doi: 10.12052/gdutxb.180023

结合ReliefF和互信息的多标签特征选择算法

陈平华[1]，黄　辉[1]，麦　淼[2]，周宏虹[3]

（1. 广东工业大学　计算机学院，广东　广州 510006；2. 广东南方报业传媒集团有限公司，广东　广州 510601；
3. 广东省科技创新监测研究中心，广东　广州 510033）

摘要： 针对传统单标签特征选择算法不能直接应用于多标签数据的问题，提出一种多标签特征选择算法——MML-RF算法. 在ReliefF的基础上，MML-RF算法提出新的类内最近邻样本查找方式，并结合多标签的贡献值改进特征权值的计算方法，能很好地适应多标签数据的特点；同时为了减少特征冗余，MML-RF算法以互信息作为特征冗余度量方式，提出一种去冗余方法，能够得到更小的特征子集. 实验表明，MML-RF多标签特征选择算法得到的特征子集规模较小，且在多标签数据集上具有很好的分类效果，能够提升多标签学习和数据挖掘工作的效率.

关键词： 特征选择；多标签学习；ReliefF；互信息；特征冗余

中图分类号： TP181　　**文献标志码：** A　　**文章编号：** 1007–7162(2018)05–0020–06

Multi-label Feature Selection Algorithm Based on ReliefF and Mutual Information

Chen Ping-hua[1], Huang Hui[1], Mai Miao[2], Zhou Hong-hong[3]

(1. School of Computers, Guangdong University of Technology, Guangzhou 510006, China; 2. Guangdong Nanfang Media Group, Guangzhou 510601, China; 3. Guangdong Science and Technology Innovation Monitoring and Research Center, Guangzhou 510033, China)

Abstract: In view of the problem that the traditional feature selection algorithm can not be applied to the multi-label learning context, a MML-RF algorithm is presented. The MML-RF improves the way of defining and searching nearest neighbor on the basis of the ReliefF, and introduces a new parameter to consider the contribution values of different labels. The improved weighting formula enables MML-RF to be used to the multi-label dataset. MML-RF algorithm makes use of mutual information as the measure of feature redundancy, and puts forward a solution to redundancy, which can get smaller subset of features. Experiments show that the feature subset of MML-RF is smaller, and has good classification effect on multi-label dataset, which can further enhance the efficiency of subsequent multi-label learning and data mining.

Key words: feature selection; multi-label learning; ReliefF; mutual information; feature redundancy

特征选择是机器学习和数据挖掘工作中的重要环节，可以移除不相关和冗余的特征，从而降低数据维度提高算法效率. 特征选择算法旨在找到一个小数量的特征集合用以描述整个数据集，且描述效果能够接近甚至超越原始的特征集合[1-4].

特征选择已经被广泛应用于单标签数据的学习中，即每个实例只有一个类别与之相关联的数据集. 其中，ReliefF是经典的单标签特征选择算法[5]，其核心思想是利用特征与类别之间的相关性来评判特征的分类能力，通过各特征的相关性大小得到平均权重. 特征的权重越大，表示该特征的分类能力越强；反之，表示该特征分类能力越弱. ReliefF算法运行效率高，且特征选择能力优秀[6-8].

然而在实际问题中，一个实例通常能够被同时标记为多个类别，称之为多标签分类问题[9-10]. 比如一部电影可以被同时标记为“剧情片”和“动作片”. 目

收稿日期： 2018-01-29
基金项目： 国家自然科学基金资助项目(61572144)；广东省科技计划项目(2013B091300009, 2014B070706007, 2017B030307002)
作者简介： 陈平华(1967–)，男，教授，主要研究方向为大数据与推荐系统.
通信作者： 黄辉(1991–)，男，硕士研究生，主要研究方向为数据挖掘、机器学习，E-mail：daniel_allo@163.com

前，多标签数据集上的机器学习和数据挖掘等领域的研究和工作已经成为热点，越来越多的研究领域需要应用到多标签分类算法，如生物信息学、情感分析、文本情感注释和文本挖掘等[11-12]. 因此开展多标签特征选择算法的研究，具有重要的理论意义和应用价值.

