结合强化学习和用户短期行为的新闻推荐算法

姚 楠[1] 何 山[1] 赵 越[1] 李任花[2]

[1](西南石油大学计算机科学学院 四川 成都 610500)

[2](成都先进功率半导体股份有限公司 四川 成都 611730)

摘 要 针对传统的协同过滤推荐算法仅根据用户历史评分矩阵进行推荐,存在矩阵稀疏和无法动态观察用户兴趣变化的问题,提出一种将用户短期行为和强化学习相结合的新闻推荐方法。将新闻文本向量化后,通过聚类提取类别特征,再根据强化学习中的状态、动作和奖励等概念,以 Double DQN 算法为框架来建立推荐模型,利用循环神经网络近似动作值函数来进行计算。最后在财新网的真实新闻浏览数据集上对提出的算法进行验证,对比传统算法,实验结果表明,提出的算法在推荐准确率、召回率等指标上都有明显提高,能够更加有效地进行推荐。

关键词 推荐系统 强化学习 新闻推荐 神经网络 聚类

中图分类号 TP391 **文献标志码** A DOI:10.3969/j.issn.1000-386x.2024.04.042

NEWS RECOMMENDATION ALGORITHM COMBINING REINFORCEMENT LEARNING AND USER SHORT-TERM BEHAVIOR

Yao Nan[1] He Shan[1] Zhao Yue[1] Li Renhua[2]

[1](*School of Computer Science, Southwest Petroleum University, Chengdu 610500, Sichuan, China*)

[2](*Chengdu Advanced Power Semiconductor Co., Ltd., Chengdu 611730, Sichuan, China*)

Abstract The traditional collaborative filtering recommendation algorithm only makes recommendations based on the user history score matrix, which has the problems of sparse matrix and inability to dynamically observe user interest changes. A news recommendation method that combines user short-term behavior and reinforcement learning is proposed. After the news text was vectorized, the category features were extracted through clustering. Based on the concepts of state, action and reward in reinforcement learning, the Double DQN algorithm was used as the framework to establish a recommendation model, and the recurrent neural network was used to approximate the action value function for calculation. The proposed algorithm was verified on the real news browsing data set of Caixin. Compared with the traditional algorithm, the experimental results show that the proposed algorithm has significantly improved the recommendation precision rate, recall rate and other indicators, and can perform more effectively.

Keywords Recommender system Reinforcement learning News recommendation Neural network Clustering

0 引 言

互联网的发展改变了传统的信息传播方式,新闻媒介由传统的纸质媒介扩展到了现代化媒体,使人们能够更加快速地接收即时消息。与此同时,网络上的数据量越来越多,对于用户和网站运营方来说,都产生了信息过载[1]的问题。推荐系统作为一种能够有效解决信息过载的信息过滤手段,在互联网各领域中得到了普遍应用,是如今不可或缺的一门个性化服务技术[2]。

推荐系统的传统主流方法包括三类[3]:基于协同过滤的方法、基于内容的方法以及混合推荐方法。协同过滤推荐根据用户和物品之间的历史交互评分矩

收稿日期:2020-12-16。**姚楠**,硕士生,主研领域:推荐系统。**何山**,副教授。**赵越**,硕士生。**李任花**,硕士生。

阵,来计算用户或者物品之间的相似度,并由此分为基于用户的协同过滤算法(User-based CF)和基于物品的协同过滤算法(Item-based CF),适用于各种非结构化数据[4]。基于内容的方法通过充分挖掘物品的特征信息,不局限于评分,来计算物品相似性进行推荐[5]。混合推荐方法则是综合利用不同推荐方法的特点来获得最终的推荐结果[6]。其中协同过滤方法是应用最为广泛的方法,其适用性较广,不用对用户和物品特征信息进行处理,但是存在矩阵稀疏的问题[7],并且忽略了其他特征信息对推荐结果的影响,也不能够动态对用户的兴趣特征变化做出反应。

