

基于模糊粗糙集的公共文化服务效能评估方法

郭帅,高全力,邵连合

(西安工程大学 计算机科学学院,陕西 西安 710048)

摘要:公共文化服务效能评估是衡量公共文化机构所提供服务的成效与供给能力的重要依据。实际评估中由于场馆基础设施建设不统一、评估方法多以主观性评估为主,难以对文化机构服务效能进行客观准确的评估。为解决上述问题,该文提出一种融合模糊粗糙集与多输入神经网络的公共文化服务评估模型。结合陕西地区公共文化机构发展水平,基于国家公共图书馆定级标准构建了效能评估指标体系。通过模糊粗糙集模型对指标集进行约简,在保证信息完备的情况下消减冗余和稀疏指标项。通过CRITIC法计算场馆评估指标的客观权重,将其与专家主观经验赋权同时输入网络模型,最终生成融合主客观数据的公共文化机构评估模型。实验结果表明,模型能较好克服数据集稀疏和指标集冗余的问题,且评估准确性较其他方法有所提升。

关键词:公共文化服务;效能评估;模糊粗糙集;CRITIC法;多输入神经网络

中图分类号:TP18

文献标识码:A

文章编号:1673-629X(2022)12-0129-07

doi:10.3969/j.issn.1673-629X.2022.12.020

Evaluation Method of Public Cultural Service Efficiency Based on Fuzzy Rough Set

GUO Shuai, GAO Quan-li, SHAO Lian-he

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

Abstract: The efficiency evaluation of public cultural services is an important basis to measure the effectiveness and supply capacity of services provided by public cultural institutions. In the actual evaluation, it is difficult to evaluate the service efficiency of cultural institutions objectively and accurately because the venue infrastructure is not unified and the evaluation methods are mostly subjective. In order to solve the above problems, we propose a public cultural service evaluation model integrating fuzzy rough set and multi input neural network. Combined with the development level of public cultural institutions in Shaanxi, we construct an efficiency evaluation index system based on the grading standard of national public libraries. The index set is reduced by fuzzy rough set model to reduce redundant and sparse index items under the condition of complete information. The objective weight of venue evaluation index is calculated by critical method, which is input into the network model together with the subjective experience weight of experts, and finally the evaluation model of public cultural institutions integrating subjective and objective data is generated. The experimental results show that the model can better overcome the problems of sparse data sets and redundant index sets, and the evaluation accuracy is improved compared with other methods.

Key words: public cultural services; efficiency evaluation; fuzzy rough set; CRITIC method; multi input neural network

0 引言

构建现代公共文化服务体系是近年来国家公共文化发展的重要战略内容,与民生、文化的发展息息相关^[1]。对公共文化机构的服务效能进行评估是提高各级场馆服务能力、改善场馆效益的有效手段。如何对

不同级别、不同领域的文化机构进行客观合理的评估对提高公共文化服务质量有着重要的意义^[2]。公共文化领域的服务效能评估研究已有十余年之久,评估多以场馆基础设施建设为主要内容,对评估场馆服务供给能力的研究较少。服务效能与社会环境、文化发展密切

收稿日期:2022-01-29

修回日期:2022-05-31

基金项目:国家自然科学基金青年项目(61902300);陕西省重点产业创新链(群)(2020ZDLGY07-05);2019年度山东省重点研发计划重大科技创新工程(厅市联合)项目(2019TSLH0209);陕西省教育统计数据研究中心专项研究课题(21JTY010);国家重点研发计划课题“公共文化资源服务效能评估及大数据智能分析平台构建研究”项目(2019YFC1521404)

作者简介:郭帅(1995-),男,硕士研究生,研究方向为公共文化效能评估;高全力,通讯作者,博士,副教授,CCF会员(A7163M),研究方向为大数据处理、智能推荐算法;邵连合,博士,副教授,研究方向为量子计算和量子信息

相关,受时间、地理、人群等多个维度的影响,波动较大,常规的绩效模型难以准确评估,并且缺少统一的公共文化服务效能的指标体系和评估标准^[3-4]。因此,构建统一的指标体系、简化评估过程、提高评估准确度是当前亟待解决的问题之一。构建面向公共文化机构的服务效能评估模型,对服务效益给出准确评估,对文化机构优化服务效益,保障服务质量有着重要意义。

