



基于模糊不一致对的多标记属性约简

高琪, 李德玉, 王素格

引用本文:

高琪, 李德玉, 王素格. 基于模糊不一致对的多标记属性约简[J]. 智能系统学报, 2020, 15(2): 374–385.

GAO Qi, LI Deyu, WANG Suge. Multi-label attribute reduction based on fuzzy inconsistency pairs[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(2): 374–385.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201905046>

您可能感兴趣的其他文章

面向一致性样本的属性约简

Attribute reduction over consistent samples

智能系统学报. 2019, 14(6): 1170–1178 <https://dx.doi.org/10.11992/tis.201905051>

多标记学习自编码网络无监督维数约简

Unsupervised dimensionality reduction of multi-label learning via autoencoder networks

智能系统学报. 2018, 13(5): 808–817 <https://dx.doi.org/10.11992/tis.201804051>

重要度集成的属性约简方法研究

Research on ensemble significance based attribute reduction approach

智能系统学报. 2018, 13(3): 414–421 <https://dx.doi.org/10.11992/tis.20170608>

连续型数据的辨识矩阵属性约简方法

A discernibility matrix-based attribute reduction for continuous data

智能系统学报. 2017, 12(3): 371–376 <https://dx.doi.org/10.11992/tis.201704032>

广义分布保持属性约简研究

Research on attribute reduction using generalized distribution preservation

智能系统学报. 2017, 12(3): 377–385 <https://dx.doi.org/10.11992/tis.21704025>

区间值模糊决策序信息系统的部分一致约简

Partially consistent reduction in interval-valued fuzzy ordered decision information system

智能系统学报. 2016, 11(4): 469–474 <https://dx.doi.org/10.11992/tis.201606013>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201905046

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20200312.1147.002.html>

基于模糊不一致对的多标记属性约简

高琪¹, 李德玉^{1,2}, 王素格^{1,2}

(1. 山西大学 计算机科学与技术学院, 山西 太原 030006; 2. 山西大学 计算智能与中文信息处理教育部重点实验室, 山西 太原 030006)

摘要: 在实际生活当中, 存在着大量的高维多标记数据, 为解决维度灾难问题, 通常需要约简属性集。针对目前的多标记属性约简算法未考虑标记关系问题, 本文提出了一种融合标记关系的模糊不一致对多标记属性约简算法。利用相对熵(KL 散度)度量标记之间的关系, 定义标记权重, 结合标记权重, 定义模糊不一致对, 考虑到属性对于模糊不一致对的区分性, 定义属性重要性并进行属性约简。在 8 个数据集上的对比实验表明, 所提基于模糊不一致对的多标记属性约简算法优于当前的多标记属性约简算法。

关键词: 多标记数据; 属性约简; 模糊不一致对; 标记权重; KL 散度; 标记关系; 模糊粗糙集; 区分矩阵

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2020)02-0374-12

中文引用格式: 高琪, 李德玉, 王素格. 基于模糊不一致对的多标记属性约简[J]. 智能系统学报, 2020, 15(2): 374-385.

英文引用格式: GAO Qi, LI Deyu, WANG Suge. Multi-label attribute reduction based on fuzzy inconsistency pairs[J]. CAAI transactions on intelligent systems, 2020, 15(2): 374-385.

Multi-label attribute reduction based on fuzzy inconsistency pairs

GAO Qi¹, LI Deyu^{1,2}, WANG Suge^{1,2}

(1. School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China; 2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education University, Shanxi University, Taiyuan 030006, China)

Abstract: In real life, there is a large amount of multi-label data, and in multi-label data processing, attribute reduction is one of the important methods to solve the high-dimensional disaster of multi-label data. Because there is a relationship between labels, in this paper we firstly use the KL divergence metric to determine the relationship between labels, then define the label weight, and then combine the label weight to define the fuzzy inconsistency pairs. Finally, considering the distinguishing ability of attributes to the fuzzy inconsistency pairs, we propose a multi-label attribute reduction algorithm based on fuzzy inconsistency pairs. Extensive experiments carried out on eight publicly available data sets verify effectiveness of the proposed algorithm named MLAR-FL by comparing it with some state-of-the-art approaches.

Keywords: multi-label data; attribute reduction; fuzzy inconsistency pairs; label weight; Kullback-Leibler divergence; the relationship of labels; fuzzy rough sets; distinguished matrix

传统的监督学习问题中, 数据只有一个类别标记变量, 通常称为单标记学习问题^[1]。但是, 在现实世界当中, 每个数据对象可能同时具有多个语义项^[2]。例如, 一个新闻可以包含经济、体育、

股票等几个主题; 一个图片可以有蓝天、湖泊、树木、绿地等多个语义标注。因此, 对于多标记数据的学习, 成为近些年的机器学习领域关注重点之一。

在多标记学习当中, 数据的高维性会严重影响多标记分类器的性能, 因而降维是解决该问题的重要手段。多标记的特征降维的方法主要有特征提取和特征选择。特征提取是通过转换或者映射的方法, 将原始高维特征空间转换到一个新的

收稿日期: 2019-05-24. 网络出版日期: 2020-03-12.

基金项目: 国家自然科学基金项目(61672331, 61573231, 61432011, 61802237); 山西省重点研发计划项目(201803D421024, 201903D421041); 山西省高等学校优秀成果培育项目(2019SK036); 山西省高等学校青年科研人員培育计划。

通信作者: 李德玉. E-mail: lidy@sxu.edu.cn.

低维特征空间的过程。

常见的特征提取方法包括线性判别分析(LDA)^[3]、特征抽取方法(MDDM)^[4]和多标记潜在语义搜索(MLSI)^[5]。特征提取方法虽然能够提高多标记分类器的分类性能,但该方法形成的新特征空间失去了原特征的语义。不同于特征提取的方法,特征选择方法是根据一定的评价准则,从原始的特征空间中选择一个(组)最优特征子集的过程。常见的评价准则有依赖性度量^[6]、信息度量^[7-8]、距离度量^[9-12]等。例如,Spolaôr等^[12]基于二元关系方法和标签幂集法,提出了基于多标记的特征选择方法(RF-BR、RF-LP、IG-BR、IG-LP);Li等^[8]利用最大化信息增益来度量特征和标记的相关性,提出基于信息增益的多标记特征选择算法(IGML)等。

在模糊粗糙集模型中,使用模糊相似关系代替经典粗糙集模型中的等价关系,进而衡量两个对象之间的不可区分性。在模糊粗糙集理论中,模糊区分矩阵概念被提出^[13-14],Chen等^[15]提出了一种基于区分矩阵中的最小元素集的属性约简方法,提高了属性约简的计算效率。Dai等^[16]提出一种基于模糊粗糙集合的最大区分对的属性约简算法。

在多标记学习中,每个样本可能同时隶属于多个标记,标记之间也可能存在着某种关系,因而本文针对多标记学习问题,利用KL散度度量标记间的关系,并定义标记权重。考虑权重情况,定义模糊不一致对,提出一种基于模糊不一致对的多标记属性约简方法(MLAR-FL),并通过实验验证了本文算法的有效性。

1 基本概念

给定一个决策表 $\langle U, A, D \rangle$,其中 $U = \{x_1, x_2, \dots, x_n\}$ 为非空有限样本集合,称其为论域, $A = \{a_1, a_2, \dots, a_p\}$ 为非空有限的条件属性集, D 为一个决策属性。设 B 为一个条件属性子集, R_a 为由属性 a 所决定的 U 上的模糊二元关系, R_D 为由决策属性 D 决定的 U 上的二元关系, $u_{Ra}(x, y)$ 为由 R_a 导出的 U 上的某个二元关系,称为样本 x, y 在属性 a 下的模糊不可区分度。 $N(u_{Ra}(x, y)) = 1 - u_{Ra}(x, y)$ 表示样本 x, y 在属性下的模糊可区分度。

