

融合网络和文本特征的智能课程推荐方法

卢春华¹ 胡晓楠² 彭璐康³ 李翠霞¹

¹(安顺学院数理学院 贵州 安顺 561000)

²(朴茨茅斯大学会计与金融管理系 朴茨茅斯 PO1 2UP)

³(北京邮电大学网络技术研究院 北京 100876)

摘要 针对在线教育平台上资源过载和传统推荐方法输入特征单一的问题,提出基于网络拓扑和文本数据的特征融合方法,使用深度学习的表示学习方法来进行智能课程推荐。对于课程设计策略构建网络拓扑结构,并使用 Node2vec 向量化方法将网络中的节点映射到低维向量空间;之后使用卷积神经网络抽取特征并融合经 Doc2vec 表示的文本数据,拟合用户数据与目标课程之间的相似度。在进行推荐时,选择两者相似度最高的 Top N 课程。实验结果表明,该方法的性能远远超过传统的推荐方法,同时,消除实验表明该文使用特征对于最终的推荐效果都有一定的积极作用。

关键词 网络特征 文本特征 推荐系统 表示学习 智能课程推荐

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2023.02.012

AN INTELLIGENT COURSE RECOMMENDATION METHOD BASED ON NETWORK AND TEXT FEATURES

Lu Chunhua¹ Hu Xiaonan² Peng Lukang³ Li Cuixia¹

¹(School of Mathematics and Physics, Anshun University, Anshun 561000, Guizhou, China)

²(Department of Accounting and Financial Management, University of Portsmouth, Portsmouth PO1 2UP, UK)

³(Institute of Network Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract Aiming at the problem of resource overload on the online education platform and the single input feature of traditional recommendation methods, this paper proposes a feature fusion method based on network topology and text data and uses the deep learning related representation learning method to carry out intelligent course recommendation. We designed strategy and constructed the network topology for the course, and used Node2vec method to map the nodes in the network to the low-dimensional vector space. The convolutional neural network was used to extract features and fuse the text data represented by Doc2vec. We fitted the similarity between user data and the target course. When making recommendations, the method chose the Top N courses with the highest similarity. The experimental results show that the performance of the method far exceeds the traditional recommended method. And the elimination experiment shows that the feature used in this paper has a certain positive effect on the final recommendation effect.

Keywords Network feature Text feature Recommendation system Representation learning Intelligent course recommendation

0 引言

近年,随着科技的发展和“互联网+”概念的迅速

普及,基于互联网的在线学习应运而生,学习方式发生了翻天覆地的变化。在线学习由于其在资源利用、学习方法和教学形式等方面的优点,逐渐成为了一种新的趋势。然而,随着时间的推移以及用户的不断增加,

在线学习系统上的数据变得愈加庞大,资源过载问题日益凸显,给学生在课程选择方面带来了极大的困难。个性化推荐方法可以通过分析学生的历史行为数据等对其进行个性化建模,自动生成潜在兴趣课程并进行智能推荐,有效缓解课程资源过载问题,为用户个性化、精准化地提供学习资源建议,提升学习效率和用户满意度。

传统的个性化推荐算法主要分为三种:基于内容的推荐方法、协同过滤推荐方法和混合型推荐方法^[1]。这些推荐算法在某些特定的应用场景下能取得良好的推荐效果,但在某些方面还存在一些缺陷^[2]。目前,深度学习在图像识别、文本处理和语音识别等领域都取得了前所未有的突破,其方法也被不断地应用到课程推荐领域。Shen 等^[3]结合卷积神经网络和隐含因子模型,利用卷积神经网络从课程资源的文本描述中获取隐式特征,提升了模型的推荐效果。朱柳青^[4]基于神经网络设计了一种课程推荐模型,该模型使用卷积神经网络处理用户和课程的数据并建模,得到它们的隐式特征,提高了课程推荐的准确性。Khribi 等^[5]根据学生浏览记录对用户进行个性化建模,并结合学习资源等内容完成课程推荐。林木辉^[6]通过构建学习资源本体,提出一种基于学习者认知水平的个性化课程推荐模型,其中认知水平是通过知识点进行难度标注产生的。Aher 等^[7]使用半监督机器学习从用户历史数据中挖掘用户的主要兴趣,较为合理地为学生进行课程推荐。Batouche 等^[8]提出一种基于改进的人工神经网络的教学资源推荐模型,获得了较为合理的推荐结果。文孟飞等^[9]通过结合支持向量机与深度学习,将网络教学视频实时推送给学习者,提升了教学资源的利用率;Zhou 等^[10]提出了基于长短期记忆网络(Long Short-Term Memory, LSTM)的在线学习路径推荐策略。总体来说,以上模型在特征选择的时候都使用同种类型的结构化数据,而现实世界中的数据大多都是异构的、非结构化的,这大大制约了推荐模型的建模能力。基于这个角度出发,本文提出了一种基于网络拓扑和文本特征的智能课程推荐方法,旨在通过融合在线教育过程中产生的大量异构数据对在线课程进行全面的建模,进而匹配用户特征并进行智能化的推荐。具体步骤为:首先为了充分利用网络结构信息,分别以课程-学生、课程-教师和课程为主体构建网络拓扑,利用 node2vec 随机游走的方法生成节点序列;然后使用句向量将它们和相关课程的描述文本信息嵌入到更低维的表示向量中;其次通过多模块的卷积神经网络(Convolutional Neural Network, CNN)分别得到它们的