现阶段关于多标签特征选择算法的研究较少，解决多标签问题的常用方法是数据转化法，即将一条多标签数据转化为多条单标签数据或者将该多标签集合作为一个新标签，进而对处理后的数据运用单标签特征选择算法. 然而数据转化法忽略了标签之间的关系，而这恰恰是多标签分类问题的核心所在，导致该类方法得到的特征子集性能不佳，解释性略差. 另一种解决多标签问题的方法是适应型特征选择算法，即将单标签特征选择算法进行扩展与改进，综合考虑多个标签的相互关系，使该类方法更加适应于多标签数据集的特点. 算法适应型方法已经成为解决多标签问题的重点研究方向[13-16].

通过对多标签数据的研究，在传统单标签特征选择算法ReliefF的基础上，提出一种可以适用于多标签数据集的特征选择方法，即MML-RF算法(Modified Multi-Label ReliefF Algorithm). 相比于传统的数据转化方法，MML-RF不需要对数据进行标签转换；相较于传统ReliefF算法，MML-RF算法通过重新定义最近邻样本的概念，综合考虑标签对特征的贡献，能够实现多标签分类问题的特征选择. 此外，算法引入互信息作为特征冗余的度量方式，提出了一种能够有效除去冗余的方法，解决了传统ReliefF系列算法不能去冗余的缺点.

1　ReliefF特征选择算法

由Kira等在1992年提出的Relief算法只适用于二分类数据的特征选择，Kononenko在Kira等的研究成果基础上提出了ReliefF算法用于解决多类问题[17]. ReliefF的基本原理和Relief相似，前者选择k个同类最近邻样本和不同类最近邻样本的操作是其与后者的基本区别. ReliefF算法在处理多类问题时，每次从训练样本集中随机取出一个样本R，每次从样本点R的同类的样本集中找出k个近邻样本，从不同类的样本集中分别找出k个近邻样本，然后不断随机地选取多个样本点进行特征权重的更新，得到特征权重排名，并设定阈值来选择有效特征.

记训练数据为$D=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$，其中$x_i\in R^p$为数据的特征空间，$p$为特征个数；$y_i\in R^q$为数据的类标签空间，$q$为类标签个数. 记$L_j,(j=1,2,\cdots,q)$为$q$个标签中的第$j$个，则在样本空间中，若第$i$个样本属于第$L_j$个类，则记为$y_i(L_j)=1$；反之，则记为$y_i(L_j)=0$. 在单标签数据集中，每一个样本只能属于一种类别，则对于某样本点R_i，有$\sum_{j=1}^{q}y(L_j)=1$.

ReliefF算法流程见算法1，其中样本点选取个数m和近邻数k的选取由数据集实际情况决定，k的取值一般设定为10. class(R_i)表示样本点R_i所属于的标签类型，diff(A,R_1,R_2)表示样本R_1和R_2在特征A上的距离. $P(C)$表示第C类样本的概率，不同类样本的贡献使用其类$P(C)$的先验概率加权，并在总和中将样本概率除以因子$[1-P(\text{class}(R_i))]$，保证不同类样本的概率权重总和为1. $M_j(C)$表示第C类数据中的第j个样本点.

算法1　ReliefF算法

输入：训练集D，取样次数m，最近邻数k

输出：特征的贡献权重向量$\boldsymbol{W}$

1) 初始化$\boldsymbol{W}(A=1:p)=0.0$

2) for $i=1:m$

3)　在D中随机选择样本点R_i

4)　找到与R_i同类的k个最近邻样本H_j

5)　对于每个$C\neq\text{class}(R_i)$，分别找到与R_i不同类的k个最近邻样本$M_j(C)$

6)　for $A=1:p$

7)　更新每个特征的贡献权值

$$\boldsymbol{W}[A]=\boldsymbol{W}[A]-\sum_{j=1}^{k}\text{diff}(A,R_i,H_j)/(m\cdot k)+$$

$$\sum_{C\neq\text{class}(R_i)}\left[\frac{P(C)}{1-P(\text{class}(R_i))}\sum_{j=1}^{k}\text{diff}(A,R_i,M_j(C))\right]/(m\cdot k)$$

8)　end for

9) end for

算法1输出为特征的权重向量，权值越大的特征被认为对样本的分类贡献越大，随后通过设定阈值剔除无效和冗余的特征，达到特征选择和数据降维的目的. 由算法1的步骤4)和步骤5)可知，其仅适用于单标签数据，本文将针对多标签数据的特征选择问题做进一步的研究.