定义1^[13] 给定一个决策表 $\langle U, A, D \rangle$,决定一个模糊区分矩阵 M' ,其中的项定义为

$$M'(x, y) = \{a_{N(u_{Ra}(x, y))} | a \in A, \forall x, y \in U\}$$

值得注意的是,上述区分矩阵每个项中的元是一个带有模糊可区分度值的条件属性集。

设 B 是一个属性子集,相对于决策属性 D ,样本 x 和 y 在 B 下的满意度(satisfiability)定义为^[17]

$$\text{SAT}_{B,D}(M'(x, y)) = \begin{cases} \min_{a \in B} \{u_{Ra}(x, y)\} = \max_{a \in B} \{1 - u_{Ra}(x, y)\}, \\ N(u_{R_D}(x, y)) = 1; \\ 0, N(u_{R_D}(x, y)) = 0 \end{cases}$$

总满意度(total satisfiability)定义为

$$\text{SAT}(B) = \frac{\sum_{x, y \in U, x \neq y} \text{SAT}_{B,D}(M'(x, y))}{\sum_{x, y \in U, x \neq y} \text{SAT}_{A,D}(M'(x, y))}$$

性质1^[13] 给定一个决策表 $\langle U, A, D \rangle$, $B \subset A$ 是一个条件属性子集。若 B 满足以下条件:

- 1) $\text{SAT}(B) = \text{SAT}(A) = 1$;
- 2) $\forall B' \subset B, \text{SAT}(B') < \text{SAT}(B)$ 。

则 B 是决策表 $\langle U, A, D \rangle$ 的一个模糊粗糙约简。

为表述方便,在下文中,对 $a \in A$,记

$$\begin{aligned} \text{sim}_a(x, y) &= u_{Ra}(x, y) \\ \text{dis}_a(x, y) &= N(u_{Ra}(x, y)) \end{aligned}$$

定义2^[15] 给定一个决策表 $\langle U, A, D \rangle$,模糊区分矩阵中每一项对应于样本对 (x, y) ,则该样本对对应的最大模糊区分属性定义为

$$\text{MDA}_A D(x, y) = \begin{cases} \{a | \text{dis}_A(x, y) = \max_{a \in A} \text{dis}_a(x, y), a \in A\} \\ u_{R_D}(x, y) = 0 \\ \Phi, u_{R_D}(x, y) = 1 \end{cases}$$

对于 $\forall a \in A$,其对应的最大模糊区分对定义为

$$\text{MDP}_a D(U) = \{(x, y) | a \in \text{MDA}_A D(x, y), (x, y) \in U \times U\}$$

由于 $\text{MDA}_A D(x, y)$ 的对称性,则上式可定义为

$$\text{MDP}_a D(U) = \{(x_i, x_j) | a \in \text{MDA}_A D(x_i, x_j), \forall x_i, x_j \in U, i < j\}$$

性质2^[16] 给定一个决策表 $\langle U, A, D \rangle$,若 B 为一个属性子集,则有

$$\text{MDP}_B D(U) = \bigcup_{a \in B} \text{MDP}_a D(U)$$

性质3^[16] 给定一个决策表 $\langle U, A, D \rangle$,若属性子集 B 满足:

- 1) $\text{MDP}_B D(U) = \text{MDP}_A D(U)$;
- 2) $\forall B' \subset B, \text{MDP}_{B'} D(U) \subset \text{MDP}_B D(U)$ 。

则 B 为决策表的一个约简。

2 多标记属性约简方法

给定一个多标记决策表 $S = \langle U, A, L \rangle$,其中 $U = \{x_1, x_2, \dots, x_n\}$ 为样本集, $A = \{a_1, a_2, \dots, a_d\}$ 为条件属性集, $L = \{l_1, l_2, \dots, l_q\}$ 为标记集。若样本 x 拥有标记 l_i ,则 $l_i(x) = 1$,否则 $l_i(x) = 0$ 。

定义3^[18] KL散度(Kullback-Leibler divergence)。设 p 和 q 是概率空间 Ω 下的两个概率分布。

$$D_{KL}(p||q) = \sum_{x \in \Omega} p(x) \log \frac{p(x)}{q(x)} \quad (1)$$

式中 $D_{KL}(p||q)$ 称为分布 p 关于分布 q 的 KL 散度。

KL 散度常被用来度量两个随机变量的差异。其值越小表明用 p 拟合 q 差异越小。

给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对于 $\forall l_i \in L$, 将其看做 $\{0,1\}$ 上的随机变量, 对应的概率分布记为 H_i , 对 $\forall l_i, l_j \in L$, 则 l_j 相对于 l_i 的 KL 散度定义为

$$D_{ij}(H_i||H_j) = \sum_{k \in \{0,1\}} H_i(k) \log \frac{H_i(k)}{H_j(k)} \quad (2)$$

对一个多标记决策表 $S = \langle U, A, L \rangle$, 标记空间 L 中的标记的重要性各不相同。为此, 可以利用 KL 散度, 定义这些标记的权重。

定义 4 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对 $\forall l_i \in L$, 定义其在决策表上的权重为

$$w_i = \left(1 - \frac{\sum_{j=1}^q D_{ij}(H_i||H_j)}{\sum_{i=1}^q \sum_{j=1}^q D_{ij}(H_i||H_j)} \right) \cdot \frac{1}{q-1}, \quad 1 < i < q \quad (3)$$

由于 KL 散度可以用来度量标记之间的差异, 因而用一个标记去拟合所有的标记, 若总差异越小, 则总损失越小, 说明该标记在整个标记空间越重要, 则该标记的权重越大。

例 1 表 1 给定一个多标记模糊决策表 $\langle U, A, L \rangle$, $U = \{x_1, x_2, x_3, x_4, x_5\}$, $A = \{a_1, a_2, a_3\}$, $L = \{l_1, l_2, l_3\}$ 。

表 1 一个多标记决策表
Table 1 A multi-label decision table

U	a_1	a_2	a_3	l_1	l_2	l_3
x_1	0.2	0.3	-0.2	1	1	0
x_2	0.3	0.2	0.1	0	0	1
x_3	-0.1	0	0	1	1	0
x_4	-0.2	-0.1	0	1	0	0
x_5	0.7	0.4	0.3	0	0	0

由表 1 可得标记变量的概率分布: $p_1(1)=3/5$, $p_1(0)=2/5$; $p_2(1)=2/5$, $p_2(0)=3/5$; $p_3(1)=1/5$, $p_3(0)=4/5$;

根据定义 3 可得标记之间 KL 散度为

$$D_{21} = p_2(0) \log \frac{p_2(0)}{p_1(0)} + p_2(1) \log \frac{p_2(1)}{p_1(1)}$$

$$D_{23} = p_2(0) \log \frac{p_2(0)}{p_3(0)} + p_2(1) \log \frac{p_2(1)}{p_3(1)}$$

同理可得 D_{21} 、 D_{31} 、 D_{13} 、 D_{23} ; 根据定义 4, 可得标记权重:

$$w_1 = \left(1 - \frac{\sum_{j=1}^3 D_{1j}(H_1||H_j)}{\sum_{i=1}^3 \sum_{j=1}^3 D_{ij}(H_i||H_j)} \right) \cdot \frac{1}{2} = 0.2847$$

$$w_2 = \left(1 - \frac{\sum_{j=1}^3 D_{2j}(H_2||H_j)}{\sum_{i=1}^3 \sum_{j=1}^3 D_{ij}(H_i||H_j)} \right) \cdot \frac{1}{2} = 0.4136$$

$$w_3 = \left(1 - \frac{\sum_{j=1}^3 D_{3j}(H_3||H_j)}{\sum_{i=1}^3 \sum_{j=1}^3 D_{ij}(H_i||H_j)} \right) \cdot \frac{1}{2} = 0.3017$$