向量表示,通过向量首尾拼接的方法来建模用户的偏好;最后,通过计算用户查看课程信息与目标课程的相似度大小得出推荐课程序列。

1 相关理论

1.1 卷积神经网络

卷积神经网络(Convolutional Neural Network, CNN)是一个重要的深度学习模型。目前,卷积神经网络在自然语言处理中的应用已经取得了较好的结果^[11-12],框架如图 1 所示,其主要包含三层结构:卷积层、池化层和全连接层。

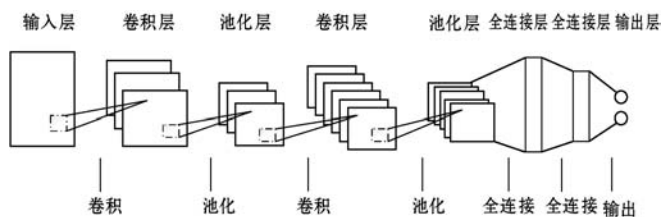


图 1 卷积神经网络结构图

卷积层:用来提取某个区域的特征,卷积层计算式如下:

$$x^j = \sigma(W^j \otimes X + b^j) \quad (1)$$

式中: x^j 表示在 j 层(卷积层)的特征图; W^j 是第 j 层的卷积核; X 表示第 j 层输入; \otimes 表示卷积操作; b^j 是第 j 层的偏置项; σ 表示激活函数。

池化层用来对不同位置的特征图进行压缩,可以减少特征图的大小,简化网络计算复杂度。池化层运算公式如下:

$$x^j = \sigma(\delta^j S(x^{j-1}) + b^j) \quad (2)$$

式中: $S(\circ)$ 表示降采样函数; δ 为乘子偏差; b 为附加偏差; σ 表示激活函数。一般情况下,池化层后接入全连接层,主要实现从特征到分类类别输出的转换。

1.2 Doc2vec

词的向量化是将自然语言中的词语映射到低维连续空间,生成一个实数向量^[13],Word2vec^[14-15]是一种使用广泛的无监督式词向量化模型,它从海量文本语料中学习富含语义信息的低维词向量,使得语义相似的单词在向量空间中距离相近^[16],Word2vec 包括两种训练模型,分别是连续词袋模型 CBOW 和 Skip-gram^[17]模型。类似 Word2vec、Doc2vec 用来进行语句或段落的向量化表示,网络结构如图 2 所示,Doc2vec 将句子和词的向量通过求平均或者首尾相连作为输入来预测文本中的下一个单词,对于给定的段落 d_i 和其词语序列 w_1, w_2, \dots, w_i ,该方法的目标是使平均对数概率最大

化,计算式如下:

$$\frac{1}{T} \sum_{t=k}^{T-k} \log p(w_t | w_{t-k}, \dots, w_{t+k}, d_i) \quad (3)$$

$$p(w_t | w_{t-k}, \dots, w_{t+k}, d_i) = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}} \quad (4)$$

$$y = b + Uh(w_{t-k}, \dots, w_{t+k}, d_i; W, D) \quad (5)$$

式中: U 和 b 是模型参数。 h 由 W 中提取的词向量与 D 中提取的段落向量联结或平均构成。

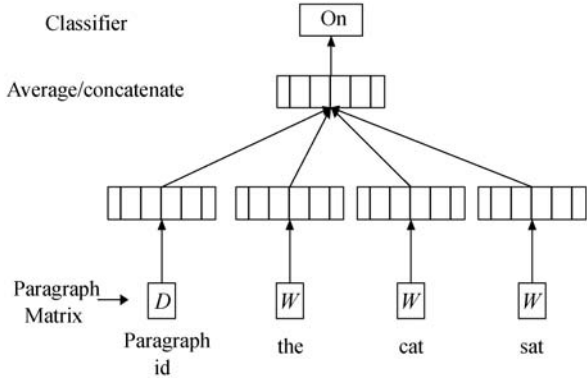


图2 Doc2vec

1.3 Node2vec

受到 Skip-gram 的启发,Vasile 等^[18]使用元数据嵌入,将项目元数据映射到潜在空间后通过计算元数据之间的相似性来得出推荐序列。后来,Perozzi 等^[19]发现自然语言处理中单词的出现频率和无标度网络中节点的出现频率均服从幂律分布^[2],并基于此提出了 Deepwalk,该方法将随机游走产生的节点序列作为文本的句子并将节点作为单词,使用 Word2vec 学习节点序列的低维表示。Grover 等^[20]在 Deepwalk 的基础上提出了 Node2vec 方法,该方法结合广度优先搜索算法(BFS)和深度优先搜索算法(DFS)这两种策略,通过调节随机游走的走向来挖掘网络拓扑。研究表明邻居节点具有相似性,而充当相同角色的节点具有结构一致性。在网络的表示中,节点间的这种性质是十分重要的。在生成序列时两个节点间的转移概率是通过如下公式产生的。