2　MML-RF多标签特征选择算法

2.1　多标签特征选择算法MLRF

在多标签问题中，每一个实例都有可能属于多

个类别，所以传统的ReliefF算法受限于其最近邻样本的定义和寻找方法，并不能适用于多标签问题.

针对ReliefF算法这一缺点，本文通过调整最近邻类内样本的寻找方式和引入多标签贡献值的分配方法，提出MLRF(Multi-label ReliefF Algorithm)多标签特征选择算法. 对于算法随机选出的样本点R_i，查找其最近邻时首先应获得该样本点所拥有的n个类标签，记为$L'=\{L_1,L_2,\cdots,L_n\}$，然后将$L'$作为样本点$R_i$的类内标签，在$L_i'(i=1,2,\cdots,n)$中分别寻找$k$个类内最近邻样本，并修改权值计算公式为

$$\boldsymbol{W}[A]=\boldsymbol{W}[A]-D(N(C_N))+D(M(C_M)). \tag{1}$$

式(1)中，$D(N(C_H))$和$D(M(C_M))$分别表示样本R_i与同类最近邻点和不同类最近邻点的平均距离.

$$D(N(C_H))=\sum_{C_N=\text{class}(R_i)}[\frac{P(C_H)}{P(\text{class}(R_i))}\sum_{j=1}^{k}\text{diff}(A,R_i,N_j(C_N))]/(m\cdot k). \tag{2}$$

$$D(M(C_M))=\sum_{C_M\neq\text{class}(R_i)}[\frac{P(C_M)}{1-P(\text{class}(R_i))}\sum_{j=1}^{k}\text{diff}(A,R_i,M_j(C_M))]/(m\cdot k). \tag{3}$$

其中，C_N和C_M分别表示样本点R_i的同类和不同类的标签，$N_j(C_N)$和$M_j(C_M)$分别表示第C_N类和第C_M类中的第j个最近邻点，其他函数和定义与算法1相同. 上述操作从改进类内点的查找方式和类内平均距离的计算上做出了改进，所得式(1)的特征权值计算公式可以很好地适应多标签数据的特点.

此外，由于在多标签数据集中拥有较多标签的实例往往被定义得更加具体，也更具解释性，本文假设拥有标签越多的实例，在多标签分类问题上的贡献越大；相反，越有较少标签的实例，可能存在定义模糊的问题，应当适当削弱其对多标签分类问题的贡献. 本文在式(1)的基础上，加入多标签的贡献参数$\omega=\Delta L'/\Delta L$参与迭代更新权值，使改进的特征权值公式能更好地适应多标签数据的特点. 其中$\Delta(\cdot)$操作表示读取全体类标签集合L和样本类内标签集合L'的元素个数，即ω表示样本点所属标签类数目占所有标签类数目的比值.

改进后的特征权值计算式如式(4)所示：

$$\boldsymbol{W}[A]=\boldsymbol{W}[A]+\omega_i[-D(N(C_N))+D(M(C_M))]. \tag{4}$$

综上所述，本文在改进类内点定义和查找方式，以及引入多标签贡献这两个方面对ReliefF算法进行了扩展与改进，得到多标签特征选择算法MLRF，算法具体流程见算法2所示：

算法2 MLRF算法

输入：训练集D，取样次数m，最近邻数k

输出：特征的贡献权重向量$\boldsymbol{W}$

1) 初始化$\boldsymbol{W}(A=1:p)=0.0$

2) for $i=1:m$

3) 在D中随机选择样本点R_i

4) 对于每个$C_H=\text{class}(R_i)$，分别找到与R_i同类的k个最近邻样本$H_j(C_H)$

5) 对于每个$C_M\neq\text{class}(R_i)$，分别找到与R_i不同类的k个最近邻样本$M_j(C_M)$

6) for $A=1:p$

7) 更新每个特征的贡献权值

$$\boldsymbol{W}[A]=\boldsymbol{W}[A]+\omega_i[-D(H(C_H))+D(M(C_M))]$$

8) end for

9) end for

算法2结合了ReliefF算法的核心思想，通过改进类内点的定义和查找方法，同时引入多标签贡献参数，使改进后的算法能很好地适用于多标签情境下的特征选择工作.