为了克服以上这些缺点,在推荐系统中引入了强化学习,通过智能体与环境的互动来得到最优策略,可以有效解决矩阵稀疏、不能动态进行推荐的问题。强化学习与深度学习相结合还产生了深度强化学习,利用深度学习极强的表征能力,可以解决实际环境中状态和动作数量非常庞大的任务,还能更高效地进行计算。在经验回放问题上,Zhao 等[8]提出了一种利用分层采样思想进行经验回放的强化学习推荐方法,通过对用户信息进行类别划分,每次回放时从经验池分层采样数据来训练网络,加速网络的收敛;Zhang 等[9]提出了基于优先经验重放技术的深度强化学习推荐方法,通过计算交叉熵来捕捉用户兴趣的动态变化用以确定回放的优先级。在状态动作定义方面,Choi 等[10]研究了一种通过 Biclustering 聚类技术将推荐系统转化为网格世界游戏的方法,不仅可以缩小状态和动作空间,还可以有效处理冷启动问题,从而提高推荐质量。在平衡探索与利用的问题上,Zheng 等[11]提出了一种基于深度强化学习的在线新闻推荐框架 DRN,同时考虑了用户点击和用户活跃度作为推荐的反馈,并采取 Dueling Bandit Gradient Descent 方法进行探索,避免经典探索方法对推荐精确度的影响。

现存的很多强化学习推荐方法依赖于充足的用户特征,而在实际场景中,很多用户在注册时不会提供完备的信息,甚至有些用户以游客的身份访问,所以本文提出了一种将用户短期行为和强化学习相结合的新闻推荐算法,在用户信息不充裕的情况下,可以避免矩阵稀疏问题并充分挖掘用户短期的浏览兴趣来动态进行推荐。首先,对新闻文本内容进行文档向量化,再进行聚类,将新闻特征用新闻 id 和聚类得到的类别特征来表示,将用户最近浏览的固定数目的新闻特征定义为用户短期行为,利用用户短期行为来对用户下一步即将浏览的新闻做出推荐,并使用经验回放技术来学习经验池中存储的样本,从而提高推荐准确性。

1　相关工作

1.1　强化学习

强化学习是一种以目标为导向的学习方法,旨在寻找连续时间序列的动态变化过程中的最优策略[12]。

强化学习的数学基础理论基于无后效性的马尔可夫决策过程,系统的下一个状态只与当前状态有关,而与之前的状态无关。马尔可夫决策过程由四元组构成,即 $MDP=(S,A,P_{sa},R)$,其中:S 表示状态空间集;A 表示动作空间集;P_{sa} 为状态转移概率,表示在状态为 s 的情况下执行动作 a 后,转移到另一个状态 s' 的概率分布;R 为奖励函数,表示在状态为 s 的情况下执行动作 a 后转移到状态 s' 所获得的即时奖励 r。

类似的,强化学习由智能体、状态、动作环境和奖励构成。智能体在时间步 t 的状态 $s_t\in S$ 下,根据策略选择动作 $a_t\in A$ 后,会在下一个时间步 $t+1$ 转移至其他状态 $s_{t+1}\in S$。同时,环境给出在 s_t 下采取动作 a_t 并到达 s_{t+1} 的即时奖励 r_t,一个简单的强化学习流程如图 1 所示。智能体的最终目标是找到最佳策略 π^* 从而使累积奖励最大化。时间步 t 的折扣未来累积奖励公式如下:

$$G_t=r_t+\gamma r_{t+1}+\gamma^2 r_{t+2}+\cdots+\gamma^{n-t}r_n \tag{1}$$

式中:$\gamma\in[0,1]$ 为折扣系数,即距离当前时间步越远的奖励,对当前奖励的影响越小。

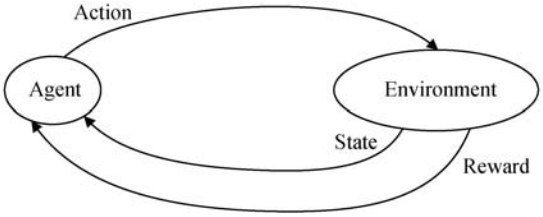

图 1　强化学习简单流程

强化学习的求解等同于利用贝尔曼方程迭代更新状态动作值函数:

$$Q(s,a)\leftarrow Q(s,a)+\alpha[r+\gamma\max_{a'}Q(s',a')-Q(s,a)] \tag{2}$$

在实际应用中,由于状态空间和动作空间通常不是一个很小的范围,为了便于求解,利用值函数近似法来逼近最优动作值函数 $Q^*(s,a)$。一般采用线性函数逼近法,也可以采用神经网络等非线性近似法,如式(3)所示。

$$Q(s,a,\theta)\approx Q^*(s,a) \tag{3}$$

式中:θ 表示用神经网络来近似时的权重参数,其中参数 θ 的值通过最小化式(4)中的损失函数来计算。

$$L(\theta)=E[(r+\gamma\max_{a'}Q(s',a',\theta)-Q(s,a,\theta))^2] \tag{4}$$

DQN 是将神经网络和 Q-learning 算法相结合的一种深度强化学习框架[13]。使用双网络结构,预测网络用来评估当前状态和动作对应的值函数,目标网络用于计算目标价值;并采取经验回放机制,在每一步将元组(s,a,r,s',T)存储在经验池中,并从经验池中随机抽取固定批量的样本参与更新计算预测网络的权重参数;每隔固定步数将目标网络的参数更新替换为预测网络的参数。这样的设置有助于去除样本间的相关性和依赖性,可以增加网络的稳定性,同时使网络快速收敛。