$$\pi_{vx} = \alpha_{pq}(t, x) \cdot \omega_{vx} \quad (6)$$

式中: ω_{vx} 表示节点 v 和节点 x 之间边的权重值, α 为转移概率,如图3所示,计算公式为:

$$\alpha_{pq}(t, x) = \begin{cases} \frac{1}{p} & d_{tx} = 0 \\ 1 & d_{tx} = 1 \\ \frac{1}{q} & d_{tx} = 2 \end{cases} \quad (7)$$

式中: v 表示当前的节点, t 是 v 的上一步所在节点,而 x 代表下一步的位置; d_{tx} 表示 t 和 x 之间的最短距离。

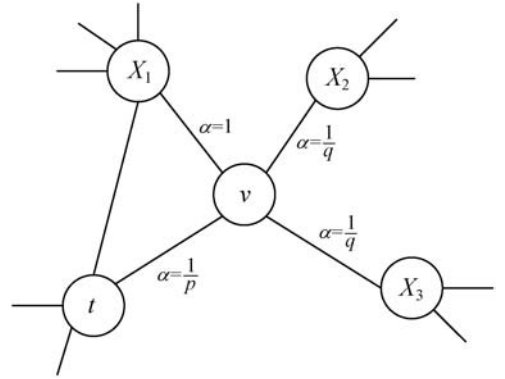


图3 Node2vec 的游走策略

2 基于课程网络和文本特征的 CNN 推荐模型

2.1 随机游走构造节点序列

在线课程网站上具有丰富的元数据信息(包括课程、老师和学生等实体以及属性信息),这些实体构成一个巨大的图结构,充分挖掘其中蕴涵的特征对于个性化推荐系统的性能提升具有重大意义。如图4所示,学生a和学生b同时学习了课程A,同时学生b也学习了课程B,而课程A和课程B属于同一个二级学科,这时课程B和学生a之间存在较强的相关性,可以给学生a推荐课程B。为了缓解数据稀疏性,本文分别以课程-学生、课程-老师和课程所属二级学科科目为依据构建课程图结构,从相关元数据的信息中得到课程的向量表示。

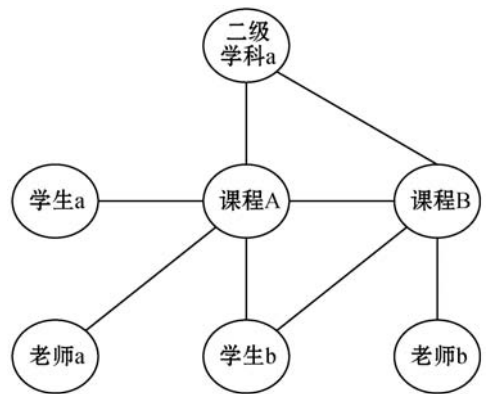


图4 课程图结构示意图

在课程-学生元数据图结构 $A_s = \{V_s, E_s, W_s\}$, V_s 表示课程节点集合, E_s 表示节点之间的边集合, W_s 是边的权重。它结合了学生社交关系(关注与被关注)、个人信息等,如具有社交关系的学生学习的课程之间、共同学生占课程总人数超过比重 P 的两个课程之间构成边。

在课程-老师元数据图结构 $A_T = \{V_T, E_T, W_T\}$ 中, V_T 表示老师节点集合, E_T 表示节点之间的边集合, W_T 是边的权重。它结合了老师之间的社交关系、个人信

息等,如具有社交关系的老师所教授课程之间、具有相同研究方向的老师教授课程之间、同一个老师所教授课程之间构成边。

在课程-课程元数据图结构 $A_C = \{V_C, E_C, W_C\}$ 中, V_C 表示课程节点集合, E_C 表示节点之间的边集合, W_C 是边的权重。课程元数据图结构主要根据二级学科进行统计的,例如计算机系统结构、计算机软件与理论、计算机应用技术属于同一个学科大类,属于它们的课程之间构成边。

关于边的权重值,以课程-课程元数据图结构为例,若其与其他课程共含有 n 种边的关系类型,课程 i 和课程 j 之间有边,若课程 i 和课程 j 之间共有 m 种关系,则课程 j 和课程 j 在图结构中边的权重值 $w_{(i,j)} \in W_S$ 为:

$$w_{(i,j)} = \begin{cases} \frac{m}{n} & e_{ij} \in E_s \\ 0 & \text{其他} \end{cases} \quad (8)$$