2.2 特征集冗余度度量

上节所提MLRF算法虽然能够用于多标签数据的特征提取，但也具有传统ReliefF算法不能去除特征冗余的缺点，即得到的特征子集仍存在冗余项，可以结合特征去冗余的方法对其进行改进. 特征的冗余性通常使用互信息来度量，互信息是信息论中的一种信息度量方式[18]，它可以看作是一个变量中所包含的关于另一个变量的信息量.

互信息的计算如式(5)所示：

$$I(\boldsymbol{X},\boldsymbol{Y})=H(\boldsymbol{X})+H(\boldsymbol{Y})-H(\boldsymbol{XY}). \tag{5}$$

式(5)中，$\boldsymbol{X}$和$\boldsymbol{Y}$为向量，$H(\cdot)$表示其信息熵. 互信息$I(\boldsymbol{X},\boldsymbol{Y})$反映了$\boldsymbol{X}$和$\boldsymbol{Y}$的相关性程度，其值越大，代表变量间的相关性越强[19]. 实际使用中，可以使用信息熵作为分母，补偿互信息对取值较多的属性的偏置，并将互信息标准化得

$$s(\boldsymbol{X},\boldsymbol{Y})=2\frac{I(\boldsymbol{X},\boldsymbol{Y})}{H(\boldsymbol{X})+H(\boldsymbol{Y})}. \tag{6}$$

将特征向量表示为$\boldsymbol{X}_i$和$\boldsymbol{X}_j$，特征间的冗余度使用互信息作为度量，则有

$$R(\boldsymbol{X}_i,\boldsymbol{X}_j)=I(\boldsymbol{X}_i,\boldsymbol{X}_j). \tag{7}$$

由式(7)可推得单个特征与特征集合$\boldsymbol{S}$之间冗余度计算式

$$R(\boldsymbol{X}_i,S)=\frac{1}{|S|}\sum_{\boldsymbol{X}_j\in S}R(\boldsymbol{X}_i,\boldsymbol{X}_j). \tag{8}$$

式(8)中，$|S|$表示特征集合S中包含的特征数量，$\boldsymbol{X}_j$为特征集合S中的特征项. 由式(5)和(8)可得特征子集S的冗余度$R(S)$为

$$R(S)=\frac{1}{|S|}\sum_{\boldsymbol{X}_i,\boldsymbol{X}_j\in S}I(\boldsymbol{X}_i,\boldsymbol{X}_j). \tag{9}$$

式(9)结合式(6)可得标准化的特征集合冗余度

$$R_s(S)=\frac{1}{|S|}\sum_{\boldsymbol{X}_i,\boldsymbol{X}_j\in S}s(\boldsymbol{X}_i,\boldsymbol{X}_j). \tag{10}$$

2.3　MML-RF多标签特征选择算法

特征选择的目标是选择出对于标签分类有益的特征，并排除冗余的特征. 为了在MLRF算法的基础上，进一步得到精简有效的特征子集，本文引入互信息作为特征冗余的度量方式，并提出了一种去冗余方法，结合所提MLRF算法后得到MML-RF(Modified Multi-Label ReliefF Algorithm)多标签特征选择算法.

