

基于深度自编码器和二次协同过滤的个性化试题推荐方法

熊慧君 宋一凡 张鹏 刘立波

(宁夏大学信息工程学院 银川 750021)

摘要 个性化试题推荐是实现高效学习的有效途径,帮助学生从“题海战术”中解脱出来,对实现适应性教学、促进教育公平具有重要意义。但目前个性化试题推荐方法大多是基于协同过滤进行试题层面的个性化推荐,没有聚焦到知识点层面,存在推荐试题定位不准确的问题。针对上述问题,对基于深度自编码器和二次协同过滤的个性化试题推荐方法进行了研究。首先考虑到学生对知识点的认知情况进行基于知识点的二次协同过滤试题推荐,然后应用项目反应理论和深度自编码器来预测学生在推荐试题上涉及推荐知识点的得分以及综合得分,最后对预测结果协同判断并控制最终个性化推荐试题的难度,产生最终的推荐试题列表。通过对比实验验证提出的推荐方法的推荐结果相对于传统试题推荐更具个性化和准确性。

关键词 个性化学习,试题推荐,协同过滤,深度学习,自编码器

中图法分类号 TP301 **文献标识码** A

Personalized Question Recommendation Based on Autoencoder and Two-step Collaborative Filtering

XIONG Hui-jun SONG Yi-fan ZHANG Peng LIU Li-bo

(School of Information Engineering, Ningxia University, Yinchuan 750021, China)

Abstract Personalized question recommendation is an effective way to improve learning efficiency. It helps students get rid of the “Massive Questions” and has important significance to achieve adaptive teaching and promote education equity. However, most of the personalized question recommendation methods are based on collaborative filtering without focusing on the knowledge points, which causes the problem that the positioning of the recommended questions are inaccurate. In order to solve this problem, a personalized question recommendation system based on deep autoencoder and a two-step collaborative filtering was adopted in this paper. Firstly, considering students’ master degree of knowledge points, the two-step collaborative filtering question recommendation based on knowledge points is realized. Secondly, item response theory and deep autoencoder are used to predict the scores and the comprehensive scores of the students involving recommended knowledge points on the recommended questions. Finally, the prediction results are synergistically decided, the difficulty of the final personalized recommendation questions is controlled, and a list of final recommended questions is generated. Comparison experiments verify that the recommended results of the proposed recommendation method are more personalized and accurate than that of traditional question recommendation methods.

Keywords Personalized learning, Personalized question recommendation, Collaborative filtering, Deep learning, Auto encoder

1 引言

大数据时代的到来,给社会各领域带来深远影响^[1]。教育大数据的本质是教育领域的大数据,其数据来源贯穿教育的全过程,不仅包括学习者的学习数据,更包括日常教育活动中的一切人类行为数据,具有多主题、多维度、多形态等特点^[2]。

美国教育部教育信息化办公室认为“自适应学习是根据学习者在学习过程中的反馈信息,动态改变学习内容以及内容的呈现方式、学习策略”^[3],不是给学生简单的评价和分数,

而是可以实现自适应地学习认知诊断、学习行为分析以及评估学生能力^[4]。

我国传统教育常常把学生放在被动接受教育的位置,学生在学习过程中缺乏学习主动性,更没有学习选择权,存在教学模式“程式化”、评价标准“单一化”等问题,并不是真正的“因材施教”^[5]。根据学生知识水平为其进行试题的个性化推荐,可以通过练习最少的题而最有效地提高学生的知识水平^[6]。

个性化试题推荐帮助学生走出信息爆炸的困境,为其提供度身定制的试题资源,实现高效学习。传统试题推荐是采

本文受自然科学基金(61862050)、2018年宁夏回族自治区重点研发项目(2018BBF02006)资助。

熊慧君(1994—),女,硕士生,主要研究方向为图形图像与智能信息处理技术;张鹏(1975—),男,博士,副教授,CCF会员,主要研究方向为智能信息处理,E-mail: pengzhang123@foxmail.com(通信作者)。

用推荐算法,寻找与学生兴趣相近的学生群体,再推荐相似学生群体中得分最高的试题资源。上述推荐算法的主要问题在于推荐试题并没有聚焦到知识点,而仅仅进行试题层面的推荐,该算法虽然一定程度上进行试题资源的筛选,但仍存在学习定位不准确的问题。学生对试题的练习实际是对知识点的练习,只有了解学生关于知识点的认知情况并由此产生推荐,才能做到个性化推荐以及高效推荐。本文在传统试题推荐算法的基础上,进行基于深度自编码器和二次协同过滤的个性化试题推荐方法研究,将深度学习引入试题推荐算法,进行基于知识点的二次协同过滤试题推荐,应用项目反应理论和深度自编码器预测学生在推荐试题上涉及推荐知识点的得分以及在推荐试题上的综合得分,以进一步控制推荐试题的难度进行个性化试题推荐。