在课程推荐系统中,热门课程通常会得到很多学生的青睐,所以在随机游走过程中考虑到边的权重可以使节点往较为热门的方向游走。在课程-学生元数据图结构 $A_S = \{V_S, E_S, W_S\}$ 中,随机选定节点 S_1 ,从 S_1 出发选定长度为 L 的游走节点序列,如图 5 所示,节点从 S_{k-2} 游走到 S_{k-1} ,则下一步的游走概率为:

$$P(k | k-1) = \alpha \times \frac{w(k-1, k)}{\sum_{v \in N(k-1)} w(k-1, v)} \quad (9)$$

式中: α 为二阶随机游走参数, v 为与 S_{k-1} 节点有边的节点。

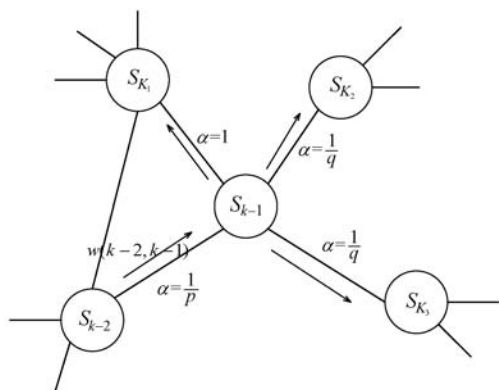


图 5 游走过程示意图

2.2 模型介绍

针对课程推荐中特征单一、结果不精准的问题,本文提出了基于边随机游走的多模块神经网络表示学习方法。该方法包括生成节点序列、使用向量嵌入学习生成图结构中节点以及课程描述文本的低维表示、训练基于异构特征的多模块 CNN 模型三个阶段。首先分别将课程-学生、课程-教师和课程-课程为主体构建

网络拓扑;其次利用 Node2vec 随机游走的方法各自生成节点序列并嵌入表示,使用 Doc2vec 将相关课程的描述文本信息嵌入到低维的表示向量中;最后通过多模块的 CNN 进行网络和文本特征的转换,并使用向量首尾拼接的方法得到的向量来用户的偏好表示 X ,计算 X 与前馈层处理的目标课程 Y (包含网络和文本特征)的相似度大小得出课程推荐序列。模型架构图如图 6 所示,课程相似度大小计算式为:

$$X^T Y = S \quad (10)$$

式中: X 表示用户的偏好向量; Y 表示目标课程向量; S 表示与目标课程(候选课程)相似度大小, S 值越大,目标课程在推荐序列中的排名越靠前。

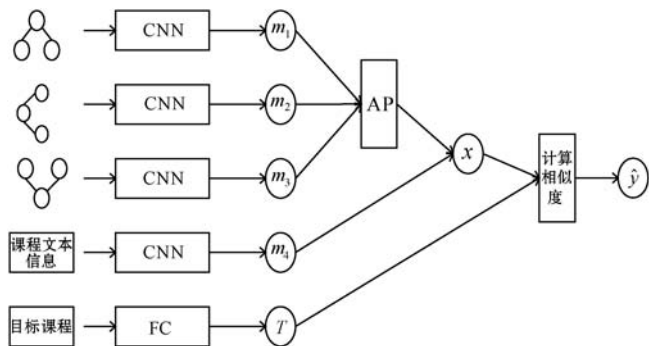


图 6 模型架构图

3 实验和结果分析

3.1 实验数据集

本文数据集来自于“爱课程”网。通过网络爬虫技术采集“爱课程_中国大学 MOOC(慕课)”上的课程,涉及的课程类别包括计算机、化学、外语、经济学、管理学、医药卫生、电子和教育教学这八个门类。采集的课程信息主要包括中文文本格式的课程名称、课程详情、课程页面链接、教师详情和学生详情等内容,实验总共收集 350 余门在线开放课程的相关元数据,共计 170 余万字。具体详情如表 1 所示。

表 1 数据集

编号	课程类别	课程数目	学生数目/千人
1	计算机	95	33.5
2	化学	28	12.7
3	外语	34	6.3
4	经济学	65	6.9
5	管理学	29	3.2
6	医药卫生	22	1.9
7	电子	64	11.7
8	教育教学	13	2.8

3.2 基准模型

本文设置的基准模型如下:

(1) 基于用户的最近邻推荐方法 (User-KNN): 将评分数据集和当前学生 ID 作为输入, 找出与当前学生的听课记录有相似偏好的其他学生, 对当前学生未学习的每个课程 A, 利用学生的近邻对课程 A 的评分进行预测; 最后, 选择评分最高的 Top N 个课程推荐给当前学生。

(2) 基于物品的最近邻推荐方法 (Item-KNN): 通过计算课程的相似度代替用户的相似度进行推荐, 但是并不以课程的内容属性计算相似度, 而是主要通过分析学生的行为记录。一般而言, 课程 A 和 B 具有很大相似度是因为喜欢物品 A 的用户大都喜欢 B。

(3) 非负矩阵分解 (NMF)^[21]: 将用户—课程矩阵分解为用户和课程矩阵相乘的形式, 矩阵中每行代表一个个体, 最终匹配得分通过这些向量表示。

(4) FGMSI^[22]: 基于社会影响力因素的图形推荐模型。该模型首先构建一个异构的社交网络, 然后计算基于元路径的相似性。最后, 从网络中提取一些特定功能, 并将其用于推荐。