MML-RF算法的基本思想是：运行多标签特征选择算法MLRF，筛选无关特征得到特征子集S，随后通过判别评价准则，进一步在S中进行去冗余操作得到S_f. 输出的特征子集S_f应尽可能与类标签相关，即对于类标签分类有益，且特征之间的冗余度应尽可能的低. 基于此原则，构建评价准则如下

$$\Phi(S)=1/\left(1+\mathrm{e}^{-\sum W(S)}\right)-R_s(S). \tag{11}$$

式(11)中，$\sum W(S)$表示运行MLRF算法后，集合S中各特征的权值之和，$R_s(S)$表示集合S的标准化冗余度，结合式(10)有

$$\Phi(S)=1/\left(1+\mathrm{e}^{-\sum W(S)}\right)-\frac{1}{|S|}\sum_{\boldsymbol{X}_i,\boldsymbol{X}_j\in S}s(\boldsymbol{X}_i,\boldsymbol{X}_j). \tag{12}$$

MML-RF算法步骤：首先运行上节提出的MLRF多标签特征选择算法，得到按MLRF评分排序的特征集和相应的权重向量，并设定阈值进行初步筛选，本文阈值设为0.8~0.85之间. 随后通过序列后向搜索的方式进行子集搜索，即每次从候选特征集中移除一个表现最差的特征，评价移除该特征后的特征集合性能，评价方式使用式(12)的评估准则. 此外，本文为了避免算法陷入局部最优解，在子集评判过程中加入容忍度ε，代表算法允许评分下降的最大范围，使得整个搜索过程尽可能可以跳出局部最优解. 算法的具体流程如算法3所示.

算法3　MML-RF算法

输入：训练集D，MLRF参数，阈值δ

输出：特征子集S_f，权重向量$\boldsymbol{W}_f$

1) 运行MLRF算法，得到权重向量$\boldsymbol{W}_0$和特征集S_0

2) 根据阈值δ，得到S，初始化S_t和S_f

3) while$|S|>0$ {

4)　for i = $|S|$: 1

5)　$S_t=S-F_i$ //按后向顺序从S中依次删除评分最低的特征

6)　if $(\Phi(S_t)>\Phi(S)+\varepsilon)$

7)　then $S_f\leftarrow S_t$ //输出结果

8)　else if $(\Phi(S_t)>\Phi(S))$

9)　then 记录S_t及$\Phi(S_t)$，$S\leftarrow S_t$

10)　else $S\leftarrow S_t$

11)　end for }

12) 查找记录中的S_t及$\Phi(S_t)$

13) 找出使得$f=\mathrm{argmax}_f(\Phi(S_t))$的$S_t$

14) 置$S_f\leftarrow S_t$，输出S_f及$\boldsymbol{W}_f$

15) end while

其中，F_i代表排名第i个的特征. 步骤6)到10)表示候选特征集在删除排名最末的特征后，评分有较大提高才能直接输出相应特征集合，而对于评分只有略微提高的这部分候选集合，等到搜索过程完毕通过比较得到其中评分最高的特征集合，如步骤13)所示，对比过程还能结合特征集合的规模进行综合分析.

传统基于互信息的去冗余方法一般只能两两比较特征冗余，不能整体地对特征集合做出评价. MML-RF算法使用改进的序列后向搜索方法进行子集的搜索和生成，可以对特征集合进行综合的考虑，且容忍度的引入可以使算法在一定程度上跳出局部最优解. 在子集评价方面，算法根据MLRF权重和冗余度度量共同构建了性能评价指标，评价过程并不依赖于分类学习器，属于过滤式方法，保持了ReliefF系列算法运行高效的优点. 最后，本文通过实验验证了MML-RF算法较之同类算法在整体性能略有提高的基础上，能够去除冗余特征，得到规模较小且有效的特征子集.

3　实验分析

3.1　实验信息

本文实验采用Mulan多标签分类库[20]中的3个数据集Emotions、Yeast和Enron（如表1所示）. 为验证MML-RF算法在多标签数据集的有效性，本文采用

MLkNN[21]多标签分类算法来评价数据集运行MML-RF算法提取特征的分类性能.

表1 数据集信息

Tab.1 Data sets

数据集	样本	特征数	类标签
Emotions	593	72	6
Yeast	2 417	103	14
Enron	1 702	1 001	53

表1中，Emotions为音乐情感分类数据集；Yeast为酵母菌基因功能分类数据集；Enron为安然邮件信息数据集. 上述3个多标签数据集样本数量较为合理，且具有不同规模的特征和标签数量.

实验采用如下5个多分类性能评价标准，对实验结果进行评估[19].