目前针对公共文化服务效能的评估多从投入产出的角度进行分析,基于数据的相对权重对机构服务效益进行评价。具有代表性的评估方法包括,曹国凤等^[5]使用六何分析法对公共数字文化服务效能的评估动因、评估主体、评估客体、评估指标、评估过程及评估优化策略进行分析,验证了服务效能评估是考核公共数字文化服务成效的重要手段。寇垠等^[6]从多个方面对公共文化服务效能评价体系的研究进行归纳,以用户满意度为核心优化绩效评估体系。周静等^[7]以数据报道和馆情评析为核心,对公共文化服务效能的评估进行了研究。得出全国公共图书馆评估逐渐转变为以效能为导向的结论,为评估提供了研究方向。司光亚等^[8]针对效能评估体系中指标之间存在影响以及评估过程主观性较强的问题,提出基于仿真大数据的构建指标体系的方法,为效能评估指标构建提供了理论依据。朱美霖等^[9]从投入产出的视角构建了公共图书馆的服务效能评估指标体系,基于因子分析法对评估模型的关键指标进行提取,与第六次全国公共图书馆评估结果进行对比,验证了模型合理性。朱剑锋等^[10]基于数据包络分析(Data Envelopment Analysis, DEA)理论,构建了公共文化服务绩效评价框架,选用C²R模型和超效率模型对公共文化服务的特点进行分析,在相关数据集上对分项和综合绩效评价进行了研究。Wang Wenling等^[11]基于DEA超效率松弛度量模型分析了影响公共文化服务效能的主要因素,根据分析结果提出了政策方面的建议。

由于各地区经济、文化发展的多样性,公共文化机构发展倾向有所不同,目前常用的评估方法主要以统一的指标体系加以权重调整得到评估分数,缺乏对场馆自身发展特色的考量。而专家经验判断主观性较强,且难以保证评估时效性。因此,该文提出基于模糊粗糙集和多输入神经网络的评估模型,通过模糊粗糙集对不同地区的场馆评估指标进行约简,保证指标数据代表性的同时消除冗余信息,使用CRITIC赋权法计算机机构采集数据的客观权重,结合多位专家针对当地公共文化发展特点所赋予的指标权重输入神经网络,训练得到机构服务效能评估模型。模糊粗糙集的约简较好地解决了网络训练中收敛速度的问题,多输入网络能够保证有效融合专家经验和客观权重数据,

得到更加准确的评估结果。

1 基础理论

1.1 SFRS 模型

相似模糊粗糙集(Similar Fuzzy Rough Set, SFRS)是一种对模糊数据和不确定性对象进行分类的理论,利用对象属性数据之间的相似度进行分类,是一种用来表示等价关系的特殊形式化语言。一个知识表达系统可以被看作一个关系数据表。

设存在非空有限对象集合 $U = \{x_1, x_2, \dots, x_n\}$ 和非空有限属性集合 $C = \{c_1, c_2, \dots, c_m\}$,集合 U 和 C 构成一个二元决策空间,在该空间中,存在矩阵 V ,表示所有属性值, V_c 为列向量,表示属性 c 的所有取值,即属性 c 的值域,对象 i 在属性 j 下的属性值为 $V_{ij}(i = 1, 2, \dots, n; j = 1, 2, \dots, m)$ 。上述关系用信息函数表示为 $f: U \times C \rightarrow V$,即对于每个 $x \in U, c \in C$,都有 $f(x, c) = V_c$ 。

定义 1: $\forall x, y \in U, \forall c_j \in C, j = 1, 2, \dots, m$, 对象间的模糊关系 R 如下:

$$R_{xy} = \{(x, y) \in U \mid \frac{1}{n} \sum_{j=1}^m |V_{xj} - V_{yj}| > \alpha\} \quad (1)$$

其中, α 表示对象 x 和 y 之间的相似性,若大于 α 则表示对象相似。

定义 2: $\forall x_s, x_t \in U, \forall c_j \in C, j = 1, 2, \dots, m$, 在模糊关系 R 下, x_s 与 x_t 之间的隶属函数值为:

$$_R(x_s, x_t) = 1 - \frac{1}{n} \sum_{j=1}^m |V_{x_s j} - V_{x_t j}| \quad (2)$$

根据定义 1 和定义 2,可证明两个对象之间存在如下性质:

①自反性:

$$R_{xx} = 1$$

②对称性:

$$R_{xy} = R_{yx}$$

故 R 为模糊相似关系。

定义 3: 所有与 x_i 模糊相似的对象构成的集合称为 x_i 的模糊相似类,表示为 F :

$$F(x_i) = \{x_j \in U \mid \frac{1}{n} \sum_{j=1}^m |V_{x_i j} - V_{x_j j}| > \gamma, \\ j = 1, 2, \dots, m\} \quad (3)$$

