Advances in Applied Mathematics 应用数学进展, 2026, 15(1), 278-292
Published Online January 2026 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2026.151028

Hans汉斯

基于模糊组合熵的不完备多标签特征选择

杨心怡

长安大学理学院，陕西 西安

收稿日期：2025年12月16日；录用日期：2026年1月9日；发布日期：2026年1月19日

摘　要

多标签数据通常具有高维特征空间与复杂的标签结构，这种高维性和复杂性易造成数据不同程度的不完备，从而影响多标签学习的性能。由此，本文提出基于模糊组合熵的不完备多标签特征选择方法。首先，在不完备多标签模糊信息系统中，通过引入特征值缺失率与调节参数定义模糊关系，进而定义模糊信息粒、模糊标签粒以及多标签模糊下上近似，建立不完备多标签模糊粗糙集。接着，在不完备多标签模糊粗糙集上引入组合熵的信息论思想，在此基础上定义模糊组合熵、模糊联合组合熵、模糊条件组合熵等信息度量，研究它们的性质和关系。最后，基于模糊组合熵分析特征的内外重要度，给出适用于不完备多标签数据的特征选择算法。实验结果表明，本文所提算法在5个多标签数据集上相较于对比方法取得了更优的分类性能：平均精度(AP)平均提升3.48%，汉明损失(HL)、排序损失(RL)、覆盖率(CV)、1-错误率(OE)分别平均降低3.02%、4.33%、2.83%和 4.64%。实验结果验证了本文所提算法的有效性。

关键词

不完备多标签模糊信息系统，模糊粗糙集，模糊组合熵，特征选择

Incomplete Multi-Label Feature Selection Based on Fuzzy Combination Entropy

Xinyi Yang

School of Science, Chang'an University, Xi'an Shaanxi

Received: December 16, 2025; accepted: January 9, 2026; published: January 19, 2026

Abstract

Multi-label data usually has high-dimensional feature Spaces and complex label structures. This high dimensionality and complexity can easily cause varying degrees of incompleteness in the data, thereby affecting the performance of multi-label learning. To address this issue, this paper proposes an incomplete multi-label feature selection method based on fuzzy combination entropy. Firstly, in

文章引用: 杨心怡. 基于模糊组合熵的不完备多标签特征选择[J]. 应用数学进展, 2026, 15(1): 278-292.
DOI: 10.12677/aam.2026.151028

the incomplete multi-label fuzzy information system, the fuzzy relationship is constructed by incorporating the feature-value missing rate together with a regulating parameter. Based on the defined fuzzy relationship, fuzzy information granule, fuzzy label granule, and multi-label fuzzy lower and upper approximation are defined to establish the incomplete multi-label fuzzy rough set. Then, the information-theoretic concept of combination entropy is introduced on the incomplete multi-label fuzzy rough set. On this basis, information metrics such as fuzzy combination entropy, fuzzy joint combination entropy, and fuzzy conditional combination entropy are defined, and their properties and relationships are studied. Finally, the intra- and extra-feature significances are analyzed based on fuzzy combination entropy, and a feature selection algorithm suitable for incomplete multi-label data is presented. The experimental results show that the algorithm proposed in this paper achieves better classification performance on five multi-label datasets compared with the comparison methods: The Average Precision (AP) is increased by an average of 3.48%, and the Hamming Loss (HL), Ranking Loss (RL), Coverage (CV), and One-Error (OE) are reduced by an average of 3.02%, 4.33%, 2.83% and 4.64% respectively. The experimental results verify the effectiveness of the algorithm proposed in this paper.

Keywords

Incomplete Multi-Label Fuzzy Information System, Fuzzy Rough Set, Fuzzy Combination Entropy, Feature Selection

1. 引言

近年来，多标签学习在众多领域得到了广泛应用[1]。随着数据规模的快速扩张，特征维度不断增多，其中包含着大量无关特征。这些特征不仅会削弱模型的预测性能，还会显著增加训练的计算成本[2]。因此，在多标签场景下采用有效的学习方法与特征选择策略，已成为应对高维多标签数据的关键途径[3]。

在实际应用中，在数据采集与标注阶段往往会出现信息缺失[4]。这不仅会导致部分特征值无法被完整获取，还会进一步增加后续学习与分析任务的不确定性与复杂性[5]。众多研究者针对不完备数据的多标签特征选择问题进行了系统的研究。Dai 等[4]考虑了特征之间的正交互作用，定义了对称耦合鉴别权评价特征和标签对之间的相关性，提出了一种用于处理特征缺失的不完备多标签特征选择方法。Dai 等[6]基于特征相关性与模糊容差关系实现缺失值与标签的恢复，提出了实例相关的不完备多标签数据特征选择方法。Li 等[7]通过学习特征相关矩阵，定义了补充特征矩阵，进而改善了多标签学习的分类性能。

粗糙集理论[8]作为处理不确定性数据的重要工具，能在无先验信息的前提下，刻画特征间的依赖关系，从而识别出关键特征，因此在特征选择研究中得到了广泛关注与应用。Lin 等[9]构造了多标签模糊粗糙集模型，提出了基于多标签模糊粗糙集的属性约简方法。Chen 等[10]构造了变精度模糊邻域粗糙集模型，提出了基于变精度模糊邻域粗糙集的多标签属性约简算法。Sun 等[11]建立了模糊多邻域粗糙集模型，提出了基于标签增强的特征选择方法。

香农熵[12]也称为信息熵，可用于衡量系统中信息的不确定性。Qian 等[13]在不完备信息系统中引入了组合熵和组合粒度的概念，并系统分析了它们的性质及关系。Zhang 等[14]提出了邻域组合熵的概念，据此提出了基于邻域组合熵的异构特征选择方法。Yang 等[15]定义了模糊熵的概念以量化多标签学习中特征的不确定性，提出了基于特征重要性和标签重要性的特征选择算法。Liao 等[16]在模糊粗糙集理论的框架下，提出了基于模糊条件熵的多标签特征选择算法。

在不完备多标签数据中，数据缺失会影响样本间相似程度的刻画，从而增加多标签特征选择过程中不确定性描述的难度。由此，本文提出基于模糊组合熵的不完备多标签特征选择方法。首先，在不完备多标签模糊信息系统中定义模糊关系，进而得到模糊信息粒、模糊标签粒以及多标签模糊下上近似，构造不完备多标签模糊粗糙集。在此基础上，引入模糊组合熵、模糊联合组合熵、模糊条件组合熵等信息度量。接着，基于模糊组合熵讨论特征的内外重要度，给出不完备多标签模糊粗糙集上的特征选择算法。最后，通过实验验证所提算法的有效性。

2. 预备知识

称 $MFIS=(U,A,f,L,Y)$ 为多标签模糊信息系统[17]，其中 $U=\{x_1,x_2,\cdots,x_n\}$ 为非空有限样本集，$A=\{a_1,a_2,\cdots,a_q\}$ 为特征集，$f:U\times A\to[0,1]$，$f(x,a)\in[0,1]$ 表示样本 x 在特征 a 下的取值。$L=\{l_1,l_2,\cdots,l_t\}$ 为标签集，标签向量集 $Y=\{y_i=(y_i^1,y_i^2,\cdots,y_i^t)\in\{0,1\}^t,i=1,2,\cdots,n\}$，$y_i$ 为与样本 x_i 关联的标签向量，且 $y_i^j=1\Leftrightarrow$ 样本 x_i 具有标签 l_j。