2 相关工作

2.1 推荐算法

试题推荐根据学生的认知状况向其推荐帮助学生高效学习的练习题。目前主流的推荐算法有协同过滤推荐算法、基于内容的推荐算法、基于模型的推荐算法和与深度学习相结合的推荐方式^[7]。近年来,随着教育大数据和智慧教育的兴起,有学者尝试将推荐算法与教育领域相结合进行个性化学习的研究,在个性化资源推荐方面取得了一定成绩。

蒋一君等^[8]将学习者要做的练习分为3种,即做错的试题、未掌握的知识点以及已掌握的知识点,并根据不同的理论采取不同的推送机制。针对做错的题,在满足一定条件时则将该题移出错题库;对于未掌握的知识点,优先选择错题库中的题,若题量不够则根据知识点权重值来选题;对于已掌握的知识点,按照艾宾浩斯遗忘曲线规律,在适当时间加以巩固练习。这种机制分类细致、考虑全面,但没有关注到各知识点下的题仍有不同难度层次,而且仍然无法摆脱做错题次数过多而记住答案的问题。郭辰^[9]通过建立知识体系进行诊断性练习,并为学习者推荐与错题具有相同知识点和题型的题目以供练习。这种机制避免了重复做错题,但仍没有考虑到题目的难度层次。王文泉^[10]主张为学习者推荐与错题具有同等难度且同属一个知识点的高质量试题。这种推荐机制避免重复做错题,也考虑到题目的难度层次,但只分为3个难度等级且仍依赖错题库,若某些知识点没有练习到,那么推荐过程中也不会出现。

申瑞民等^[11]基于最小生成树思想构建合理的学习资源序列^[12]。牛文娟通过分析学生之间的交互信息,发现学习者的隐藏信息,了解学习者的喜好以及学习习惯,并为其推荐合适的学习资源^[12]。杨超^[13]基于项目反应理论构建学习者的学习能力,并结合粒子群思想进行学习资源的推荐^[13]。Andrew等^[14]基于协同过滤推荐算法进行个性化试题推荐。Dascalua等^[15]将学习资源的推荐应用到在线学习系统上,根据学习者的学习记录,为其推荐合适的学习路径。

将深度学习应用到推荐系统领域,利用深度学习模型获得用户和项目的隐表示,并基于隐表示对用户进行项目推荐是近年推荐算法研究的热点^[16]。基于深度学习的推荐系统不仅可以利用深度学习技术融合多元异构数据,自动学习获取用户和项目的隐藏特征,更能构建用户行为的序列模式和兴趣变化,并提高推荐的准确性。

2.2 项目反应理论

认知诊断是在传统考试的基础上,为学生提供关于知识点的诊断信息,即提供关于知识点/技能的掌握情况信息,在教育教学中具有重要的作用^[16]。它是现代教育测量方法与认知心理学的结合,通过考试实现对学生学习过程中存在的问题进行诊断与反馈,这些诊断信息有助于学生、家长、教师对学生学习情况有清楚全面的了解,有助于及时查缺补漏,并促进教师的因材施教,帮助实现素质教育。

随着信息技术不断发展,越来越多心理与教育学家投入对认知诊断理论的深度研究中。认知诊断模型一般分为离散型和连续型。项目反应理论(IRT)^[18]是一种经典的连续型认知诊断模型,它认为学生能否答对题目既取决于自身能力,又受限于题目区分度等因素。不同能力的学生对不同难度的题目反应会有所不同,通过分析学生关于题目的反应就可以估计学生的考试水平,而学生的考试水平与答题正确率存在相关关系。在此假设前提下,可以根据不同能力的学生可能答对题目的情况与实际答对情况构建相应的数学模型,从而精确估计被试能力。

3 个性化试题推荐方法

传统个性化试题推荐是在协同过滤算法的基础上,基于学生做题记录找到与其做题记录相近的学生群体,进一步根据相似学生群体的做题记录推荐适合该学生的试题。此种推荐思路虽然能实现试题推荐,但并不能做到真正个性化的试题推荐。这是因为学习的根本在于对知识点的学习,试题是无限的,而知识点是有限的,只有学习薄弱知识点才能做到有的放矢。一个优秀的试题推荐应该透过学生做题记录的表象,获知学生的认知情况,即对知识点的掌握情况,并综合考虑知识点的掌握情况、难度、重要性等因素推荐真正科学合理的试题。