定义 5 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对应的标记权重集 $W = \{w_1, w_2, \dots, w_q\}$, 定义决策表上的模糊不一致样本对集合为

$$FIP_A D = \{(x, y) | \exists l_k \in L, l_k(x) \neq l_k(y), x, y \in U\} \quad (4)$$

模糊一致样本对集合为

$$FCP_A D = \{(x, y) | \forall l_k \in L, l_k(x) = l_k(y), x, y \in U\} \quad (5)$$

对于每一个模糊不一致样本对, 定义

$$LL(x, y) = \{i | l_i(x) \neq l_i(y), 1 \leq i \leq q, (x, y) \in FIP_A D\} \quad (6)$$

为模糊不一致样本对包含的不一致的标签集合。同时定义:

$$DIS(x, y) = \begin{cases} \sum_{k \in LL(x, y)} w_k, & (x, y) \in FIP_A D \\ 0, & (x, y) \in FCP_A D \end{cases} \quad (7)$$

为模糊不一致样本对的模糊度量矩阵, 若 $DIS(x, y)$ 为 0, 则表示 x, y 在标签集合上为一致的, 若 $DIS(x, y)$ 不为 0, 则表示 x, y 在标签集合上为不一致的, 且值越大, 则表示样本之间的不一致程度越大。

定义 6 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 则对于每一个属于模糊度量矩阵的样本对而言, 区分样本的最大模糊不相似属性集定义为

$$MDA_A D(x, y) = \begin{cases} \{a | dis_A(x, y) = \max_{a \in A} dis_a(x, y), a \in A\}, & DIS(x, y) \neq 0 \\ \varphi, & DIS(x, y) = 0 \end{cases} \quad (8)$$

性质 4 对于最大模糊不相似属性集有

$$MDA_A D(x, y) = MDA_A D(y, x) \quad (9)$$

证明 由于模糊区分矩阵具有对称性。根据定义 6, 性质 4 可得证。

定义 7 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 则对于 $\forall a \in A$, 可定义在该属性上的最大模糊不相似样本对为

$$MDFIP_a D(U) = \{(x, y) | a \in MDA_A D(x, y), (x, y) \in FIP_A D\} \quad (10)$$

定义 8 给定一个多标记决策表 $S = \langle U, A, L \rangle$,

$L>$, $\forall B \subseteq A$, 定义:

$$\text{MDFIP}_B D(U) = \bigcup_{\forall a \in B} \text{MDFIP}_a D(U) \quad (11)$$

性质 5 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对于属性子集 B , 则

$$\text{MDFIP}_B D(U) \subseteq \text{MDFIP}_A D(U), \forall B \subseteq A \quad (12)$$

证明 根据定义 8, 性质 5 可得证。

性质 6 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $\forall B' \subseteq B \subseteq A$, $\text{MDFIP}_{B'} D(U) \subseteq \text{MDFIP}_B D(U)$ 。

证明 由于 $B' \subseteq B$, 所以假设 $B' = \{a_1, a_2, a_3\}$, $B = \{a_1, a_2, a_3, a_4\}$, 则由定义 8 可得:

$$\begin{aligned} \text{MDFIP}_{B'} D(U) &= \text{MDFIP}_{a_1} D(U) \cup \\ &\quad \text{MDFIP}_{a_2} D(U) \cup \text{MDFIP}_{a_3} D(U) \\ \text{MDFIP}_B D(U) &= \text{MDFIP}_{a_1} D(U) \cup \text{MDFIP}_{a_2} D(U) \cup \\ &\quad \text{MDFIP}_{a_3} D(U) \cup \text{MDFIP}_{a_4} D(U) = \\ &\quad \text{MDFIP}_{B'} D(U) \cup \text{MDFIP}_{a_4} D(U) \end{aligned}$$

由并运算的性质可得, $\text{MDFIP}_{B'} D(U) \cup \text{MDFIP}_{a_4} D(U) \supseteq \text{MDFIP}_B D(U)$, 则性质 6 可证。

由上述性质可得, $|\text{MDFIP}_B D(U)|$ 在属性子集 B 上, 满足单调性。

性质 7 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对于 $\forall B' \subseteq B$, $\forall a \in C - B'$, 则

$$|\text{MDFIP}_{B'} D(U)| \leq |\text{MDFIP}_B D(U)| \quad (13)$$

$$|\text{MDFIP}_{C-B'} D(U)| \leq |\text{MDFIP}_{B' \cup \{a\}} D(U)| \quad (14)$$

性质 8 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $\text{MDFIP}_B D(U) = \text{MDFIP}_A D(U)$, 当且仅当 $|\text{MDFIP}_B D(U)| = |\text{MDFIP}_A D(U)|$ 。

证明 根据性质 5 中式 (12), 可得 $\text{MDFIP}_{B'} D(U) = \text{MDFIP}_A D(U) \Leftrightarrow |\text{MDFIP}_{B'} D(U)| = |\text{MDFIP}_A D(U)|$,

性质 9 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 若属性子集 B 满足条件:

- 1) $\text{MDFIP}_B D(U) = \text{MDFIP}_A D(U)$
- 2) $\forall B' \subset B, \text{MDFIP}_{B'} D(U) \subset \text{MDFIP}_B D(U)$

则 B 为该决策表的一个约简。

证明

$$\begin{aligned} \text{SAT}(B) &= \frac{\sum_{x,y \in U, x \neq y} \text{SAT}_{B,L}(M'(x,y))}{\sum_{x,y \in U, x \neq y} \text{SAT}_{A,L}(M'(x,y))} = \\ &\quad \frac{\sum_{x,y \in U, x \neq y} \begin{cases} \max_{a \in B} \{1 - u_{Ra}(x,y)\}, & \text{DIS}(x,y) \neq 0 \\ 0, & \text{DIS}(x,y) = 0 \end{cases}}{\sum_{x,y \in U, x \neq y} \begin{cases} \max_{a \in A} \{1 - u_{Ra}(x,y)\}, & \text{DIS}(x,y) \neq 0 \\ 0, & \text{DIS}(x,y) = 0 \end{cases}} \end{aligned}$$

若 B 为一个约简, 则相当于

$$\begin{aligned} \text{SAT}(B) = \text{SAT}(A) &\Leftrightarrow \sum_{x,y \in U, x \neq y} \begin{cases} \max_{a \in B} \{1 - u_{Ra}(x,y)\}, & \text{DIS}(x,y) \neq 0 \\ 0, & \text{DIS}(x,y) = 0 \end{cases} = \\ &\quad \sum_{x,y \in U, x \neq y} \begin{cases} \max_{a \in A} \{1 - u_{Ra}(x,y)\}, & \text{DIS}(x,y) \neq 0 \\ 0, & \text{DIS}(x,y) = 0 \end{cases} \Leftrightarrow \\ &\quad \sum_{x,y \in U, x \neq y} \text{dis}_a(x,y), \forall a \in \text{MDA}_B D = \\ &\quad \sum_{x,y \in U, x \neq y} \text{dis}_a(x,y), \forall a \in \text{MDA}_A D \Leftrightarrow \\ &\quad \text{MDFIP}_B D(U) = \text{MDFIP}_A D(U) \end{aligned}$$

同理可得:

$$\begin{aligned} \text{SAT}(B') < \text{SAT}(B) &\Leftrightarrow \\ \text{MDFIP}_{B'} D(U) &\subset \text{MDFIP}_B D(U) \end{aligned}$$

定义 9 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $\forall a \in A$, 简化后的最大模糊相似对定义为

$$\text{MDFIP}'_a D(U) = \{(x_i, x_j) | a \in \text{MDA}_A D(x_i, x_j), \forall x_i, x_j \in U, i < j\} \quad (15)$$

$$\text{MDFIP}^*_a D(U) = \{(x_i, x_j) | a \in \text{MDA}_A D(x_i, x_j), \forall x_i, x_j \in U, i > j\} \quad (16)$$