Double DQN 的思想是对动作的选择和动作状态值估计进行解耦,预测网络负责动作的选择,目标网络负责 Q 值的估算,避免产生过高估计。

1.2 循环神经网络

循环神经网络是一种可以处理序列数据的深度学习网络。与密集连接网络、卷积神经网络比,它具有记忆的模块,是一种具有内部循环特性的神经网络[14]。在面对序列数据时,不再单独处理每个输入,而是遍历所有的序列元素,并在每一步保存一个状态,其中包含与已处理过的内容相关的信息。

如图 2 所示是一个简单循环网络的示意图,在每一步不仅对当前输入的信息进行计算,还要加入前期保留的状态信息,一起计算得到当前步的输出值。

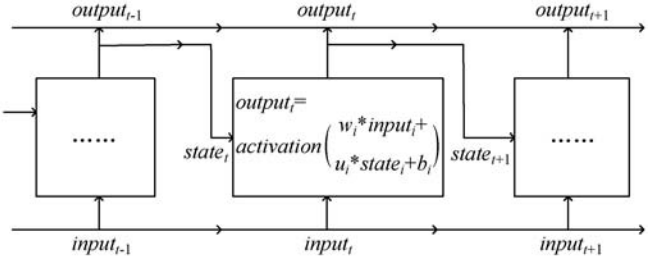

图 2 简单循环神经网络示意图

针对简单循环网络在某些情况下会出现梯度消失的问题,出现了 LSTM 和 GRU 等简单循环网络的变体。LSTM 模型含有记忆模块,每个记忆模块包含了多个自相关的核心信元 cell 和门控结构:输入门、输出门和遗忘门,增加了一种携带信息跨越多个时间步的方法,可以有效解决梯度消失的问题,本文将采用 LSTM 网络来处理用户短期内浏览的新闻序列数据。

1.3 新闻推荐算法

1.3.1 新闻特征预处理

在推荐之前,先对新闻文本数据进行预处理。

新闻文档向量化使用 word2vec 与 tf-idf 结合[15]的方法。文档向量化具体步骤如下:首先使用 jieba 分词对新闻文本进行分词处理;再使用 word2vec 方法对分词结果进行词向量化得到每个词对应的向量 $\boldsymbol{w}_t$;最后结合词的 tf-idf 值来计算得到文档向量。

词 t 的 idf 计算公式如下:

$$idf(t) = \log\left(\frac{M}{n_t} + 0.01\right) \tag{5}$$

式中:M 表示新闻文档总数量;n_t表示所有文档中出现词 t 的文档数量。

词 t 的 tf-idf 计算公式如下:

$$K(t,D_i) = \frac{tf(t,D_i) \times idf(t)}{\sqrt{\sum_{t\in D_i}[tf(t,D_i) \times idf(t)]^2}} \tag{6}$$

式中:$tf(t,D_i)$表示词 t 在第 i 篇文档中的词频。

则每篇新闻文档可用如向量表示:

$$\boldsymbol{d}_i = \sum_{t\in D_i} \boldsymbol{w}_t K(t,D_i) \tag{7}$$

获得新闻文档向量之后,利用 k-means 聚类方法对文档向量进行聚类,得到新闻的类别特征。在不同的聚类数目下,计算每个变量点到其所属类别中心的距离平方作为畸变程度,利用畸变程度之和作为成本函数评估聚类效果,并选择畸变程度改善幅度最大的位置对应的肘部拐点处作为最终的 k 值[16]。本文计算了 k 值从 2 到 50 的聚类结果畸变程度,选择 7 作为最终的 k 值。

1.3.2 构建强化学习推荐框架

对新闻信息进行预处理之后,构建以 Double DQN 算法框架为基准的新闻推荐算法。

Q 网络对应强化学习中的 agent,其结构如图 3 所示。Q 网络的输入由长度为 H 的用户短期浏览信息组成,浏览信息包括两部分:新闻 id 和新闻类别按照用户浏览时间的先后顺序依次排列。将用户短期浏览信息输入 Q 网络之后,输出数据集中的每条新闻在下一个时间步被推荐的概率,分别代表通过在当前状态下采取不同的动作而得到的 Q 值。

