45.45	MCB-FWSVM
segment	88.61	84.89	95.20	94.20	79.07	KNN
iris	94.00	95.32	97.53	94.70	96.79	KNN
wine	69.12	71.37	72.15	78.68	93.15	Nai. Bay.
Average	78.11	79.04	81.05	81.54	76.93	MCB-FWSVM

表 4 几种算法的分类准确率比较(组合 2)

Table 4 Comparison of classification accuracy of several algorithms(set 2)

(单位:%)

数据集	SVM	FWSVM	KNN	MCB-FWSVM	Nai. Bay.	Best result
Wisconsin Breast Cancer	90.48	96.92	94.79	97.31	95.35	MCB-FWSVM
Letter-recognition	61.93	91.19	95.20	95.32	64.42	MCB-FWSVM
ionosphere	94.30	94.02	87.76	94.60	88.09	MCB-FWSVM
Spambase	73.48	75.55	83.65	83.72	81.93	MCB-FWSVM
Pima	65.10	65.10	65.52	64.32	75.19	Nai. Bay.
Bupa	59.12	61.76	60.41	61.87	57.07	MCB-FWSVM
german	70.00	67.60	71.42	72.90	71.97	MCB-FWSVM
magicGammaTelescope	65.87	65.79	79.76	65.84	72.53	KNN
waveform	33.84	37.00	78.43	84.46	79.17	MCB-FWSVM
glass	69.12	70.07	71.93	69.62	45.45	KNN
segment	30.13	26.97	95.20	63.07	79.07	KNN
iris	94.70	95.36	97.53	96.67	96.79	KNN
wine	39.92	39.92	72.15	49.95	93.15	Nai. Bay.
Average	65.23	68.25	81.05	76.67	76.93	KNN

表 5 几种算法的分类准确率比较(组合 3)

Table 5 Comparison of classification accuracy of several algorithms(set 3)

(单位:%)

数据集	SVM	FWSVM	KNN	MCB-FWSVM	Nai. Bay.	Best result
Wisconsin Breast Cancer	95.32	95.90	94.79	95.91	95.35	MCB-FWSVM
Letter-recognition	96.04	98.17	95.20	98.79	64.42	MCB-FWSVM
ionosphere	94.01	94.01	87.76	94.59	88.09	MCB-FWSVM
Spambase	75.85	82.00	83.65	85.09	81.93	MCB-FWSVM
Pima	65.10	65.49	65.52	70.44	75.19	Nai. Bay.
Bupa	58.86	64.08	60.41	64.48	57.07	MCB-FWSVM
german	70.20	70.40	71.42	72.50	71.97	MCB-FWSVM
magicGammaTelescope	65.81	65.81	79.76	65.42	72.53	KNN
waveform	43.06	86.64	78.43	86.76	79.17	MCB-FWSVM
glass	73.88	73.42	71.93	75.07	45.45	MCB-FWSVM
segment	39.91	36.23	95.20	83.64	79.07	KNN
iris	95.91	94.70	97.53	97.30	96.79	KNN
wine	42.69	42.22	72.15	70.23	93.15	Nai. Bay.
Average	70.51	74.54	81.05	81.86	76.93	MCB-FWSVM

表 6 几种算法的分类准确率比较(组合 4)

Table 6 Comparison of classification accuracy of several algorithms(set 4)

(单位:%)

数据集	SVM	FWSVM	KNN	MCB-FWSVM	Nai. Bay.	Best result
Wisconsin Breast Cancer	95.91	95.90	94.79	96.78	95.35	MCB-FWSVM
Letter-recognition	97.98	97.91	95.20	98.38	64.42	MCB-FWSVM
ionosphere	92.29	92.89	87.76	93.62	88.09	MCB-FWSVM
Spambase	83.22	85.16	83.65	86.79	81.93	MCB-FWSVM
Pima	65.23	62.63	65.52	71.74	75.19	Nai. Bay.
Bupa	64.63	61.44	60.41	64.03	57.07	MCB-FWSVM
german	70.20	73.70	71.42	74.30	71.97	MCB-FWSVM
magicGammaTelescope	65.40	65.43	79.76	66.98	72.53	KNN
waveform	86.20	84.52	78.43	87.10	79.17	MCB-FWSVM
glass	72.95	70.10	71.93	73.53	45.45	MCB-FWSVM
segment	65.06	57.92	95.20	93.03	79.07	KNN
iris	97.30	96.67	97.53	95.36	96.79	KNN
wine	51.06	52.30	72.15	83.15	93.15	Nai. Bay.
Average	77.49	76.66	81.05	82.75	76.93	MCB-FWSVM

表 7 几种算法的分类准确率比较(组合 5)