其中, γ 表示模糊相似性,为 $[0, 1]$ 区间内的取值,越靠近 1 则对象间的相似性越强。

粗糙集模型的属性约简依赖于下近似关系,即模糊相似类的集合包含关系,经典粗糙集模型认为分类必须完全同时属性才可进行约简,导致在实际计算中分类结果极易受到异常数据影响。因此在模糊粗糙集的经典算法上进行推广,综合考虑了模糊决策的下

近似和上近似,将模糊粗略近似与自我信息的概念相结合,使模糊粗糙集算法适用于属性约简^[12-14]。同时结合变精度粗糙集理论,认为在分类比例区间内的近似关系即为正确分类^[15]。因此在此基础上建立集合 X 的上、下边界,表示为 β 和 α 。通常上下界的关系为 $0 < \beta < \alpha < 1$ 。显然, α 下边界是包含度相对比较大的等价类之并,而 β 上边界是包含度相对较小的等价类之并。尽管约定 $0 < \beta < \alpha < 1$,但事实上,为了达到多数包含的现实需要,往往取 α 为一个非常接近 1 且小于 1 的正数,而 β 则取一个非常接近 0 且大于 0 的正数。

定义 4: $\forall X \in U, \forall R \in C$, 则 $\forall X = \{x_1, x_2, \dots, x_n\} \subseteq U$ 下近似集和上近似集分别为:

$$\overline{R_U}(X) = U \{x \in U \mid \frac{|X \cap F(x)|}{|F(x)|} > \alpha\} = \{\overline{R_U}(X_1), \overline{R_U}(X_2), \dots, \overline{R_U}(X_n)\} \quad (4)$$

$$\underline{R_U}(X) = U \{x \in U \mid \frac{|X \cap F(x)|}{|F(x)|} \geq \beta\} = \{\underline{R_U}(X_1), \underline{R_U}(X_2), \dots, \underline{R_U}(X_n)\} \quad (5)$$

上近似表示最大相似范围,是知识库中集合的并集。下近似集表示最小相似范围,是知识库中集合的交集,通常用 $POS_U(X)$ 表示,属性约简主要是以下近似集作为依据。

定义 5: X 为属性集合 C 产生的模糊相似类集合, $X = U/C = \{X_1, X_2, \dots, X_n\}$, 设删除属性 c_i 后的指标集为 $R_i \subseteq U$, 以 R_i 为属性集计算得到的模糊相似类与 $F(x)$ 的差异称为近似分类质量, 定义为:

$$V_{R_i}(X) = \sum_{i=1}^n |POS_U(X_i)| / |U| \quad (6)$$

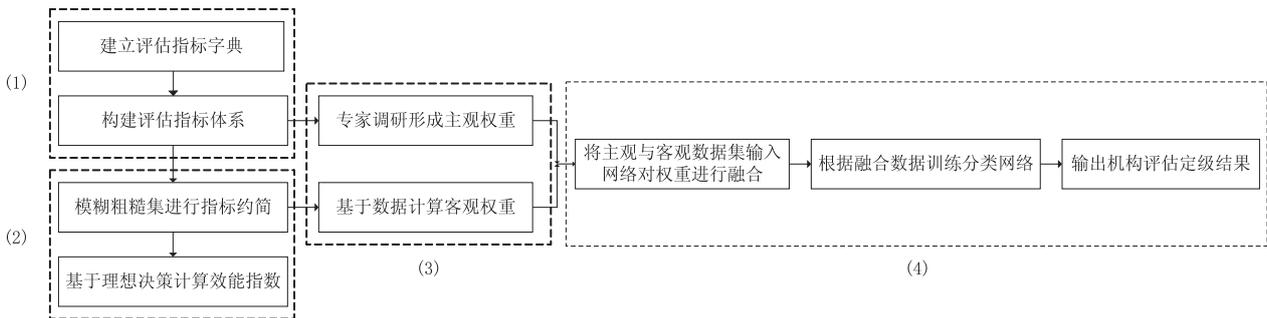


图 1 公共文化效能评估模型评估过程

(1) 构建模型评估的指标体系。指标体系直接影响评估的过程与结果,为确保指标的科学性与合理性,指标维度和指标项的确定借鉴了国际图书馆绩效评价标准、全国第六次县级以上公共图书馆评估定级标准等权威机构颁布的指标体系,与多家公共文化部门专家咨询和分析,确定效能评估的主要影响因素。