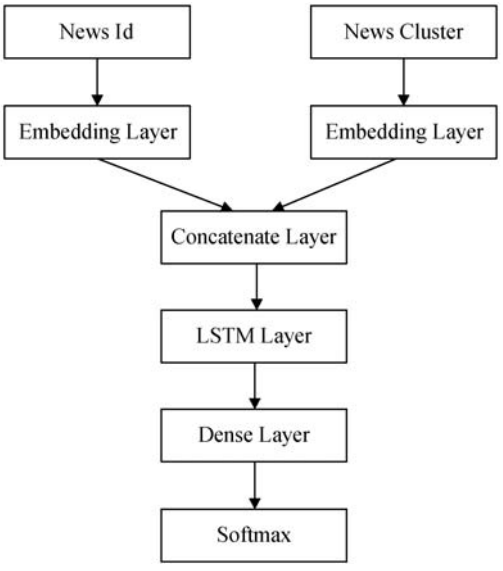

图 3 Q 网络结构

对于输入数据,将新闻的 id 和类别均用数值索引来表示,在输入网络之后,首先对每个数值用 Embedding 层嵌入为 $1\times V$ 的向量。在嵌入层之后,使用 Concatnate 层连接每组新闻嵌入的 id 向量和类别向量,每条新闻的维度为 $1\times 2V$,从而获得用每条新闻特征表示的维度为 $H\times 2V$ 的浏览信息。然后,使用 LSTM 层处理这些向量。

最后将 LSTM 层输出的结果利用全连接层进行处理,通过 Softmax 函数输出每条新闻下一时刻被推荐的概率。

状态设计:在构建的推荐模型中,当前状态信息包括用户最近的 H 条浏览记录,其中每条浏览记录包括被浏览新闻的 id 和类别。强化与推荐系统结合应用中的状态通常定义为用户某一次的行为,本文引入用户短期内浏览的多条新闻特征作为状态,使智能体在决策时可以通过参考更多的用户行为,来更好地学习用户的短期动态兴趣,从而对其做出更准确的推荐。以 $H=5$ 为例,状态 s 如下:

$$s=\begin{bmatrix}(NewsId_0,NewsCategory_0)\\(NewsId_1,NewsCategory_1)\\\vdots\\(NewsId_4,NewsCategory_4)\end{bmatrix} \tag{8}$$

下一个状态将添加用户实际观看的下一条新闻进入当前状态,并删除当前状态中最早浏览的那条新闻,保持状态中只包含 H 条浏览的新闻信息。对于处于状态 s 的浏览记录,推荐系统据此向用户推荐下一时刻的新闻,根据实际数据中用户下一时刻是否观看此条新闻,来训练网络。

动作和奖励设计:动作值对应于每条新闻的索引,选择的新闻索引即为下一时刻推荐系统选择推荐给用户的新闻。在选择动作时采取 ε-greedy 贪心策略,由于动作空间较大,所以在探索中以一定的概率 μ 直接选取用户实际下一条浏览的新闻作为动作,加速模型的学习和收敛。如果推荐正确,推荐的新闻即为用户下一时刻实际浏览的新闻,则奖励值为 1,否则为 -1。

$$r=\begin{cases}1 & \text{推荐的新闻 = 实际下一条浏览的新闻}\\-1 & \text{其他}\end{cases} \tag{9}$$

基于 Double DQN 框架的推荐算法流程由算法 1 给出。

算法 1　结合强化学习和用户短期行为的推荐算法新闻推荐算法训练过程

预处理新闻特征
初始化经验池 D,最大容量为 N
初始化预测网络权重参数为 θ
初始化目标网络权重参数 $\theta'=\theta$
初始化探索概率 ε
初始化探索中直接推荐实际下一条浏览新闻的概率 μ
For episode = 1, M do
　初始化状态 s_1
　For $t=1$, T do
　　以概率 $\varepsilon\times\mu$ 选择实际下一条记录为动作 a_t
　　以概率 $\varepsilon\times(1-\mu)$ 随机选择动作 a_t
　　否则依据公式 $a_t=\max_a Q(s_t,a,\theta)$ 选择动作 a_t
　　执行动作 a_t,获得奖励 r_t,转移到状态 s_{t+1}
　　存储 (s,a,r,s',T) 到经验池 D 中
　　从经验池 D 随机采样 minibatch 大小的样本
　　设

$$y_i=\begin{cases}r_j & \text{当前经验轨迹终止在时间步 } j+1\\r_j+\gamma Q(s_{j+1},\arg\max_{a'}Q(s_{j+1},a',\theta),\theta') & \text{其他}\end{cases}$$