定义 1 [17]设 $MFIS=(U,A,f,L,Y)$ 为多标签模糊信息系统，$\forall B\subseteq A$，B 的模糊关系 R_B 定义为：

$$R_B(x,y)=\exp\left\{-\frac{1}{2\sigma^2}\sum_{a\in B}\left(f(x,a)-f(y,a)\right)^2\right\},\ \forall x,y\in U \tag{1}$$

其中 σ 为高斯核宽度参数。则 R_B 为 U 上的模糊相似关系。$\forall x\in U$，$\delta\in[0,1]$，x 的模糊信息粒 $[x]_B^\delta$ 定义为：

$$[x]_B^\delta(y)=\begin{cases}R_B(x,y), & R_B(x,y)\geq\delta,\\ 0, & R_B(x,y)<\delta,\end{cases}\ \forall y\in U \tag{2}$$

$\forall l_j\in L$，定义标签粒 $L_j=l_j^*=\{x_i\in U\mid y_i^j=1\}$，标签粒的全体 $L=\{L_1,L_2,\cdots,L_t\}$ 构成 U 的覆盖。标签粒 L_j 的模糊标签粒 $\tilde{L}_j$ 定义为：

$$\tilde{L}_j(x)=\frac{\left|[x]_A^\delta\cap L_j\right|}{\left|[x]_A^\delta\right|},\forall x\in U,j=1,2,\cdots,t \tag{3}$$

称 $\tilde{L}=\{\tilde{L}_1,\tilde{L}_2,\cdots,\tilde{L}_t\}$ 为标签粒集 $L=\{L_1,L_2,\cdots,L_t\}$ 关于 A 的多标签模糊粒覆盖。

定义 2 [17]设 $MFIS=(U,A,f,L,Y)$ 为多标签模糊信息系统，$\tilde{L}=\{\tilde{L}_1,\tilde{L}_2,\cdots,\tilde{L}_t\}$ 为 U 上的多标签模糊粒覆盖。$\forall\delta\in[0,1]$，$B\subseteq A$，$\tilde{L}$ 关于 B 的多标签模糊下、上近似分别定义为：

$$\underline{R}_B^\delta(\tilde{L})=\{\underline{R}_B^\delta(\tilde{L}_1),\underline{R}_B^\delta(\tilde{L}_2),\cdots,\underline{R}_B^\delta(\tilde{L}_t)\},\ \overline{R}_B^\delta(\tilde{L})=\{\overline{R}_B^\delta(\tilde{L}_1),\overline{R}_B^\delta(\tilde{L}_2),\cdots,\overline{R}_B^\delta(\tilde{L}_t)\}, \tag{4}$$

其中 $\tilde{L}_j$ 关于 B 的模糊下近似 $\underline{R}_B^\delta(\tilde{L}_j)$ 和上近似 $\overline{R}_B^\delta(\tilde{L}_j)$ 分别定义为：

$$\underline{R}_B^\delta(\tilde{L}_j)(x)=\inf_{y\in U}\max\{1-[x]_B^\delta(y),\tilde{L}_j(y)\},\ \overline{R}_B^\delta(\tilde{L}_j)(x)=\sup_{y\in U}\min\{[x]_B^\delta(y),\tilde{L}_j(y)\}. \tag{5}$$

3. 不完备多标签模糊粗糙集与信息度量

针对不完备数据，文献[18]通过定义相似度函数构造了模糊关系，以刻画样本之间的模糊相似程度。本节借鉴文献[18]中相似度函数的构造思想，在不完备多标签模糊信息系统中引入特征值缺失率并考虑调节参数，定义新的模糊关系，从而建立不完备多标签模糊粗糙集，在此基础上给出不完备多标签模糊粗糙集上的信息度量。

3.1. 不完备多标签模糊粗糙集

定义 3 设 $MFIS=(U,A,f,L,Y)$ 为多标签模糊信息系统，若存在 $x\in U$，$a\in A$ 使得 $f(x,a)=*$，则称

该信息系统为不完备多标签模糊信息系统，记作 $IMFIS=(U,A,f,L,Y)$。

定义 4 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统，$\forall a\in A$，$\beta\in[0,1]$，a 诱导的 U 上的模糊关系为 $R_a^I(x_i,x_j)=(r_{ij}^a)_{n\times n}$，$r_{ij}^a$ 的具体定义如下：

$$r_{ij}^a=\begin{cases}(1-\beta)^2, & f(x_i,a)=*\wedge f(x_j,a)=*\\ \mu_a(1-\beta)^2, & f(x_i,a)=*\wedge f(x_j,a)\neq *\\ \dfrac{1}{1+\dfrac{1}{\sigma}\sum\limits_{a\in B}\left(f(x_i,a)-f(x_j,a)\right)^2}, & f(x_i,a)\neq *\wedge f(x_j,a)\neq *\end{cases} \tag{6}$$

其中 β 为调节参数，$\mu_a=\dfrac{\left|\{x_i\in U\mid f(x_i,a)=*\}\right|}{|U|}$ 表示 a 的特征值缺失率，σ 为平滑参数。$\forall B\subseteq A$，B 的模糊关系 R_B^I 定义为：

$$R_B^I(x_i,x_j)=\underset{a\in B}{\wedge}R_a^I(x_i,x_j). \tag{7}$$

在样本存在缺失特征值时，通过引入特征值缺失率与调节参数 β，模糊关系能够自适应调节样本对的相似度，并对其取值进行约束；在样本不存在缺失特征值时，相似度采用柯西核函数进行刻画，其平缓的衰减特性能够有效减少异常值对相似度计算的影响，从而提高结果稳定性。

定义 5 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统，$\forall x\in U$，$\delta\in[0,1]$，x 的模糊信息粒 $[x]_B^{I,\delta}$ 定义为：

$$[x]_B^{I,\delta}(y)=\begin{cases}R_B^I(x,y), & R_B^I(x,y)\geq\delta,\\ 0, & R_B^I(x,y)<\delta,\end{cases}\quad \forall y\in U \tag{8}$$

$\forall L_j\in L$，标签粒 L_j 的模糊标签粒 $\tilde{L}_j^I$ 定义为：

$$\tilde{L}_j^I(x)=\frac{\left|[x]_A^{I,\delta}\cap L_j\right|}{\left|[x]_A^{I,\delta}\right|},\forall x\in U,j=1,2,\cdots,t \tag{9}$$

称 $\tilde{L}^I=\{\tilde{L}_1^I,\tilde{L}_2^I,\cdots,\tilde{L}_t^I\}$ 为标签粒集 L 关于 A 的不完备多标签模糊粒覆盖。

定义 6 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统，$\tilde{L}^I=\{\tilde{L}_1^I,\tilde{L}_2^I,\cdots,\tilde{L}_t^I\}$ 为 U 上的不完备多标签模糊粒覆盖。$\forall\delta\in[0,1]$，$B\subseteq A$，$\tilde{L}^I$ 关于 B 的多标签模糊下、上近似分别定义为：

$$\underline{R}_B^{I,\delta}(\tilde{L}^I)=\left\{\underline{R}_B^{I,\delta}(\tilde{L}_1^I),\underline{R}_B^{I,\delta}(\tilde{L}_2^I),\cdots,\underline{R}_B^{I,\delta}(\tilde{L}_t^I)\right\},\ \overline{R}_B^{I,\delta}(\tilde{L}^I)=\left\{\overline{R}_B^{I,\delta}(\tilde{L}_1^I),\overline{R}_B^{I,\delta}(\tilde{L}_2^I),\cdots,\overline{R}_B^{I,\delta}(\tilde{L}_t^I)\right\}, \tag{10}$$