(2) 通过模糊粗糙集理论进行指标约简,提取主要影响因素,剔除稀疏指标项。指标集的约简增强了

近似分类质量可以表示指标 c_i 的重要度,若不等于 1 则表示该属性不可被约简。在实际计算中,近似分类质量越小,表示属性重要性越高,因此可根据近似分类质量值计算约简属性集权重。设约简后属性集 $R = \{c_1, c_2, \dots, c_n\}$, 则属性权重向量 W 表示为:

$$W_R = \left\{ \frac{V_{R_i}(x)_{\max} - V_{R_i}(x)}{V_{R_i}(x)_{\max} - V_{R_i}(x)_{\min}}, i = 1, 2, \dots, n \right\} \quad (7)$$

在对象集合中,所有属性值的最优集合成为理想决策 D , 理想决策的属性值 $d_j = V(j = 1, 2, \dots, n)$, 对象 x 与理想决策 D 的距离被称作贴适度,贴适度越小,对象方案越理想,其定义为:

$$\text{dist}(x_j, D) = \sqrt{\sum_{j=1}^n (V_{ij} - d_j)^2} \quad (8)$$

1.2 DNN 神经网络

神经网络是使用数学模型模拟人脑对数据的处理过程,分为训练和评估过程,能够根据数据自主学习,逼近任意的非线性函数。在数据挖掘、模式识别、信号处理等多个领域都得到广泛应用。DNN 神经网络也称作深度神经网络,与 BP 网络不同,每个节点和下一层所有节点都有运算关系,因此也叫全连接神经网络,由输入层、输出层和隐藏层构成,全连接神经网络通常有多个隐藏层,增加隐藏层可以更好分离数据的特征。

2 公共文化服务效能评估模型

为更好地评估公共文化机构的服务供给能力,该文提出一种公共文化服务效能评估模型,如图 1 所示,由以下四个部分构成。

数据的表征能力,简化模型网络结构,提高模型泛化能力。同时可根据评估对象与理想决策的贴适度计算出评估对象的排名和相对效能指数。

(3) 指标权重的确定受多种因素影响,在评估时主观性指标强调服务效益与用户主观感受,客观性指标体现建设与发展能力,主客观权重的结合能有效提高评估的准确性与合理性。主观权重由专家参照相关评估标准和地方文化发展特点确定,客观权重采用

CRITIC 法计算得到。

(4)使用多输入神经网络构建评估模型的数据融合和结果输出。神经网络能够以较高的精度拟合函数模型和提取数据特征,因此使用两个输入网络分别对使用主客观权重计算出的评估数据样本进行融合。然后通过分类网络对机构的服务供给能力进行定级。

2.1 文化服务效能评估指标框架构建

评估指标体系是评估模型研究的基础,经典评估方法中为了保证评估的合理性和适应性,通常会构建规模庞大的指标字典来保证评估结果的综合性。在实际评估中,由于场馆自身发展受多种因素共同作用,不同场馆的基础建设存在差异,指标数据的可获得性无法保证,导致最终的采集数据存在数据集稀疏、误差较大等问题。同时,主观性指标通常需要专家进行标度,难以适应强调实时性的数字化评估。因此,借鉴《全国第六次县级以上公共图书馆评估定级标准》、ISO11620 绩效指标等标准化指标体系,与公共文化领域多位专家共同探讨,以数据代表性和可获得性为前提,构建了适用于数字化评估的公共文化服务效能指标体系(见表 1),指标从资源建设、资源使用、效率、发展和影响力五个维度来评价场馆服务效能。

2.2 基于模糊粗糙集理论的指标约简

场馆的服务效能受社会因素和文化发展的影响,不同地区评估指标的倾向存在一定程度上的差异,且部分冗余的指标信息会对评估结果造成影响。因此需要根据不同地区场馆的采集数据对指标集进行约简,在保证指标评估精度的同时,确保指标体系为最小指标集。在该模型中,参与评估的场馆为对象集合 U ,场馆三级指标为属性集合 C , V_{ij} 表示场馆在某一个指标上的数据,使用全国第六次公共图书馆评估定级作为决策依据。

算法 1:基于 SFRS 模型的指标约简。

输入:场馆与指标集决策矩阵;