　　计算损失函数更新参数 θ:

$$L(\theta)=(y_i-Q(s,a,\theta))^2$$

　　每隔 C 步更新 $\theta'=\theta$
　　End For
　End For

在训练网络得到推荐模型后,再按照时间步顺序依次获得推荐结果,直至形成长度为 top N 的推荐列表,具体推荐流程由算法 2 给出。

算法 2　结合强化学习和用户短期行为的推荐流程新闻推荐算法推荐流程

输入:用户在训练集中最后浏览的 H 条新闻的特征 SI
训练得到的模型 Model
推荐列表长度 top N
空的推荐列表 RL
输出:推荐列表 RL
For i = 1, top N do
　输入 SI 用 Model 得到当前时间步的输出
　选择输出中概率最高的新闻 n * 作为当前时间步的推荐新闻,并加入 RL
　删除 SI 中最前的一条新闻特征
　加入新闻 n * 的特征在 SI 末尾
　更新 SI,且保持长度为 H
End For

2　实　验

2.1　数据集

本次实验使用的数据集是 DataCastle 竞赛官方网站给出的新闻浏览数据集,包括在财新网上随机抽取的 10 000 名用户在 2014 年 3 月期间的所有新闻浏览记录,每条记录包括用户 id、新闻 id、时间戳格式的浏览时间、新闻标题、新闻详细正文内容和发布时间。数据真实可靠,可以保证实验结果的准确性。

这次实验筛选出数据集中新闻浏览总数不少于40条的用户新闻浏览数据。这样可以确保用户是连续动态地使用新闻系统,从而更好地捕捉用户兴趣的动态变化,并且使推荐系统可以充分利用历史浏览数据学习如何进行下一时间步的推荐,从而保证推荐结果有更高的准确性。将每个用户的浏览记录按照浏览时间的顺序排列,以每个用户最后浏览的15条新闻作为测试集,其余的浏览记录作为训练集。

在生成推荐结果时,采用每个用户在训练数据中的最后一部分新闻浏览数据作为模型输入,得到下一时间步的推荐新闻,加入推荐新闻列表,直至推荐数量到达top N。

实验以基于用户的协同过滤推荐方法、基于物品的协同过滤推荐方法、增量更新协同过滤算法IUIBCF[17]、基于VSM和bisecting K-means的推荐算法[18]作为实验参照对象,对推荐结果进行了对比分析。

2.2 实验环境

实验硬件平台如下:处理器Intel(R) Core(TM) i7-10710U CPU @ 1.10 GHz,内存16 GB。实验软件平台如下:操作系统Windows 10(64位),开发语言Python 3.7,软件工具Keras、Numpy。

2.3 评价标准

由于本数据集中不包含用户评分等级信息,只有用户浏览的记录,所以在给用户生成个性化的top N推荐列表后,采取常用的准确率(Precision)、召回率(Recall)、F1值(F1 measure)和多样性(Diversity)四个指标[19]来对推荐结果进行评价。

推荐结果准确率的计算公式如下:

$$P_{\text{recision}} = \frac{(R(u) \cap T(u))}{R(u)} \tag{10}$$

推荐结果召回率的计算公式如下:

$$R_{\text{ecall}} = \frac{(R(u) \cap T(u))}{T(u)} \tag{11}$$

式中:$R(u)$表示模型根据用户在训练集中最后H条浏览记录对用户做推荐所生成的长度为top N的推荐新闻列表;$T(u)$则表示用户在测试集中实际浏览的新闻列表。

受推荐列表长度变化的影响,有时候只使用准确率和召回率不能有效评估推荐系统的性能,所以在与其他算法的对比实验中加入了平衡二者的F1值作为评估指标。推荐结果F1值的计算公式如下:

$$F_1 = \frac{2 \times P_{\text{recision}} \times R_{\text{ecall}}}{P_{\text{recision}} + R_{\text{ecall}}} \tag{12}$$

多样性描述了推荐列表中新闻两两之间的不相似性,通过计算所有用户推荐列表内新闻的不相似性指数,来衡量推荐模型满足用户多样化兴趣的能力。单个用户的多样性计算公式如下:

$$Diversity\_user = 1 - \frac{\sum_{i,j \in R(u), i \neq j} sim(i,j)}{\frac{1}{2}|R(u)|(|R(u) - 1|)} \tag{13}$$

$$sim(i,j) = \frac{\boldsymbol{v}_i \cdot \boldsymbol{v}_j}{\|\boldsymbol{v}_i\| \|\boldsymbol{v}_j\|} \tag{14}$$

