

# 面向智慧知识服务的多案例诱导适配机制

张建华, 温丹丹, 曹子傲, 高亚瑞

(郑州大学 管理学院 河南 郑州 450001)

**摘要:** 为提升智慧知识服务中案例知识适配的自主性和精确性,提出一种基于决策属性间相关关系的多案例诱导适配机制。首先,基于概率论判断决策属性的相关性;然后,根据决策属性独立或相关状况分别选择不同的适配方法。在决策属性相关情形下,根据决策属性个数和值域规模的差异分别选择基于幂集、基于适配器链和基于适配器链组合的适配方法。在 Student Performance-MAT 数据集上的仿真实验表明,所提机制具有良好的性能。

**关键词:** 智慧知识服务; 决策属性相关; 多案例诱导; 适配机制

中图分类号: G203

文献标志码: A

文章编号: 1671-6841(2025)04-0080-08

DOI: 10.13705/j.issn.1671-6841.2024008

## Multi-case Derivational Adaptation Mechanism for Intelligent Knowledge Service

ZHANG Jianhua, WEN Dandan, CAO Ziao, GAO Yarui

(School of Management, Zhengzhou University, Zhengzhou 450001, China)

**Abstract:** The proposed mechanism aimed to enhance autonomy and accuracy in adapting case knowledge within intelligent knowledge services. It utilized a multi-case induced adaptation approach that hinged on the correlation between decision attributes. Initially, the mechanism assessed the relevance of decision attributes using probability theory. Then, it employed different adaptation methods based on the independence or relevance of these attributes. When decision attributes were correlated, the mechanism selected adaptation techniques such as power set-based adaptation, adapter chain-based adaptation, and adapter chain combination-based adaptation. These choices depended on variations in the number of decision attributes and the size of the value domain. Simulation experiments conducted on the Student Performance-MAT dataset demonstrated the effectiveness of the mechanism.

**Key words:** intelligent knowledge service; relevant decision attributes; multi-case derivational; adaptation mechanism

### 0 引言

人工智能的加持促使知识服务不断升级服务的广度和深度,知识服务进入智慧化服务阶段<sup>[1]</sup>。智慧知识服务是指服务提供者基于自身的知识能力,根据用户面临的知识情境实时组织知识资源、指导用户知识实践、帮助用户实现知识创新的过程<sup>[2]</sup>。

智慧知识服务的本质是供需的动态匹配,知识服务的内容与用户能力、需求的匹配度决定着用户的服务体验<sup>[3]</sup>。随着用户需求的个性化、情境化、综合化发展,智慧知识服务面临着用户知识能力不足、知识应用情境复杂等挑战<sup>[4]</sup>。能呈现知识应用情境并辅助知识实践的案例知识可以有效应对上述挑战,并逐渐成为智慧知识服务领域的重要研究议题<sup>[2,5-6]</sup>。学者围绕案例知识的结构设计<sup>[2,7]</sup>、检索

收稿日期:2024-01-11

基金项目:国家社会科学基金项目(19BTQ035)

第一作者:张建华(1975—),男,教授,主要从事知识服务与知识创新研究,E-mail:tjzhangjianhua@163.com。

通信作者:温丹丹(1986—),女,博士研究生,主要从事隐性知识服务研究,E-mail:mydwdd@126.com

和匹配<sup>[8]</sup>、适配和重用<sup>[5]</sup>开展了大量研究。案例适配基于对“相似性假设”的质疑,谋求基于检出案例的解决方案能产生所遇新问题的解决方案,也是实现知识服务过程中低领域知识依赖、自助化的重要抓手<sup>[2,9]</sup>。

案例适配的相关研究可归纳为两类:一是设计案例决策属性值的修改方法<sup>[10]</sup>,即转换适配;二是设计由条件属性到决策属性的推导方法,即诱导适配<sup>[11]</sup>。早期,学者对转换适配方法的研究较多,随着知识服务对象向大众转变,诱导适配方法逐渐得到关注<sup>[12]</sup>。张建华<sup>[11]</sup>认为多案例诱导适配对领域知识的依赖更少、适配结果更为良好,并详细阐述了多案例诱导适配的内涵及适配机制,为研究多案例诱导适配提供了框架。后续学者在此基础上提出了简单案例诱导适配方法<sup>[12-14]</sup>。随着知识应用情境逐渐复杂,部分学者开始关注复杂案例的诱导适配问题,研究多决策属性、相关属性的案例诱导适配<sup>[15-16]</sup>。学者普遍认为,属性之间的依赖关系影响案例诱导适配结果<sup>[17]</sup>,但当前的研究大多将决策属性集视为整体来探讨适配方法,忽略了决策属性集内部相关关系对适配结果的影响<sup>[13,15]</sup>。也有研究通过对具体分类算法的改进来提升适配效率,适配效果高度依赖具体算法,制约了算法的泛化能力<sup>[11,13-14]</sup>。