输出:约简指标集 T 。

1. 对数据进行 z-score 标准化处理
2. 计算场馆的模糊相似类 F
3. 初始化约简指标集 $T = \Phi$
4. 从原始指标集删除指标 $C - \{c_i\}$,得到指标集

R_i

5. 计算约简后的场馆模糊相似类 F_{R_i}
6. 计算近似分类质量 V_{R_i}

7. 若近似分类质量 $V_{R_i} >$ 下近似集精确度,则该指标不影响分类结果,可被约简,否则该指标不能被约简,放入约简指标集 T

8. 重复步骤 4 ~ 7,直到遍历完所有指标

表 1 公共文化服务效能评估指标

一级指标	二级指标	三级指标
x_1 资源建设	x_{11} 馆藏 资源建设	x_{111} 人均馆藏量
		x_{112} 古籍保护计划
	x_{12} 资源 可获得性	x_{121} 读者所需资源可获得性
		x_{122} 放置准确性
		x_{123} 成功的馆际互借比例
		x_{124} 政府信息公开查询服务
	x_{13} 基础服务建设	x_{131} 建筑面积
		x_{132} 书库面积
		x_{133} 公共服务设施保障
		x_{141} 阅览室终端数
		x_{142} 电子馆藏数量
	x_{14} 数字化网络 资源服务	x_{143} 电子资源库建设
		x_{144} 网络基础设施保障
		x_{151} 馆内员工数量
x_{15} 人员建设	x_{152} 专业技术人员	
	x_{153} 服务人员数量	
	x_{211} 馆藏流通人次	
x_2 资源使用	x_{21} 使用率	x_{212} 网站访问量
		x_{213} 外借文献比例
	x_{22} 举办活动	x_{214} 非注册用户比例
		x_{221} 活动参加人次
x_3 效率	x_{31} 馆藏成本	x_{222} 培训人次
		x_{311} 新增藏量购置费
	x_{32} 人员成本	x_{312} 馆藏使用成本
		x_{313} 新增数据资源购置费
		x_{321} 服务人员规模
		x_{322} 工资福利支出
x_{33} 服务成本	x_{331} 用户人均服务成本	
	x_{332} 图书馆平均访问成本	
x_4 发展	x_{41} 创新能力	x_{411} 服务创新推广
		x_{421} 基层培训组织
	x_{42} 人才培养	x_{422} 本员工培训时间占比
		x_{423} 参与科研项目情况
	x_{43} 持续支持	x_{431} 获得额外财政资源政策投入
x_{432} 政策投入		
x_5 影响力	x_{51} 文化功能	x_{511} 公共文化影响力
		x_{512} 弘扬社会文化活动建设
	x_{52} 宣传推广	x_{513} 传统文化的弘扬
		x_{521} 文创产品开发种类
x_{53} 社会教育	x_{531} 特殊人群基础设施建设	
	x_{532} 未成年人基础设施建设	

2.3 CRITIC 客观权重计算

CRITIC 客观权重赋权法是一种基于指标间的对比强度和指标冲突性来描述指标客观权重的计算方法。该方法在计算时考虑了指标间的数据关联和数据波动程度,从数据本身的客观属性计算权重,较常见的熵权法、标准离差法在实际应用中效果更好。数据的对比强度是指单一指标在不同对象之间的差异性,由于衡量该指标波动的程度,波动越大则相对权重会越高。指标冲突性是指两个指标之间的相关性,若两个指标之间相关性较低,表明冲突性较强,相对权重会更高,相关性较强,表明冲突性较弱,相对权重会更低。

2.4 多输入 DNN 神经网络模型

在公共文化效能评估计算过程中,需要综合考虑专家主观经验与指标客观权重,通常专家经验与客观数据对于场馆的评估重心并不完全一致,在训练时所提取的数据特征也存在差异,因此选择多输入神经网络作为权重融合模型,结构如图 2 所示。

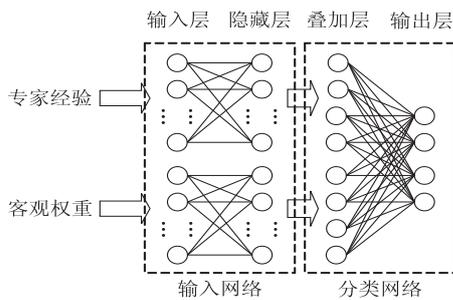


图 2 多输入神经网络结构

训练过程如下:

