

项目结题验收单

1 专家验收表（主持人所在单位组织 3-5 名专家对项目进行验收、自评。）

项目名称	高校图书馆用户学科咨询需求与服务匹配模型构建			
主持人	卢璐	职务/职称	助理馆员/初级	
所在单位	 (加盖公章) 南京林业大学图书馆			
专家意见	<p>很多高校加大“双一流”学科建设的力度，图书馆参与学科研究项目的机会也越来越多，学科馆员的服务重心逐渐向基于用户需求的知识服务倾斜。然而传统的图书馆用户需求与学科馆员是图书馆随机匹配，这种模式效率差，读者满意度低，缺乏个性化服务。为了提高知识服务质量，需要根据图书馆用户学科咨询需求推荐相对合适的学科馆员。因此图书馆个性化推荐服务，成为了图书馆服务领域的热点。</p> <p>本项目为提高图书馆馆员学科咨询服务的精准度和适用度，解决相关应用普遍存在覆盖度低、精度欠佳等问题，提出基于 ELM 算法的图书馆用户学科咨询需求与服务匹配模型。</p> <p>通过课题组成员深入地学习和研究，实现了预定研究目标。对后续相关研究提出三点建议：1、完善访谈内容，尽量降低采集的用户信息受主观引导的概率，提高信息数据的精准性；2、加强各指标间的关联分析，通过实证分析得出各项指标对结果影响的强弱关系，使研究更具说服力；3、进一步优化卷积神经网络和深度学习模型，加强拓扑网络数据采集整合试验，提高匹配精度。</p> <p>评审专家组一致认为，该课题已完成研究任务，达到预期的研究目标，同意结题。</p>			
专家签字	刘刚	胡兵	王斌	周磊
职务/职称	研究员	副研究员	教授	副馆长



项目编号：2020024
注：项目编号请查看立项
通知，也可缺省

CALIS 全国农学文献信息中心研究项目 结题报告

项目名称：高校图书馆用户学科咨询需求与服务匹配模型构建

项目关键词：学科咨询服务；用户咨询；卷积神经网络；推荐系统

项目单位(盖章)： 南京林业大学图书馆

通信地址：
江苏省南京市玄武区龙蟠路 159 号南京林业大学
(210037)

项目主持人：卢璐

联系电话：18260092931

电子邮件：2462866161@qq.com

提交日期：2021 年 4 月 25 日

题目:基于 ELM 算法的图书馆用户学科咨询需求与服务匹配模型研究

关键词: 学科咨询服务; 用户匹配; ELM 算法; 推荐系统

1 研究背景、目的及意义

1.1 研究背景

在高校“双一流”建设背景下,图书馆参与学科研究项目中的机会越来越多,学科馆员的服务重心逐渐向基于用户需求的知识服务倾斜。传统的图书馆用户需求与学科馆员是图书馆随机匹配,这种模式效率差,读者满意度低,缺乏个性化服务。为了提高知识服务质量,需要根据图书馆用户学科咨询需求推荐相对合适的学科馆员。因此图书馆个性化推荐服务,成为了图书馆服务领域的热点。

1.2 目的及意义

1.2.1 研究现状及目的

近年来,数据挖掘、机器学习、推荐系统等技术发展迅速,广泛的应用在医疗、商务、教育等诸多领域,在图书馆服务方面也出现诸多应用。如任萍萍¹基于用户特性画像进行学科服务的顶层优化设计,学科知识服务模型维度全面,但是缺乏基础数据佐证,模型体系有待完善;曹静仁²建立“用户学习需求—学科服务供给”回归模型,探讨高校的学科服务、学习习惯和学习空间三者关系,但是挖掘相关因素过于单一;林淑贞³等人使用 LRSM 模型完成读者画像信息与图书馆管理服务相匹配的模型,该模型取得了较好的满意度,但该模型需要读者画像数据,数据收集和挖掘十分困难,不适用于动态调整的学科馆员团队;[Aisyah Larasati](#)⁴利用线上和线下结合调查方法收集图书馆使用者的诉求,通过决策树构建影响顾客访问的因素模型,并预测使用者访问图书馆的频率;Dharna⁵等人通过对读者的借阅数据进行挖掘,根据挖掘出的频繁项集建立了一款图书推荐系统,取得了非常好的匹配度,