综上,多案例诱导适配能提升知识服务过程的自助性和服务内容的精准性,但是相关的研究尚未形成完整的体系。同时,针对复杂结构案例的诱导适配研究尚不够深入。鉴于此,本文拟基于智慧知识服务视角,探讨决策属性不同相关情形的多案例诱导适配机制。首先,基于概率理论定量地描述决策属性间的相关关系;其次,基于多标签分类框架、问题分解思想和模块思想深入分析决策属性间的相关性对多案例诱导适配的影响,并建模决策属性值的推导过程,以期在保证适配效果的同时,还能提升适配效率。

## 1 问题描述

多案例诱导适配的主要思想是基于相似案例集挖掘条件属性集对决策属性集的决定关系,利用该关系求解目标案例。通常将案例知识纳入简化的知识表达系统  $S$  中,  $S = (U, A, V, f)$ , 其中:  $U$  为适配案例集;  $A$  为案例的属性集, 包含条件属性集  $C$  和决策属性集  $D$ ;  $V$  为属性的值域;  $f$  为信息函数, 表达了  $A \rightarrow V$  的决定关系。  $U = \{U_1, U_2, \dots, U_e\}$ ,  $U_i$  表示第  $i$

个适配案例,  $U_0$  为目标案例;  $C = \{C_1, C_2, \dots, C_h\}$ ,  $C_m$  为第  $m$  个条件属性,  $(c_{i1}, c_{i2}, \dots, c_{ih})$  为第  $i$  个案例的条件属性向量,  $C_0 = (c_{01}, c_{02}, \dots, c_{0h})$  表示目标案例的条件属性向量;  $D = \{D_1, D_2, \dots, D_k\}$ ,  $D_j$  表示第  $j$  个决策属性,  $(d_{i1}, d_{i2}, \dots, d_{ik})$  表示第  $i$  个案例的决策属性向量。本文目标是给定适配案例集  $U$ , 针对目标案例  $U_0$  的条件属性向量  $C_0$ , 运用可行的方法确定目标案例的决策属性向量  $D_0$ 。

目标导向式问题分解是解决复杂问题的基本方式, 分解获得的子问题按分解顺序构成问题链, 各子问题的求解模型构成模型链, 将子问题的解按顺序集成可得到原复杂问题的解。该思路为解决复杂问题提供了更大的灵活性, 因为分解后不仅任何解决原子问题的方法皆可用, 而且易于实现特定算法或集成框架。决策属性间不同的相关性决定了问题的分解方式。基于上述思想, 多案例诱导适配的求解步骤如下。

- 1) 分析决策属性间的相关关系。
- 2) 确定问题分解方式, 构建问题链。
- 3) 选择原子问题适配解的合适求解模型, 构建模型链。
- 4) 依次求解原子问题的适配解。
- 5) 集成所有原子问题的解, 得到目标问题的解。

## 2 多案例诱导适配框架

### 2.1 决策属性相关性分析

相关性分析是研究变量之间关系常见的统计方法, 相关性直观反映了属性之间的相互影响关系及程度。案例知识中存在条件属性间相关、条件属性与决策属性间相关、决策属性间相关三种属性相关关系。前两种相关关系对案例适配的影响可以分别通过属性约简和权重配置解决<sup>[14-16,18-19]</sup>, 本文重点关注决策属性间的相关。Wang 等<sup>[20]</sup>从概率论视角定义了标签之间的相关关系, 本文根据该定义对决策属性间的相关做如下定义。

**定义 1** 对于任意的决策属性对  $(D_m, D_n) \in D, m \neq n, P(d_{im}, d_{in})$  是联合概率分布,  $P(d_{im})$  和  $P(d_{in})$  是边缘概率分布, 当且仅当

$$P(d_{im}, d_{in}) = P(d_{im}) * P(d_{in})$$

时, 认为决策属性  $D_m$  和  $D_n$  相互独立。推广到所有的决策属性, 当且仅当

$$P(d_{i1}, d_{i2}, \dots, d_{ik}) = P(d_{i1}) * P(d_{i2}) * \dots * P(d_{ik}), i = \{1, 2, \dots, e\}$$

时, 认为所有决策属性是相互独立的。

## 2.2 问题分解及适配模型

**2.2.1 决策属性独立的多案例诱导适配** 当决策属性相互独立时,决策属性集对条件属性集的依赖关系由每个决策属性对条件属性集依赖关系的并集构成。此时,案例诱导适配过程等价于多个单决策属性案例诱导适配过程的重复和叠加。针对每个决策属性可以分别构造相同或不同的适配解求解模型,然后将每个模型的解组合构成目标问题的决策属性向量。此时,多个适配模型可并行运行,从而降低运行时耗。具体流程见算法1。