考虑到上述情况,本文进行了基于深度自编码器和二次协同过滤的个性化试题推荐方法研究。

3.1 试题推荐方法整体框架

本文提出的基于深度自编码器和二次协同过滤的个性化试题推荐方法的框架如图1所示。

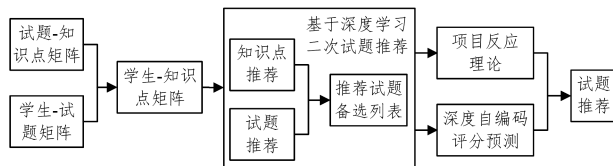


图1 个性化试题推荐系统框架图

从图1中可以看出,本文提出的个性化试题推荐方法的主要流程是:1)分析处理数据集,获取学生与试题的交互记录,构造试题与知识点的关联矩阵。2)进行基于知识点的二次协同过滤试题推荐,首先基于学生-知识点矩阵进行知识点的推荐,其次在推荐知识点的基础上进行包含推荐知识点的试题推荐。3)基于项目反应理论构造预测模型,判断学生在推荐试题上涉及该知识点的步骤是否正确。4)对推荐试题进行基于评分预测,判断学生在每道试题上的正确率。5)控制推荐试题的难度范围,确定最终的试题推荐列表。

3.2 试题推荐方法的具体实现

3.2.1 构造学生-知识点矩阵

对学习日志数据进行处理分析,构造学生-知识点矩阵,

同时进行可视化展示,方便学生细致全面地了解自己的学习现状,并为后续进行基于知识点的个性化试题推荐进行数据准备。

首先根据学生的学习日志,获得学生-试题交互记录,数据格式为“student_id::problem_id::rating::timestamp”,并用学生在试题上的正确率表示学生对于该试题的掌握情况;再根据试题往往包含多个知识点的特征“student_id::kc_id::rating::timestamp”,将学生进行的试题练习归类为学生关于知识点的练习,用学生在包含某个知识点上的全部试题的正确率表示学生对该知识点的掌握情况,将学生-试题交互记录进一步转化为学生-知识点交互记录“student_id::kc_id::rating::timestamp”;最后按知识点所在的章节进行划分,并对处理后的结果使用 echarts 进行展示,方便学生对学习的知识点进行梳理,对整体掌握情况有清楚的了解。

以某一学生为例进行个性化的能力评估,具体分析结果如图 2 所示。其中,叶子结点表示知识点,其大小表示知识点的重要性(重要性用知识点出现在题库中的次数表示),颜色表示学生对于该知识点的掌握情况(红色表示包含该知识点的试题正确率在 0.4 以下;绿色表示正确率在 0.4 以上;浅绿色表示尚未学习该知识点),与叶子结点相连的结点表示该知识点所在的单元。通过构建学生的知识图谱,可以直观清楚地了解学生对知识点的学习进度以及掌握情况。

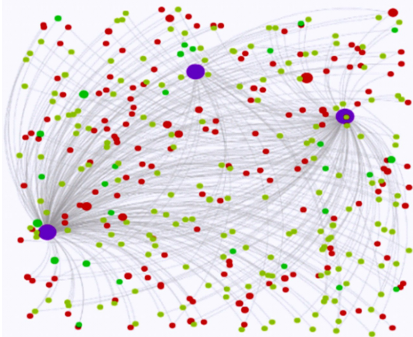


图 2 某一学生学习过程能力分析图(电子版为彩色)

3.2.2 基于知识点二次协同过滤的试题推荐

传统的试题推荐方法主要局限在试题层面,而学生学习过程本质是对知识点的学习,为了使推荐试题更具有针对性,本文采用深度学习的方法进行基于知识点的二次协同过滤试题推荐,该推荐算法的原理如图 3 所示。

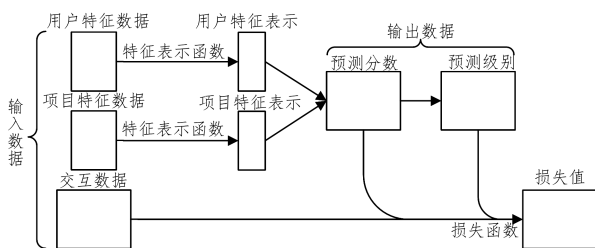


图 3 基于深度学习的二次协同过滤推荐算法原理图

首先进行基于学生的知识点推荐,输入数据为学生特征数据 User Features、知识点特征数据 Item Features 和学生-知识点交互数据 Interactions,其中学生特征数据 User Features 包括学生的 ID、名字,知识点特征数据 Item Features 包括知识点 ID、名字、所在章节,Interactions 表示学生在知识点上的错误率。推荐算法采用局部线性嵌入(LLE)方法对输入