(1)神经网络输入维度取决于指标维度,训练指标由模糊粗糙集理论对原始指标体系约简得到,指标约简能够降低数据维度,同时保证网络模型的收敛速度和泛化能力。约简指标集的数量为两个网络的输入节点数量。

(2)网络输入由两部分组成,一个输入接收由专家经验评估的权重向量,另一个接收由 CRITIC 客观赋权法根据场馆数据计算得到的指标权重向量。

(3)将计算得到的数据结果分别输入神经网络,由于两部分数据在权重表示上倾向不同,客观权重更倾向于场馆基础建设和业务数据,专家经验倾向于文化传播和服务数据,因此两部分网络的隐藏层和节点数的选择并不相同,不同的网络结构能更好地提取各自部分的数据特征,保证分析模型的综合判别能力。

(4)两个输入网络的处理结果会在叠加层进行连接,将输出合并在一起由最后的分类网络对场馆的评估结果进行分类。

(5)分类网络的结果表明了场馆在经过主客观权重融合计算所评估的等级,模型评估结果与国家公共

图书馆定级标准进行对比,误差大于阈值时,会计算模型损失,调整网络参数继续训练直到输出满足要求。

(6)输出部分为 DNN 分类网络,训练标签为 4 种定级标准的 one-hot 数据,在训练中发现对于分类结果相对独立的评估模型中,输出层激活函数和损失函数选择 sigmoid 和 binary_crossentropy 的组合在分类准确度和模型损失结果上优于 softmax 和 categorical_crossentropy。

3 仿真实验

3.1 SFRS 模型参数优化

选择《2019 中国公共图书馆事业发展基础数据概览》中陕西地区 111 家公共图书馆的统计数据作为公共文化服务效能评估模型数据集,样本数据涵盖场馆基础建设、人员构成、特色文化发展等多个维度共计 86 项统计指标,满足模型评估的基本数据需求。由于统计数据单位不统一,且各地区基础建设和发展差异较大,为避免预处理后小数据特征被抹除,因此选择零均值标准化方法对数据进行标准化处理:

$$x' = \frac{x - \bar{x}}{\sigma}$$

评估模型中模糊粗糙集和分类网络都涉及数据降维和距离计算,该方法能较好地避免距离计算过程中极端值和异常值对数据整体的影响。模糊粗糙集模型计算时需要预置模糊相似度和分类精度两个参数,在不同维度数据中,参数的选择对计算结果有较大影响。为保证获得较好的约简效果,通过控制变量,逐一调整参数值获得最佳计算结果,参数变化对结果的影响如图 3 所示。

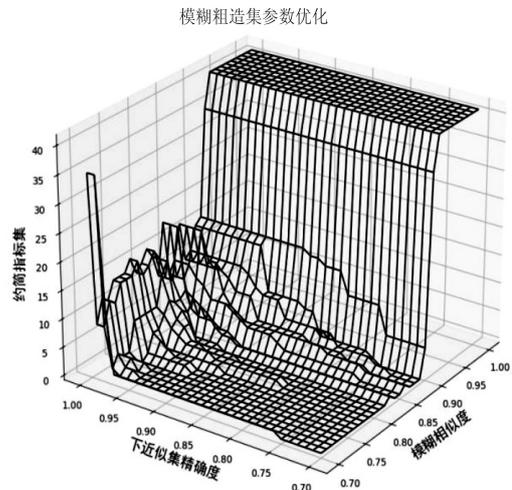


图 3 模糊粗糙集参数

从图中可以看出模糊相似度在 0.9,下近似精确度为(0.85, 1.0]时可以获得较为稳定的约简结果。使用该参数对原始指标集 42 个指标进行约简得到 16

项三级指标,可表示原始指标集 90% 以上精度。使用 CRITIC 法计算客观权重得到原始输入数据,如表 2 所示(保留三位有效数字)。

表 2 约简指标集

权重 分类	指标体系															
	x_1						x_2		x_3			x_4	x_5			
	x_{11}	x_{12}	x_{13}	x_{14}			x_{15}	x_{21}	x_{213}	x_{31}	x_{32}		x_{41}	x_{51}	x_{53}	
	x_{111}	x_{121}	x_{134}	x_{141}	x_{143}	x_{144}	x_{153}	x_{211}	x_{213}	x_{312}	x_{321}	x_{322}	x_{411}	x_{512}	x_{531}	x_{532}
主观	0.081	0.073	0.065	0.069	0.069	0.069	0.069	0.069	0.012	0.069	0.040	0.052	0.069	0.061	0.061	0.073
客观	0.012	0.063	0.122	0.108	0.008	0.125	0.094	0.090	0.063	0.021	0.024	0.049	0.019	0.098	0.036	0.065