### 算法1 决策属性独立的多案例诱导适配

输入:  $h$  个条件属性,  $k$  个决策属性,  $e$  个相似案例的适配案例集  $U$ , 以及目标案例的条件属性向量  $C_0$ 。

输出: 目标问题的决策属性向量  $D_0$ 。

Step1 将案例适配问题划分为  $k$  个独立的单决策属性案例适配问题。

Step2 分别构造  $k$  个适配解轨迹求解模型。

Step3 组合  $k$  个决策属性值构造目标案例的决策属性向量  $D_0$ 。

### 2.2.2 决策属性相关的多案例诱导适配

1) 基于决策属性幂集的多案例诱导适配。处理决策属性相关的简单思路是:构造一个新的复合决策属性  $D'$ , 以隐藏决策属性间的相关关系。 $D'$  是原决策属性集的幂集,即  $D'$  的值域是原决策属性值的所有可能组合,此时可将多决策属性的案例诱导适配问题转化为单决策属性案例诱导适配<sup>[20-21]</sup>。给定一个待解问题,基于幂集的适配器输出最有可能的决策属性值为  $d'$ ,  $d'$  实际上是一组决策属性值,具体流程见算法2。由此可见,基于幂集的适配方法隐式地、充分地考虑原决策属性间的相关性。算法复杂度为  $O(N * K * W)$ , 其中:  $N$  为适配案例集决策属性的个数;  $K$  为所有决策属性值域规模之积;  $W$  为基算法的复杂度。基于幂集的适配方法对决策属性域较小的问题比较有效,但是随着决策属性个数和取值范围的扩大,复合决策属性将面临空间爆炸,且构建复合决策属性大大增加了数据稀疏性。

### 算法2 基于幂集的多案例诱导适配

输入:  $h$  个条件属性,  $k$  个决策属性,  $e$  个相似案例的适配案例集  $U$ , 以及目标案例的条件属性向量  $C_0$ ;

输出: 目标问题的决策属性向量  $D_0$ 。

Step1 构造新的决策属性  $D'$ , 建立  $d'$  与  $\{d_1, d_2, \dots, d_k\}$  的映射关系。

Step2 根据映射关系,用  $D'$  的值替代原案例集中每个案例的决策属性向量。

Step3 构造  $D'$  的适配解求解模型,并求解  $D'_0$ 。

Step4 根据映射关系,确定  $D_0$ 。

2) 基于适配器链的多案例诱导适配。当决策属性之间存在相关关系时,其中一个决策属性可以通过其他决策属性预测得到。链型分类器(classifier chain, CC)算法是解决多标签分类问题的基本方法。对于给定的标签序列,算法每次预测一个标签值,在预测当前标签时将之前所有的标签作为扩展的条件属性<sup>[21]</sup>。尽管附加属性只占全部条件属性空间的小部分,但如果属性之间存在强相关,那么这些属性将赋予基分类器更强的预测能力<sup>[22]</sup>。研究显示,基于CC的分类方法在解决标签相关的分类问题中具有诸多优点,比如低时间复杂性、高准确率等<sup>[20]</sup>。基于CC算法构建多案例诱导适配器链的具体思路是:对于给定的决策属性序列  $\{D_1, D_2, \dots, D_k\}$ , 依次分析决策属性对条件属性集的依赖关系,且在分析当前决策属性的依赖关系时,须将之前所有的决策属性作为条件属性的扩展,具体流程见算法3。决策属性间的相关关系依托适配器链进行传导,虽然每个案例平均增加了  $|k|/2$  个决策属性,但是当  $k$  的值较小时,增加的算法复杂性可以忽略不计<sup>[22]</sup>。每个链条节点可采用不同的算法构建适配器;若所有节点采用相同算法,则既可以有效节约内存,亦可有效降低算法复杂度。确定决策属性序列是基于适配器链适配方法的重点,现有研究常使用随机法、多条链组合法等。与算法2相比,适配器链算法受决策属性域的约束较小,当决策属性的值域规模较大时,算法会有更好的表现。

### 算法3 基于适配器链的多案例诱导适配

输入:  $h$  个条件属性,  $k$  个决策属性,  $e$  个相似案例的适配案例集  $U$ , 以及目标案例的条件属性向量  $C_0$ , 决策属性序列  $\{D_1, D_2, \dots, D_k\}$ 。