其中 $\tilde{L}_j^I$ 关于 B 的多标签模糊下近似 $\underline{R}_B^{I,\delta}(\tilde{L}_j^I)$ 和上近似 $\overline{R}_B^{I,\delta}(\tilde{L}_j^I)$ 分别定义为：

$$\underline{R}_B^{I,\delta}(\tilde{L}_j^I)(x)=\inf_{y\in U}\max\left\{1-[x]_B^{I,\delta}(y),\tilde{L}_j^I(y)\right\},\ \overline{R}_B^{I,\delta}(\tilde{L}_j^I)(x)=\sup_{y\in U}\min\left\{[x]_B^{I,\delta}(y),\tilde{L}_j^I(y)\right\}. \tag{11}$$

3.2. 不完备多标签模糊粗糙集上的信息度量

本节将文献[13]提出的组合熵引入不完备多标签模糊粗糙集，以不完备多标签模糊信息粒作为信息刻画的基本粒度，定义模糊组合熵等信息度量，研究其性质和关系。

定义 7 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。$\forall\delta\in[0,1]$，$B\subseteq A$，B 的模糊组合熵定义为：

$$FCE_\delta^I(B)=\frac{1}{n}\sum_{i=1}^{n}\frac{C_n^2-C_{\left|[x_i]_B^{I,\delta}\right|}^2}{C_n^2}, \tag{12}$$

$\forall B,C\subseteq A$， B 和 C 的模糊联合组合熵定义为：

$$FCE_{\delta}^{I}(B,C)=\frac{1}{n}\sum_{i=1}^{n}\frac{C_{n}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap[x_i]_C^{I,\delta}\right|}^{2}}{C_{n}^{2}}, \tag{13}$$

C 相对于 B 的模糊条件组合熵定义为：

$$FCE_{\delta}^{I}(C\mid B)=\frac{1}{n}\sum_{i=1}^{n}\frac{C_{\left|[x_i]_B^{I,\delta}\right|}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap[x_i]_C^{I,\delta}\right|}^{2}}{C_{n}^{2}}, \tag{14}$$

$\tilde{L}^{I}$ 相对于 B 的模糊条件组合熵定义为：

$$FCE_{\delta}^{I}(\tilde{L}^{I}\mid B)=\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{t}\frac{C_{\left|[x_i]_B^{I,\delta}\right|}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap\tilde{L}_j^{I}\right|}^{2}}{C_{n}^{2}}. \tag{15}$$

定理 1 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。$\forall B,C\subseteq A$，下列结论成立：

(1) $FCE_{\delta}^{I}(B,C)=FCE_{\delta}^{I}(C,B)$；

(2) $FCE_{\delta}^{I}(C\mid B)=FCE_{\delta}^{I}(B,C)-FCE_{\delta}^{I}(B)$.

证明 易证结论(1)成立，下证(2)。

$$\begin{aligned}FCE_{\delta}^{I}(C\mid B)&=\frac{1}{n}\sum_{i=1}^{n}\frac{C_{[x_i]_B^{I,\delta}}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap[x_i]_C^{I,\delta}\right|}^{2}}{C_{n}^{2}}\\&=\frac{1}{n}\sum_{i=1}^{n}\frac{C_{n}^{2}+C_{[x_i]_B^{I,\delta}}^{2}-C_{n}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap[x_i]_C^{I,\delta}\right|}^{2}}{C_{n}^{2}}\\&=\frac{1}{n}\sum_{i=1}^{n}\frac{C_{n}^{2}-C_{\left|[x_i]_B^{I,\delta}\cap[x_i]_C^{I,\delta}\right|}^{2}}{C_{n}^{2}}-\frac{1}{n}\sum_{i=1}^{n}\frac{C_{n}^{2}-C_{\left|[x_i]_B^{I,\delta}\right|}^{2}}{C_{n}^{2}}\\&=FCE_{\delta}^{I}(B,C)-FCE_{\delta}^{I}(B).\end{aligned}$$

4. 基于模糊组合熵的不完备多标签特征选择

本节基于模糊组合熵定义特征重要度，进而提出适用于不完备多标签场景的特征选择算法。

定义 8 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。若 $FCE_{\delta}^{I}(\tilde{L}^{I}\mid A)<FCE_{\delta}^{I}(\tilde{L}^{I}\mid A-\{a\})$，称 a 在 A 中是必要的；否则，称 a 在 A 中是冗余的。$\forall B\subseteq A$，若 $FCE_{\delta}^{I}(\tilde{L}^{I}\mid B)\le FCE_{\delta}^{I}(\tilde{L}^{I}\mid A)$，且 $\forall a\in B$，$FCE_{\delta}^{I}(\tilde{L}^{I}\mid B-\{a\})>FCE_{\delta}^{I}(\tilde{L}^{I}\mid B)$，称 B 是 A 的特征约简。A 中所有必要特征构成的集合称为 $IMFIS$ 的核，记为 $Core(A)$。

定义 9 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。$\forall B\subseteq A,\ a\in B$，$a$ 关于 B 和 $\tilde{L}^{I}$ 的特征内重要度定义为：

$$S_{in}^{I}(a,B,\tilde{L}^{I})=FCE_{\delta}^{I}(\tilde{L}^{I}\mid B-\{a\})-FCE_{\delta}^{I}(\tilde{L}^{I}\mid B), \tag{16}$$

$\forall a\in A-B$， a 关于 B 和 $\tilde{L}^{I}$ 的特征外重要度定义为：

$$S_{out}^{I}(a,B,\tilde{L}^{I})=FCE_{\delta}^{I}(\tilde{L}^{I}\mid B)-FCE_{\delta}^{I}(\tilde{L}^{I}\mid B\cup\{a\}). \tag{17}$$

定理 2 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。$\forall a\in A$，

(1) $Core(A)=\{a\in A\mid FCE_{\delta}^{I}(\tilde{L}^{I}\mid B)<FCE_{\delta}^{I}(\tilde{L}^{I}\mid B-\{a\})\}$；

(2) $a\in Core(A)\Leftrightarrow S_{in}^{I}(a,A,\tilde{L}^{I})>0$.

证明 易证结论(1)成立，下证(2)。 $\forall a \in Core(A)$，则 $FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A\right) < FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A-\{a\}\right)$，因此 $S_{in}^{I}\left(a,B,\tilde{L}^{I}\right)>0$。若 $\forall a \in A$，有 $S_{in}^{I}\left(a,B,\tilde{L}^{I}\right)>0$，则 $FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A\right) < FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A-\{a\}\right)$，可得 a 在 A 中是必要的，因此 $a \in Core(A)$。

推论 设 $IMFIS=(U,A,f,L,Y)$ 为不完备多标签模糊信息系统。$\forall a \in A$，$a \notin Core(A) \Leftrightarrow S_{in}^{I}\left(a,A,\tilde{L}^{I}\right) \le 0$。