式中:$sim(i.j) \in [0,1]$表示推荐列表$R(u)$中,新闻i的文档表示向量$\boldsymbol{v}_i$和新闻j的表示向量$\boldsymbol{v}_j$之间的余弦相似度。

推荐模型的整体多样性则是所有用户推荐列表多样性的平均值,如式(15)所示。

$$Diversity = \frac{1}{|U|} \sum_{u \in U} Diversity\_user(R(u)) \tag{15}$$

2.4 结果与分析

1)确定窗口期参数(H)。在不同的窗口期参数($H=5,10,15,20$)下,通过不同长度的top N推荐来对比推荐性能,分别在top N取1、5、10、15、20、25的长度下进行推荐。实验得到不同窗口长度值下的准确率和召回率,如表1和表2所示。

表1 不同窗口期参数在不同推荐长度下的推荐准确率(%)

窗口长度	推荐长度					
	1	5	10	15	20	25
5	32.60	32.16	27.62	21.25	16.34	13.45
10	32.97	32.67	29.19	24.25	18.70	15.09
15	31.87	31.21	27.29	21.86	17.07	13.96
20	32.97	32.16	27.80	21.90	17.12	14.07

表2 不同窗口期参数在不同推荐长度下的推荐召回率(%)

窗口长度	推荐长度					
	1	5	10	15	20	25
5	2.17	10.72	18.41	21.25	21.78	22.42
10	2.20	10.89	19.46	24.25	24.93	25.15
15	2.12	10.40	18.19	21.86	22.76	23.27
20	2.20	10.72	18.53	21.90	22.83	23.44

不同窗口期参数的设置影响了模型在输入层接收到的信息数量。通过以上两个表格中的推荐性能对比,可以看出在窗口期长度取10时,在不同的推荐长度下,推荐的准确率和召回率都明显高于其他情况下的值,所以确定窗口参数H的值为10。

2)对比实验。图4和图5显示了基于用户的协

同过滤算法、基于物品的协同过滤算法、增量更新协同过滤算法、基于 VSM 和 bisecting K-means 的推荐算法以及本文算法在不同推荐长度下的准确率曲线和召回率曲线。这 5 种算法的准确率都呈现由高到低的趋势,且召回率都呈现由低到高的趋势,这是因为推荐列表长度逐渐增加而测试集长度固定,根据准确率和召回率的计算公式可以解释此趋势。在不同推荐长度下,本文算法的两种性能基本都高于其他算法,在推荐列表长度较长时,所提出算法的性能与较新的 VSM + KM 算法基本持平。