输出: 目标问题的决策属性向量  $D_0$ 。

Step1 将案例适配问题划分为  $k$  个单决策属性案例适配问题。

Step2 令  $n=0$ 。

Step3  $n=n+1$ , 判断  $n$  是否大于  $k$ , 如果是, 转 step6。

Step4 构造  $D_n$  的适配模型并求解。

Step5 将  $D_n$  作为扩展的条件属性, 转 Step3。

Step6 组合  $k$  个决策属性值, 构造决策属性向量  $D_0$ 。

3) 基于适配器链组合的多案例诱导适配。相

关关系并不总是存在于所有的决策属性之间,充分利用决策属性间的独立和相关关系对决策空间降维是改进适配算法的有效途径<sup>[20]</sup>。基于相关性分析,将初始决策属性集划分为低相关的多个子集,每个子集可能包含多个决策属性,针对每个决策属性子集构建适配器或适配器链。然后,利用无监督方法将决策属性子集划分为多个相关性较低的次子集,并将每个次子集作为适配器链的节点。节点内部充分考虑决策属性间的相关性,基于幂集方法构建适配器。如此,可充分利用决策属性间的独立和相关关系。适配器链组合虽然需要更多的计算资源,但仍然比直接考虑所有决策属性的适配器链方法有更低的时耗<sup>[20]</sup>,具体流程见算法4。

**算法4** 基于适配器链组合的多案例诱导适配  
输入:  $h$  个条件属性,  $k$  个决策属性,  $e$  个相似案例的适配案例集  $U$ , 以及目标案例的条件属性向量  $C_0$ 。

输出: 目标问题的决策属性向量  $D_0$ 。

Step1 分析决策属性的相关性,将初始决策属性集划分为低相关的决策属性子集。

Step2 分别用聚类算法对每个决策属性子集进行聚类分析,得到相关性较低的多个决策属性次子集。

Step3 对每个次子集调用算法2。

Step4 对每个决策属性子集调用算法3。

Step5 集结所有适配器的输出,构成目标案例的决策属性向量。

**2.2.3 多案例诱导适配集成框架** 鉴于决策属性间相关关系的复杂性,在实践中往往需要采用不同的策略对案例进行诱导适配,有必要把不同的适配策略集成在一起。用问题转换及分解的方法为构建多案例诱导适配的集成框架提供了便捷。如此,在不同诱导适配策略的基础上,建立一个通用的集成框架,其具体流程见算法5。

**算法5** 多案例诱导适配集成框架

输入:  $h$  个条件属性,  $k$  个决策属性,  $e$  个相似案例的适配案例集  $U$ , 目标案例的条件属性向量  $C_0$ 。

输出: 目标案例的决策属性向量  $D_0$ 。

Step1 基于概率理论判断决策属性之间是否相互独立,如果是,调用算法1。

Step2 判断决策属性的值域规模是否大于给定的阈值,如果是,转Step3,否则调用算法2。

Step3 判断决策属性个数是否大于给定的阈值,如果是,调用算法4,否则调用算法3。

Step4 集结所有算法的输出构成目标案例的

决策属性向量。

Step5 判断对算法输出的目标案例决策属性向量是否满意,如果是,转Step6,否则进入人机交互进行调整。

Step6 输出目标案例的决策属性向量。

### 3 算例

选取UCI数据库中的“Student Performance-MAT”数据集验证本文方法的有效性。该数据集的每个案例包含30个条件属性和3个决策属性(即第一、第二、第三学年的成绩)。对初始数据做离散化处理构建简化的知识表达系统。采用协方差法对条件属性与决策属性进行相关性分析,选取与决策属性高度相关的Medu、Fedu、Fjob、traveltime、studytime、famsup、nursery、internet、health 9个属性作为条件属性集。设定值域规模阈值为10,决策属性个数阈值为5<sup>[22]</sup>。离散化的待解问题条件属性向量为(4,3,5,2,2,2,1,2,3),采用本文算法求解待解问题的决策属性向量,为知识用户提供非零基求解基础。

首先,基于属性独立性定义对决策属性的相关性进行判断。根据概率理论求得三个决策属性的联合概率分布( $P(D_1, D_2, D_3)$ )和边缘概率分布( $P(D_1) * P(D_2) * P(D_3)$ ),如表1所示。结合独立性定义和表1数据发现,三个决策属性均不独立,则该问题不适用算法1。

其次,判断案例的决策属性值域规模。三个决策属性皆有五个取值,则值域规模为125,大于给定阈值,因此该问题不适用算法2。

第三步,判断决策属性个数。决策属性的个数小于给定阈值,采用算法3对该案例进行适配。将决策空间划分为三个单决策属性案例适配空间。为了简化计算,三个适配链节点均将加权朴素贝叶斯算法作为基算法,采用海林格距离方法分别确定三个决策属性的适配案例视图。