(1) Hamming Loss(HL)：汉明损失，用于考察样本在单个标记上的误分类情况，该项取值越小，算法性能越好.

(2) One-Error(OE)：一类错误，用于考察样本类别排序最前端的标记不属于标记集合的情况，该项取值越小，算法性能越好.

(3) Coverage(CV)：覆盖率，用于考察样本标签排序序列覆盖所有相关标签的搜索深度情况，该项取值越小，算法性能越好.

(4) Ranking Loss(RL)：排列损失，用于考察样本的类标记排序中出现排序错误的情况，该项取值越小，算法性能越好.

(5) Average Precision(AP)：平均精度，用于考察样本的预测标签排序序列中的相关标签仍为样本标签的情况，该项取值越大，算法性能越好.

3.2 实验结果及分析

实验使用MML-RF多标签特征选择算法和ReliefF-ML算法[22]，分别得到不同数据集下的特征排序序列，并使用MLkNN多标签分类模型作为分类器，k值取10，平滑参数为1，测试采用5层交叉验证. 得到平均分类精度等各项参数，并通过比较两种算法取到最佳平均分类精度时的多项性能指标及其特征子集规模，说明算法在特征选择和数据降维方面的能力.

由表2可得，MML-RF算法在Emotions、Yeast和Enron 3个数据集上选取的特征子集规模较之ReliefF-ML算法，分别减少36.2%、41.5%和33.8%.

在选取特征比例分别为80%、80%和50%时，ReliefF-ML算法在数据集上得到的特征子集分类效果达到最好；而MML-RF算法在3个数据集上特征子集占比分别为51.4%、46.6%和33.1%左右时即能达到最佳的分类效果. 可以得知MML-RF算法在Emotions和Yeast数据集上特征降维能力优秀，能得到更小比例的特征子集；而在Enron数据集原本千维级别特征的基础上，算法仅使用33%左右比例的特征即可达到最佳分类性能. 综上所述，MML-RF所得到的特征子集规模更小，能够很好地减小数据规模.

表2 算法所选特征子集数量

Tab.2 The selected feature sets

数据集	原始特征	ReliefF-ML		MML-RF	
		所选特征	选择比例/%	所选特征	选择比例/%
Emotions	72	58	80	37	51.4
Yeast	103	82	80	48	44.6
Enron	1 001	500	50	331	33.1

两种算法在多标签分类器下的分类性能，如表3所示. 在Emotions数据集上，MML-RF算法在各项度量参数上较之ReliefF-ML算法均有提高；在Yeast数据集上，MML-RF算法的分类精度基本持平于ReliefF-ML，而其他4项度量参数均有所提升；在Enron数据集上，MML-RF在分类精度上较之ReliefF-ML有较大提升，RL和OE指标也均有所提高，只有HL和CV值略高于ReliefF-ML. 由上述结果可知，经过MML-RF算法进行多标签特征提取后的数据多项分类性能得到了优化，整体表现优于ReliefF-ML算法.

表3 两种算法性能参数对比

Tab.3 Performance comparison of two algorithms

算法	HL	OE	CV	RL	AP
MML-RF(Emotions)	0.215 3	0.268 1	1.880 3	0.168 1	0.796 8
ReliefF-ML(Emotions)	0.217 6	0.268 5	1.885 6	0.170 8	0.794 3
MML-RF(Yeast)	0.210 8	0.257 1	6.462 9	0.181 3	0.763 9
ReliefF-ML(Yeast)	0.211 2	0.259 3	6.505 8	0.186 7	0.764 3
MML-RF(Enron)	0.066 4	0.453 4	26.423 4	0.243 8	0.501 8
ReliefF-ML(Enron)	0.063 7	0.473 6	25.334 0	0.257 7	0.491 3

综上所述，MML-RF算法能够有效地去除特征冗余，获得规模更小的特征子集，并且在整体性能上略优于同类的适应型算法，经其输出的特征子集更加精简和有效.

