doi:10.3969/j.issn.1001 - 893x.2013.03.020

基于属性选择法的朴素贝叶斯分类器性能改进*

焦 鹏[1,**],王新政[1],谢鹏远[2]

(1.海军航空工程学院,山东 烟台 264001;2.解放军 91055 部队,浙江 台州 318050)

摘 要:为提高朴素贝叶斯(Naive Bayesian)分类器的分类准确率,对朴素贝叶斯属性选择算法及假设属性概率值先验分布中的参数设置问题进行分析,提出将属性先验分布的参数设置加入到属性选择的过程中,并研究当先验分布服从 Dirichlet 分布及广义 Dirichlet 分布情况下的具体调整步骤。以 UCI 数据库为例进行仿真实验,结果表明当先验分布服从广义 Dirichlet 分布时,该方法可提高分类的准确率,如 Parkinsons 数据集,效率可提升 13.32%。

关键词:朴素贝叶斯分类器;先验分布;属性选择法;广义 Dirichlet 分布

中图分类号:TP181 **文献标志码**:A **文章编号**:1001 - 893X(2013)03 - 0329 - 06

Performance Improvement of Naive Bayesian Classifier Based on Feature Selection

JIAO Peng[1], WANG Xin-zheng[1], XIE Peng-yuan[2]

(1.Naval Aeronautical Engineering University, Yantai 264001, China; 2.Uint 91055 of PLA, Taizhou 318050, China)

Abstract: In order to improve the accuracy of the naive Bayesian classifier(NBC), the selective naive Bayesian (SNB) method and the attributes' prior distribution are studied. A method for combining prior distribution and feature selection together is proposed, which finds out the best prior for each attribute after all attributes have been determined by the SNB algorithm. The experimental result on 10 data sets form UCI data repository shows that this method with the general Dirichlet prior generally achieves higher classification accuracy, such as the the efficiency of the data sets of Parkinson's can be enhanced by 13.32%.

Key words: naive Bayesian classifier; prior distribution; feature selection algorithm; generalized Dirichlet distribution

1 引 言

在数据挖掘领域,朴素贝叶斯分类器(Naive Bayesian Classifier, NBC)由于运算速度快、分类准确率较高,得到了广泛的应用。NBC 假设一个属性值对给定类的影响独立于其他属性值,这样的假设有助于提高运算效率,然而现实中往往不能满足。研究者提出各种方法对 NBC 的分类性能进行改进,如树增强朴素贝叶斯(Tree Augmented Naive Bayes, TAN)[1]、惰性贝叶斯规则(Lazy Bayesian Rules)[2]、特征加权(Weighted Naive Bayes, WNB)[3]等方法。这些方法与 NBC 相比通常具有较好的分类精度,在一定程度上改进了 NBC 的性能。研究显示,当数据样本属性之间相关程度很高时会降低分类准确率,因此希望 NBC 使用的属性集合尽可能地服从条件独立,即需要一个属性选择机制[4]。在众多的属性选择方法中,朴素贝叶斯属性选择算法(Selective Naive Bayesian Algorithm, SNB)能有效剔除多余或影响分类结果的属性,因此常被用于 NBC 中[5]。另外,为了改善 NBC 的分类效果,通常假设属性的可能值服从某种先验分布,一般是 Dirichlet 分布或广义

* **收稿日期**:2012 - 08 - 02;**修回日期**:2012 - 11 - 12 Received date:2012 - 08 - 02; Revised date:2012 - 11 - 12
** **通讯作者**:Jiaopeng_NEAU@hotmail.com **Corresponding author**: Jiaopeng_NEAU@hotmail.com

Dirichlet 分布。针对先验分布的参数设置已有很多学者提出各种设定方法[1]。在以往的研究中,属性选择结束后一般会直接进行分类而不考虑先验分布。本文将各属性先验分布的参数调整加入到属性选择的过程中使之成为一个整体,即首先运用 SNB 算法对样本数据集合进行属性选择,再根据其得出的属性选择顺序对选出的属性群进行参数的个别调整。通过对 UCI 数据库中的 10 个样本数据集合进行分析,仿真实验结果表明与以往的方法相比,本文提出的方法可提高分类准确性。