定义 10 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 对于一个属性子集 $B \subseteq A$, 则其对应的约简后的最大模糊相似对定义为

$$\text{MDFIP}'_B D(U) = \bigcup_{a \in B} \text{MDFIP}'_a D(U) \quad (17)$$

$$\text{MDFIP}^*_B D(U) = \bigcup_{a \in B} \text{MDFIP}^*_a D(U) \quad (18)$$

性质 10 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $\forall B' \subseteq B \subseteq A$, $\text{MDFIP}'_{B'} D(U) \subseteq \text{MDFIP}'_B D(U)$ 。

证明 根据定义 10, 显然, $\forall a \in B$, $\text{MDFIP}'_B D(U)$ 是由 $\text{MDFIP}'_a D(U)$ 组成的, 因而可得, $\text{MDFIP}'_{B'} D(U) \subseteq \text{MDFIP}'_B D(U)$, $\forall B' \subseteq B$, 即证。

性质 11 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $\text{MDFIP}'_B D(U) = \text{MDFIP}'_A D(U)$, 当且仅当 $|\text{MDFIP}'_B D(U)| = |\text{MDFIP}'_A D(U)|$

证明 根据定义 10 和性质 7, 性质 11 可得证。

性质 12 给定一个多标记决策表 $S = \langle U, A, L \rangle$, 属性子集 B 是满足条件:

- 1) $\text{MDFTP}'_B D(U) = \text{MDFTP}'_A D(U)$
- 2) $\forall B' \subset B, \text{MDFTP}'_{B'} D(U) \subset \text{MDFTP}'_B D(U)$

则 B 是该决策表的一个约简。

证明 根据定义 10 和性质 7, 性质 12 可得证。

定义 11 给定一个多标记决策表 $S = \langle U, A, L \rangle$, $B \subseteq A$, $\forall a \in A - B$, 在条件属性子集 B 的基础上, 属性 a 相对于标签集 L 的重要度定义为

$$\text{sig}(a, B, L) = \sum_{x,y \in U, x \neq y} \text{DIS}(x,y), (x,y) \in \text{MDFIP}'_a D(U) \quad (19)$$

根据前文所述, 本文所设计的基于模糊不一致对的属性约简算法如下所示。

算法1 基于模糊不一致对的属性约简算法

(MLAR-FL)

输入 决策表 $\langle U, A, L \rangle$;

输出 约简后的属性子集。

1) $\text{red}=\emptyset$; $R=A$; $z=1$;2) for $l \in L$

3) 根据式(3)计算标记的权重;

4) end for

5) 初始化: DIS , $\text{MDFIPA}'D(U)$ 6) while $z=1$ 7) $B=\text{red}$;8) for $\forall a \in R$

9) 根据式(19)计算属性的重要性;

10) end for

11) 取最大属性重要性的属性值 y_k , 对应属性为 k ;12) if $y_k > 0$ 13) $\text{red}=\text{red}+k$; $R=R-k$;14) $\text{DIS}(x,y)=0$, $(x,y) \in \text{MDFIP}_k'D(U)$;

15) else

16) $z=0$;

17) end if

18) end while

在实验中, 对于数值型数据的属性 a , 样本 x,y 之间的模糊二元关系的度量为

$$u_{R_a}(x,y) = \max(\min(\frac{f(a,y)-f(a,x)+\sigma_a}{\sigma_a}, \frac{f(a,x)+\sigma_a-f(a,y)}{\sigma_a}), 0)$$

对于符号型数据的属性 a , 样本 x,y 之间的模糊二元关系的度量为

$$u_{R_a}(x,y) = \begin{cases} 0, & f(a,y) \neq f(a,x) \\ 1, & f(a,y) = f(a,x) \end{cases}$$

3 实验及结果分析

3.1 实验设置

采用8个Mulan数据集进行实验, 数据集描述如表2所示。

表2 数据集描述

Table 2 Description of datasets

数据集	实例数	属性数	标签数	数据类型	训练样本数	测试样本数
arts	5 000	462	26	数值型	2 000	3 000
business	5 000	438	30	数值型	2 000	3 000
computer	5 000	681	33	数值型	2 000	3 000
education	5 000	550	33	数值型	2 000	3 000
health	5 000	612	32	数值型	2 000	3 000
recreation	5 000	606	22	数值型	2 000	3 000
reference	5 000	793	33	数值型	2 000	3 000
science	5 000	743	40	数值型	2 000	3 000

本文采用十折交叉验证, 将提出的方法基于模糊不一致对的属性约简算法 (MLAR-FL) 与下列算法进行对比实验。基于标记关系的模糊粗糙模型提出的属性约简算法 (FRMFS)^[18]; 基于模糊粗糙集的多标签属性约简算法 (MLFRS)^[19], MLFRS 算法对标记不同的样本进行抽样, 重新定义上下近似, 从而得到新的特征选择的模型; 基于标记权重的多标签属性约简算法 (LWMF)^[20]; 基于 ReliefF 的多标签属性约简算法 (RF-ML)^[21], 其利用分类间隔赋予特征权重, 再利用特征对于标记的区分行进行特征选择。将5种属性约简方法在多标签分类的评价指标上进行比较。由于 LWMF 算法和 RF-ML 算法为排序算法, 最后分类器选择特征的个数由 MLFRS 最后得到的特征个数决定。

3.2 评价指标

在多标签分类当中, 我们常使用6种指标来评价分类学习算法的好坏, 即 Average Precision (AP)、Ranking Loss (RL)、Hamming Loss (HL)、Coverage (CV)、One-error (OE)、Micro-F1 (F1)^[19]。

令测试集为 $Z = \{(x_i, Y_i)\}_{i=1}^n \subset R^d \times \{+1, -1\}^q$, 根据预测函数 $f_i(x)$ 可定义排序函数 $\text{rank}_f(x, l) \in \{1, 2, \dots, q\}$ 。

Average Precision (AP): 用于考察所有样本的预测标记排序中位置排在该样本标记前面的标记仍属于该样本标记的概率的平均, 数值越大, 说明算法的性能越好, 定义为

$$\text{avgPre}(f) = \frac{1}{n} \sum_{i=1}^n \frac{1}{|R_i|} \sum_{l \in R_i} \frac{\{k | \text{rank}_f(x_i, k) \leq \text{rank}_f(x_i, l), k \in R_i\}}{\text{rank}_f(x_i, l)}$$

Ranking Loss (RL): 用来考察所有样本的不相

关标记的排序在相关标记前面的概率的平均,数值越小,说明算法的性能越好,定义为

$$\text{rLoss}(f) = \frac{1}{n} \sum_{i=1}^n \frac{1}{|R_i| \left| \bar{R}_i \right|}$$

$$\{|(l, k) | \text{rank}_f(x_i, l) \geq \text{rank}_f(x_i, k), (l, k) \in R_i \times \bar{R}_i\}$$

Hamming Loss(HL):用于度量样本在单个标记上的误分类的情况,数值越小,说明算法的性能越好,定义为

$$\text{hLoss}(h) = \frac{1}{n} \sum_{i=1}^n \frac{1}{L} \sum_{l=1}^L [h_l(x_i) \neq Y_{il}]$$

Coverage(CV):用于度量样本遍历所有与其相关的类别标记平均所需要的步数,数值越小,说明算法的性能越好,定义为

$$\text{coverage}(f) = \frac{1}{n} \sum_{i=1}^n \max_{l \in R_i} (\text{rank}_f(x_i, l)) - 1$$