证明 由定理 2 易证结论成立。

根据上述结论，可以构造基于模糊互补熵的不完备多标签特征选择算法(IMFSFCE)。首先，计算 a 关于 A 和 $\tilde{L}^{I}$ 的特征内重要度，将内重要度为正的特征子集作为初始特征子集。然后，分别计算 $\tilde{L}^{I}$ 相对于特征子集 B 和全集 A 的模糊条件组合熵，若 $FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid B\right) < FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A\right)$，则输出特征子集；否则，从未选特征中挑选外重要度最大的特征加入 B，直到 $\tilde{L}^{I}$ 相对于特征子集 B 的模糊条件组合熵小于 $\tilde{L}^{I}$ 相对于全集 A 的模糊条件组合熵。具体特征选择过程见算法 1。

算法 1 基于模糊互补熵的不完备多标签特征选择(IMFSFCE)

输入：不完备多标签模糊信息系统 $IMFIS=(U,A,f,L,Y)$，$\delta \in [0,1]$。

输出：特征子集 B。

1 初始化 $B \leftarrow \varnothing$；

2 $\forall a \in A$，计算模糊关系 R_{A}^{I}、$R_{A-\{a\}}^{I}$，模糊信息粒 $[x_i]_{A}^{I,\delta}$、$[x_i]_{A-\{a\}}^{I,\delta}$；

3 计算模糊标签粒 $\tilde{L}_{j}^{I}$ 和多标签模糊粒覆盖 $\tilde{L}^{I}$；

4 由式(24)计算 a 的特征内重要度 $S_{in}^{I}\left(a,A,\tilde{L}^{I}\right)=FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A-\{a\}\right)-FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A\right)$，若 $S_{in}^{I}\left(a,A,\tilde{L}^{I}\right)>0$，则 $B \leftarrow B \cup \{a\}$；

5 计算 $FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid B\right)$，若 $FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid B\right) < FCE_{\delta}^{I}\left(\tilde{L}^{I} \mid A\right)$，执行步骤 7；否则执行步骤 6；

6 $\forall b \in A-B$，由式(25)计算 b 的特征外重要度 $S_{out}^{I}\left(a,B,\tilde{L}^{I}\right)$，若 $S_{out}^{I}\left(a,B,\tilde{L}^{I}\right)=\max S_{out}^{I}\left(a,B,\tilde{L}^{I}\right)$，则 $B \leftarrow B \cup \{b\}$，并执行步骤 5；

7 输出特征子集 B。

5. 实验

5.1. 实验环境

为验证 IMFSFCE 算法的有效性，选取 Mulan 数据库 5 个多标签数据集进行实验分析。表 1 列出了 5 个多标签数据集的相关信息。实验采用多标签 K 最近邻(Multi-Label K-Nearest Neighbor, ML-KNN)分类器，近邻数量设置为 10，平滑参数设置为 0.1 [19]。

Table 1. Information of multi-label datasets
表 1. 多标签数据集信息

数据集	样本数	特征数	标签数	领域	训练样本数	测试样本数
Flags	194	19	7	Images	129	65
Emotions	593	72	6	Music	391	202

续表

Cal500	502	68	174	Music	251	251
Water quality	1060	16	14	Chemistry	530	530
Virus	207	440	6	Biology	124	83

本文采用平均精度(Average Precision, AP)、汉明损失(Hamming Loss, HL)、排序损失(Ranking Loss, RL)、覆盖率(Coverage, CV)、1-错误率(One Error, OE)作为分类评价指标[20]。其中，AP 值越高，表明分类性能越好；HL、RL、CV、OE 值越低则分类性能越好。后续实验用符号“↑”表示“值越大分类性能越优”，符号“↓”表示“值越小分类性能越优”。在实验结果的呈现中，最优值以粗体形式突出显示。

评价指标的具体定义[21]如下：

设不完备多标签模糊信息系统 $IMFIS=(U,A,f,L,Y)$。训练集 $D=\{(x_i,y_i)\mid 1\le i\le n, x_i\in U, y_i\in Y\}$，$f(x_i,l_j)$ 为样本 x_i 具有标签 l_j 的概率，y_i 为样本 x_i 的真实标签向量，y_i' 为多标签分类器预测样本 x_i 的标签向量，$rank(\cdot,\cdot)$ 为 $f(\cdot,\cdot)$ 的排序函数。

(1) 平均精度(AP)用于衡量模型预测标签集合中的标签排序在整体排序中的表现。AP 值越大，说明模型对相关标签的识别与排序更准确，分类性能越优：

$$\mathrm{AP}=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|y_i|}\sum_{l\in y_i}\frac{\left|\{l'\in y_i\mid rank(x_i,l')\le rank(x_i,l)\}\right|}{rank(x_i,l)}, \tag{18}$$

(2) 汉明损失(HL)用于衡量模型在标签空间上产生的错误预测比例。该指标反映了样本标签被误分类的次数。HL 值越小，说明模型在标签判断上产生的错误更少，分类性能越优：

$$\mathrm{HL}=\frac{1}{n}\sum_{i=1}^{n}\frac{|y_i\oplus y_i'|}{k}, \tag{19}$$

(3) 排序损失(RL)用于衡量模型在排序过程中将无关标签排在相关标签之前的次数。该指标反映了排序错误的程度。RL 值越小，说明模型在区分相关与无关标签时的排序能力越强，分类性能越优：

$$\mathrm{RL}=\frac{1}{n}\sum_{i=1}^{n}\frac{1}{|y_i||\bar{y}_i|}\left|\{(l_1,l_2)\mid rank(x_i,l_1)>rank(x_2,l_2),(l_1,l_2)\in y_i\times\bar{y}_i\}\right|, \tag{20}$$

(4) 覆盖率(CV)用于衡量预测结果中，为包含全部相关标签所需在排序列表上向下遍历的平均距离。CV 值越小，说明模型更容易在较前的位置找到所有相关标签，分类性能越优：

$$\mathrm{CV}=\frac{1}{n}\sum_{i=1}^{n}\max_{l\in y_i} rank(x_i,l)-1, \tag{21}$$

(5) 1-错误率(OE)用于统计预测排名第一的标签未包含在样本真实标签集合中的次数。OE 值越小，说明模型对最相关标签的识别越可靠，分类性能越优：

$$\mathrm{OE}=\frac{1}{n}\sum_{i=1}^{n}\left[\left[\arg\max_{l\in L} f(x_i,l)\right]\notin y_i\right]. \tag{22s}$$

5.2. 参数分析

由于原始多标签数据集是完整的，本文采用随机缺失方法使数据集不完整，δ 以步长 0.1 在 $[0.1,0.5]$ 内取值，取缺失率 10%、20%、30%、40%、50%分别进行实验。5 个数据集在不同取值下的评价指标如图 1~5 所示。