数据 User Features 和 Item Features 进行处理,寻找学生 X 的 K 个近邻点 $\{X_1, X_2, \dots, X_k\}$,再根据均方差计算每个近邻点的损失函数,计算公式为:

$$J(w) = \sum_{i=1}^N |x_i - w_{ij} \sum_{j=1}^K x_j|^2 \quad (1)$$

其中, N 是学生样本总数, K 是近邻点总数。通过下式:

$$\sum_{j=1}^K w_{ij} = 1 \quad (2)$$

进行权重归一化处理,计算出该元数据的局部重建权值矩阵 $W_i = \{w_{i1}, w_{i2}, \dots, w_{ik}\}$,最后根据该学生特征元数据的局部重建权值矩阵和近邻点计算得出该学生的特征表示:

$$x' = w_{i1} X_1 + w_{i2} X_2 + \dots + w_{ik} X_k \quad (3)$$

知识点的特征表示采用同样的方法计算,只是输入数据变为知识点特征数据。通过上述方法建立两个低维度的向量 User Representation 和 Item Representation,再对有效信息用户积预测函数计算:

$$Prediction = user_repr * item_repr \quad (4)$$

其中, $user_repr$ 是学生特征表示值, $item_repr$ 是知识点特征表示值。求出预测值并根据预测值进行知识点推荐。获得推荐知识点之后,采用上述方法再次进行基于知识点的试题推荐,输入数据为用户特征数据与试题特征数据,再依次经过特征数据提取、预测函数计算来确定基于知识点的推荐试题列表。

在进行二次协同过滤的试题推荐时,首先基于学生-知识点矩阵进行知识点推荐,然后在推荐知识点上,基于学生-试题矩阵进行包含该知识点的试题推荐,获得试题的推荐列表。下一步将获取的推荐试题列表与学生对于推荐知识点的认知情况相结合,运用项目反应理论,判断学生在推荐试题上涉及推荐知识点的关键步骤上的表现。

3.2.3 项目反应理论模型的应用

项目反应理论是一种连续型认知诊断模型。它认为学生能否答对题目既取决于自身能力水平,又受限于题目区分度等因素,通过分析学生的反应就能估计学生的能力水平,预测学生的答题正确率。本文提出了一种基于知识点的试题推荐策略,一道试题往往包含多个知识点,试题作答正确需要掌握所有知识点,而试题回答错误只需有一个知识点错误。为了对目标知识点进行针对性练习,则需要保证学生在涉及该知识点的步骤有较为合适的正确率,使学生既达到练习的目的,又不因为试题过于简单而丧失兴趣。构造项目反应理论模型,根据学生能力水平、试题难度、知识点难度等因素预测学生在目标知识点试题的对应步骤的正确率。

在原始数据集中保留学生、试题、知识点、正确率部分的数据,构造认知诊断数据集,数据格式为“student_id::problem_id::kc_id::rating”(rating 值为 0/1,即学生 S 在试题 P 的步骤 K 能否做对,将正确率大于 α 的记录标记为 1,反之标记为 0)。使用 Adaboost 训练模型时,认知诊断数据集的训练数据中每个样本都被赋予一个权重,这些权重构成向量 D 。首先这些权重被初始化为相同的值,在训练数据上训练出一个弱分类器并计算该分类器的错误率,错误率的定义为:

$$\epsilon = \frac{incorrectsamples}{samples} \quad (5)$$

其中, $incorrectsamples$ 表示未正确分类的样本数, $samples$ 表示所有样本数。再根据错误率计算权重,使分类错误的样本

权重增加,分类正确的样本权重减少。其中,错误样本权重更改公式为:

$$D_i^{(t+1)} = \frac{D_i^{(t)} e^{\infty}}{\text{Sum}(D)} \quad (6)$$

正确样本更改公式为:

$$D_i^{(t+1)} = \frac{D_i^{(t)} e^{-\infty}}{\text{Sum}(D)} \quad (7)$$

其中, t 指当前的分类器, i 指数据集中第 i 个样本。然后用更新后的样本权重训练数据集获得新分类器,再重复上述步骤,直到分类器的错误率为0或者达到最大迭代次数。最后将所有弱分类器进行加权求和得到分类结果,分类器的权重计算公式为:

$$\infty = \frac{1}{2} \ln \left(\frac{1-\epsilon}{\epsilon} \right) \quad (8)$$

使用训练好的模型预测学生在推荐试题上涉及推荐知识点的关键步骤能否做对。对于目标学生 S 的推荐知识点 K 和包含推荐知识点的推荐试题 P 使用训练好的 Adaboost 模型预测。预测结果为1的试题表示学生在该推荐试题上涉及推荐知识点的步骤能够答对,可以起到练习的目的,并结合后续评分预测协同作为参与最后试题推荐决策的评判;反之则认为学生在该推荐试题上关于推荐知识点的步骤不能做对,则不能起到较好的推荐效果。