One-error(OE):用于表示排名靠前的标签不在正确的标签集中的样本比例,数值越小,说明算

法的性能越好,定义为

$$\text{OE} = \frac{1}{n} \sum_{i=1}^n [(\arg \max_{y_i \in L} f(x_i, y_i)) \notin Y_i]$$

Micro-F1(F1):用于衡量在不同标签下的预测情况,数值越大,说明算法的性能越好,定义为

$$\text{F1} = \frac{2 \times \sum_{i=1}^n \|Y_i \cap y_i\|_1}{\sum_{i=1}^n \|y_i\|_1 + \sum_{i=1}^n \|Y_i\|_1}$$

3.3 实验结果与分析

本文采用 MLKNN^[22] 分类器进行实验,设置 K 取 10,平滑系数取 1。

表 3~8 列出了在 8 个数据集上,MLAR-FL 算法与其余 4 种算法在 6 种评价指标上面的实验结果。对于给定的评价指标,符号“↓”表示该评价指标的取值越小,分类性能越优;符号“↑”表示该评价指标的取值越大,分类性能越优。当中,最优结果使用加粗表示。

表 3 AP 评价指标下各算法的性能比较 (↑)

Table 3 Performance comparison of algorithms under AP evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	0.424 5±0.019 4	0.463 0±0.016 3	0.460 6±0.021 2	0.435 5±0.019 4	0.499 3±0.026 3
business	0.860 4±0.019 8	0.861 2±0.019 7	0.860 3±0.020 9	0.862 1±0.019 3	0.870 4±0.016 0
computer	0.599 0±0.017 0	0.613 5±0.017 4	0.615 6±0.022 9	0.603 1±0.019 3	0.634 7±0.018 4
education	0.479 0±0.030 1	0.510 8±0.029 1	0.528 0±0.029 9	0.486 6±0.024 7	0.542 1±0.029 1
health	0.614 8±0.018 8	0.634 3±0.011 9	0.629 4±0.025 5	0.619 6±0.016 6	0.667 4±0.016 5
recreation	0.372 3±0.014 9	0.405 0±0.017 3	0.339 2±0.013 6	0.382 9±0.018 9	0.443 6±0.015 1
reference	0.563 6±0.035 9	0.583 5±0.034 0	0.583 6±0.033 6	0.566 3±0.033 9	0.609 2±0.030 8
science	0.400 5±0.018 2	0.410 6±0.016 7	0.403 9±0.017 2	0.396 9±0.020 0	0.446 9±0.027 6
平均分类结果	0.539 2	0.560 2	0.552 5	0.544 1	0.589 2

表 4 RL 评价指标下各算法的性能比较 (↓)

Table 4 Performance comparison of algorithms under RL evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	0.181 1±0.007 7	0.169 8±0.006 9	0.169 2±0.008 4	0.180 2±0.008 6	0.155 5±0.010 3
business	0.049 1±0.009 5	0.048 7±0.009 4	0.049 7±0.010 1	0.048 0±0.008 7	0.042 3±0.007 4
computer	0.104 5±0.006 4	0.099 1±0.006 5	0.099 8±0.008 2	0.103 8±0.007 1	0.090 4±0.005 4
education	0.109 7±0.009 3	0.099 9±0.009 9	0.096 5±0.010 6	0.107 5±0.008 1	0.091 5±0.011 1
health	0.081 2±0.006 2	0.075 2±0.004 9	0.079 1±0.005 8	0.081 0±0.005 9	0.066 0±0.004 8
recreation	0.224 1±0.008 3	0.208 6±0.007 0	0.214 9±0.008 3	0.218 9±0.012 2	0.195 2±0.007 6
reference	0.106 4±0.010 8	0.099 1±0.009 6	0.100 9±0.010 0	0.105 3±0.010 0	0.088 6±0.009 0
science	0.154 5±0.010 2	0.151 2±0.009 8	0.150 8±0.008 7	0.152 7±0.010 0	0.142 6±0.010 7
平均分类结果	0.126 3	0.118 9	0.120 1	0.124 6	0.109 0

表 5 HL 评价指标下各算法的性能比较 (↓)

Table 5 Performance comparison of algorithms under HL evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	0.062 7±0.000 8	0.061 2±0.001 0	0.062 0±0.001 0	0.062 1±0.000 8	0.060 1±0.001 4
business	0.028 7±0.003 0	0.028 7±0.002 9	0.028 7±0.002 9	0.028 5±0.003 0	0.028 1±0.002 8
computer	0.044 2±0.002 4	0.041 8±0.002 1	0.042 6±0.002 9	0.044 1±0.002 5	0.039 6±0.002 4
education	0.044 2±0.001 2	0.043 6±0.001 4	0.042 6±0.001 1	0.044 2±0.001 2	0.041 8±0.001 6
health	0.050 8±0.001 1	0.047 8±0.001 5	0.048 8±0.001 9	0.050 4±0.001 2	0.044 3±0.001 6
recreation	0.065 3±0.002 5	0.064 7±0.002 4	0.064 1±0.002 5	0.064 9±0.002 6	0.063 8±0.002 3
reference	0.036 3±0.001 3	0.033 4±0.001 5	0.035 6±0.001 8	0.035 7±0.001 0	0.030 6±0.001 4
science	0.035 7±0.000 9	0.035 6±0.000 9	0.035 7±0.000 9	0.035 6±0.000 9	0.035 0±0.000 8
平均分类结果	0.045 9	0.044 6	0.045 0	0.081 2	0.042 9

表 6 CV 评价指标下各算法的性能比较 (↓)

Table 6 Performance comparison of algorithms under CV evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	6.231 7±0.240 6	5.937 0±0.213 0	5.930 7±0.298 1	6.208 0±0.273 6	5.541 7±0.317 3
business	2.628 0±0.442 4	2.617 3±0.415 1	2.638 3±0.450 0	2.583 0±0.407 5	2.368 7±0.370 9
computer	4.908 7±0.342 8	4.676 3±0.287 3	4.734 3±0.420 4	4.891 7±0.346 6	4.341 3±0.279 7
education	4.516 0±0.353 9	4.150 7±0.367 3	4.106 0±0.376 3	4.472 0±0.321 0	3.893 7±0.409 0
health	4.032 7±0.269 4	3.834 3±0.230 5	3.953 3±0.245 1	3.993 3±0.265 6	3.487 7±0.228 0
recreation	5.737 7±0.275 4	5.398 7±0.247 3	5.570 7±0.282 6	5.655 7±0.355 8	5.132 3±0.270 0
reference	4.014 7±0.360 6	3.779 3±0.334 7	3.835 0±0.346 2	3.973 0±0.339 5	3.423 3±0.344 6
science	7.626 3±0.482 0	7.503 0±0.450 8	7.494 3±0.382 9	7.527 7±0.451 4	7.159 0±0.487 1
平均分类结果	4.961 9	4.737 0	4.782 8	4.913 0	4.418 4

表 7 OE 评价指标下各算法的性能比较 (↓)

Table 7 Performance comparison of algorithms under OE evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	0.761 7±0.289	0.692 0±0.032 2	0.702 7±0.029 8	0.740 3±0.027 0	0.637 3±0.032 9
business	0.136 3±0.026 4	0.136 7±0.026 0	0.136 3±0.026 4	0.135 3±0.025 3	0.132 0±0.021 8
computer	0.479 7±0.022 0	0.463 3±0.027 5	0.460 7±0.032 1	0.474 0±0.027 1	0.439 7±0.025 2
education	0.681 0±0.038 1	0.643 3±0.038 2	0.621 7±0.039 1	0.673 3±0.031 3	0.608 0±0.030 9
health	0.496 0±0.024 1	0.463 7±0.020 9	0.472 3±0.043 7	0.488 3±0.019 2	0.425 7±0.020 4
recreation	0.810 3±0.025 9	0.771 3±0.031 4	0.778 3±0.025 9	0.797 0±0.032 1	0.720 3±0.029 4
reference	0.530 7±0.048 7	0.514 3±0.049 5	0.515 0±0.045 6	0.530 3±0.046 9	0.493 0±0.039 7
science	0.755 7±0.022 0	0.736 7±0.019 8	0.749 0±0.028 6	0.761 0±0.024 1	0.687 0±0.043 1
平均分类结果	0.581 4	0.552 6	0.554 5	0.574 9	0.517 8