2 基础理论分析

2.1 朴素贝叶斯分类器

NBC 利用贝叶斯准则(Bayesian Decision Rule)以及属性间条件独立假设作为分类的依据[6]。根据贝叶斯定理,假设有 n 个属性 $X_1, X_2, \cdots, X_n$,其中一笔数据 $x=(x_1, x_2, \cdots, x_n)$属于第 j 个类别值的概率为 C_j:

$$p(C_j \mid x)=\frac{p(C_j, x)}{p(x)}=\frac{p(x \mid C_j)}{p(x)} \times p(C_j) \quad (1)$$

其中,$p(C_j \mid x)$表示在给定某项样本数据下分类到类别值 C_j 的概率,称为后验概率(Posterior Probability);$p(x)$表示数据 x 出现的概率。比较不同类别值的后验概率时,式(1)可以简化为

$$p(C_j \mid x) \propto p(x \mid C_j) \times p(C_j) \quad (2)$$

根据 NBC 的属性条件独立假设,将式(2)展开:

$$p(C_j \mid x) \propto p(x_1 \mid C_j) \times \cdots p(x_n \mid C_j) \times p(C_j)$$

$$=\prod_{i=1}^{n} p(x_i \mid C_j) \times p(C_j)$$

因此,若某一类别值 C_j 的后验概率最大,NBC 可预测该笔数据 x 的类别值为 C_j。

2.2 朴素贝叶斯属性选择算法

文献[7]将属性选择法分成 Filter 和 Wrapper 两种。Filter 方法根据统计测度分析属性之间的关系选择属性,不考虑选择的属性是否影响特定分类器的表现。而 Wrapper 方法在选择过程中使用分类器的表现评估属性的重要性。由于 Wrapper 方法考虑分类器的表现来筛选属性,对未知的数据有较好的分类准确率,而 Filter 方法不需要反复地求得分类结果,因此执行速度较快。

SNB 算法是由文献[5]提出,属于 Wrapper 方法,其运行流程如图 1 所示。初始阶段设定 S 为空集合,从样本数据中选取一个属性到朴素贝叶斯分类器中做分类并计算分类准确率。重复这个过程直到样本数据中所有属性各自对应的分类准确率都已知,选择使分类准确率最高的属性 X_i 加入 S 中,此时 X_i 就是 NBC 选择的第一个属性。从样本数据中计算未被选择的属性配合 X_i 得到的分类准确率,选出与之配合能使分类准确率最高的属性 X_j,X_j 即为第二个选择的属性,重复以上步骤直到分类准确率不再提升。SNB 算法的特点是采用前向搜索属性的方法,即初始的属性集合不包含任何属性,一次选择一个属性直到分类准确率不再提升为止。经过实验发现,SNB 算法使用较少的属性但分类准确率较高,说明该分类方法有过滤冗余属性的功能。

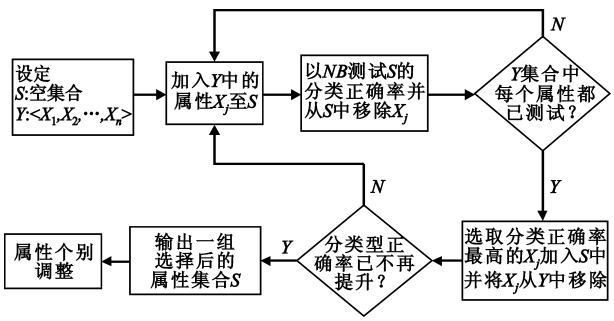

图 1 SNB 算法流程图

Fig.1 Flowchart of SNB algorithm

3 先验分布的参数设定

3.1 先验分布的参数设定

使用 NBC 进行分类时使用数据样本集合中的所有属性,并通过调整先验分布中的参数来提高分类准确率。而 SNB 算法则是通过选择对分类准确率真正有帮助的属性供 NBC 使用。由于筛选的过程耗费时间,因此先验分布的参数常使用 Laplace Estimator 用以满足无信息性,使得属性可能值的出现概率期望值相同。