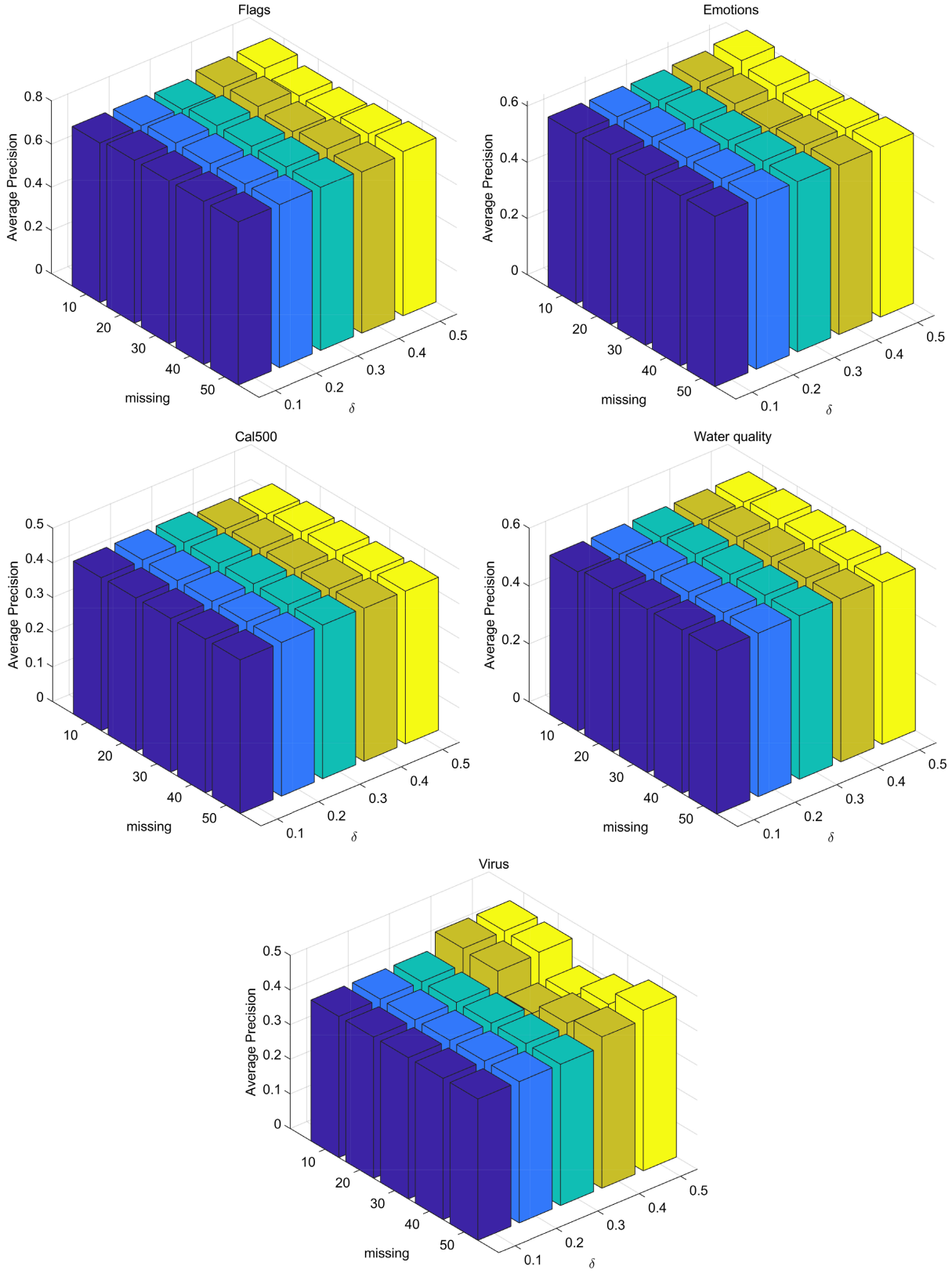

Figure 1. The AP index results of multi-label datasets under different values
图 1. 多标签数据集在不同取值下的 AP 指标结果

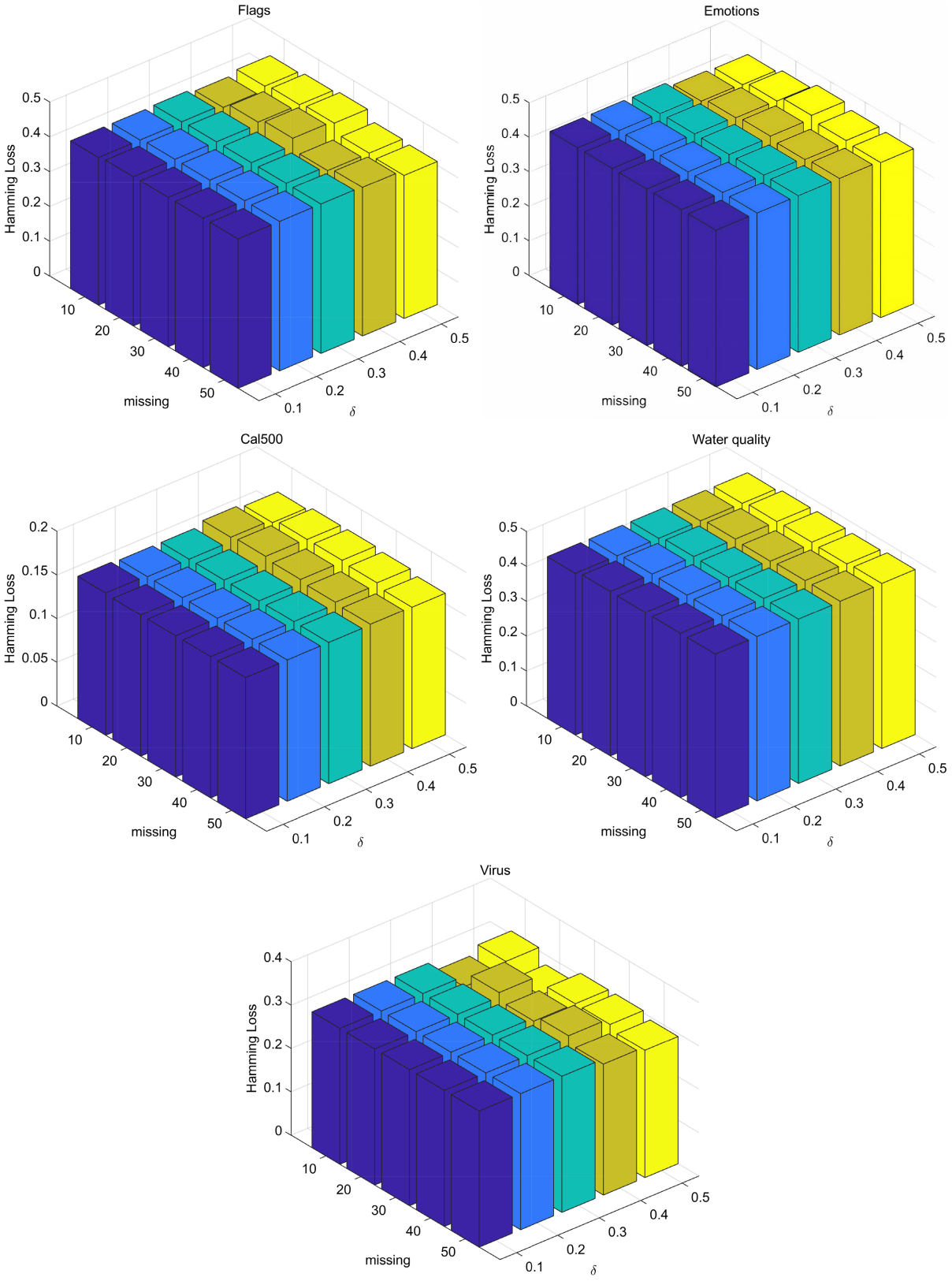

Figure 2. The HL index results of multi-label datasets under different values
图 2. 多标签数据集在不同取值下的 HL 指标结果