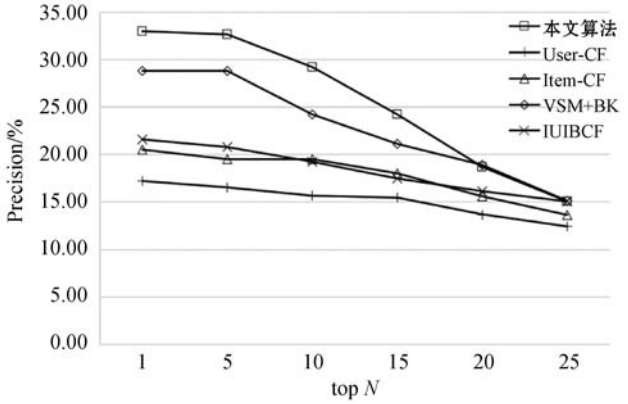

图 4　不同算法在不同推荐长度下的推荐准确率对比

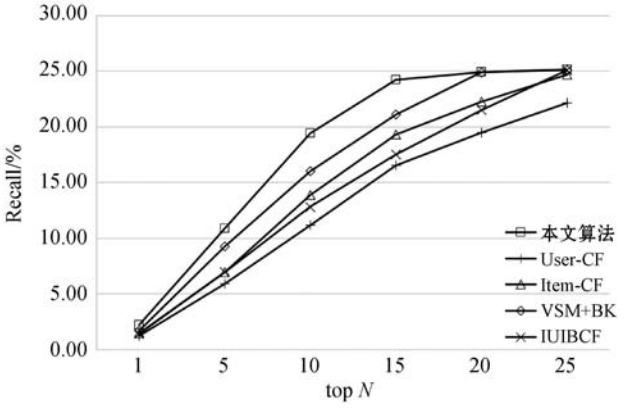

图 5　不同算法在不同推荐长度下的推荐召回率对比

5 种算法的 F1 值对比如图 6 所示。所有 F1 值均呈现先上升后下降的趋势,这取决于准确率和召回率的变化。可以看出本文算法在推荐长度较短时有更好的表现。由于本文的推荐方法是不断加入新推荐的新闻特征来继续进行下一时间步的推荐,随着推荐列表的增长,模型输入数据中包含的用户真实浏览记录越来越少,推荐新闻占比越来越多,所以可以解释推荐效果在推荐列表长度较大时随着长度的增加下降得略快,符合本文算法按照时间步向用户逐步推荐的预期,因此对于每一步产生的推荐结果的筛选还需要进一步深入研究。总体来说本文方法的 F1 值与其他 4 种方法相比具有明显优势,验证了根据用户短期兴趣进行动态推荐的有效性。

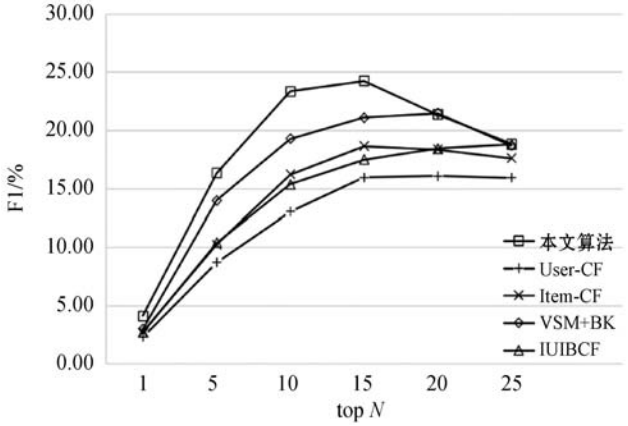

图 6　不同算法在不同推荐长度下的 F1 值对比

图 7 显示了 5 种推荐算法在不同推荐长度下的多样性对比结果。可以看出总体多样性指标表现最好的是 IUIBCF 增量协同过滤算法,本文方法随着推荐列表的增长逐渐表现出优势,在后期与 Item-CF 算法的多样性基本持平,且明显优于其他两种算法,体现了强化学习中探索策略的有效性。

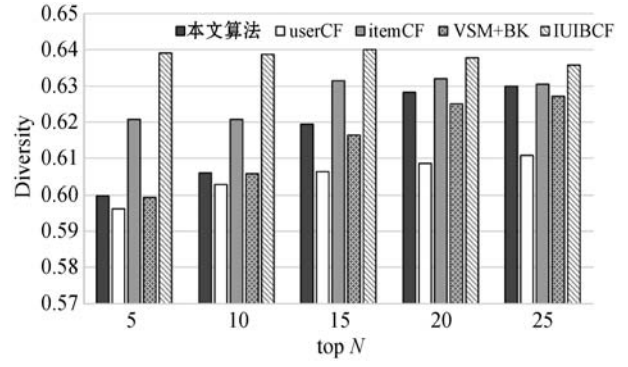

图 7　不同算法在不同推荐长度下的推荐多样性对比

综合多种对比指标来说,本文提出的方法在提升了 F1 值的同时保持了一定的多样性,优于文中其他对比方法。但是在推荐列表长度增加时优势逐渐减弱,每一时间步推荐新闻的选择依据还需要进一步研究进行完善和扩充。

3　结　语

本文从用户短期内的浏览记录出发,针对用户兴趣的动态变化,结合循环神经网络和深度强化学习框架来进行学习,并应用到新闻个性化推荐上,可以避免传统推荐方法中矩阵稀疏的问题,并适用于用户信息不充足的场景。实验结果表明,与传统推荐和较新的方法相比,本文方法在 top N 推荐下具有更好的效果,证明了该方法的有效性。下一步,将考虑新闻文本更多其他维度的特征信息,并考虑和用户的长期兴趣相结合,更准确地挖掘用户兴趣,来进一步提升推荐性能。

参 考 文 献

[1] Ferdaous H, Bouchra F, Brahim O, et al. Recommendation using a clustering algorithm based on a hybrid features selection method[J]. Journal of Intelligent Information Systems, 2018,51(1):183 - 205.

[2] Li L, Zheng L, Yang F, et al. Modeling and broadening temporal user interest in personalized news recommendation [J]. Expert Systems with Applications,2014,41(7):3168 - 3177.

[3] 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.

[4] Zhang Z, Peng T, Shen K. Overview of collaborative filtering recommendation algorithms[J]. IOP Conference Series: Earth and Environmental Science,2020,440(2):22063.