文献[8]在实验研究中针对 Dirichlet 部分依序测试 α_i,发现 $\alpha_i=60$ 后分类准确率与 α_i 成反比。因此对每个属性做参数设定时,将参数 α_i 的范围设定为[1,60],并取其中的整数。除了限定范围外,调整参数时需满足无信息性的限制,即在此条件下评估各事件的发生概率,都应给定相同的估计值且总和为 1。此时各个变量对应的期望值应相同,但是不同的先验分布推导出的期望值公式不同,参数的限制也有所差异。对 Dirichlet 分布而言,推导出的

结果为各参数需设定相同的值($\alpha_1 = \alpha_2, \cdots, = \alpha_{k+1}$),以此限制在[1,60]作调整满足无信息性。广义 Dirichlet 分布在无信息性的限制满足所有参数必须满足式(4):

$$\frac{\alpha_i}{\alpha_i + \beta_i} = \frac{1}{k - i + 2}, i = 1,2,\cdots,k \tag{4}$$

由此可知,只要 α_i 已知,便可通过式(4)求得 β_i 的值,因此无论是 Dirichlet 分布还是广义 Dirichlet 分布都只需调整参数 α_i。对于整个样本数据的所有属性做参数设定时,先选择的属性对接下来其他属性的最优参数设定有很大的影响。因此首先判断各属性对分类的重要性,重要性最高的属性优先调整参数和决定先验分布。而 SNB 算法在挑选属性的过程可看作属性重要性的排序过程,根据 SNB 算法选择的顺序作为属性参数调整的顺序。

3.2　Dirichlet 分布

定义 1:随机向量 $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$满足 $\theta_1 + \theta_2 + \cdots + \theta_k \leqslant 1$ 且 $\theta_j > 0 (j = 1,2,\cdots,k)$,如果其概率密度函数为

$$f(\theta) = \frac{\Gamma(\alpha)}{\prod_{j=1}^{k+1}\Gamma(\alpha_j)} \prod_{j=1}^{k} \theta_j^{\alpha_j - 1} (1 - \theta_1 - \theta_2 - \cdots - \theta_k)^{\alpha_{k+1} - 1} \tag{5}$$

其中,$\alpha_j > 0 (j = 1,2,\cdots,k+1)$且 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_{k+1}$,则随机向量 θ 服从 k 维 Dirichlet 分布,记作 $\theta \sim D_K(\alpha_1, \alpha_2, \cdots, \alpha_K; \alpha_{K+1})$。

假设数据样本中某个属性有 $k+1$ 个可能值,令随机向量 $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$为该属性前 k 个可能值对应的出现概率,且 $\theta \sim D_K(\alpha_1, \alpha_2, \cdots, \alpha_K; \alpha_{K+1})$。令随机向量 $y = (y_1, y_2, \cdots, y_{k+1})$代表此属性 $k+1$ 个可能值分别的发生次数,y_j 表示此属性第 j 个可能值出现的次数,如果 $y \mid \theta$服从多项分配,根据 Dirichlet 分布的共轭性质可推得后验概率分布 $\theta \mid y \sim D_K(\alpha'_1, \alpha'_2, \cdots, \alpha'_K; \alpha'_{K+1})$,其中 $\alpha'_j = \alpha_j + y_j, j = 1,2,\cdots,k+1$ 表示根据收集样本数据推导出属性可能值的出现概率同样服从 Dirichlet 分布,不过参数值有所改变。如果 θ_m 为 θ 的一个变量,则 θ_m 在给定 y 的条件下,期望值为

$$E(\theta_m \mid y) = \frac{y_m + \alpha_m}{y + \alpha} \tag{6}$$

其中 $y = y_1 + y_2 + \cdots + y_{k+1}$。与 NBC 使用后验概率进行分类相比,此处将后验分布的期望值作为后验概率进行计算。当一项新的样本数据出现时,就可利用式(6)计算在给定类别下属性的某个可能值发生的概率 $E(\theta_m \mid y)$,再利用式(3)找出使后验概率最大的类别值,作为此项样本数据的预估类别值。