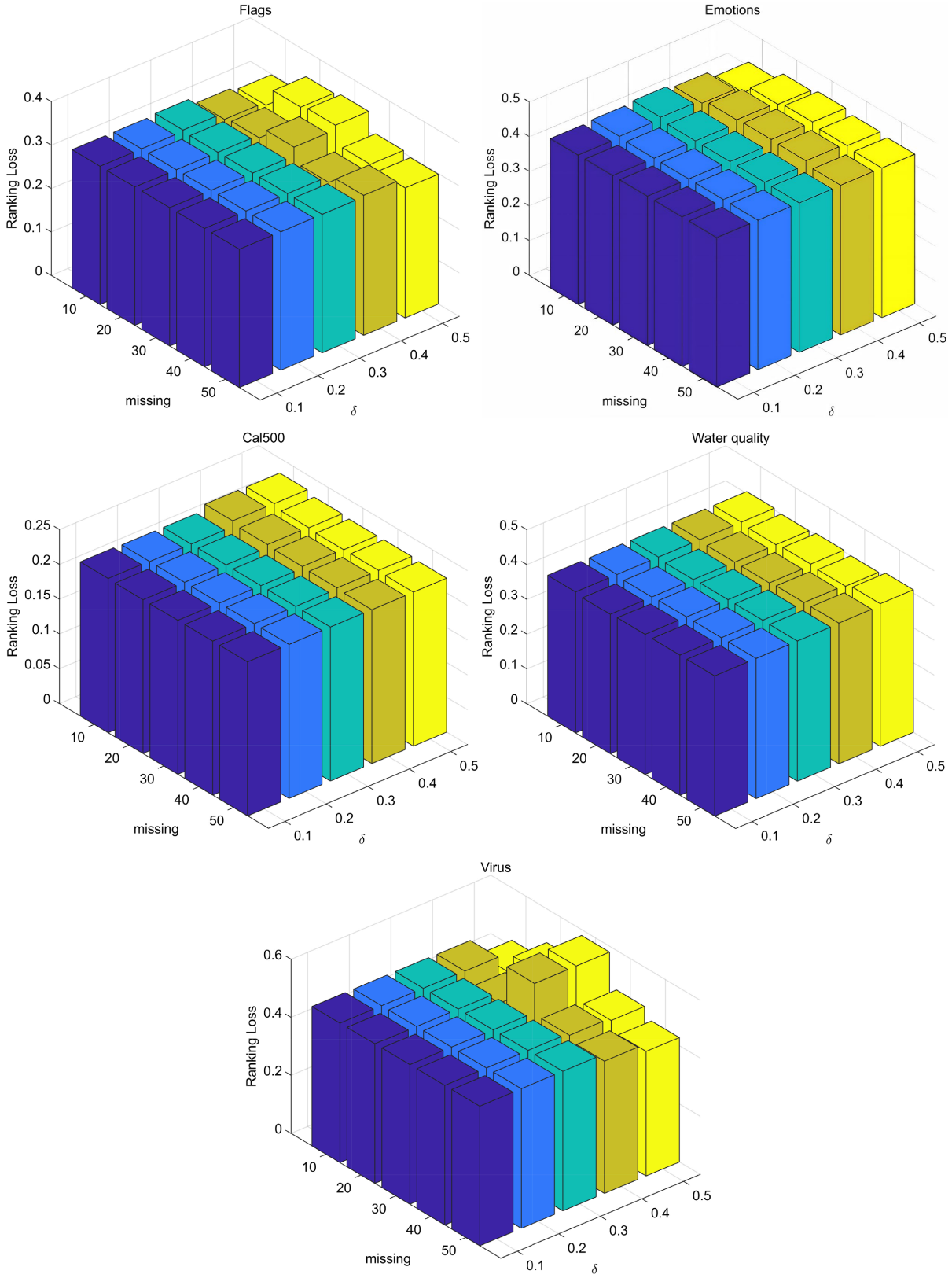

Figure 3. The RL index results of multi-label datasets under different values
图 3. 多标签数据集在不同取值下的 RL 指标结果

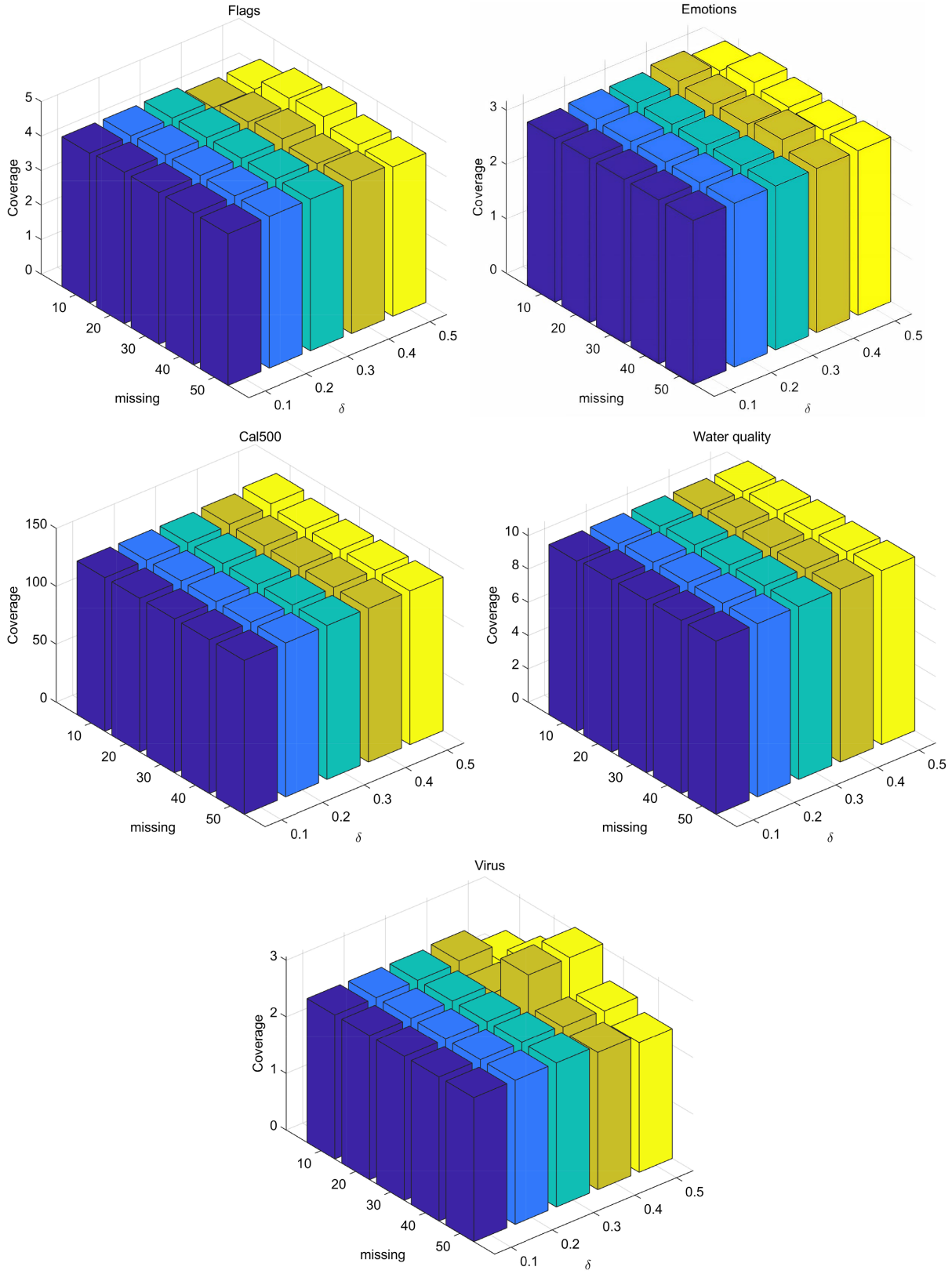

Figure 4. The CV index results of multi-label datasets under different values
图 4. 多标签数据集在不同取值下的 CV 指标结果