但该方法效率较低，规模大不使用普通图书馆的服务推荐系统。Chien⁶等人通过对读者长期行为模式的挖掘，构建了图书推荐系统，该系统需要长期对读者行为进行跟踪，缺乏时效性。

现有的图书馆服务推荐模型很多都是基于对大规模数据进行深度挖掘推广出个性化服务的模型。但实际生活中，很多图书馆规模较小，学科馆员数量少，特别是社区、普通高校等馆员工作分配结构单一，上述方法在小规模图书馆因数据集不够庞大而精度过低无法进行有效性推荐。范翠玲等学者⁷为了解决此种问题，提出建立阶梯式学科化服务团队满足不同需求的用户，但是目前仅限于理论探究没有构建可操作模型。

1.2.2 研究意义

本项目为提高图书馆馆员学科咨询服务的精准度和适用度，解决相关应用普遍存在覆盖度低、精度欠佳等问题，提出基于 ELM 算法的图书馆用户学科咨询需求与服务匹配模型。对前期样本数据进行分类、构建和测试，并将训练样本输入极限学习机(Extreme Learning Machine, ELM)网络，构建分类预测模型，通过交叉验证等方式，评价模型性能。实验结果表明，利用该方案为用户匹配相对合适的学科咨询馆员是科学可行的。

2 研究内容及方法

2.1 研究思路与方法

本文提出了以用户感知为导向，在合理精度范围内匹配相对合适学科馆员的设想，通过建立一种基于 ELM 算法⁸的读者需求和服务匹配的模型，首先通过搜集用户学术研究基本数据和对学科服务馆员满意度等相关数据，并对这些数据进行预处理，获取用户和馆员服务满意的数据。其次，对这些数据进行分类，构建训练样本数据和测试样本数据。然后，将训练样本输入极限学习机(Extreme Learning Machine, ELM)网络⁹，构建分类预测模型，结合测试样本数量调节相关参数。最后，通过交叉验证等方式，评价模型性能。实验结果表明，我们的方案是科学可行的。

前期工作：

- 1、检索、归纳和整理了“学科馆员”、“学术支撑”和“数字技术”等相关文献；
- 2、对本校图书馆主要负责学科支持的 6 位老师进行知识背景、性格画像和工作方式几个维度的初步分析；
- 3、对本校 50 名不同学科不同年级的本科生进行“关于实际学术课题研究中的学术支撑”的话题进行深度访谈并整理。

2.2 研究内容

2.2.1 数据和方法

(1) 数据来源

我们通过对某高校图书馆 214 名用户借书卡的 id 收集并筛选了用户最近半年借阅书籍情况和咨询馆员进行学科服务信息，并对样本用户进行深入访谈，获取了相关用户进行最近一次学术研究过程中具体措施，比如使用专业数据库、创客空间的辅助、团队合作意识以及高水平论文的关注度等，并结合该读者对学科馆员的满意度，构建了原始的数据¹⁰。相关数据如图 1 所示。

182	73	8.93	3.65	0.65	0.06	0	0	3
183	72.77	8.6	3.66	1.11	0.11	0	0	3
184	71.79	9.65	3.76	0.58	0.11	0	0	3
185	72.14	8.99	3.34	1.54	0.56	0	0	3
186	71.94	8.81	3.36	1.63	0.57	0	0.09	3
187	72.61	8.79	3.58	1.31	0.61	0	0	3
188	72.89	8.53	3.52	1.35	0.57	0	0	3
189	72.64	8.65	3.39	1.28	0.52	0	0	3
190	72.48	8.38	3.45	1.76	0.6	0	0.17	3
191	72.7	8.44	3.57	1.38	0.56	0	0.1	3
192	73.01	8.58	3.4	1.26	0.52	0	0	3
193	71.36	9.14	3.78	0.91	0.23	0	0.37	3
194	72.69	10.09	1.85	1.86	0.6	0	0	4
195	70.7	6.93	0	3.02	6.21	0	0	4
196	73.88	10.17	0	1.83	0.97	0	0	4
197	73.39	11.27	0.33	1.51	0.13	0	0.28	4
198	73.75	11.53	0	1.65	0.38	0	0	4
199	72.18	9.7	1.61	2.17	0.76	0.24	0.51	4
200	73.03	11.32	0	1.76	0.47	0	0	4
201	72.86	11.41	1.88	1.56	0.47	0	0	4
202	70.48	6.96	0	3.04	6.21	0	0	4
203	73.44	11.62	1.71	1.56	0.58	0	0	4
204	72.25	12.5	0	1.4	0.33	0	0	4
205	72.22	12.24	0	1.58	0.32	0	0	4