调整步骤如下:

(1)将所有属性的参数值设定为 Laplace Estimator,利用 SNB 算法选择一组属性,假设共选择 m 个属性;

(2)针对选择属性中的第一个属性参数,计算 $\alpha_1 = \alpha_2, \cdots, = \alpha_k$ 在[1,60]之间各整数的分类准确率,选择使分类准确率最高的参数值为 α_1^*,表示第一个属性的最优参数值;

(3)设定第一个属性的 $\alpha_1 = \alpha_2, \cdots, = \alpha_{k+1} = \alpha_1^*$,再针对第二个选择的属性参数计算在[1,60]之间整数的分类准确率,选择使分类准确率最高的参数值设定为 α_2^*,使用同样的方法找出 $\alpha_3^*, \alpha_4^*, \cdots, \alpha_m^*$。

3.3　广义 Dirichlet 分布

定义 2:随机向量 $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$满足 $\theta_1 + \theta_2 + \cdots + \theta_k \leqslant 1$ 且 $\theta_j > 0 (j = 1,2,\cdots,k)$,如果其概率密度函数为

$$f(\theta) = \frac{\Gamma(\alpha_j + \beta_j)}{\prod_{j=1}^{k+1}\Gamma(\alpha_j)\Gamma(\beta_j)} \theta_j^{\alpha_j - 1} (1 - \theta_1 - \theta_2 - \cdots - \theta_k)^{\lambda_j} \tag{7}$$

其中,参数 $\alpha_j, \beta_j, \lambda_j$ 满足 $\alpha_j > 0 (j = 1,2,\cdots,k)$,$\beta_j > 0 (j = 1,2,\cdots,k)$,$\lambda_k = \beta_k - 1$ 及 $\lambda_j = \beta_j - \alpha_{j+1} - \beta_{j+1} (j = 1,2,\cdots,k-1)$,则随机变量 θ 服从 k 维广义 Dirichlet 分布。记作 $\theta \sim GD_K(\alpha_1, \alpha_2, \cdots, \alpha_K; \beta_1, \beta_2, \cdots, \beta_K)$。

与 Dirichlet 分布在朴素贝叶斯分类器的作用一样,当假设样本数据中某属性服从广义 Dirichlet 分布,如果该属性有 $k+1$ 个可能值,令随机向量 $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$为该属性前 k 个可能值的概率,且 $\theta \sim GD_K(\alpha_1, \alpha_2, \cdots, \alpha_K; \beta_1, \beta_2, \cdots, \beta_K)$。令随机向量 $y = (y_1, y_2, \cdots, y_{k+1})$表示该属性的 $k+1$ 个可能值分别发生的次数,y_j 表示此属性第 j 个可能值出现的次数。如果 $y \mid \theta$ 服从多项分配,由于广义 Dirichlet 分布也具有共轭性质,可知后验概率分布

$\theta \mid y \sim GD_K(\alpha'_1, \alpha'_2, \cdots, \alpha'_K; \beta'_1, \beta'_2, \cdots, \beta'_K)$。

如果 θ_m 是 θ 的一个变量,则 θ_m 在给定 y 的条件下,期望值为

$$E(\theta_m \mid y) = \frac{y_m + \alpha_m}{\alpha_m + \beta_m + n_m} \prod_{i=1}^{m-1} \frac{\beta_i + n_{i+1}}{\alpha_i + \beta_i + n_i} \tag{8}$$