3.2.4 得分预测

深度学习可以对数据进行表征学习,在评分预测方面有较好的效果,其中深度自编码器通过构建一个恒等函数,可以使得模型的输出尽可能地接近模型输入值。当隐藏层单元个数小于输入的维数,可以让深度自编码网络学习输入数据的压缩表示,使得该表示可以以较小的误差重构回原函数。深度自编码器具有善于发现输入数据特征相关性的特点,因此本文采用深度自编码器评分预测的方法预测学生在试题上的正确率,并控制推荐试题的难度。深度自编码器在自编码器的基础上增加了隐藏层的数量,可以从原始数据中学习更为复杂高阶的特性。深度自编码器的网络结构如图4所示。

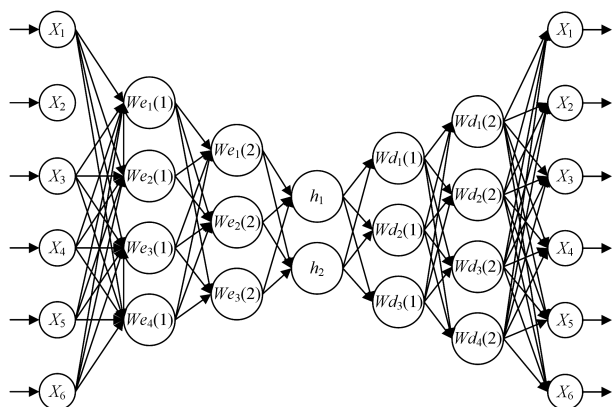


图4 深度自编码器的网络结构

输入数据是学生与试题交互记录的数据集,其格式为“student_id::problem_id::rating”,其中rating是学生在该试题上的正确率。评分预测将交互记录数据集处理成 U 个学生与 M 个试题的矩阵。每行 i 唯一表示一个学生,每列 j 唯一表示一个试题,rating为0表示学生没有练习过此道试题,深度自编码器的输入层和输出层的大小对应数据集中试题的数量。训练时,给定一个输入数据 X (学生-试题矩阵的一行),数据正向传递依次进行编码解码过程,将最后输出数据

同原始学生数据进行判断,根据结果调整各层网络权重,使得经过深度自编码处理后的数据尽可能表示输入学生数据,输出数据的结果也包括输入数据得分为0的值。将目标学生的做题得分记录作为一组输入数据,调用训练好的模型可以预测学生关于试题的得分。

3.2.5 确定试题推荐列表

将项目反应理论预测结果与评分预测结果协同进行判断,确定最终的试题推荐列表。

美国教育学家布鲁纳认为学生学习过程中存在3种内在动机:好奇心、胜任感、互助欲,正确调用3种内在学习动机,可以提高学生的学习效率。在推荐试题时要考虑“具有最适合的不确定性”的试题,模棱两可的情况最可能引起学生的好奇心,但也要考虑试题难度,要使学生相信成功的可能超过失败,才能保持学习的积极性。因此在确定最终推荐时,要综合考虑学生在推荐试题与推荐试题上涉及推荐知识点的表现。确定试题推荐列表的具体流程如图5所示。首先给定阈值 α, β_1, β_2 ,对于基于项目反应理论预测学生在试题上关键步骤的正确率大于 α 的试题,则认为学生能做对该试题,将该试题放入备选试题列表1,反之放弃推荐该试题。再对使用深度自编码预测学生在试题上得分在 $[\beta_1, \beta_2]$ 内的试题放入备选试题列表2,反之放弃推荐该试题。最后对备选试题列表1、备选试题列表2进行协同判断,取其交集进行推荐,确定最终推荐试题。在进行试题推荐时,可以通过设定阈值 α, β_1, β_2 对试题难度进行限制。

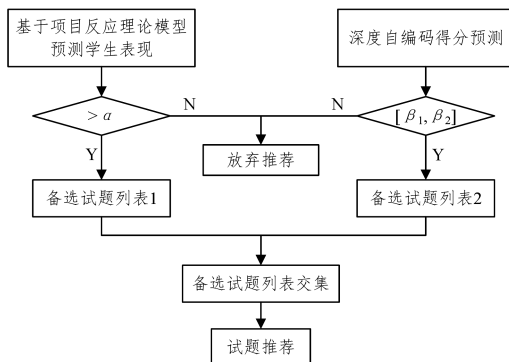


图5 确定试题推荐列表流程图

4 实验与结果分析

4.1 实验条件与数据集

本实验所采用的数据集是 KDD-CUP2010^[19]数据集。KDD-CUP2010的日志数据包括学生ID(Anon Student Id)、问题层次(Problem Hierarchy)、问题名字(Problem Name)、步骤名字(Step Name)、错误次数(Incorrects)、正确次数(Corrects)、已定义的知识点(KC Model Name)、遇到知识点的次数(Opportunity)等一系列做题记录信息。如何基于学生日志信息获取学生对于知识点的认知情况、知识点的难度、知识点的重要性、试题难度等相关信息是进行个性化试题推荐的重要内容。