图 1 部分原始数据

上述数据共计 214 个，前 9 列依次为近半年借阅书籍数量、咨询馆员一对一学科服务时长、查阅课题所需基本文献时长、文献阅读时长、学习课题检索、定题和申报时长、为课题学习资源检索和相关软件利用时长、进行小组合作研究时长、学习数据挖掘分析时长、借助虚拟空间完成研究时长。最后一列为图书馆对应的学科馆员。其中对应学科馆员 6 人，分别用编号 0~5 表示，根据这些数据分析预测，建立服务推荐模型。

(2) 数据预处理

数据预处理是为了方便模型对数据的处理，包对馆员的编码，编号为 0 的馆员对应 1, 0, 0, 0, 0, 0, 0; 编号为 1 的馆员对应 0, 1, 0, 0, 0, 0, 0; 编号为 2 对应 0, 0, 1, 0, 0, 0, 0; 编号为 3 对应 0, 0, 0, 1, 0, 0, 0; 编号为 4 对应 0, 0, 0, 0, 1, 0, 0; 编号为 5 对应 0, 0, 0, 0, 0, 1, 0。下图 2 为编码后的馆员。

0	0	1	0	0	0
0	1	0	0	0	0
1	0	0	0	0	0
0	0	1	0	0	0
0	1	0	0	0	0
0	1	0	0	0	0
0	0	0	1	0	0
0	0	1	0	0	0
0	0	1	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
1	0	0	0	0	0
1	0	0	0	0	0
0	0	1	0	0	0
0	1	0	0	0	0

图 2 馆员编码

2.2.2 ELM 模型结构与应用流程

(1) ELM 模型结构

我们提出了一个基于 ELM 算法的图书馆用户学科咨询需求与服务匹配模型，包括对数据的预处理以及构建 ELM 网络并调节网络参数和比较不同分类模型，最后通过交叉验证验证模型性能。模型结构框图如图 3 所示。

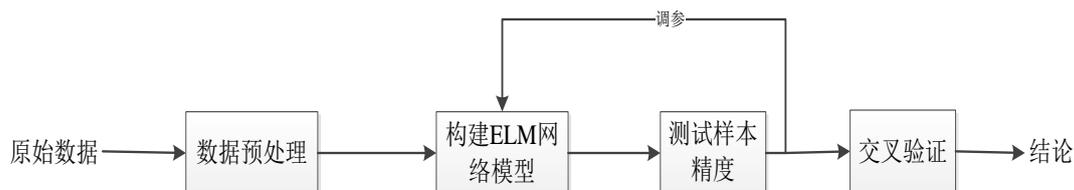


图 3 模型结构框图

(2) ELM 网络模型流程

极限学习机 (Extreme Learning Machine, ELM) 是一类基于前馈神经网络构建的机器学习模型, 与传统的 BP 神经网络模型不同, 该模型隐含层节点的权重为随机或人为给定的, 不需要更新, 因此 ELM 模型具有效率高, 使用方便等优点。

ELM 算法过程如下:

假设 N 为样本个数, 样本 $X_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$, L 为隐含层神经元个数, 则对于单隐含层神经网络可表示为:

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = o_j, j = 1, \dots, N$$

其中 $g(x)$ 为激活函数, 即 $X_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$, $W_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ 为输入权重, β_i 为输出权重, b_i 是第 i 个隐层单元的偏置。 $W_i \cdot X_j$ 表示 W_i 和 X_j 的内积。 ELM 网络隐含层学习

的目标是使得输出的误差最小, 可以表示为:

$$\sum_{j=1}^N \|o_j - t_j\| = 0$$

$$\text{即存在 } \beta_i, W_i \text{ 和 } X_j \text{ 使得 } \sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = t_j, j = 1, \dots, N$$