其中,$j = 1,2,\cdots,k$,$m = 1,2,\cdots,k$。

$$E(\theta_{k+1}|y) = \prod_{i=1}^{k} \frac{\beta_i + n_{i+1}}{\alpha_i + \beta_i + n_i} \quad (9)$$

由式(8)和式(9)可知,在给定类别值下,除了估计某属性的最后一个可能值的发生概率使用式(9)之外,估算其他可能值的发生概率都可采用式(8)计算。最后再利用式(3)找出具有最大后验概率的类别值,作为此项样本数据的预估类别值。

调整步骤如下:

(1)将所有属性的参数值设定为 Laplace Estimator 后由 SNB 算法挑选出 m 个属性;

(2)针对第一个属性的首个参数 α_1 计算在[1,60]之间整数的分类准确率,挑选使分类准确率最高的参数值为 α_1^*,表示第一个属性的最优参数值;

(3)假设 $\alpha_1 = \alpha_1^*$,针对 α_2 计算在[1,60]之间整数的分类准确率,选择使分类准确率最高的参数值设定为 α_2^*,用同样的方法找出 $\alpha_3^*, \alpha_4^*, \cdots, \alpha_{k+1}^*$,并通过式(4)计算相应的 β_i;

(4)针对 SNB 挑选的第二个属性,以步骤 1 ~ 2 的方式调整其参数,并采用相同的方式找出第 3 ~ m 个属性的最优参数值。

4 实验验证

4.1 模式评估

参考文献[9],整理得出以下两个指标。

(1)分类准确率:采用 K-fold 交互认证,将样本数据中的数据分成 K 个集合,一个集合称为一个 fold。当其中一个 fold 作为测试的集合时,其他 K-1 个 fold 结合成一个训练数据,之后重复进行 K 次,直到 K 个 fold 都作为测试的集合,最后取 K 次分类准确率的平均值作为指标。

(2)属性个数:比较选择的属性和通过属性选择法减少的属性个数,如果只需少量属性即可获得良好的分类准确率,表示该属性选择法确实能有效地过滤冗余属性。

4.2 实例验证

本节针对 UCI[10] 上的 10 个样本数据集合进行计算并评估其性能。

表 1 为样本数据集合的相关属性。将 K-fold 交互式认证法的 K 值设定为 5,使得样本数据集合最小的 tae 集合在每个 fold 平均有 30 项,因此不会因为测试项数量过小导致结果无统计意义。另外,如果样本数据集合中某些属性出现遗漏值则忽略,只使用其他没有遗漏值的属性作运算。由于 NBC 无法直接使用连续型属性,应将数据离散化。在离散化的方法中,ten-bin 是将连续型属性分成 10 个等区间,并按照属性值大小放入这 10 个区间,即变成有 10 个可能值的离散属性。本文选用的资料文件包含连续属性和离散属性,样本数量从 151 ~ 8 124不等,目的就是研究本文提出的方法在各种情况下的分类准确率,得出较为客观的结论。

表 1 实验样本数据属性及属性选择结果

Table 1 List of sample attributes and selected results

Dataset	Instances	Atributes	Disctete	Continuous	NBC	SNB
chess	3 196	36	0	36	36	14
dermatology	366	33	1	32	33	11
Hepatitis	155	19	19	19	19	7
Image segmentation	2 310	18	18	0	18	9
liver	345	6	0	6	6	5
Mushroom	8 124	21	21	0	21	3
Parkinson	197	22	0	22	22	3
Sonar	208	60	0	60	60	5
tae	151	5	4	1	5	4
vote	435	16	0	16	16	5