4 结论

本文对传统ReliefF单标签特征选择算法进行了

适应型改进，定义了多标签数据类内最近邻的查找方式，从而得到适应于多标签情境的类内距离计算方法，并结合多标签贡献值，对特征权值的计算方法作出了进一步改进，得到的扩展算法MLRF能很好地适应多标签数据特点. 随后，在MLRF的基础上，以互信息作为特征冗余的度量，提出一种可以去除特征冗余的多标签特征选择算法——MMF-RF算法，其解决了传统ReliefF算法不能应用于多标签情境和不能去除特征冗余的问题. 实验表明MML-RF算法能够在提升算法整体性能的前提下，去除特征冗余，获得更加精简有效的特征子集，可以提升多标签学习的效率. 今后的工作将集中在如何提升复杂数据集的分类精度[23]和结合wrapper式[24-25]方法进行特征去冗余这两个方向.

参考文献：

[1] O'LEARY D, KUBBY J. Feature selection and ANN solar power prediction[J/OL]. Journal of Renewable Energy, 2017, 2437387[2017-12-05]. https://doi.org/10.1155/2017/2437387.

[2] CHANDRASHEKAR G, SAHIN F. A survey on feature selection methods [J]. Computers & Electrical Engineering, 2014, 40(1): 16-28.

[3] 姚旭, 王晓丹, 张玉玺, 等. 特征选择方法综述[J]. 控制与决策, 2012, 27(2): 161-166.
YAO X, WANG X D, ZHANG Y X, *et al*. Summary of feature selection algorithms [J]. Control and Decision, 2012, 27(2): 161-166.

[4] 徐峻岭, 周毓明, 陈林, 等. 基于互信息的无监督特征选择[J]. 计算机研究与发展, 2012, 49(2): 372-382.
XU J L, ZHOU Y M, CHEN L, *et al*. An unsupervised feature selection approach based on mutual information [J]. Journal of Computer Research and Development, 2012, 49(2): 372-382.

[5] ROBNIK-ŠIKONJA M, KONONENKO I. Theoretical and empirical analysis of ReliefF and RReliefF [J]. Machine Learning, 2003, 53(1-2): 23-69.

[6] XIE Y, LI D, ZHANG D, *et al*. An improved multi-label Relief feature selection algorithm for unbalanced datasets[C]//Advances in Intelligent Systems and Interactive Applications.[S.l.]: Springer, 2017: 141-151.

[7] FU Z, LU G, TING K M, *et al*. A survey of audio-based music classification and annotation [J]. IEEE Transactions on Multimedia, 2011, 13(2): 303-319.

[8] TANG J, ALELYANI S, LIU H. Feature selection for classification: a review [J]. Documentación Administrativa, 2014: 313-334.

[9] ZHANG M L, ZHOU Z H. A review on multi-label learning algorithms [J]. IEEE Transactions on Knowledge and Data Engineering, 2014, 26(8): 1819-1837.

[10] KANJ S, ABDALLAH F, DENOEUX T, *et al*. Editing training data for multi-label classification with the k-nearest neighbor rule [J]. Pattern Analysis and Applications, 2016, 19(1): 145-161.

[11] QIU W R, ZHENG Q S, SUN B Q, *et al*. Multi-iPPseEvo: a multi-label classifier for identifying human phosphorylated proteins by incorporating evolutionary information into Chou's General PseAAC via Grey System Theory[J/OL]. Molecular Informatics, 2016, 36(3) [2017-11-25]. https://doi.org/10.1002/minf.201600085.

[12] 贺科达, 朱铮涛, 程昱. 基于改进TF-IDF算法的文本分类方法研究[J]. 广东工业大学学报, 2016, 33(5): 49-53.
HE K D, ZHU Z T, CHENG Y. A research on text classification method based on improved TF-IDF algorithm [J]. Journal of Guangdong University of Technology, 2016, 33(5): 49-53.

[13] ZHAO K, CHU W S, DE L T F, *et al*. Joint patch and multi-label learning for facial action unit detection[C]//Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2015: 2207-2216.

[14] WU B, ZHONG E, HORNER A, *et al*. Music emotions recognition by multi-label multi-layer multi-instance multi-view learning[C]//ACM International Conference on Multimedia.[S.l.]: ACM, 2014: 117-126.