3.3 评价指标

在本文实验中, 选取准确率 (Precision) 和折扣累计利润及其归一化 (Normalized Discounted Cumulative Gain, NDCG) 作为推荐结果的评估指标^[14]。

Precision 表示推荐方法推荐的课程被用户喜欢的比例, 其计算式如下:

$$precision = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|} \quad (11)$$

式中: $R(u)$ 表示推荐模型认为用户喜欢的课程; $T(u)$ 表示用户真正喜欢的课程。

NDCG 指标通过比较推荐课程结果列表中位置与课程指定所在位置, 评估出推荐结果有效性, 其计算式如下:

$$DCG = \sum_{i=1}^n \frac{2^{r(i)-1}}{\log_2^{(1+i)}} \quad (12)$$

$$NDCG = \frac{DCG}{iDCG} \quad (13)$$

式中: $r(i)$ 是第 i 个课程的结果得分; $iDCG$ 是查询的理想值, 即查询结果好的状态下计算出来的 DCG 值。

3.4 模型超参数设置

对于神经网络模型, 其超参数的初始化设置对模型性能是极其重要的。本文对所构建网络拓扑利用 Node2vec 随机游走的方法进行嵌入表示, 对将相关课程

的描述文本信息使用 Doc2vec 的方法进行嵌入表示, 其中 Node2vec 和 Doc2vec 参数设置如表 2 和表 3 所示。

表 2 Node2vec 参数设置

名称	含义	大小
p	控制随机游走采样的节点更靠近	1
q	或远离起始节点	1
Size	特征向量维度	200
Epoch_number	训练迭代次数	100
window	当前词与预测词的最大距离	3
min_count	字典截断	5
sample	随机采样配置阈值	1e-5
workers	训练的并行数	32

表 3 Doc2vec 参数设置

名称	含义	大小
Size	特征向量维度	200
Epoch_number	训练迭代次数	100
window	当前词与预测词的最大距离	3
min_count	字典截断	5
sample	随机采样配置阈值	1e-5
workers	训练的并行数	32

最后通过多模块的 CNN 进行网络和文本特征的转换, 为了训练得到性能最好的模型, Epoch 大小设定为 250, 且在每 20 个 Epoch 后保存一次模型, 最后选择在测试集上效果最好的模型作为本文最后的推荐模型。表 4 给出了本本提出模型的参数设置情况。

表 4 模型超参数设置

名称	含义	大小
Embedding_dim	嵌入维度	200
Batch_size	批处理大小	128
Filters	卷积核大小	3 × 200
Kernel_size	卷积核个数	32
Dropout	Dropout 比例	50%
Learning rate	初始学习率	0.001
Epoch_number	迭代次数	250

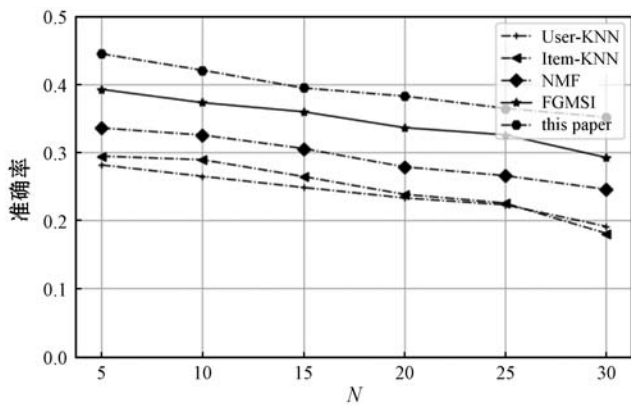
3.5 模型对比结果分析

Top N 推荐中, N 表示在推荐列表中的前 N 个推荐课程。实验结果如表 5 和图 7 所示, 本文推荐模型较其他模型表现出了最好的推荐效果。当 $N = 10$ 时, 在推荐指标准确率 Precision 指标上, 本文模型比基于用户的最近邻推荐方法有约 16.4 个百分点的性能提升, 比基于物品的最近邻推荐方法有约 10.2 个百分点的提

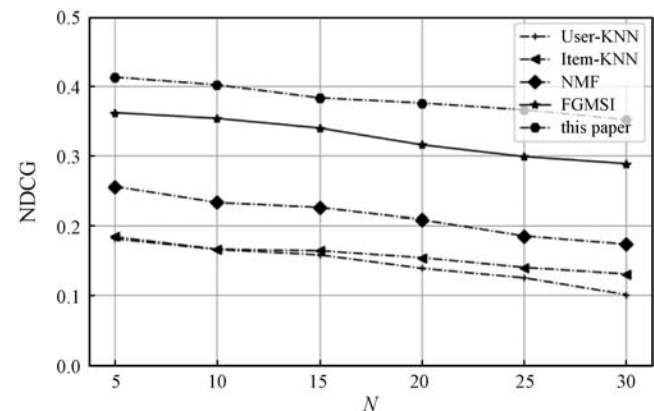
升,比 NMF 推荐方法有约 9.5 个百分点的提升,比基于神经网络模型的 FGMSI 推荐方法有约 5.2 个百分点的提升;在更注重推荐列表排序的指标 NDCG 方面,本文模型比基于用户的最近邻推荐方法有约 23.6 个百分点的提升,比基于物品的最近邻推荐方法有约 23.5 个百分点的提升,比 NMF 推荐方法有约 16.9 个百分点的提升,比基于神经网络模型的 FGMSI 推荐方法有约 4.86 个百分点的提升。可以得出,本文以课程、学生和教师为主体构建网络拓扑,利用 Node2vec 随机游走的方法获取了课程与课程、学生与课程、老师与课程之间的关系,并通过多模块的神经网络结构将它们有机地融合在一起,有效地提升了模型的推荐性能。