基于上述个性化试题推荐方法,本文选用 Tensorrec^[19]进行基于知识点的二次试题推荐,使用 Adaboost 构建项目反应理论模型预测学生在推荐知识点的表现,利用深度自编码器进行评分预测学生在推荐试题上的表现,最终实现个性化试题推荐。

实验条件: E5-2620 v4 双 CPU, 64GB 内存, 500G SSD 硬盘, Quadro P5000, Ubuntu 18.04 操作系统。

4.2 评价指标

4.2.1 准确率、召回率、F1

一个优秀的推荐算法应该满足: 1) 为使用者推荐适合的资源; 2) 各种资源能会涉及到; 3) 根据使用者的反馈不断提高推荐的质量^[17]。本次实验采用离线实验的思想, 选择预测准确率、召回率、F1 作为评价指标。

准确率(Precision Rate)是预测正确的项目占实际推荐项目的比例。召回率(Recall Rate)是指推荐的资源和测试集资源交集占有推荐资源的比率。F1 是准确率和召回率的调和平均值, 用于综合反映推荐结果整体的指标。其具体定义形式为:

$$\begin{aligned} Precision &= \frac{TR}{TR+FR} \\ Recall &= \frac{TR}{TR+NR} \\ F1 &= \frac{2 \times Precision \times Recall}{Precision + Recall} \end{aligned} \quad (9)$$

其中, TR 表示推荐结果中推荐正确的项, FR 表示推荐结果中错误的项, NR 表示正确的结果但是没有进行推荐。

4.2.2 平均得分 AR

试题推荐与传统的电影推荐或者音乐推荐并不相同, 其本质是根据学生对知识点的实际掌握情况进行相应难度试题的推荐, 既不能过于简单也不能过于困难。所以仅使用传统基于 TOP-N 或者评分预测思想进行推荐的评估参数(准确率、召回率、F1)进行实验的评估结果并不十分全面, 因此本文提出了平均得分 AR 来作为试题推荐结果的一个评估参

数, 计算方式为 $AR = \frac{\sum_{i=1}^n R}{n}$, 其中 R 是每个试题的学生得分, n 是推荐试题的数量。AR 表示学生在推荐试题上的平均得分, 若果 AR 值偏大, 则表明推荐试题较为简单, 无法达到练习的目的; 如果 AR 值偏小, 则表明推荐试题难度过大, 学生正确率偏低。推荐结果的 AR 值应保持在适当范围内。

4.3 实验结果与分析

为了对实验结果进行评估, 本文引入了以下方法进行对比分析。

1) 基于项目的协同过滤方法。根据学生的学习记录, 找出跟学生表现较好的试题最相似的项目, 然后根据项目的相似分数从高到低进行推荐。

2) 基于 KNN 推荐方法。通过学生的做题记录, 计算学生之间的相似度, 并根据最相似的 K 个学生的得分情况预测目标学生在试题上的得分情况, 并进行推荐。

3) 基于矩阵分解的推荐方法。将学生与试题得分矩阵 $R_{m \times n}$, 分解为学生矩阵 $S_{m \times k}$ 和试题矩阵 $P_{n \times k}^T$, 其中学生矩阵 $S_{m \times k}$ 表示学生 S 的 K 维潜在因子, 即学生的内部特征; 试题矩阵 $P_{n \times k}$ 表示试题 P 的 K 维潜在因子, 即试题的内部特征。通过矩阵分解可以挖掘学生和试题的潜在因子, 预测学生对试题的得分情况, 并产生推荐。

实验采用 KDD-CUP2010 数据集, 其中包括 3310 个学生关于 188368 道试题包含 2128 个知识点的 8918043 条做题记录。首先筛选交互记录较少的学生日志, 再在实验中分别选取训练集比例 0.6, 0.7, 0.8, 计算 3 种方法的准确率、召回率、F1 值、平均得分 AR。划分后的数据集规模如表 1 所列。

表 1 实验数据集

训练集比例	0.6			0.7			0.8		
	Full	Train	Test	Full	Train	Test	Full	Train	Test
Users	2914	1873	1041	2914	2120	794	2914	2363	551
Items	15045	15029	12985	15045	15043	11806	15045	15045	9921
Interaction	418290	265416	152874	418290	302816	115474	418290	338272	80018

实验采用的方法是根据学生的做题记录进行基于知识点的二次协同过滤, 并对推荐试题列表采用项目反应理论思想和基于深度自编码的评分预测判断学生在对知识点的掌握情况下能否完成对试题的练习以及对试题中知识点的关键步骤, 最终确定试题推荐列表。