Table 7 Comparison of classification accuracy of several algorithms(set 5)

(单位:%)

数据集	SVM/%	FWSVM/%	KNN/%	MCB-FWSVM/%	Nai. Bay./%	Best result
Wisconsin Breast Cancer	95.15	95.48	94.79	95.90	95.35	MCB-FWSVM
Letter-recognition	97.81	97.96	95.20	97.98	64.42	MCB-FWSVM
ionosphere	91.73	92.59	87.76	92.88	88.09	MCB-FWSVM
Spambase	79.53	84.68	83.65	87.18	81.93	MCB-FWSVM
Pima	65.36	63.41	65.52	70.96	75.19	Nai. Bay.
Bupa	62.36	64.02	60.41	64.03	57.07	MCB-FWSVM
german	69.60	72.50	71.42	74.00	71.97	MCB-FWSVM
magicGammaTelescope	65.64	65.66	79.76	65.29	72.53	KNN
waveform	85.86	85.06	78.43	86.68	79.17	MCB-FWSVM
glass	66.31	69.62	71.93	72.93	45.45	MCB-FWSVM
segment	56.41	49.05	95.20	91.04	79.07	KNN
iris	96.02	97.37	97.53	97.33	96.79	KNN
wine	47.18	46.07	72.15	84.14	93.15	Nai. Bay.
Average	75.30	75.65	81.05	83.10	76.93	MCB-FWSVM

由表 3—表 7 可知,尽管设置不同参数会影响分类准确度,但相对来说,MCB-FWSVM 在 13 个数据集上的表现基本一致,在大多数的数据集上都获得了最高的准确率,且在 13 个数据集的平均准确率上也占优势。相对于其他分类算法,MCB-FWSVM 在不同的数据集上的表现更稳定,表明 MCB-FWSVM 具有较好的鲁棒性。同时,相比于 FWSVM 算法,MCB-FWSVM 在 13 个数据集上都取得了较好的分类精度,尤其是在 segment 和 wine 这两个数据集上,MCB-FWSVM 的分类性能获得了显著的提升,这从图 1 的相关系数热力图可以得到解释。如图 1 所示,segment 和 wine 的特征内相关

系数值代表的颜色变化较明显,并且覆盖面积的占比较大,表明这两个数据集中各个特征的冗余度存在较大区别,使得 MCB-FWSVM 根据不同特征的冗余度得出的相应的惩罚因子的差别相对较大,获得的特征权重与 FWSVM 方法中获得的特征权重相比,在一定程度上减小了冗余信息和噪音对分类精度的负面影响,更能体现出特征对分类的真实贡献,也因此使得分类精度获得了极大的提升。其他数据集的相关系数热力图颜色大多数相近,少数变化明显的覆盖面积的占比比较小,因此总的来说对特征权重值的影响较小,故 MCB-FWSVM 获得的分类精度提升得不明显。

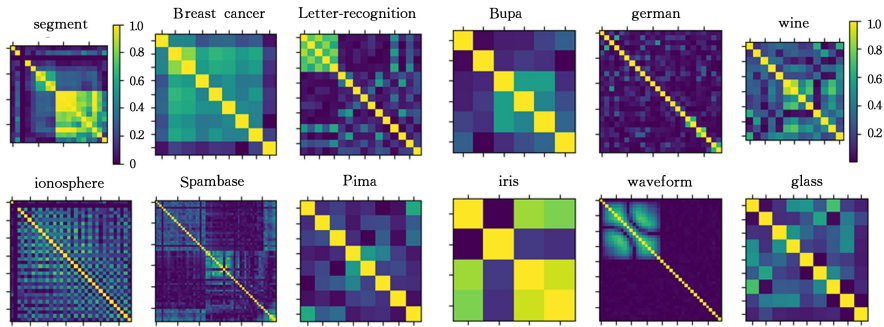


图1 相关系数热力图

Fig. 1 Heatmaps of correlation coefficients

4.3 显著性水平检验

为了进一步说明 MCB-FWSVM 在分类性能上的提升,用 python 中 scipy 包中的 ttest_rel 函数做了显著性水平检验,检验 FWSVM 和 MCB-FWSVM 这两种方法在 13 个数据集上的泛化能力是否存在明显差异。令原假设为显著性差异不存在,显著性水平 $\alpha = 0.05$,待检验数据为 FWSVM 在 13 个数据集上的分类准确率,以及 MCB-FWSVM 在 13 个数据集上的分类准确率。5 种参数组合下的 p 值如表 8 所列。

表 8 5 种参数组合下的显著性检验($\alpha = 0.05$)Table 8 Significance test under five parameter sets($\alpha = 0.05$)