表 5 N 为 10 时推荐模型的性能

方法	Precision@ 10	NDCG@ 10
User-KNN	0.264 5	0.165 9
Item-KNN	0.288 9	0.166 2
NMF	0.325 1	0.232 5
FGMSI	0.372 5	0.353 1
本文模型	0.420 2	0.401 7



(a) 准确率



(b) NDCG

图 7 各模型在数据集上的性能比较

3.6 推荐实例

为了能直观地说明本文模型的推荐机制,本文选

取了数据集中的一条推荐实例加以说明。如图 8 所示,对于特定学生 A,推荐目标是根据目前学习记录(计算机组成原理、数据结构、人工智能原理等),及其相关的图结特征和文本描述特征,将目标课程“操作系统”进行精准推荐。在本实例中,由于“操作系统”与已学习的三门课程具有类似的网络结构,同时,它们的描述文本也具有相似之处(如:计算机、基础课程、大学生必修等),因而本文提出的模型会将目标课程排在相对靠前的位置(第二名),它与排名第一的课程“计算机网络”具有相似的特征。这说明本文提出的模型能够进行精准推荐,同时对于推荐结果是可解释的,显示出其在大规模在线教育平台应用中存在极大的潜力。

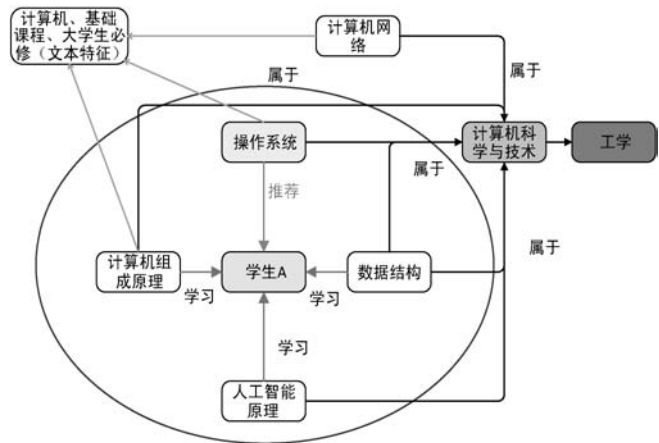
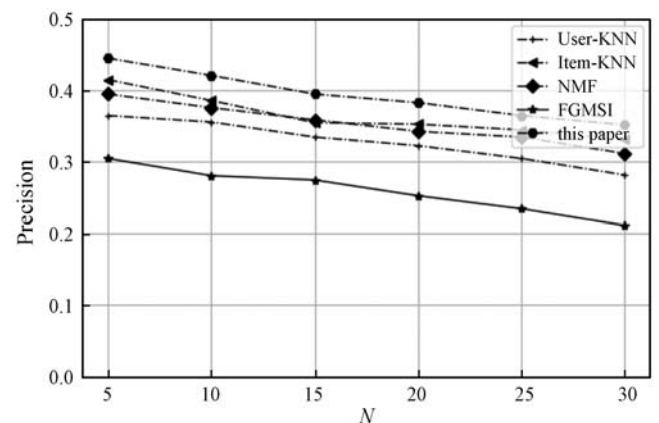


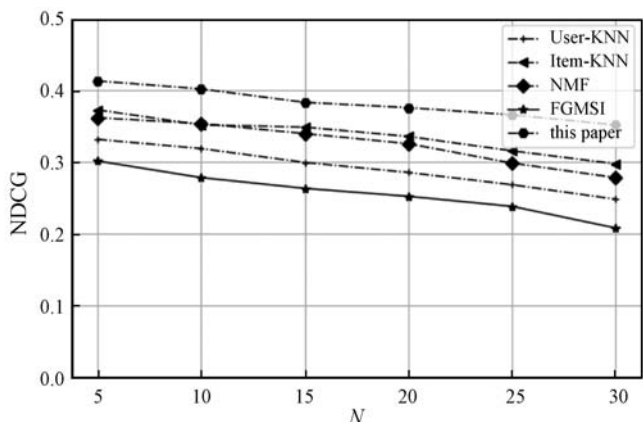
图 8 推荐实例

3.7 特征重要性分析

除了上述实验,文本通过依次将各特征模块消除的方法来探索各特征对模型性能的影响。实验结果如图 9 所示,图中,C_Student 表示在原模型的基础上消除课程-学生元数据信息模块,同理,C_Teacher、C_Course、C_Description 表示在原模型的基础上分别消除课程-教师、课程-课程和相关课程的文本描述信息模块。