根据约简结果可以看出,不同场馆的基础性建设,如实体和电子馆藏建设、场馆建筑面积等方面与政策投入存在较强关联,发展差别不大;但在个性化服务方面,如特殊群体基本设施建设、公共文化弘扬等方面则呈现差异,这类指标反映了场馆在面向公众服务方面所进行的投入,更能体现出公共文化机构在履行职能方面的能力。

3.2 多输入神经网络模型

约简后的指标集包含 16 个指标,则神经网络的两个输入层节点数均为 16,输入网络对数据进行降维,拼接层将两个输入网络的输出合并在一起,进行分类训练,输出结果为场馆的定级。公共文化定级分为 4 个等级,所以输出层节点数为 4。在训练中发现,部分等级场馆数据样本较少,训练中对该类场馆特征提取不足,难以准确评估。在数据中增加高斯白噪声是常见的数据增强方式,一定程度的噪声数据加入能够提高分类或预测模型的稳定性^[16]。因此,为解决样本数据不足的问题,为原始数据添加信噪比为 40 的高斯白噪声对数据进行增强,提高模型泛化能力,同时对原始数据集划分 70% 为训练数据集,20% 为校验数据集,10% 为测试数据集。在数据输入网络选择 Relu 函数为激活函数,分类网络选择 Sigmoid 函数为激活函

数。binary_crossentropy 作为损失函数,epoch 为 5 000 次。模型分类的训练准确度和校验准确度如图 4 所示,使用测试数据集对评估模型进行测试,实验结果使用 F1-score 进行评价,其中 Precision 指精准率,表示模型预测的某一类别中,预测正确的比例;Recall 指召回率,表示模型对某一类别的样本数据正确预测的比例;F1-score 是对精准率和召回率的加权调和平均,聚焦在较低的值上,能够更加准确地评价模型的整体性能;Accuracy 指模型精确度,表示模型在所有样本数据中预测结果正确的比例,与其他模型的对比测试结果如表 3 所示。

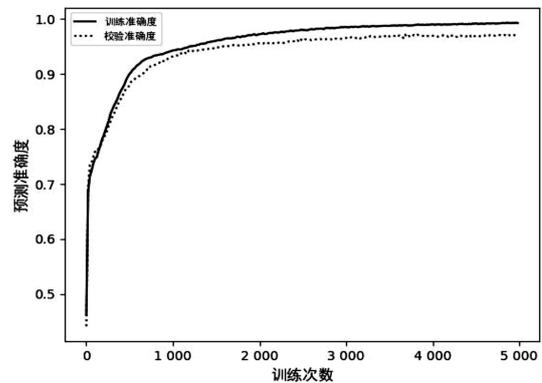


图 4 训练准确度趋势

表 3 模型评估定级准确度对比实验结果

	Model	Precision	Recall	F1-score	Accuracy
原始 指标集	CCR 模型	0.565 6	0.562 8	0.527 2	0.711 7
	BCC 模型	0.518 2	0.545 5	0.502 6	0.671 2
	DEA 超效率模型	0.659 9	0.496 3	0.560 7	0.707 2
	模糊粗糙集-BP 神经网络	0.753 6	0.768 9	0.761 1	0.782 6
	模糊粗糙集-多输入神经网络	0.867 7	0.686 2	0.731 1	0.842 7
约简 指标集	CCR 模型	0.674 6	0.682 4	0.634 4	0.747 7
	BCC 模型	0.537 4	0.567 2	0.523 0	0.765 8
	DEA 超效率模型	0.666 0	0.499 1	0.568 1	0.734 2
	模糊粗糙集-BP 神经网络	0.875 7	0.888 9	0.882 3	0.890 7
	模糊粗糙集-多输入神经网络	0.966 4	0.975 7	0.970 8	0.972 6

从实验结果中可以看出,经过模糊粗糙集约简的指标集能够显著提高评估准确性。此外,融合神经网络的评估模型能够有效结合主客观权重,使评估定级的结果更加准确。DEA模型使用约简指标集能够提高3%~7%的准确度,融合神经网络的评估模型能够提高10%~13%的准确度。多输入神经网络较BP神经网络准确度提升了8%。由对比实验可以得出,指标集约简和多输入网络结构在实验数据集中能够有效提高评估模型精度。