实验对比结果如图 6 所示。从图中可以看出, 在训练集所占的比例为 0.6, 0.7, 0.8 的实验中, 本文提出的基于深度自编码器和二次协同过滤的个性化试题推荐方法具有较为明显的优势, 在推荐准确率、召回率、F1 值上具有好的表现。如图 6(a) 所示, ItemCF 和 KNN 的准确率随着训练集的增大而缓慢降低, NMF 的准确率一直保持在较低水平且变化不大, 但是本文提出的推荐方法始终保持在较好水平, 且在训练集比例为 0.7 时准确率达到最好。4 种推荐方法的召回率均不断升高, 且升高趋势相同。如图 6(c) 所示, 本文提出的推荐方法、ItemCF、KNN 的 F1 值均随着训练集的增加而升高, 且 ItemCF、KNN 和 NMF 的变化趋势相同, 但 NMF 的 F1 值一直较小, 而本文推荐方法的 F1 值随着训练集比重的增大而出现明显的升高。

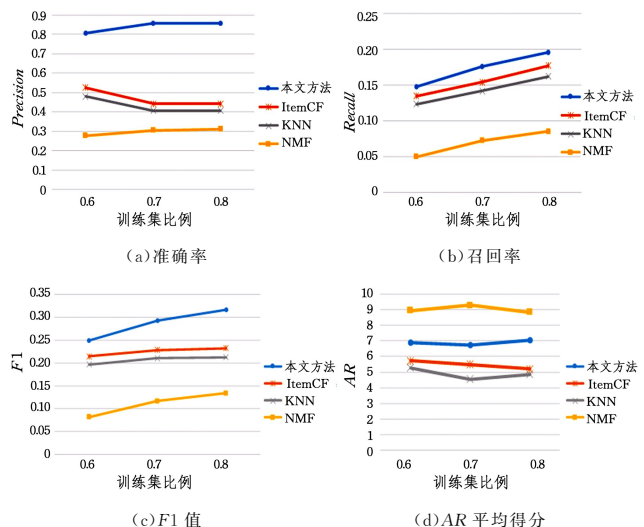


图 6 实验结果评估

平均得分 AR 的实验结果如表 2 所列, 对于传统推荐方法 ItemCF 和 KNN 推荐的试题, 随着数据集划分的不同及训

练数据的增大,学生的平均得分出现波动,但是始终稳定在 5 分左右,表明推荐的试题相对较难,对于学生的做题心理产生较大负担。当推荐试题难度较大,而学生在进行练习时,保持较低的准确率时,会严重打击学生学习的自信心,不利于学习的进步。对于 NMF 推荐的试题,学生得分一直处于较高水平,虽然可以提高学生信心,但是从教育角度看,推荐的试题较为容易,并不能达到学习知识点的目的。而使用本文提出的个性化试题推荐算法进行试题推荐时,学生的平均得分保持在 7 分左右,既能对学生起到激励作用,又能使学生在训练过程中学习和巩固所学知识。

表 2 平均得分 AR

比例	0.6	0.7	0.8
本文方法	6.9	6.7	7.1
ItemCF	5.7	5.5	5.2
KNN	5.3	4.6	4.9
NMF	8.9	9.2	8.8

综上所述,本文提出了一种新的基于深度自编码器和二次协同过滤的个性化试题推荐方法,其在进行试题推荐时有较好的准确性和稳定性,且符合教育理念,对提升学生能力具较好的表现。

结束语 本文在大数据背景下,对智慧教学中的适应性测试的方法进行了研究,不同于传统的 TOP-N 试题推荐,本文进行了基于知识点的二次协同过滤试题推荐,并将项目反应理论与深度自编码评分预测融入推荐算法中,进行了基于深度自编码器和二次协同过滤的个性化试题推荐方法的研究。与其他传统的单一的推荐方法相比,本文所提方法在推荐试题的难度控制上更为合理,且在各难度范围的试题推荐上取得了较好的推荐结果。