基于熵权法和欧氏距离法确定案例视图相似度,对决策属性1求待解问题的适配案例集。基于加权朴素贝叶斯算法求解 $D1$ 的概率分布, $D1$ 的属性值为1,2,3,4,5的概率分别为0.0448,0.1325,0.2979,0.1477,0.0140,将概率最大的值确定为决策属性 $D1$ 的值,即 $D1$ 的值为3;将 $D1$ 作为条件属性的扩展,结合案例视图求解 $D2$ 的概率分布,并确定 $D2$ 的值, $D2$ 的属性值为1,2,3,4,5的概率分别为0.0143,0.1380,0.3962,0.0801,0.0143,确定 $D2$

表 1 决策属性的联合概率及边缘概率分布

Table 1 Joint probability and edge probability distribution of decision attributes

		$P(D_1, D_2, D_3)$					$P(D_1) * P(D_2) * P(D_3)$				
$D_3$	$D_1$	$D_2$					$D_2$				
		1	2	3	4	5	1	2	3	4	5
1	1	0.007 7	0	0	0	0	0.000 1	0.000 5	0.001 0	0.000 6	0.000 1
	2	0.017 9	0.043 6	0.002 6	0	0	0.000 8	0.005 3	0.011 1	0.006 9	0.001 3
	3	0.005 1	0.002 6	0.015 4	0	0	0.001 0	0.006 1	0.012 8	0.008 0	0.001 5
	4	0.002 6	0	0	0	0	0.001 0	0.006 3	0.013 1	0.008 2	0.001 5
	5	0	0	0	0	0	0.000 3	0.002 1	0.004 4	0.002 7	0.000 5
2	1	0	0.012 8	0	0	0	0.000 1	0.000 8	0.001 6	0.001 0	0.000 2
	2	0	0.094 9	0.023 1	0	0	0.001 4	0.008 7	0.018 2	0.011 3	0.002 1
	3	0	0.020 5	0.005 1	0	0	0.001 6	0.010 0	0.021 0	0.013 1	0.002 5
	4	0	0	0	0	0	0.001 6	0.010 3	0.021 5	0.013 4	0.002 5
	5	0	0	0	0	0	0.000 5	0.003 4	0.007 1	0.004 5	0.000 8
3	1	0	0	0	0	0	0.000 3	0.002 0	0.004 1	0.002 6	0.000 5
	2	0	0.020 5	0.059 0	0	0	0.003 5	0.022 3	0.046 8	0.029 2	0.005 5
	3	0	0.012 8	0.207 7	0.002 6	0	0.004 1	0.025 8	0.054 2	0.033 8	0.006 3
	4	0	0	0.092 3	0.015 4	0	0.004 2	0.026 4	0.055 4	0.034 6	0.006 5
	5	0	0	0	0	0	0.001 4	0.008 8	0.018 4	0.011 5	0.002 2
4	1	0	0	0	0	0	0.000 2	0.001 3	0.002 7	0.001 7	0.000 3
	2	0	0	0	0	0	0.002 4	0.014 9	0.031 3	0.019 5	0.003 7
	3	0	0	0.005 1	0.025 6	0	0.002 7	0.017 3	0.036 2	0.022 6	0.004 2
	4	0	0	0.020 5	0.176 9	0	0.002 8	0.017 7	0.037 0	0.023 1	0.004 3
	5	0	0	0.002 6	0.043 6	0	0.000 9	0.005 9	0.012 3	0.007 7	0.001 4
5	1	0	0	0	0	0	0	0.000 3	0.000 6	0.000 4	0.000 1
	2	0	0	0	0	0	0.000 5	0.003 2	0.006 7	0.004 2	0.000 8
	3	0	0	0	0	0	0.000 6	0.003 7	0.007 8	0.004 9	0.000 9
	4	0	0	0	0.002 6	0	0.000 6	0.003 8	0.008 0	0.005 0	0.000 9
	5	0	0	0	0.005 1	0.051 3	0.000 2	0.001 3	0.002 6	0.001 7	0.000 3

注:表格中的“0”为保留小数点后 4 位四舍五入后得到。

的值为 3;  $D_3$  的属性值为 1, 2, 3, 4, 5 的概率分别为 0.0673, 0.1535, 0.2950, 0.1118, 0.0158, 确定  $D_3$  的值为 3。将适配器链的输出进行集结, 构成待解问题的决策属性值向量。综上, 根据本文算法得到待解问题的决策属性向量是 (3, 3, 3)。知识用户可根据自身知识基础和实际情况判断输出解的质量, 若不满意, 可进入人机交互系统进行调整, 否则, 输出目标问题的解为 (3, 3, 3)。至此, 多案例诱导适配过程结束。最后结合知识用户的应用反馈将该案例交由知识管理系统, 由系统判断其是否进入案例库。