4 结束语

在公共文化服务效能评估中,常见的评估指标体系无法对发展方向各异的场馆服务效益进行综合评估,且指标规模大导致评估过程繁杂、数据采集困难等问题。该文借鉴全国公共图书馆评估定级标准和ISO11620国际图书馆绩效评价标准体系,与公共文化领域相关专业人员结合各级场馆实际发展水平,共同研讨总结了适用于数字化评估的指标体系,能较好地保证在不同等级场馆中的数据获得性问题。

针对实际评估中场馆采集数据稀疏、指标信息冗余的问题,通过模糊粗糙集理论基于原始数据集进行指标约简,消除数据样本不完全、信息冗余的指标项。由于模糊粗糙集理论存在不确定性参数,通过实验模拟得到适用于数据样本的参数值对指标集进行约简。同时基于指标项的最优集计算场馆的相对效能指数,为评估决策提供参考依据。

在服务效能评估中,专家经验得到的主观数据与采集场馆基础建设所得客观数据对场馆的服务效能有不同的反馈,为全面的评估场馆效益,提出采用多输入神经网络模型,使用两个输入网络分别处理主客观数据,然后合并网络输出进行场馆定级预测。该方法较传统评估方法结合了主观评判与客观计算,评估维度更高,可应用于公共文化相关效能评估系统中,提供面向场馆服务数据的实时评估,具有一定的应用前景。

由于可供研究的数据样本较少,目前仅使用陕西地区公共图书馆相关数据进行验证,应用于其他地区时模型中的部分参数适用性需要基于数据样本进行确定,在之后研究中会在其他地区的数据样本上继续优化,提高模型适用性。

参考文献:

- [1] 只莹莹. 新时代公共文化服务共建共享推进策略探析[J]. 图书馆, 2021(2): 8-14.
- [2] 吴高, 林芳, 韦楠华. 公共数字文化服务绩效评价现状、问题及对策分析[J]. 图书情报工作, 2019, 63(2): 60-67.
- [3] ZHAO Yuan, WAN Yi, CHUN Jiao. An unbalanced and inadequate development of the Chinese public libraries' public culture services: an investigation of 31 senior library specialists[J]. Libri, 2021, 71(3): 293-306.
- [4] 李艾芸. 公共文化服务效能评价指标体系构建与陕西实证研究[D]. 西安: 长安大学, 2020.
- [5] 曹国风. 公共数字文化服务效能评估探析——基于六何分析框架[J]. 河南图书馆学刊, 2021, 41(6): 79-81.
- [6] 寇垠, 黄凤. 我国公共文化服务绩效评价体系研究现状与展望[J]. 文化软实力研究, 2016, 1(4): 68-75.
- [7] 周静. 基于馆情媒体报道统计分析的公共图书馆服务效能评估[J]. 四川图书馆学报, 2018(2): 10-14.
- [8] 司光亚, 高翔, 刘洋, 等. 基于仿真大数据的效能评估指标体系构建方法[J]. 大数据, 2016, 2(4): 57-68.
- [9] 朱美霖. 投入产出视角下的公共图书馆文化服务效能评估模型研究[J]. 图书馆研究与工作, 2021(2): 14-21.
- [10] 朱剑锋. 基于DEA方法的公共文化服务绩效评价实证研究[D]. 武汉: 武汉大学, 2014.
- [11] WANG Wenling, CHEN Tong. Efficiency evaluation and influencing factor analysis of China's public cultural services based on a super-efficiency slacks-based measure model[J]. Sustainability, 2020, 12(8): 3146.
- [12] WANG C, HUANG Y, DING W, et al. Attribute reduction with fuzzy rough self-information measures[J]. Information Sciences, 2021, 549(12): 68-86.
- [13] BAI Juncheng. Neighborhood rough set-based multi-attribute prediction approach and its application of gout patients[J]. Applied Soft Computing Journal, 2022, 114: 108127.
- [14] 张漫, 李晶莹, 严胡勇, 等. 基于粗糙集和Petri网的油层含油识别研究[J]. 计算机技术与发展, 2020, 30(1): 174-178.
- [15] WANG Guoqiang. Double-local rough sets for efficient data mining[J]. Information Sciences, 2021, 571: 475-498.
- [16] 李晓燕, 李骏, 马尽文. 高斯过程混合模型在含噪输入预测策略下的煤矿瓦斯浓度柔性预测[J]. 信号处理, 2021, 37(11): 2031-2040.