参考文献

- [1] 杨现民,唐斯斯,李冀红. 发展教育大数据:内涵、价值和挑战[J]. 现代远程教育研究,2016(1):50-61.
- [2] 杜娟敏,方海光,李维杨,等. 教育大数据研究综述[J]. 中国教育信息化,2016(19):1-3.
- [3] OXMAN S, WONG W, INNOVATIONS D V X. WhitePaper: Adaptive Learning Systems[EB/OL]. [2015-01-02]. <http://www.integratededsolution.com/wp-content/uploads/2014/10/DVx-Adaptive-Learning-White-Paper-February-20131.pdf>.
- [4] 徐立芳,莫宏伟,李金,等. 智能教育与教育智能化技术研究[J]. 教育现代化,2018,4(3):115-117,119.
- [5] 刘邦奇. 智慧教育:新时代的教育变革与转型[N]. 中国教育报,2018-01-27(3).
- [6] LORD F. Some test theory for tailored testing[C]// Computer assisted instruction, testing, and guidance. New York: Harper&Row,1970:139-183.
- [7] BATMAZ Z, YUREKLI A, BILGE A, et al. A review on deep learning for recommender systems: challenges and remedies[J]. Artificial Intelligence Review,2018:1-37.
- [8] 蒋一君,邱飞岳,刘迎春,等. 基于错题库的个性化练习生成模型研究[J]. 中国教育信息化,2011(8):73-74.
- [9] 郭辰. 高中英语诊断性练习系统的需求分析、模块设计和知识体系构建[D]. 合肥:中国科学技术大学,2013:22.
- [10] 王文泉. 错题管理系统中个性化推荐练习算法的设计与实现[J]. 中国教育信息化,2016(11):67-70.
- [11] 申瑞民,汤跃阳,韩鹏. 基于概念图的教学内容智能调整模型及算法实现[J]. 上海交通大学学报,2002(5):102-105.
- [12] 牛文娟. 基于协同过滤的学习资源个性化推荐研究[D]. 北京:北京理工大学,2014.
- [13] 杨超. 基于粒子群优化算法的学习资源推荐方法[J]. 计算机应用,2014,34(5):1350-1353.
- [14] ADACHE I, FOURNIER S, CHIFU A G. Harnessing Ratings and Aspect-Sentiment to Estimate Contradiction Intensity in Temporal-Related Reviews [J]. Procedia Computer Science, 2017,112(C):1711-1720.
- [15] DASCALU M I, BODEA C N, MOLDOVEANU A, et al. A recommender agent based on learning styles for better virtual collaborative learning experiences[J]. Computers in Human Behavior, 2015,45(45):243-253.
- [16] 黄立威,江碧涛,吕守业,等. 基于深度学习的推荐系统研究综述[J]. 计算机学报,2018,41(7):1619-1647.
- [17] 孙琳琳. 认知诊断评估对实现有效诊断教学的促进作用[J]. 世界最新医学信息文摘,2018,18(20):243.
- [18] RAMI S, BENNANI S, IDRISSE M K. Towards a method for analyzing learning style using item response theory[C]// International Conference on Information Technology Based Higher Education and Training. IEEE,2017:1-5.
- [19] SIGKDD:KDD Cup 2010:Student performance evaluation [OL]. <http://www.kdd.org/kdd-cup/view/kdd-cup-2010-student-performance-evaluation/Data>.
- [20] KIRK J. jfkirk/tensorrec[OL]. <https://github.com/jfkirk/tensorrec>.
- [1] 杨现民,唐斯斯,李冀红. 发展教育大数据:内涵、价值和挑战[J]. 现代远程教育研究,2016(1):50-61.
- [2] MASSA P, AVESANI P. Trust-Aware Collaborative Filtering for Recommender Systems[M]// On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Springer Berlin Heidelberg,2004:492-508.
- [3] 李聪,梁昌勇. 基于 n 序访问解析逻辑的协同过滤冷启动消除方法[J]. 系统工程理论与实践,2012(7):1537-1545.
- [4] 李小浩. 协同过滤推荐算法稀疏性与可扩展性问题研究[D]. 重庆:重庆大学,2015.
- [5] 孙小华. 协同过滤系统的稀疏性与冷启动问题研究[D]. 杭州:浙江大学,2005.
- [6] 徐键. 协同过滤中数据稀疏问题与推荐实时性的研究[D]. 兰州:兰州大学,2016.
- [7] TAO Z, JIE R, MATÚS M, et al. Bipartite network projection and personal recommendation [J]. Physical Review E, 2007, 76(4):70-80.
- [8] 周波,杨朝峰. 发送者和接受者能力的二分网络推荐算法研究[J]. 情报工程,2016,2(2):71-80.
- [9] 何平凡. 基于排序学习的 Top-N 推荐算法研究[D]. 北京:北京理工大学,2016.
- [10] 赵向宇. Top-N 协同过滤推荐技术研究[D]. 北京:北京理工大学,2014.
- [11] 陈嘉颖,于炯,杨兴耀,等. 基于复杂网络节点重要性的链路预测算法[J]. 计算机应用,2016(12):3251-3255,3268.
- [12] 荣莉莉,郭天柱,王建伟. 复杂网络节点中心性[J]. 上海理工大学学报,2008,30(3):227-230.
- [13] 姚尊强,尚可,许可. 加权网络的常用统计量[J]. 上海理工大学学报,2012,34(1):18-26.

(上接第 166 页)