为验证算例所用方法的适用性和优越性, 基于算例数据集, 将算例方法与基于经典算法的算法 1 和算法 3 进行对比实验。针对三个决策属性对比实

验的适配准确率、精确率、召回率和  $F1$  值分别如表 2、表 3 所示。由表 2、表 3 可知, 基于各经典算法的适配器链的准确率、精确率、召回率和  $F1$  值均高于算法 1, 且在同一链条上, 三个决策属性的每个指标呈顺序递增状态。以决策树算法为例, 基于算法 3 的  $D_2$  和  $D_3$  的准确率明显高于算法 1, 且在该链条上,  $D_1$ 、 $D_2$  和  $D_3$  的适配准确率依次递增。上述结论在随机森林及随机森林链、KNN 及 KNN 链、朴素贝叶斯及朴素贝叶斯链、算例方法上同样成立。观察表 2、表 3 发现, 适配精确率、召回率和  $F1$  值基本呈现与准确率一致的规律。

由算例和实验可以发现本文方法具有如下特点。

表2 适配准确率和精确率对比

Table 2 Comparison of the adaptation accuracy and precision

方法	准确率			精确率		
	$D_1$	$D_2$	$D_3$	$D_1$	$D_2$	$D_3$
决策树	0.275	0.425	0.450	0.281	0.465	0.326
决策树链	0.275	0.675	0.825	0.281	0.671	0.834
随机森林	0.269	0.420	0.429	0.319	0.376	0.328
随机森林链	0.269	0.580	0.790	0.319	0.59	0.773
KNN	0.300	0.450	0.450	0.290	0.370	0.350
KNN 链	0.300	0.675	0.800	0.290	0.691	0.793
朴素贝叶斯	0.375	0.275	0.350	0.387	0.316	0.322
朴素贝叶斯链	0.375	0.575	0.825	0.387	0.594	0.845
算例方法	0.375	0.650	0.872	0.387	0.678	0.876

表3 适配召回率和 F1 值对比

Table 3 Comparison of the adaptation recall and F1 score

方法	召回率			F1 值		
	$D_1$	$D_2$	$D_3$	$D_1$	$D_2$	$D_3$
决策树	0.275	0.425	0.450	0.260	0.434	0.378
决策树链	0.275	0.675	0.825	0.260	0.663	0.826
随机森林	0.269	0.420	0.429	0.259	0.381	0.349
随机森林链	0.269	0.580	0.790	0.259	0.569	0.764
KNN	0.300	0.450	0.450	0.280	0.380	0.350
KNN 链	0.300	0.675	0.800	0.280	0.663	0.785
朴素贝叶斯	0.375	0.275	0.350	0.365	0.290	0.335
朴素贝叶斯链	0.375	0.575	0.825	0.365	0.575	0.833
算例方法	0.375	0.650	0.872	0.365	0.640	0.870

1) 充分利用决策属性间的相关关系。标签之间的相关性对多标签分类的影响已经获得了充分证实<sup>[20]</sup>,同理推断,决策属性间的相关性亦影响案例诱导适配。本文方法首先基于概率理论判断决策属性间的相关性,根据决策属性独立或相关,分别设计不同的多案例诱导适配方法。针对决策属性相关的情形,设计了基于幂集(算法2)、基于适配器链(算法3)和基于适配器链组合(算法4)的三种多案例诱导适配方法。算法2隐式、充分地利用决策属性间的相关关系,适用于决策属性值域规模较小,且所有决策属性间存在强相关关系的情形;算法3局部地利用决策属性间的相关关系,适用于决策属性值域规模较大,且存在序列相关的情形;算法4同时利用决策属性间的独立和相关关系,适用于决策属性个数较多,且同时存在独立和相关的情形。从利用决策属性相关关系视角分析,算法3和算法4是算法2的折中。对比实验表明,算法1(基于某个经典算法)中三个决策属性的适配准确率、精确率、召回率和F1值几乎相等,而算法3的同一链条上三个决策属性的上述指标值呈递增状态,这充分说明适配器链