[5] Lops P, Jannach D, Musto C, et al. Trends in content-based recommendation[J]. User Modeling and User-Adapted Interaction,2019,29(2):239 - 249.

[6] Walek B, Fojtik V. A hybrid recommender system for recommending relevant movies using an expert system[J]. Expert Systems with Applications,2020,158:113452.

[7] Huynh H X, Phan N Q, Pham N M, et al. Context-similarity collaborative filtering recommendation[J]. IEEE Access, 2020,8:33342 - 33351.

[8] Zhao Z R, Chen X L. Deep reinforcement learning based recommend system using stratified sampling[J]. IOP Conference Series Materials Science and Engineering,2018,466: 12110.

[9] Zhang Y, Su X Y, Yong L. A novel movie recommendation system based on deep reinforcement learning with prioritized experience replay [C]//19th International Conference on Communication Technology,2019:1496 - 1500.

[10] Choi S, Ha H, Hwang U, et al. Reinforcement learning based recommender system using bi-clustering technique[EB]. arXiv:1801.05532,2018.

[11] Zheng G J, Zhang F Z, Zheng Z H, et al. DRN: A deep reinforcement learning framework for news recommendation [C]//World Wide Web Conference,2018:167 - 176.

[12] Sutton R, Barto A, Williams R. Reinforcement learning is direct adaptive optimal control[J]. IEEE Control Systems Magazine,1992,12(2):19 - 22.

[13] Sutton R, Barto A. Reinforcement learning: An introduction [M]. Cambridge: MIT Press,2017.

[14] 弗朗索·瓦肖莱. Python 深度学习[M].张亮,译.北京:人民邮电出版社,2018.

[15] 唐明,朱磊,邹显春.基于 Word2Vec 的一种文档向量表示[J].计算机科学,2016,43(6):214 - 217,269.

[16] 毕曦文,纪明宇,吴鹏,等.个性化高校新闻分类推荐的应用研究[J].计算机应用与软件,2019,36(7):218 - 223.

[17] Jia Z Y, Yang Y T, Gao W, et al. User-based collaborative filtering for tourist attraction recommendations[C]//IEEE International Conference on Computational Intelligence & Communication Technology,2015:22 - 25.

[18] 袁仁进,陈刚,李锋.面向新闻推荐的用户兴趣模型构建与更新[J].计算机应用研究,2019,36(12):3593 - 3596.

[19] 项亮.推荐系统实践[M].北京:人民邮电出版社,2012: 23 - 26.

(上接第 255 页)

[13] Schubotz M, Grigorev A, Leich M, et al. Semantification of identifiers in mathematics for better math information retrieval[C]//39th International ACM SIGIR Conference on Research and Development in Information Retrieval,2016:135 - 144.

[14] Pathak A, Pakray P, Sarkar S, et al. Mathirs: Retrieval system for scientific documents[J]. Computación y Sistemas,2017,21(2):253 - 265.

[15] Kristianto G, Topic G, Aizawa A. MCAT math retrieval system for NTCIR-12 MathIR task[C]//12th NTCIR Conference on Evaluation of Information Access Technologies, 2016:323 - 330.

[16] Lin X, Gao L, Hu X, et al. A mathematics retrieval system for formulae in layout presentations[C]//37th International ACM SIGIR Conference on Research and Development in Information Retrieval,2014:697 - 706.

[17] 俞琰,陈磊,姜金德,等.基于依存句法分析的中文专利候选术语选取研究[J].图书情报工作,2019,63(18):109 - 118.

[18] Schmitz M, Bart R, Soderl S, et al. Open language learning for information extraction[C]//2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning,2012:523 - 534.

[19] Mausam M. Open information extraction systems and downstream applications[C]//25th International Joint Conference on Artificial Intelligence,2016:4074 - 4077.

[20] 赵华茗,钱力,余丽.依存句法特征的科研命名实体识别算法[J].图书情报工作,2020,64(11):108 - 115.

[21] Manning C, Surdeanu M, Bauer J, et al. The Stanford CoreNLP natural language processing toolkit[C]//52nd Annual Meeting of the Association for Computational Linguistics: System,2014:55 - 60.

[22] Falenska A, Kuhn J. The (Non-) Utility of structural features in BiLSTM-based dependency parsers[C]//57th Annual Meeting of the Association for Computational Linguistics,2019:117 - 128.

[23] Wang Y, Gao L, Wang S, et al. WikiMirs 3.0: A hybrid MIR system based on the context, structure and importance of formulae in a document[C]//15th ACM/IEEE-CS Joint Conference on Digital Libraries,2015:173 - 182.