表8 F1评价指标下各算法的性能比较(†)

Table 8 Performance comparison of algorithms under F1 evaluation index

数据集	FRMFS	MLFRS	LWMF	RF-ML	MLAR-FL
arts	0.004 0±0.000 8	0.047 1±0.007 0	0.038 8±0.014 1	0.026 4±0.003 0	0.086 9±0.014 7
business	0.733 4±0.027 7	0.732 5±0.027 0	0.733 3±0.027 5	0.735 0±0.027 4	0.737 7±0.027 8
computer	0.379 3±0.029 5	0.245 2±0.053 6	0.330 2±0.037 3	0.411 9±0.023 4	0.334 0±0.036 1
education	0.000 7±0.001 2	0.033 1±0.009 3	0.067 4±0.006 2	0.008 2±0.005 4	0.086 3±0.014 3
health	0.297 3±0.032 8	0.251 5±0.028 7	0.318 1±0.043 2	0.406 7±0.018 1	0.330 4±0.020 5
recreation	0.001 0±0.001 8	0.013 8±0.004 7	0.035 3±0.009 1	0.006 5±0.004 4	0.053 7±0.005 4
reference	0.068 4±0.097 8	0.153 5±0.044 9	0.265 1±0.074 2	0.004 8±0.004 7	0.231 1±0.028 0
science	0.001 3±0.002 3	0.022 2±0.007 8	0.012 9±0.008 0	0.001 0±0.002 2	0.064 5±0.017 2
平均分类结果	0.158 0	0.187 3	0.225 1	0.200 0	0.240 5

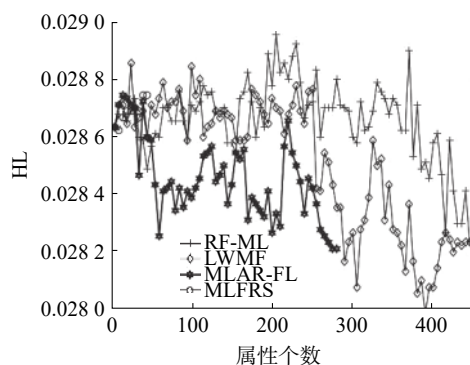
根据表3~8可以得出以下结论:

1) 对于AP、RL、HL、CV、OE指标上MLAR-FL算法在8个数据集上得到的分类精度都高于其余4种算法;对于F1指标, computer、health、reference数据集上, MLAR-FL算法低于其余算法。因而,对于6个评价指标而言,MLAR-FL在前5个评价指标上的表现很好,在F1指标上,得到的结果并非最好。

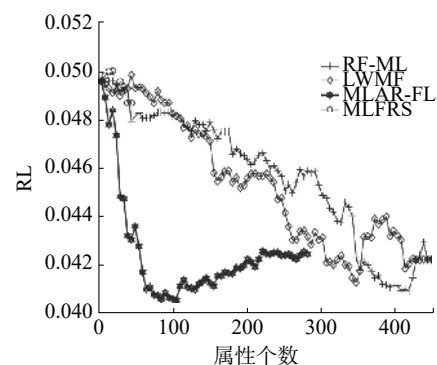
2) 从统计的8个数据集,6个评价指标,总共48个对比结果可以看出,MLAR-FL的胜率为93.75%,总体而言,MLAR-FL的分类情况较好。

总之,在8个数据集上,MLAR-FL的分类性能比其余4种算法的分类性能好。但是即使得到的特征子集的分类性能好,也不能从整体上了解算法的分类性能在特征数目变化时的变化情况。

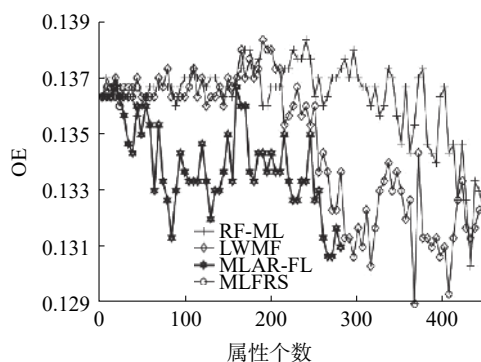
为了能够从整体上直观地对比各个算法的分类性能随着特征数目的变化情况,图1~3分别给出了在数据集business、reference、science上面6种性能评价指标AP、HL、RL、OE、CV、F1下,分类性能随着特征数目的变化趋势。由于FRMFS约简时得到的属性子集数目较少,因而比较分类特征随属性数目的变化趋势时,只考虑MLFRS、LWMF、RF-ML、MLAR-FL 4种算法。



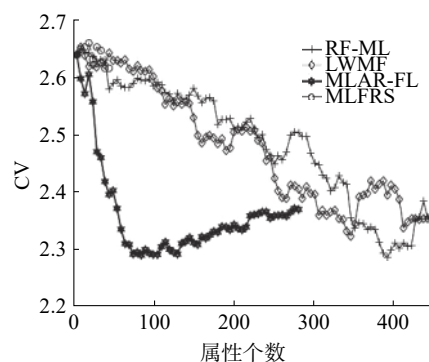
(a) 汉明损失(HL)



(b) 排序损失(RL)



(c) 1-错误率(OE)



(d) 覆盖率(CV)

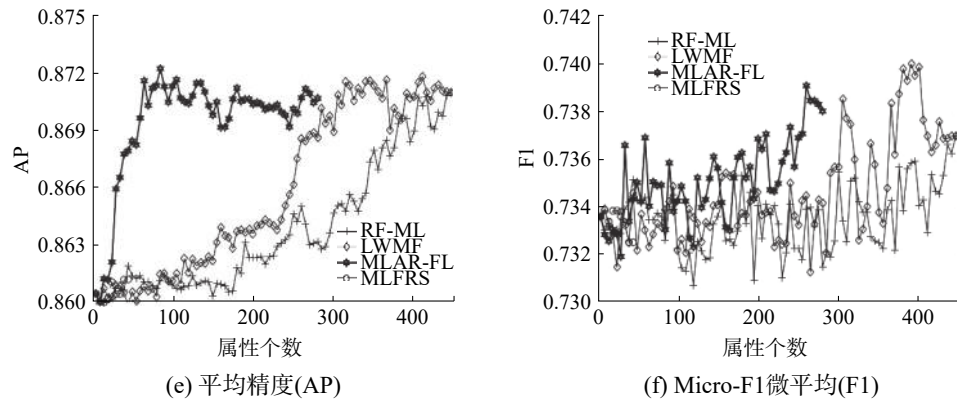


图1 数据集 business 上6种评价指标下各算法的分类性能的变化情况

Fig. 1 Changes in classification performance of each algorithm under six evaluation indicators on the business data set