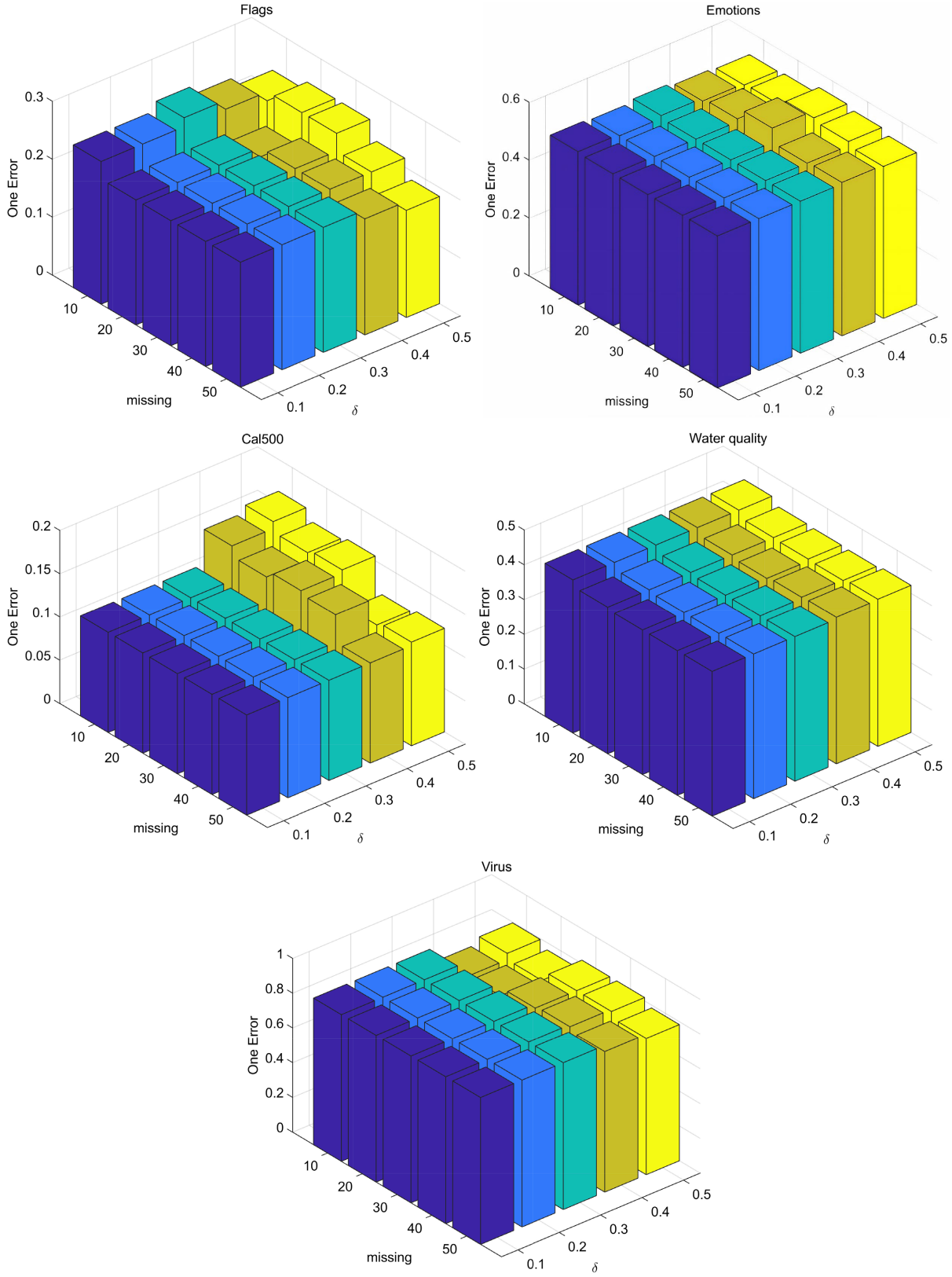

Figure 5. The OE index results of multi-label datasets under different values
图 5. 多标签数据集在不同取值下的 OE 指标结果

5.3. 实验结果

本文选取能够使评价指标达到最优的缺失率与 δ 作为最终的参数设置。在此基础上，对所提IMFSFCE 算法、使用全部特征的方法以及不同消融设置(无 β 、无 μ_a 、无 β 与 μ_a)进行对比实验。实验结果如表 2 所示。实验结果中，最优指标值以粗体形式表示。

Table 2. Experimental results of different methods on five evaluation metrics for each datasets
表 2. 不同方法在各数据集上的五种评价指标实验结果

数据集	方法	AP (↑)	HL (↓)	RL (↓)	CV (↓)	OE (↓)
Flags	IMFSFCE	**0.7596**	**0.4110**	**0.2987**	**4.3231**	**0.1846**
	使用全部特征	0.7536	0.4286	0.3179	4.3846	0.2154
	无 β	0.7536	0.4286	0.3179	4.3846	0.2154
	无 μ_a	0.7473	0.4198	0.3185	4.3846	0.2154
	无 β 、 μ_a	0.7536	0.4286	0.3179	4.3846	0.2154
Emotions	IMFSFCE	**0.6172**	**0.4356**	**0.4110**	**2.9851**	0.5198
	使用全部特征	0.6047	0.4530	0.4311	3.0149	0.5248
	无 β	0.6002	0.4513	0.4303	3.0743	0.5297
	无 μ_a	0.6123	0.4480	0.4244	3.1436	**0.5149**
	无 β 、 μ_a	0.6040	0.4513	0.4323	3.0248	0.5248
Cal500	IMFSFCE	**0.4421**	**0.1614**	**0.2208**	**132.5339**	**0.1155**
	使用全部特征	0.4421	0.1614	0.2208	132.5339	0.1155
	无 β	0.4388	0.1654	0.2217	132.4303	0.1155
	无 μ_a	0.4421	0.1614	0.2208	132.5339	0.1155
	无 β 、 μ_a	0.4421	0.1614	0.2208	132.5339	0.1155
Water quality	IMFSFCE	**0.5641**	**0.4673**	**0.4000**	**10.3736**	**0.4133**
	使用全部特征	0.5632	0.4704	0.4020	10.3830	0.4152
	无 β	0.5632	0.4704	0.4020	10.3868	0.4152
	无 μ_a	0.5612	0.4714	0.4052	10.4245	0.4152
	无 β 、 μ_a	0.5632	0.4704	0.4020	10.3830	0.4152
Virus	IMFSFCE	**0.4633**	**0.2932**	**0.4283**	**2.2892**	**0.7831**
	使用全部特征	0.4071	0.3133	0.4802	2.5422	0.8434
	无 β	0.4071	0.3133	0.4802	2.5422	0.8434
	无 μ_a	0.4048	0.3173	0.4775	2.5422	0.8554
	无 β 、 μ_a	0.4071	0.3133	0.4802	2.5422	0.8434