(a) Precision



(b) NDCG

图9 各实验在数据集上的性能比较

由图9可知,对于本文模型四个输入模块(课程-学生元数据信息、课程-老师元数据信息、课程-课程元数据信息和课程的文本描述信息),去掉某种输入特征之后模型的性能都会下降,说明本文使用的三个网络拓扑结构和文本特征对于模型都有正面的积极作用。在三个网络拓扑结构中,课程-教师网络对模型推荐结果影响最小,课程-课程网络次之,课程-学生网络影响最大,表明对于学生与课程关系的挖掘具有重大潜力。而课程文本描述信息相较三种网络特征对模型的推荐效果影响都大,说明文本特征对推荐模型效果最重要。这进一步表明本模型融合课程-学生、课程-教师和课程-课程的网络拓扑结构,并结合课程描述文本信息来建模学生的偏好,能为学生提供个性化、精准化的学习资源推荐服务。

4 结 语

本文提出了一种融合网络特征和文本特征的智能课程推荐方法。首先构建在线课程的网络拓扑结构,并将其通过 Node2vec 映射到低维向量空间中,同时将文本数据通过 Doc2vec 表示,使用多模块 CNN 神经网络拟合网络和文本特征并计算用户和目标课程之间的相似度。在进行推荐时,选择匹配度最高的 Top N 课程。实验结果表明,该方法的性能远远超过四种基准模型性能,同时消除实验表明本文使用的每种特征对于推荐性能的提升都有积极作用。在以后的工作中,将挖掘更多形式的异构特征(例如学习序列、课程的多模态数据),并探讨不同特征融合方法对于智能课程推荐的影响。

参 考 文 献

[1] Lin Q, Niu Y, Zhu Y, et al. Heterogeneous knowledge-based attentive neural networks for short-term music recom-

mendations[J]. IEEE Access, 2018, 6: 58990 - 59000.

- [2] 何瑾琳,刘学军,徐新艳,等.融合 node2vec 和深度神经网络的隐式反馈推荐模型[J].计算机学报,2019,46(06):41 - 48.
- [3] Shen X, Yi B, Zhang Z, et al. Automatic recommendation technology for learning resources with Convolutional neural network[C]//2016 International Symposium on Educational Technology (ISET). IEEE, 2016.
- [4] 朱柳青.基于深度学习的课程推荐与学习预测模型研究[D].杭州:浙江工商大学,2018.
- [5] Khribi MK, Jemmi M, Nasraoui O. Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval[J]. Journal of Educational Technology and Society, 2009, 12(4):30 - 42.
- [6] 林木辉.融合学习者时序行为和认知水平的个性化学习资源推荐算法[J].计算机系统应用,2018,27(10):219 - 225.
- [7] Aher S, Lobo L. Combination of machine learning algorithms for recommendation of courses in E-Learning system based on historical data[J]. Knowledge-Based Systems, 2013, 51(1):1 - 14.
- [8] Batouche B, Brun A, Boyer A. Unsupervised machine learning based on recommendation of pedagogical resources[C]//European Conference on Technology Enhanced Learning. Springer, 2014: 548 - 549.
- [9] 王孟飞,胡超,于文涛.一种基于特征提取的教育视频资源推送方法[J].现代远程教育研究,2016(3):104 - 112.
- [10] Zhou Y, Huang C, Hu Q, et al. Personalized learning full-path recommendation model based on LSTM neural networks[J]. Information Sciences, 2018, 444: 135 - 152.
- [11] Wen Y, Zhang W, Luo R, et al. Learning text representation using recurrent convolutional neural network with highway layers [EB/OL]. <https://arxiv.org/abs/1606.06905>.
- [12] Kim, Yoon. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. ACL, 2014: 1746 - 1751.
- [13] 汪静,罗浪,王德强.基于 Word2Vec 的中文短文本分类问题研究[J].计算机系统应用,2018,27(5):209 - 215.
- [14] Mikolov T, Chen K, Corrado G, et al. Efficient estimation of word representations in vector space [EB/OL]. <https://arxiv.org/abs/1301.3781v2>.
- [15] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality [C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. ACM, 2013:3111 - 3119.
- [16] 唐明,朱磊.基于 word2vec 的一种文档向量表示[J].计算机学报,2016,43(6):214 - 217.