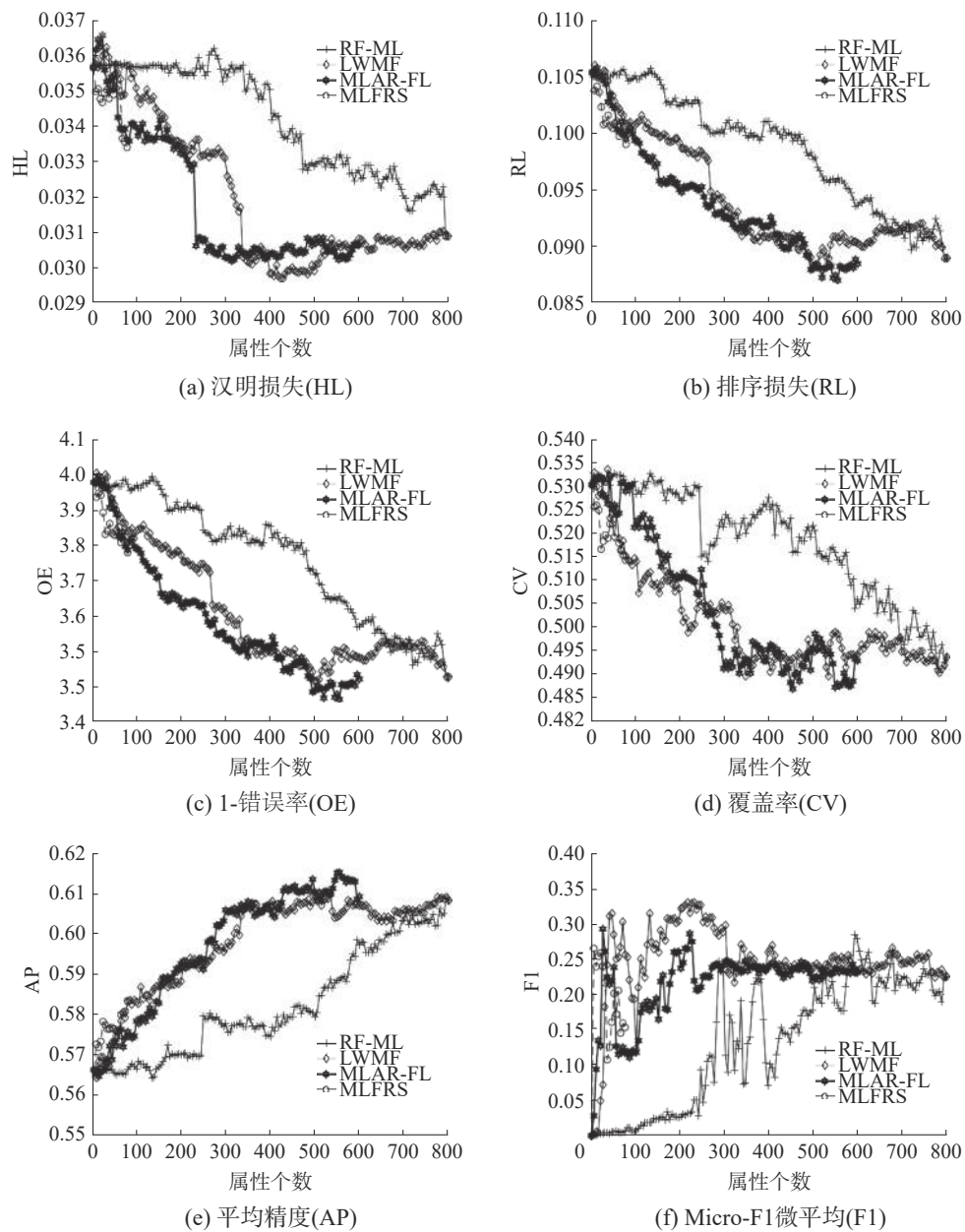


图2 数据集 reference 上6种评价指标下各算法的分类性能的变化情况

Fig. 2 Changes in classification performance of each algorithm under six evaluation indicators on the reference data set

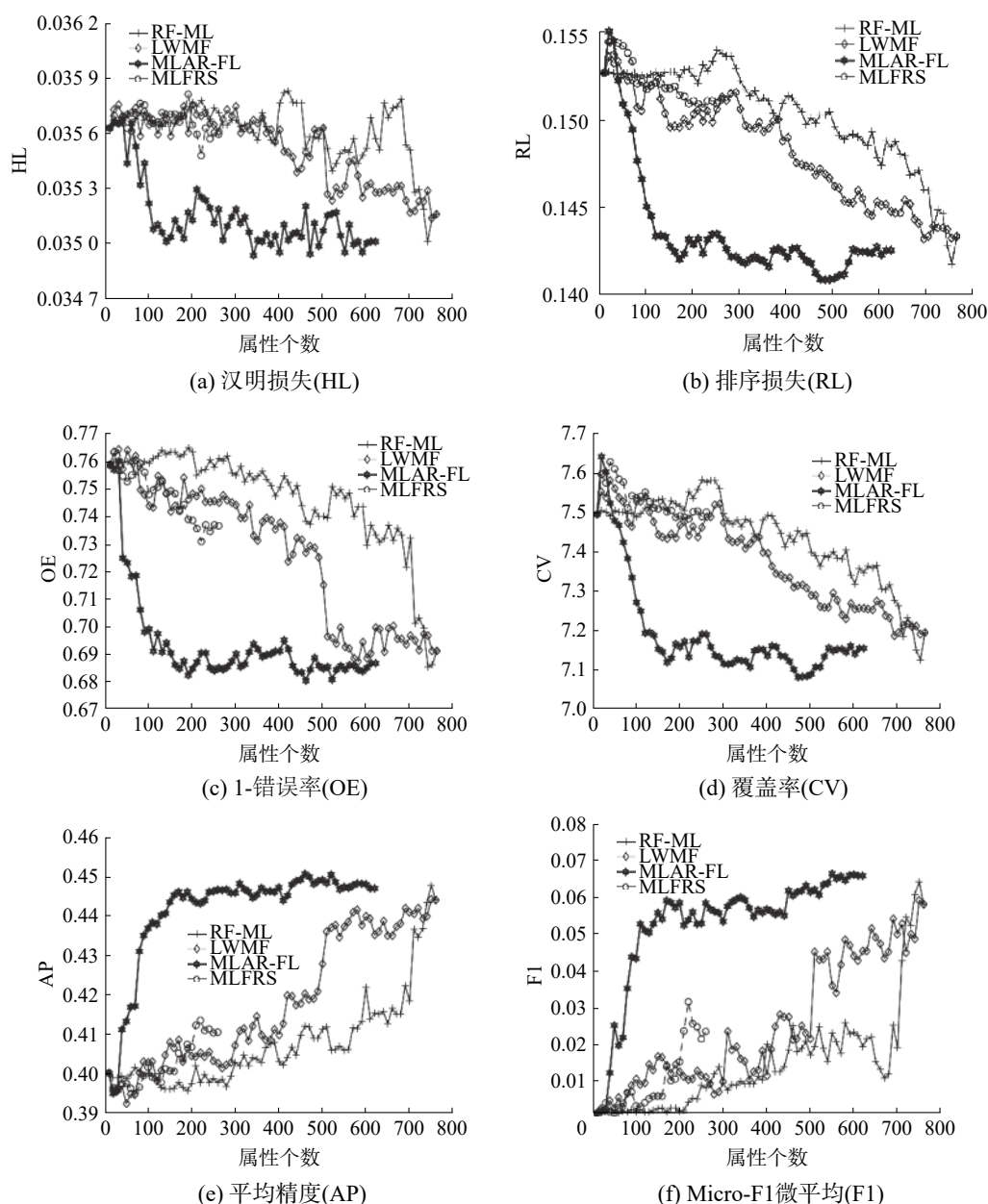


图3 science数据集上6种评价指标下各算法的分类性能的变化情况

Fig. 3 Changes in classification performance of each algorithm under six evaluation indicators on the science data set

对于这3个数据集,针对各种评价指标,由图1~3可以发现:

1) 在science数据集上,对于6种评价指标而言,MLAR-FL的分类性能比其余3种算法的分类性能好,并且在很大程度上优于其余算法;在business数据集上,MLAR-FL的分类性能在AP、RL、CV 3种评价指标上,在小于225个属性个数的情况下,分类情况较大程度上优于其余3种算法,在OE、HL、F1 3种评价指标上,分类性能存在波动,但是仍然在小于225个属性个数的情况下,优于其余3种算法;在reference数据集上,对于AP、CV、OE、RL、HL 5种评价指标而言,MLAR-FL的分类性能很明显地优于RF-ML算