由表 2 可知，对于 AP 指标，IMFSFCE 算法在所有数据集上的实验结果均不低于其余四种对比方法；对于 HL、RL、CV 指标，IMFSFCE 算法在所有数据集上的实验结果均不高于其余四种对比方法；对于 OE 指标，IMFSFCE 算法在 Flags、Cal500、Water quality、Virus 数据集上的实验结果均不高于其余四种对比方法，在 Emotions 数据集上与最优方法相差 0.0049。相较于其余四种对比方法，IMFSFCE 算法在 5 个数据集上的 AP 平均提升 3.48%，HL、RL、CV 和 OE 平均下降 3.02%、4.33%、2.83%和 4.64%。上述结果表明，IMFSFCE 算法在保证平均精度提升的同时，有效控制了多种误差相关指标的增幅，整体上获得了更优的分类性能。

为进一步评估 IMFSFCE 算法在数据维度压缩方面的表现，本文统计了该算法在各数据集上的约简率，如表 3 所示。

Table 3. Feature reduction ratios of each dataset
表 3. 各数据集的特征约简率

数据集	原始特征数	选择特征数	约简率
Flags	19	4	78.95%
Emotions	72	22	69.44%
Cal500	68	32	52.94%
Water quality	16	11	31.25%
Virus	440	3	99.32%

由表 3 可知，IMFSFCE 算法在各数据集上均能有效降低数据维度，约简率介于 31.25%~99.32%之间。综上，IMFSFCE 算法能够在有效压缩数据维度的同时保持良好的分类性能，实验结果验证了本文算法的有效性，适用于不完备多标签特征选择场景。

6. 结论

本文提出了基于模糊组合熵的不完备多标签特征选择。在不完备多标签模糊信息系统中，通过改进模糊关系，定义了模糊信息粒、模糊标签粒以及多标签模糊下近似与上近似，建立了不完备多标签模糊粗糙集。进一步地，将组合熵引入不完备多标签模糊粗糙集，定义了模糊组合熵、模糊联合组合熵、模糊条件组合熵等信息度量，给出了基于模糊组合熵的特征内外重要度。最后，基于信息度量与重要度分析，提出了不完备多标签模糊粗糙集上的多标签特征选择算法。在 5 个多标签数据集上的实验结果表明，本文所提特征选择算法能够在不完备场景下改善多标签分类性能，验证了该方法的可行性与有效性。

参考文献

[1] Zhang, P., Liu, G. and Song, J. (2023) MFSJMI: Multi-Label Feature Selection Considering Join Mutual Information and Interaction Weight. *Pattern Recognition*, **138**, Article ID: 109378. https://doi.org/10.1016/j.patcog.2023.109378

[2] Li, Y., Hu, L. and Gao, W. (2024) Multi-Label Feature Selection with High-Sparse Personalized and Low-Redundancy Shared Common Features. *Information Processing & Management*, **61**, Article ID: 103633. https://doi.org/10.1016/j.ipm.2023.103633

[3] Sheikhpour, R., Mohammadi, M., Berahmand, K., Saberi-Movahed, F. and Khosravi, H. (2025) Robust Semi-Supervised Multi-Label Feature Selection Based on Shared Subspace and Manifold Learning. *Information Sciences*, **699**, Article ID: 121800. https://doi.org/10.1016/j.ins.2024.121800

[4] Dai, J. and Wang, J. (2025) Multi-Label Feature Selection with Missing Features by Tolerance Implication Granularity

Information and Symmetric Coupled Discriminant Weight. *Pattern Recognition*, **162**, Article ID: 111365. https://doi.org/10.1016/j.patcog.2025.111365

[5] Han, Y., Sun, G., Shen, Y. and Zhang, X. (2018) Multi-Label Learning with Highly Incomplete Data via Collaborative Embedding. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 1494-1503. https://doi.org/10.1145/3219819.3220038

[6] Dai, J., Chen, W., Qian, Y. and Pedrycz, W. (2025) Instance-Dependent Incomplete Multi-Label Feature Selection by Fuzzy Tolerance Relation and Fuzzy Mutual Implication Granularity. *IEEE Transactions on Knowledge and Data Engineering*, **37**, 5994-6008. https://doi.org/10.1109/tkde.2025.3591461

[7] Li, J., Li, P., Zou, Y. and Hu, X. (2021) Multi-Label Learning with Missing Features. 2021 *International Joint Conference on Neural Networks* (*IJCNN*), Shenzhen, 18-22 July 2021, 1-8. https://doi.org/10.1109/ijcnn52387.2021.9533967

[8] Pawlak, Z. (1982) Rough Sets. *International Journal of Computer & Information Sciences*, **11**, 341-356. https://doi.org/10.1007/bf01001956

[9] Lin, Y., Li, Y., Wang, C. and Chen, J. (2018) Attribute Reduction for Multi-Label Learning with Fuzzy Rough Set. *Knowledge-Based Systems*, **152**, 51-61. https://doi.org/10.1016/j.knosys.2018.04.004

[10] Chen, P., Lin, M. and Liu, J. (2020) Multi-Label Attribute Reduction Based on Variable Precision Fuzzy Neighborhood Rough Set. *IEEE Access*, **8**, 133565-133576. https://doi.org/10.1109/access.2020.3010314

[11] Sun, L., Du, W., Ding, W., Long, Q. and Xu, J. (2025) Granular Ball-Based Fuzzy Multineighborhood Rough Set for Feature Selection via Label Enhancement. *Engineering Applications of Artificial Intelligence*, **145**, Article ID: 110191. https://doi.org/10.1016/j.engappai.2025.110191

[12] 苗夺谦. Rough Set 理论及其在机器学习中的应用研究[Z]. 北京: 中国科学院自动化研究所, 1997.

[13] Qian, Y. and Liang, J. (2006) Combination Entropy and Combination Granulation in Incomplete Information System. In: *Proceedings of the Rough Sets and Knowledge Technology*, Springer, 184-190. https://doi.org/10.1007/11795131_27

[14] Zhang, P., Li, T., Yuan, Z., Luo, C., Liu, K. and Yang, X. (2024) Heterogeneous Feature Selection Based on Neighborhood Combination Entropy. *IEEE Transactions on Neural Networks and Learning Systems*, **35**, 3514-3527. https://doi.org/10.1109/tnnls.2022.3193929

[15] Yang, T., Wang, C., Chen, Y. and Deng, T. (2025) A Robust Multi-Label Feature Selection Based on Label Significance and Fuzzy Entropy. *International Journal of Approximate Reasoning*, **176**, Article ID: 109310. https://doi.org/10.1016/j.ijar.2024.109310

[16] Liao, C. and Yang, B. (2025) A Novel Multi-Label Feature Selection Method Based on Conditional Entropy and Its Acceleration Mechanism. *International Journal of Approximate Reasoning*, **185**, Article ID: 109469. https://doi.org/10.1016/j.ijar.2025.109469

[17] 陈曦, 马建敏, 刘权芳. 基于模糊依赖决策熵的多标签特征选择[J]. 昆明理工大学学报(自然科学版), 2024, 49(2): 62-72.

[18] Dai, J. (2013) Rough Set Approach to Incomplete Numerical Data. *Information Sciences*, **241**, 43-57. https://doi.org/10.1016/j.ins.2013.04.023

[19] Zhang M.L., Zhou Z.H. (2014) A Review on Multi-Label Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, **26**, 1819-1837. https://doi.org/10.1109/tkde.2013.39

[20] Zhang M.-L. and Zhou Z.-H. (2007) ML-KNN: A Lazy Learning Approach to Multi-Label Learning. *Pattern Recognition*, **40**, 2038-2048. https://doi.org/10.1016/j.patcog.2006.12.019

[21] 陈曦. 三种多标签数据表上的特征选择方法[D]: [硕士学位论文]. 西安: 长安大学, 2024.