4.3 测试结果及分析

表 2 列出各模式下选择的属性群及个数,其中 NBC 表示朴素贝叶斯分类器使用的属性个数,即所有属性个数,SNB 表示本文方法使用的属性个数。

表 2 属性选择结果

Table 2 List of attribute selection results

Dataset	NBC	SNB
chess	36	14
dermatology	33	11
Hepatitis	19	7
Image segmentation	18	9
liver	6	5
Mushroom	21	3
Parkinson	22	3
Sonar	60	5
tae	5	4
vote	16	5

表 3 为所有样本数据集合在各模式下的分类准确率,其中的粗体数值表示各样本数据集合最高的

分类准确率。NBC 表示朴素贝叶斯分类器,使用 Laplace Estimator。SNB 表示使用 SNB 选出的属性做预测的分类准确率,使用 Laplace Estimator。MD 表示先验分布为 Dirichlet 分布,调整出最优参数的分类准确率。MG 表示先验分布为广义 Dirichlet 分布,调整最优参数的分类准确率。分类效果如图 2 所示。

表 3　分类准确率汇总表

Table 3 List of classification accuracy

Dataset	NBC	SNB	MD	MG
chess	90.24	91.18	**91.36**	**91.36**
dermatology	93.36	94.63	94.63	**94.94**
Hepatitis	80.90	83.92	**93.99**	85.11
Image segmentation	85.00	87.12	87.48	**87.66**
liver	57.93	60.66	61.37	**64.08**
Mushroom	91.52	**95.67**	**95.67**	**95.67**
Parkinson	73.91	87.23	87.23	**87.67**
Sonar	73.40	74.99	75.37	**75.74**
tae	53.69	55.78	56.90	**62.37**
vote	87.14	92.27	**92.47**	**92.47**

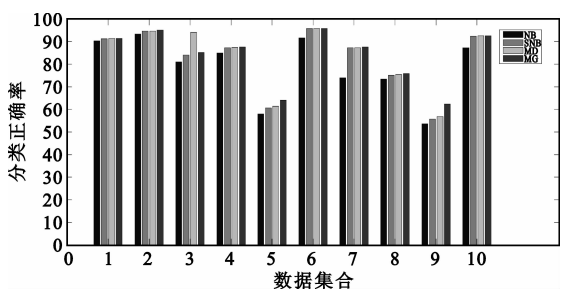

图 2　分类结果对比图

Fig.2 Comparison chart of classification results

分析以上结果可知,在各个样本数据集合中,SNB 的准确率要优于 NBC。例如 Parkinsons 数据集,NBC 使用 22 个属性的分类准确率为 73.91%,而 SNB 筛选出 3 个属性的准确率为 87.23%,效率提升 13.32%。这说明在实际样本数据集合中并非每个属性都具有分类价值,属性间也不完全服从条件独立的假设,以上两点都会影响 NBC 的分类准确率。

SNB 算法筛选出的属性数量与原始属性数量无关,这说明分类准确率不再提升时 SNB 算法即终止,因此选择的属性数量只与各阶段已选入的属性群计算的分类准确率有关,与原始属性无关。

当先验分布服从广义 Dirichlet 分布时,准确率在多数样本数据集合中最高。一般而言,广义 Dirichlet 分布比 Dirichlet 分布更能提升 NBC 的准确率,但受样本数据集合内噪声的影响,有可能使得服从 Dirichlet 分布时准确率更高。本文研究的属性集合经过 SNB 算法选择,可基本滤除干扰属性的影响,在这样的属性集合下可使广义 Dirichlet 充分发挥其效用。

5　小　结

本文运用 SNB 算法对样本数据集合进行属性选择,在属性选择的过程中针对已选出的属性,分析各个属性的特点加入最适合该属性的先验分布后再做选择,并根据选择过程中的分类准确率调整先验分布的参数,最终产生的一组具有适合先验分布的属性集合以提高分类准确率。仿真实验结果表明该方法可在保证分类效率的前提下提高分类准确率。在分析 SNB 算法的分类结果时,发现准确率不再上升即停止选择的准则过于严格,即使目前的准确率下降,继续选择若干个属性后仍有可能进一步提高分类准确率。在下一步研究中考虑设置一个缓冲区间,当准确率下降在某个范围内时仍可以进行选择。这样可避免由于选择属性过少对分类准确率的影响,使得分类效果得到进一步的改善。

参考文献:

[1] Friedman N,Geiger D,Goldszmidt M. Bayesian network classifiers[J]. Machine Learning,1997,29(2/3):131 – 163.