- [17] Le Q, Mikolov T. Distributed representations of sentences and documents [C]//International conference on machine learning. IEEE, 2014:1188 – 1196.
- [18] Vasile F, Smirnova E, Conneau A. Meta-Prod2Vec: Product embeddings using side-information for recommendation [C]//Proceedings of the 10th ACM Conference on Recommender Systems. ACM, 2016: 225 – 232.
- [19] Perozzi B, Al-Rfou R, Skiena S. Deepwalk: Online learning of social representations [C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014: 701 – 710.
- [20] Grover A, Leskovec J. node2vec: Scalable feature learning for networks [C]//Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2016: 855 – 864.
- [21] Lee D D, Seung H S. Algorithms for non-negative matrix factorization [C]//International Conference on Neural Information Processing Systems. MIT Press, 2001: 556 – 562.
- [22] Chen J, Ying P, Zou M. Improving music recommendation by incorporating social influence [J]. *Multimedia Tools & Applications*, 2018, 78(3):2667 – 2687.
- ~~~~~
- (上接第 53 页)
- [2] Biswas S K, Marbaniang L, Purkayastha B, et al. Rainfall forecasting by relevant attributes using artificial neural networks—a comparative study [J]. *International Journal of Big Data Intelligence*, 2016, 3(2):111 – 121.
- [3] Goswami B N, Venugopal V, Sengupta D, et al. Increasing trend of extreme rain events over India in a warming environment [J]. *Science*, 2006, 314(5804):1442 – 1445.
- [4] Seo J H, Lee Y H, Kim Y H. Feature selection for very short-term heavy rainfall prediction using evolutionary computation [J]. *Advances in Meteorology*, 2014(1–4):1 – 15.
- [5] Lin Y H, Chiu C C, Lin Y J, et al. Rainfall prediction using innovative grey model with the dynamic index [J]. *Journal of Marine Science and Technology*, 2013, 21(1):63 – 75.
- [6] Somvanshi V K, Pandey O P, Agrawal P K, et al. Modeling and prediction of rainfall using artificial neural network and ARIMA techniques [J]. *Journal of Indian Geophysical Union*, 2006, 10(2):141 – 151.
- [7] Geetha A, Nasira G M. Time-series modelling and forecasting: Modelling of rainfall prediction using ARIMA model [J]. *International Journal of Society Systems Science*, 2016, 8(4):361 – 372.
- [8] 刘剑. 基于 ARIMA 模型在降水量趋势分析与预测中的应用 [J]. *水利科技与经济*, 2018, 24(9):67 – 69.
- [9] Sumi S, Monira M, Faisal Z, et al. A rainfall forecasting method using machine learning models and its application to the Fukuoka city case [J]. *International Journal of Applied Mathematics and Computer Science*, 2012, 22(4):841 – 854.
- [10] Parmar A, Mistree K, Sompura M. Machine learning techniques for rainfall prediction: A review [C]//International Conference on Innovations in Information Embedded and Communication Systems, 2017.
- [11] 王汉涛, 张潇潇. 仿生算法优化 BP 神经网络在降雨空间插值中的应用 [J]. *水资源与水工程学报*, 2019, 30(3):106 – 112.
- [12] 郭尚瓚, 肖达, 袁行远. 基于神经网络和模型集成的短时降雨预测方法 [J]. *气象科技进展*, 2017(1):107 – 113.
- [13] Chen S T, Yu P S, Tang Y H. Statistical downscaling of daily precipitation using support vector machines and multivariate analysis [J]. *Journal of Hydrology*, 2010, 385(1–4):13 – 22.
- [14] Li X, Du Z, Song G. A method of rainfall runoff forecasting based on deep convolution neural networks [C]//International Conference on Advanced Cloud and Big Data (CBD), 2018.
- [15] Frederik K, Daniel K, Claire B, et al. Rainfall – runoff modelling using Long Short-Term Memory (LSTM) networks [J]. *Hydrology and Earth System Sciences*, 2018, 22(11):6005 – 6022.
- [16] 刘乐, 王洪国, 王宝伟. 基于 PCA 和改进 BP 网络的降雨预报模型研究 [J]. *计算机工程与应用*, 2008, 44(12):234 – 237.
- [17] Zhang P C, Zhang L, Leung H, et al. A Deep-Learning based precipitation forecasting approach using multiple environmental factors [C]//IEEE International Congress on Big Data, 2017.
- [18] 张鹏程, 贾阳阳. 一种基于多层感知器的动态区域联合短时降水预报方法 [J]. *计算机应用与软件*, 2018, 35(11):153 – 158, 183.
- [19] Shi X, Gao Z, Lausen L, et al. Deep learning for precipitation nowcasting: A benchmark and a new model [C]//Advances in Neural Information Processing Systems, 2017.
- [20] Wang Y B, Long M S, Wang J M, et al. PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs [C]//Advances in Neural Information Processing Systems, 2017.
- [21] Ma Y, Peng H, Cambria E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM [C]//32nd AAAI Conference on Artificial Intelligence, 2018.
- [22] Cho K, Courville A, Bengio Y. Describing multimedia content using attention-based encoder-decoder networks [J]. *IEEE Transactions on Multimedia*, 2015, 17(11):1875 – 1886.
- [23] Mckee T B, Doesken N J, Kliest J. The relationship of drought frequency and duration to time scales [C]//8th Conference on Applied Climatology, 1993.