可以有效传导决策属性间的相关关系,也说明本文方法在决策属性相关情形下的有效性和优越性。

2) 具有较强的泛化能力。本文方法可以根据适配流程的不同阶段差异化选择具体的算法,从而改进多案例诱导适配的效率和效果。针对算例数据集的对比实验显示,算例方法在 $D_3$ 的适配准确率、精确率、召回率和F1值上的表现最为优异;而基于KNN的适配器链在 $D_2$ 相关的指标值上表现最优; $D_1$ 相关的指标值显示朴素贝叶斯、朴素贝叶斯链和算例算法并无太大差异。因此,在具体的案例适配过程中,可以分别利用朴素贝叶斯、KNN链和算例算法为 $D_1$ 、 $D_2$ 、 $D_3$ 构建适配器,以期最优化适配结果。上述分析说明,本文方法不仅可以应用领域的特点选择多种技术算法,而且可以根据适配流程的不同阶段差异化地选择技术算法,从而使适配性能最优。

## 4 结语

在人工智能技术的支撑下,充分利用知识自身

蕴含的隐性规律提高知识服务的自主性和智能性是智慧知识服务的实现路径。针对智慧知识服务中的案例适配问题,本文提出一个考虑属性相关性的多案例诱导适配集成框架。该框架集合了不同决策属性相关性、不同决策属性个数及不同决策属性值域规模的案例适配策略及方法。实验表明,本文方法在复杂结构案例的诱导适配中具有良好的性能,且能结合多种技术算法从适配策略和流程上提升案例适配效果。本文仅提出了不同相关情形的算法原理并进行初步验证,尚未深入探讨各个算法的应用边界。在未来的研究中,应结合具体应用领域和场景深入探讨框架中不同算法的边界条件。

## 参考文献:

- [1] 李川,朱学芳,刘子溪. 基于5G+AI的知识服务智慧应用生态分析[J]. 情报科学,2024,42(1):2-9,50.  
LI C, ZHU X F, LIU Z X. Analysis of knowledge service smart application ecology based on 5G&AI[J]. Information science,2024,42(1):2-9,50.
- [2] 马海云,薛翔. 面向知识服务的领域知识结构研究[J]. 情报学报,2022,41(1):73-82.  
MA H Y, XUE X. Research on domain knowledge structure oriented to knowledge services[J]. Journal of the China society for scientific and technical information, 2022, 41(1): 73-82.
- [3] 马海云,曹思源,薛翔. 面向用户需求的健康领域知识组织与服务框架研究[J]. 情报资料工作,2022,43(2):84-92.  
MA H Y, CAO S Y, XUE X. Research on knowledge organization and service framework in health field oriented to user needs[J]. Information and documentation services, 2022, 43(2): 84-92.
- [4] 张翠娟,柯平,姚伟. 后知识服务时代的知识管理:从数字赋能到知识赋能[J]. 情报理论与实践,2020,43(9):17-23,6.  
ZHANG C J, KE P, YAO W. Knowledge management in the post-knowledge service era: from digital empowerment to knowledge empowerment[J]. Information studies: theory & application, 2020,43(9):17-23,6.
- [5] 顾天阳,赵旺,曹林. 跨组织医疗健康大数据聚合与案例知识推理方法研究[J]. 情报科学,2022,40(3):40-44.  
GU T Y, ZHAO W, CAO L. Cross-organizational medical and health big data fusion and knowledge recommendation method [J]. Information science, 2022, 40(3): 40-44.
- [6] 张建华,郭启迪,曹子傲,等. 基于匹配度和谱聚类的知识推荐研究[J]. 郑州大学学报(理学版),2023,55(4):1-7.  
ZHANG J H, GUO Q D, CAO Z A, et al. Knowledge recommendation based on matching degree and spectral clustering [J]. Journal of Zhengzhou university (natural science edition), 2023,55(4):1-7.
- [7] 顾东晓,李晓玥,谢懿,等. 基于体测健康大数据的灰色案例推理与个性化知识服务方法[J]. 情报科学,2021,39(12):3-8.  
GU D X, LI X Y, XIE Y, et al. Case-based reasoning and personalized service based on physical testing big data [J]. Information science, 2021, 39(12): 3-8.
- [8] 姚鑫,郭海湘,顾明赟,等. 基于案例推理的滑坡灾害应急相似案例智能生成研究[J]. 系统工程理论与实践,2021,41(6):1570-1584.  
YAO X, GUO H X, GU M Y, et al. Intelligent generation of similar case of landslide disaster emergency based on case-based reasoning[J]. Systems engineering-theory & practice, 2021, 41(6): 1570-1584.
- [9] DAS R, GODBOLE A, NAIK A, et al. Knowledge base question answering by case-based reasoning over sub-graphs[EB/OL]. (2022-02-22) [2023-11-30]. <https://arxiv.org/pdf/2202.10610v2>.
- [10] 郑晶,张恺. 考虑决策者后悔规避的瓦斯爆炸案例决策方法[J]. 浙江大学学报(理学版),2020,47(3):337-344.  
ZHENG J, ZHANG K. Case-based decision method for gas explosion considering decision maker's regret aversion [J]. Journal of Zhejiang university (science edition), 2020,47(3):337-344.
- [11] 张建华. 知识管理自学习案例多案例诱导型适配机制[J]. 系统管理学报,2014,23(6):845-853.  
ZHANG J H. Multi-case derivational adaptation mechanism of KML-case for knowledge management[J]. Journal of systems & management, 2014, 23(6): 845-853.
- [12] 张恺,杨隆浩,高建清,等. 基于累积置信规则库推理的台风灾害直接经济损失预测[J]. 灾害学,2024,39(1):64-68,74.  
ZHANG K, YANG L H, GAO J Q, et al. Typhoon disaster direct economic losses prediction based on cumulative belief rule-based system[J]. Journal of catastrophology, 2024, 39(1): 64-68, 74.
- [13] ZHANG J H, LI L C, BOAMAH F A, et al. A multi-case induction adaptation study of tacit knowledge based on NRS and CBR [J/OL]. Kybernetes, (2023-06-08). <http://www.emerald.com/insight/content/doi/10.1108/K-01-2023-0049/full/html>.
- [14] 张建华,叶建文,李方方. 隐性知识外显案例之多案例诱导型适配研究[J]. 情报理论与实践,2020,43(1):98-103.