[15] CHEN G, YE D, XING Z, *et al*. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization[C]//International Joint Conference on Neural Networks. [S.l.]: IEEE, 2017: 2377-2383.

[16] 陈平华, 周鹏. 一种应用于噪声点分布密集环境下的噪声点识别算法[J]. 广东工业大学学报, 2014, 31(3): 39-43.
CHEN P H, ZHOU P. A recognition algorithm of noise applied to environments with intensive noise-data distribution [J]. Journal of Guangdong University of Technology, 2014, 31(3): 39-43.

[17] 黄莉莉, 汤进, 孙登第, 等. 基于多标签ReliefF的特征选择算法[J]. 计算机应用, 2012, 32(10): 2888-2890.
HUANG L L, TANG J, SUN D D, *et al*. Feature selection algorithm based on multi-label ReliefF [J]. Journal of Computer Applications, 2012, 32(10): 2888-2890.

[18] VERGARA J R, ESTÉVEZ P A. A review of feature selection methods based on mutual information [J]. Neural Computing and Applications, 2014, 24(1): 175-186.

[19] 胡学钢, 许尧, 李培培, 等. 一种过滤式多标签特征选择算法[J]. 南京大学学报(自然科学版), 2015, 51(4): 723-730.
HU X G, XU Y, LI P P, *et al*. A fillter multi-label feature selection algorithm [J]. Journal of Nanjing University (Natural Sciences), 2015, 51(4): 723-730.

(下转第50页)

[16] WEI X M, GUO C. Global Existence for a mathematical model of the immune response to canner [J]. Nonlinear Analysis: Real World Applications, 2010(11): 3903-3911.

[17] WEI X M, CUI S B. Existence and uniqueness of the global solution for a mathematical model of the use of macrophages in tumor medicine [J]. Nonlinear Analysis: Real World Applications, 2007, 8: 922-938.

[18] WEI X M, CUI S B. Existence and uniqueness of global solutions for a mathematical model of antiangiogenesis in tumor growth [J]. Nonlinear Analysis: Real World Applications, 2008, 9(5): 1827-1836.

[19] WEI X M. Global existence for a free boundary problem modeling the growth of necrotic tumors in the presence of inhibitors [J]. Inter Pure Appl. Math, 2006, 280: 321-338.

[20] 崔尚斌. 偏微分方程现代理论引论[M]. 北京: 科学出版社, 2015.

[21] LADYZENSKAJA O A, SOLONNIKOV V A, URALCEVA N N. Linear and quasilinear partial differential equations of parabolic type, translations of mathematical monographs [M]. Amer Math Soc, 1968, 23.

(上接第25页)

[20] TSOUMAKAS G, SPYROMITROS-XIOUFIS E, VILCEK J, *et al*. MULAN: a Java library for multi-label learning [J]. Journal of Machine Learning Research, 2011, 12(7): 2411-2414.

[21] CHERMAN E A, VALVERDE-REBAZA J, MONARD M C. Lazy multi-label learning algorithms based on mutuality strategies [J]. Journal of Intelligent & Robotic Systems, 2015, 80(1): 261-276.

[22] REYES O, MORELL C, Ventura S. Scalable extensions of the ReliefF algorithm for weighting and selecting features on the multi-label learning context [J]. Neurocomputing, 2015(161): 168-182.

[23] LEE J, KIM D W. Mutual information-based multi-label feature selection using interaction information [J]. Expert Systems with Applications, 2015, 42(4): 2013-2025.

[24] RODRIGUES D, PEREIRA L A M, NAKAMURA R Y M, *et al*. A wrapper approach for feature selection based on bat algorithm and optimum-path forest [J]. Expert Systems with Applications, 2014, 41(5): 2250-2258.

[25] 张浩荣, 陈平华, 熊建斌. 基于蚁群模拟退火算法的云环境任务调度[J]. 广东工业大学学报, 2014, 31(3): 77-82.
ZHANG H R, CHEN P H, XIONG J B. Task scheduling algorithm based on simulated annealing ant colony algorithm in cloud computing environment [J]. Journal of Guangdong University of Technology, 2014, 31(3): 77-82.