[2] Zheng Zijian, Webb G I, Ting Kaiming. Bayesina rules: A lazy semi – semi – Naïve Baesian learning technique competitive to boosting decision trees[C]// Proceeding of the 16th International Conference on Machine Learning. Bled, USA: IEEE,1999:493 – 502.

[3] Webb G I,J Pazzani M J. Adjusted probability naïve Bayesian induction[C]//Proceeding of the 11th Australian jiont Conference on Artificial Intelligence. Adelaide, Australia: IEEE, 1998:285 – 295.

[4] Ioan Pop. An approach of the naive Bayesian classifier for the document classification[J]. General Mathematics, 2006, 14 (4):135 – 138.

[5] Langley P, Sage S, Induction of selective bayesian classifiers [C]//Proceedings of UAI-94 10th International Conference on Uncertainty in Artificial Intelligence. Seattle, WA: IEEE, 1994:399 – 406.

[6] John G H, Kohavi R, Pfleger K. Irrelevant features and the subset selection problem[C]//Proceedings of the 11th International Conference on Machine Learning. New Brunswick, NJ:IEEE,1994: 121 – 129.

[7] Kim Chanju, Hwang Kyu-Baek. Naive Bayes classifier learning with feature selection for spam detection in social bookmarking[C]//Proceeding of Europe Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. Antwerp, Belgium: IEEE, 2008: 184 – 198.

[8] 余芳,姜云飞.一种基于朴素贝叶斯分类的特征选择方法[J].中山大学学报(自然科学版),2004,43(5):118 – 120.
YU Fang, JIANG Yun-fei. Selection method based on the characteristics of the naive Bayesian classifier[J]. Zhongshan University University (Natural Science Edition), 2004, 43(5):118 – 120.(in Chinese)

[9] 秦锋, 任诗流, 程泽凯,等. 基于属性加权的朴素贝叶斯分类算法[J]. 计算机工程与应用, 2008, 44(6): 107 – 109.
QIN Feng, REN Shi-liu, CHENG Ze-kai, et al. Naive Bayes classification algorithm based on attribute weighting[J]. Computer Engineering and Applications, 2008, 44(6): 107 – 109.(in Chinese)

[10] Frank A, Asuncion A. UCI Machine Learning Repository [EB/OL]. (2010)[2012 – 07 – 15]. http://archive.ics.uci.edu/ml.

作者简介:

焦　鹏(1980—),男,陕西西安人,2009 年获硕士学位,现为博士研究生,主要从事智能信息处理、复杂设备故障预测及诊断研究;

JIAO Peng was born in Xi′an, Shaanxi Province, in 1980. He received the M.S. degree in 2009. He is currently working toward the Ph.D. degree. His research concerns intelligent information processing and prognostics and diagnosis of complex equipment.

Email: Jiaopeng_NEAU@hotmail.com

王新政(1949—),男,陕西汉中人,海军航空工程学院教授、博士生导师,主要从事信息对抗技术、智能设备检测研究;

WANG Xin-zheng was born in Hanzhong, Shaanxi Province, in 1949. He is now a professor and also the Ph.D. supervisor. His research concerns information warfare and intelligent test technology.

谢鹏远(1980—),男,陕西安康人,2009 年获硕士学位,现为工程师,主要研究方向为智能信息处理及电子对抗。

XIE Peng-yuan was born in Ankang, Shaanxi Province, in 1980. He received the M.S. degree in 2009. He is now an engineer. His research interests include intelligent information processing and electronic countermeasures.