法,与LWMF和MLFRS算法的分类性能比较接近,但仍在很小程度上优于这两种算法,对于F1评价指标而言,LWMF的分类性能与MLAR-FL的分类性能比较接近,且在小于300个属性个数的情况下,LWMF的分类性能优于这两种算法。

2) 对于不同的数据集,得到的约简的属性个数不同,但是可以发现,在小于250个属性个数的情况下,MLAR-FL的分类性能普遍优于其余3种算法,同时,随着属性个数的增加,分类性能越来越优,而在达到最高值后,会趋于平稳或者变差,因为越来越多的属性加入到特征空间当中,可能会与原先的属性之间存在一定的影响,

影响样本相关性的判断,从而影响分类结果。

以上结果表明,融合标记关系定义模糊不一致对的 MLAR-FL 算法的分类性能优于其余 4 种算法,得到的属性子集在一定程度上提高了多标记分类的分类性能。

4 结束语

针对多标记数据属性约简算法中为考虑标记关系的问题,本文提出了基于模糊不一致对的多标签属性约简算法,利用 KL 散度定义标记权重,结合标记权重定义模糊不一致对,使用最大区分对方法进行属性约简。在 8 个数据集,使用 6 个评价指标的实验验证了本文所提算法 MLAR-FL 的有效性,本文不仅使用新的方式度量标记关系,也证明融合该种标记关系在多标记属性约简算法中可以更好地提高约简结果的效能。同时,实验发现基于样本对的属性约简算法普遍存在时间及空间开销过大问题,因而如何提升此类算法的效率将是下一步需要研究的工作。

参考文献:

- [1] 李华, 李德玉, 王素格, 等. 基于粗糙集的多标记专属特征学习算法 [J]. 小型微型计算机系统, 2015, 36(23): 2730–2734.
LI Hua, LI Deyu, WANG Suge, et al. Multi-label learning with label-specific features based on rough sets[J]. Journal of Chinese computer systems, 2015, 36(23): 2730–2734.
- [2] ZHANG Minling, ZHOU Zhihua. A review on multi-label learning algorithms[J]. *IEEE transactions on knowledge and data engineering*, 2014, 26(8): 1819–1837.
- [3] HOTELLING H. Relations between two sets of variates[J]. *Biometrika*, 1936, 28(3/4): 321–377.
- [4] SUN Liang, JI Shuiwang, YE Jieping, et al. Multi-label dimensionality reduction[M]. Florida: CRC Press, 2013: 20–22.
- [5] YU Kai, YU Shipeng, TRESP V. Multi-label informed latent semantic indexing[C]//Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, USA, 2005: 258–265.
- [6] ZHANG Lingjun, HU Qinghua, DUAN Jie, et al. Multi-label feature selection with fuzzy rough sets[C]//Proceedings of 9th International Conference on Rough Sets and Knowledge Technology. Shanghai, China, 2014: 121–128.
- [7] LIN Yaojin, HU Qinghua, LIU Jinghua, et al. Multi-label feature selection based on neighborhood mutual information[J]. *Applied soft computing*, 2016, 38: 244–256.
- [8] LI Ling, LIU Huawen, MA Zongjie, et al. Multi-label feature selection via information gain[C]//Proceedings of the 10th International Conference on Advanced Data Mining and Applications. Guilin, China, 2014: 345–355.
- [9] LI Yuwen, LIN Yaojin, LIU Jinghua, et al. Feature selection for multi-label learning based on kernelized fuzzy rough sets[J]. *Neurocomputing*, 2018, 318: 271–286.
- [10] 李京政, 杨习贝, 王平心, 等. 模糊粗糙集的稳定约简方法 [J]. 南京理工大学学报, 2018, 42(1): 68–75.
LI Jingzheng, YANG Xibei, WANG Pingxin, et al. Stable attribute reduction approach for fuzzy rough set[J]. Journal of Nanjing University of Science and Technology, 2018, 42(1): 68–75.
- [11] SPOOLAÔR N, CHERMAN E A, MONARD M C, et al. Filter approach feature selection methods to support multi-label learning based on ReliefF and information gain[C]//Proceedings of 21th Brazilian Symposium on Artificial Intelligence on Advances in Artificial Intelligence. Curitiba, Brazil, 2012: 72–81.
- [12] ESCALERA S, RADEVA P, PUJOL O. Complex salient regions for computer vision problems[C]//Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA, 2007: 18–23.
- [13] JENSEN R, SHEN Qiang. New approaches to fuzzy-rough feature selection[J]. *IEEE transactions on fuzzy systems*, 2009, 17(4): 824–838.
- [14] JENSEN R, SHEN Qiang. Fuzzy-rough sets assisted attribute selection[J]. *IEEE transactions on fuzzy systems*, 2007, 15(1): 73–89.
- [15] CHEN Degang, ZHANG L, ZHAO Suyun, et al. A novel algorithm for finding reducts with fuzzy rough sets[J]. *IEEE transactions on fuzzy systems*, 2012, 20(2): 385–389.
- [16] DAI Jianhua, HU Hu, WU Weizhi, et al. Maximal-discernibility-pair-based approach to attribute reduction in fuzzy rough sets[J]. *IEEE transactions on fuzzy systems*, 2018, 26(4): 2174–2187.
- [17] 郭荣超, 李德玉, 王素格. 基于标记关系的模糊粗糙集模型 [J]. 模式识别与人工智能, 2017, 30(10): 952–960.
GUO Rongchao, LI Deyu, WANG Suge. Fuzzy rough set model based on label relations[J]. Pattern recognition and artificial intelligence, 2017, 30(10): 952–960.
- [18] 李晓艳, 张子刚, 张逸石, 等. 一种基于 KL 散度和类分离策略的特征选择算法 [J]. 计算机科学, 2012, 39(12): 224–227.

- LI Xiaoyan, ZHANG Zigang, ZHANG Yishi, et al. KL-divergence based feature selection algorithm with the separate-class strategy[J]. Computer science, 2012, 39(12): 224–227.
- [19] LIN Yaojin, LI Yuwen, WANG Chenxi, et al. Attribute reduction for multi-label learning with fuzzy rough set[J]. Knowledge-based systems, 2018, 152: 54–61.
- [20] 林梦雷, 刘景华, 王晨曦, 等. 基于标记权重的多标记特征选择算法 [J]. 计算机科学, 2017, 44(10): 289–295, 317.
- LIN Menglei, LIU Jinghua, WANG Chenxi, et al. Multi-label feature selection algorithm based on label weighting[J]. Computer science, 2017, 44(10): 289–295, 317.
- [21] SPOLAÔR N, CHERMAN E A, MONARD M C, et al. ReliefF for multi-label feature selection[C]//Proceedings of 2013 Brazilian Conference on Intelligent Systems. Fortaleza, Brazil, 2013: 6–11.
- [22] ZHANG Minling, ZHOU Zhihua. ML-KNN: a lazy learning approach to multi-label learning[J]. [Pattern recognition](#), 2007, 40(7): 2038–2048.

作者简介:



高琪, 硕士研究生, 主要研究方向为粗糙集、多标记学习。



李德玉, 教授, 博士, 主要研究方向为粒计算、机器学习, 多标记学习。主持国家自然科学基金项目 2 项, 参加过 3 项国家 863 计划项目等。出版著作 2 部, 发表学术论文 80 余篇。



王素格, 教授, 博士, 主要研究方向为自然语言处理、文本挖掘。主持国家自然科学基金项目 2 项, 山西省自然科学基金 1 项。发表学术论文 80 余篇。