, 用 H 表示隐层节点的输出, β 表示输出权重, T 表示期望输出, 则其矩阵表达式为:

$$H\beta = T$$

其中 $H(W_1, W_2, \dots, W_L, b_1, b_2, \dots, b_L, X_1, X_2, \dots, X_L)$,

$$\beta = \begin{bmatrix} \beta_1^T \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \beta_L^T \end{bmatrix}_{L \times m} \quad T = \begin{bmatrix} T_1^T \\ \cdot \\ \cdot \\ \cdot \\ T_L^T \end{bmatrix}_{N \times m}$$

(3) 构建 ELM 网络

如图 4 为 ELM 算法拓扑图, 构建 ELM 网络模型步骤为:

1. 本模型输入层为 9 个特征数据, 因此输入层神经元为 9;
2. 确定隐含层神经元大小, 训练前随机产生连接权值 ω 和阈值;

3.计算输出层权值 β 。

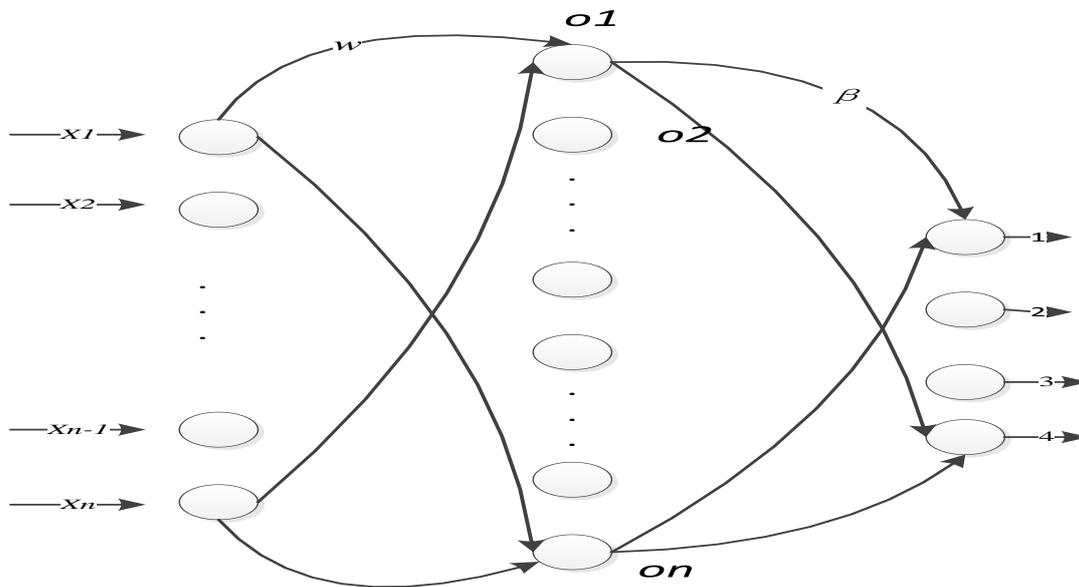


图 4 ELM 算法拓扑图

将所处理后的数据分成训练和测试两个部分，通过训练数据调整 β_i ， W_i 和 X_j ，使得输出的误差最小，并通过测试数据，计算精度，建立图书馆服务人员推荐模型。

2.2.3 实验结果验证和分析

(1) 实验验证

将 214 条数据前 164 条作为训练数据，后 50 条为测试数据。依次将 164 条训练数据输入 ELM 网络中，并通过测试数据评估模型的分类精度。图 5 为隐含层为 30 时，不同激活函数对应的分类精度。

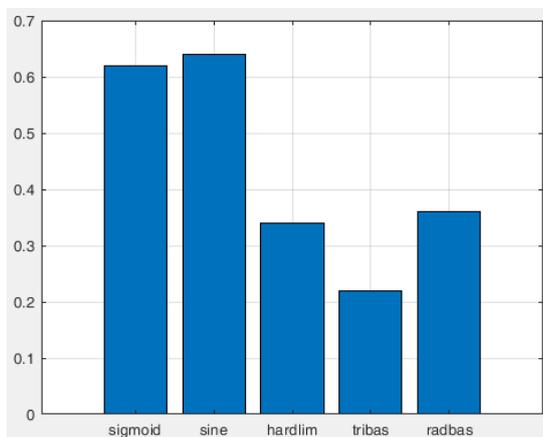


图 5 不同激活函数对应精度图

通过比较结果，我们选择 sine 激活函数时，分类精度较高，可以到 66%左右，也证明我们的方案是可行的。同时，我们对比不同隐含层神经元个数，通过交叉验证的方式确定了模型分类的进度。