组号	1	2	3	4	5
P	0.0731	0.0533	0.1166	0.0756	0.0689
$P > \alpha?$	是	是	是	是	是

由表 8 可知,5 种参数组合下的 p 值均大于 0.05,因此可以认为 MCB-FWSVM 算法相对于 FWSVM 在泛化能力上获得了显著提升。

结束语 FWSVM 将特征对分类的互信息作为相应特征权重,对 SVM 算法的分类性能有一定提升,然而在特征间或多或少都会存在冗余信息,对分类精度造成负面影响。MCB-FWSVM 将特征间的相关系数作为衡量冗余度的一个指标,并据此设定一个惩罚因子来尽可能地消除特征间的冗余,以获得更真实的特征权重。由对 UCI 中 13 个数据集的实验可以看出,MCB-FWSVM 算法相对于 FWSVM 以及几种主要的传统分类算法具有更强的全局适用性,尤其是当样本特征间的冗余度波动越大时,其对泛化能力的提升就越多。同时,该算法可以得出较为真实的特征对分类的贡献度,这可以用于特征裁剪,使得在尽可能保证分类精度的前提下降低数据维度,提高分类速度,这也将是下一步研究的目标。

参考文献

- [1] BUTT K J. A study of feature selection algorithms for accuracy estimation[J]. *Ajr American Journal of Roentgenology*, 2012, 149(6):1155-1160.
- [2] IQBAL R A. Empirical Learning Aided by Weak Domain Knowledge in the Form of Feature Importance[C]// *International Conference on Multimedia and Signal Processing*. IEEE Computer Society, 2011: 126-130.
- [3] ZHANG L, WANG Z. Ontology-based Clustering Algorithm with Feature Weights[J]. *Journal of Computational Information Systems*, 2010, 6(9): 2959-2966.
- [4] WANG T H, TIAN S F, HUANG H K. Feature Weighted Support Vector Machine[J]. *Journal of Electronics & Information Technology*, 2009, 31(3): 514-518. (in Chinese)
汪廷华, 田盛丰, 黄厚宽. 特征加权支持向量机[J]. *电子与信息学报*, 2009, 31(3): 514-518.
- [5] WALTON S, HASSAN O, MORGAN K, et al. Modified cuckoo search: A new gradient free optimisation algorithm[J]. *Chaos, Solitons & Fractals*, 2011, 44(9): 710-718.
- [6] SHANNON C E. A mathematical theory of communication[J]. *Bell System Technical Journal*, 1948, 27(3): 379-423, 623-656.
- [7] MALIK H H, FRADKIN D, MOERCHEN F. Single pass text classification by direct feature weighting[J]. *Knowledge & Information Systems*, 2011, 28(1): 79-98.
- [8] GIVEKI D, SALIMI H, BAHMANYAR G R, et al. Automatic Detection of Diabetes Diagnosis using Feature Weighted Support Vector Machines based on Mutual Information and Modified Cuckoo Search[J]. *Computer Science*, 2012, abs/1201. 2173.
- [9] XING H J, HA M H, HU B G, et al. Linear feature-weighted support vector machine[J]. *Fuzzy Information and Engineering* 2009, 1(3): 289-305.
- [10] SHANNON C E, WEAVER W. *The Mathematical Theory of Communication*[M]. Urbana: University of Illinois Press, 1949.
- [11] CHEN Y, HAO Y. A Feature Weighted Support Vector Machine and K-Nearest Neighbor Algorithm for Stock Market Indices Prediction[J]. *Expert Systems with Applications*, 2017, 80: 340-355.
- [12] IQBAL R A. Using Feature Weights to Improve Performance of Neural Networks[J/OL]. <http://arXiv:1101.4918>.
- [13] LIU L, ZHANG J, LI P, et al. A Label Correlation Based Weighting Feature Selection Approach for Multi-label Data[M]. Berlin: Springer International Publishing, 2016: 369-379.
- [14] WANG Y, LI T. Feature and Sample Weighted Support Vector Machine [M] // *Knowledge Engineering and Management*. Springer Berlin Heidelberg, 2011: 365-371.
- [15] GAO Y L, LIU Y X. An improved feature-weighted method based on K-NN[C] // *Control Conference*. IEEE, 2016: 6950-6956.
- [16] WOLFEL M, EKENEL H K. Feature weighted mahalanobis distance: Improved robustness for Gaussian classifiers[C] // *13th European Signal Processing Conference*. IEEE, 2005: 2018-2021.
- [17] JIA G, ZHAO H, PAN Z, et al. Local Feature Weighting for Data Classification[M] // *Transactions on Education XIII*. Springer Berlin Heidelberg, 2017: 293-302.