- ZHANG J H, YE J W, LI F F. Research on multi-case derivational adaptation of tacit knowledge explicit cases [J]. *Information studies: theory & application*, 2020, 43(1):98-103.
- [15] 郑晶,王应明. 面向多部门多属性的群决策应急方案调整方法[J]. *运筹与管理*, 2020, 29(2):66-72.
- ZHENG J, WANG Y M. Group emergency alternative adaptation method oriented characteristic of multi-department and multi-attribute [J]. *Operations research and management science*, 2020, 29(2):66-72.
- [16] ZHAO J F, HUO Y L, XIONG J, et al. Case retrieval strategy of turning process based on grey relational analysis [J]. *Computer modeling in engineering & sciences*, 2024, 138(2):1663-1678.
- [17] MOUGOUIE B, BERGMANN R. Similarity assessment for generalized cases by optimization methods [M]// *Lecture Notes in Computer Science*. Berlin: Springer, 2002.
- [18] 谭睿璞,张文德,陈圣群,等. 异质信息环境下基于案例推理的应急决策方法 [J]. *控制与决策*, 2020, 35(8):1966-1976.
- TAN R P, ZHANG W D, CHEN S Q, et al. Emergency decision-making method based on case-based reasoning in heterogeneous information environment [J]. *Control and decision*, 2020, 35(8):1966-1976.
- [19] 李兴国,蒋梦鑫,杨雪洁,等. 基于案例推理的重大公共卫生事件协同决策方法 [J]. *情报科学*, 2023, 41(5):10-16, 25.
- LI X G, JIANG M X, YANG X J, et al. A collaborative decision-making method for grave public health emergencies based on case-based reasoning [J]. *Information science*, 2023, 41(5):10-16, 25.
- [20] WANG R, YE S H, LI K, et al. Bayesian network based label correlation analysis for multi-label classifier chain [J]. *Information sciences*, 2021, 554:256-275.
- [21] 丁家满,周蜀杰,李润鑫,等. 基于关联规则和拓扑序列的分类器链方法 [J]. *软件学报*, 2023, 34(9):4210-4224.
- DING J M, ZHOU S J, LI R X, et al. Classifier chains method based on association rules and topological sequences [J]. *Journal of software*, 2023, 34(9):4210-4224.
- [22] MISHRA N K, SINGH P K. Feature construction and smote-based imbalance handling for multi-label learning [J]. *Information sciences*, 2021, 563:342-357.

(上接第46页)

- [16] ZHENG C P, FAN X L, WANG C, et al. GMAN: a graph multi-attention network for traffic prediction [C]// *Proceedings of the AAAI conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2020: 1234-1241.
- [17] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: a new framework for spatial-temporal network data forecasting [C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2020: 914-921.
- [18] LI Y G, YU R, SHAHABI C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting [EB/OL]. (2018-02-22) [2023-05-23]. <http://arxiv.org/abs/1707.01926>.
- [19] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting [EB/OL]. (2017-09-17) [2023-05-23]. <http://arxiv.org/abs/1709.04875>.
- [20] GUO S N, LIN Y F, WAN H Y, et al. Learning dynamics and heterogeneity of spatial-temporal graph data for traffic forecasting [J]. *IEEE transactions on knowledge and data engineering*, 2022, 34(11):5415-5428.
- [21] CHOI J, CHOI H, HWANG J, et al. Graph neural controlled differential equations for traffic forecasting [C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2022: 6367-6374.