采用 k 折交叉验证，将总样本 S 分成 k 份不同的子集，每份样本个数为 S/k，若 k=5,则每份样本个数为 92，假如把第一份作为测试集，则第二、三、四、五份为训练集，以此类推五份中的每份都会作为测试集被调用一次。最后的预测结果为：

$$\bar{x} = \frac{1}{k} \sum_{i=1}^k x_i$$

其中 x_i 代表 k 折交叉验证中第 1、2、3、4 次的结果， \bar{x} 为 k 折交叉验证的最终结果。

下表为采用采用 sine 激活函数，在 4 倍交叉验证情况下不同隐含层神经元个数对应的精度表。

隐含层神经元 个数	精度
10	56%
20	61%%
30	67.21%%
40	63.32%
50	62.13%

结果表明在隐含层神经元为 30 时，分类效果最好，预测的精度在 67.21%。

(2) 结果分析

由于数据调查具有一定的客观因素在里面，因此不需要过多在意实验精度，在 ELM 网络的输出层，我们根据每个读者的特征与对应 6 个输出的管理员进行概率排序，输出值越大，则说明对应的管理员越适合。图 6 为测试数据对应的管理员输出。

-0.2438	0.7140	0.4761	-1.1088	-0.1809	1.4085
0.9671	-0.0181	-0.1570	0.1014	0.0098	0.1022
-0.0186	1.5254	-0.2588	-0.5030	0.1773	0.0853
-0.0159	0.2856	0.5800	0.1232	0.0497	-0.0523
-0.0276	0.4137	0.6741	-0.0670	0.0028	-0.0176
0.0334	0.7029	0.4450	0.0786	-0.1475	0.0094
0.0080	1.0994	-0.0215	-0.0395	-0.0800	0.0728
-0.0505	-0.5035	0.4461	-0.3116	1.1554	0.1272
0.0173	0.8235	0.2762	-0.0915	-0.0397	0.0329
-0.2368	0.2438	1.8899	-1.0310	0.5531	-0.5145
-0.0487	0.7293	0.2700	0.0835	-0.0315	0.0267
0.3672	1.5612	-0.6132	0.0035	-0.3402	0.1520
-0.0085	0.4271	0.4063	0.1956	-0.0165	-0.0058
0.0124	0.4377	0.7924	-0.2326	0.0482	-0.0874
-0.0241	0.1333	0.7594	-0.0471	-0.0109	0.1261
-0.0493	0.3128	0.5199	0.1151	0.0076	0.0625
-0.0118	0.6658	0.4419	-0.1156	-0.0239	0.0299
-0.0528	0.6598	0.4258	-0.0584	-0.0040	0.0027
-0.0090	-0.1190	0.8778	0.1948	-0.0097	0.1254

图 6 输出管理员概率

上图中每一行对应一个用户输出的 6 个馆员概率, 值越大表示该名馆员越适合服务这名用户。

3 结论与建议

本文提出了一种基于 ELM 算法的图书馆用户学科咨询需求与服务匹配模型, 该模型克服了图书馆规模小, 图书馆馆员数量少, 数据量不充分导致的分析困难和无法精确匹配等问题, 具有效率高、使用方便等优点。实验结果表明, 我们的方案是可行的。

从目前的应用效果来看, 该模型还需在以下几方面进一步扩展: 首先, 增多调查样本数量, 并完善访谈内容, 增设隐性问题选择, 替代主体的显性表达, 尽量降低采集的用户信息受主观引导的概率, 提高信息数据的精准性; 其次, 优化图书馆学科服务多样性, 整合学科馆员个人特色和学科偏向, 为不同用户学术需求提供更为精准和深度的资源服务; 最后, 进一步优化 ELM 网

络算法，加强拓扑网络数据采集整合试验。有关图书馆馆员为用户进行学科服务的精准服务问题将在后续研究中深入讨论。

4 项目成果（发表的文章、开发的软件、取得的实践效果等）

[1]卢璐.学科馆员在大学生科研活动中的角色定位和作用——以南京林业大学本科生调研为例[J].江苏科技信息,2020,37(23):26-28.

本文通过深度访谈了南京林业大学本科生在科研活动中的动机、途径、优势和困难，探究了学科馆员的角色定位：传授基本技能的引导者，培养高阶思维能力的同理者以及优势资源的整合者。从而勾勒出新时代要求下学科馆员的素质轮廓,学科馆员需要集学科服务联络人、数据管理分析和科研项目协调人等多重角色于一身，为继续探究图书馆用户学科咨询需求与服务匹配模型打下基础。

5 参考文献

- [1] 任萍萍.“双一流”驱动下高校图书馆学科知识服务能力体系建设研究[J].情报科学,2019,37(12):93-97.
- [2] 曹静仁.高校图书馆学习共享空间与学科服务实证研究[J].图书馆工作与研究,2019(07):15-19.
- [3] 林淑贞.图书馆读者需求与服务匹配模型研究[J].图书馆研究,2018,48(06):106-111.
- [4] Aisyah Larasati,Muhammad Farhan,Puji Rahmawati,Nabila Azzahra,Apif Miftahul Hajji,Anik Nur Handayani. Designing Classification Models of Patron Visits to an Academic Library using Decision Tree[P]. Proceedings of the 2019 1st International Conference on Engineering and Management in Industrial System (ICOEMIS 2019),2019.
- [5] Patel D, Dangra J. A Book Recommendation System for Cloud Computing Framework[J]. International Journal of Computer Applications, 2017, 11(5):167-183.
- [6] Chien, Tzu-Chao|Chen, Zhi-Hong|Chan, Tak-Wai. Exploring Long-term Behavior Patterns in a Book Recommendation System for Reading[J]. Journal of Educational Technology & Society, 2017, 20(2):27-36.
- [7] 范翠玲,郑云武,叶文伟.高校图书馆阶梯式学科化服务团队隐性知识共享模式研究[J].图书馆理论与实践,2019(07):12-16.
- [8] 徐建新, 侯振华. 基于集成 ELM 模型的航空发动机性能参数预测[J]. 中国民航大学学报, 2017, 35(2):20-23.
- [9] Li Z P, Fan X H, Chen G, et al. Optimization of iron ore sintering process based on ELM model and multi-criteria evaluation[J]. Neural Computing & Applications, 2017, 28(8):2247-2253.
- [10] 向宏华.“互联网+”思维下图书馆服务创新研究[J]. 图书馆工作与研究, 2017, 1(4):5-10.
- [11] 柯平. 21 世纪的图书馆员[J]. 图书馆建设, 2004 (1): 5-9.
- [12] The Association of Research Libraries. Strategic Thinking and Design Initiative: Extended and Updated Report[R/OL]. [2018-12-18]. <https://www.arl.org/storage/documents/publications/arl-strategic-thinking-and-design-initiative-extended-and-updated-report-june2016.pdf>.
- [13] 郑建敏, 吴宝瑞. 刍议新建本科院校学科布局问题 [J]. 中国高教研究, 2006(4):50-51.
- [14] 刘丽娟, 袁曦临, 杨明芳. 锚定用户需求痛点的学科服务钩状模型研究 [J]. 图书情报工作, 2019, 63 (15): 60-67.
- [15] 黄如花,黄雨婷.全球视野下我国图书馆学教育变革之思考[J].图书情报知识,2019(05):4-14.
- [16] 洛林·W·安德森. 布鲁姆教育目标分类学: 分类学视野下的学与教及其测评(完整版, 修订版) [M]. 蒋小平, 张琴美, 罗晶晶, 译. 北京: 外语教学与研究出版社, 2009: 23—24.
- [17] 钟志贤. 论学习环境设计 [J]. 电化教育研究, 2005, (07): 35—41.
- [18] 陈树莲. 学生学科竞赛功利化的现状分析与对策 [J]. 教学与管理, 2007 (36): 53-54.
- [19] 吴建中. 从未来看现在: 图书馆发展的下一个十年[J]. 图书馆建设, 2016(1):4-9.
- [20] 黄如花,黄雨婷.全球视野下我国图书馆学教育变革之思考[J].图书情报知识,2019(05):4-14.
- [21] 李梅. 大学图书馆的研究数据服务创新与数据馆员新角色——英国爱丁堡大学范例研究[J]. 图书与情报, 2019(03):122-130.
- [22] 杨沉. 图书馆新型能力: 涵义、驱动因子、框架及其建构路径 [J]. 国家图书馆学刊, 2017 (5):59-67.