Journal of Computer Applications
计算机应用, 2020, 40(9): 2613 - 2621
ISSN 1001-9081
CODEN JYIIDU
2020-09-10
http://www.joca.cn

文章编号:1001-9081(2020)09-2613-09
DOI:10.11772/j.issn.1001-9081.2019112002

# 基于改进的深度Q网络结构的商品推荐模型

傅　魁,梁少晴*,李　冰
(武汉理工大学 经济学院,武汉 430070)
(*通信作者电子邮箱colorfulsq@foxmail.com)

**摘　要**:传统推荐方法存在数据稀疏和特征识别差等问题,为了解决这些问题,根据隐式反馈构建具有时序性的正负反馈数据集。由于正负反馈数据集和商品购买具有强时序性特征,引入长短期记忆(LSTM)网络作为模型构件。考虑用户自身特征和用户动作选择回报由不同的输入数据决定,对竞争架构的深度Q网络进行改进,融合用户正负反馈和商品购买时序性,设计了基于改进的深度Q网络结构的商品推荐模型。模型对正负反馈数据进行区分性训练,对商品购买的时序性特征进行提取。在Retailrocket数据集上,与因子分解机(FM)模型、W&D模型和协同过滤(CF)模型中表现最好的相比,所提模型的准确率、召回率、平均准确率(MAP)和归一化折损累计增益(NDCG)分别提高了158.42%、89.81%、95.00%和67.57%。同时,使用DBGD作为探索方法,改善了推荐商品多样性低的缺陷。

**关键词**:深度强化学习;正负反馈数据集;竞争网络架构;长短期记忆网络;商品推荐
**中图分类号**:TP181　　**文献标志码**:A

## Commodity recommendation model based on improved deep Q network structure

FU Kui, LIANG Shaoqing*, LI Bing
(*School of Economics*, *Wuhan University of Technology*, *Wuhan Hubei* 430070, *China*.)

**Abstract:** Traditional recommendation methods have problems such as data sparsity and poor feature recognition. To solve these problems, positive and negative feedback datasets with time-series property were constructed according to implicit feedback. Since positive and negative feedback datasets and commodity purchases have strong time-series feature, Long Short-Term Memory (LSTM) network was introduced as the component of the model. Considering that the user's own characteristics and action selection returns are determined by different input data, the deep Q network based on competitive architecture was improved: integrating the user positive and negative feedback and the time-series features of commodity purchases, a commodity recommendation model based on the improved deep Q network structure was designed. In the model, the positive and negative feedback data were trained differently, and the time-series features of the commodity purchases were extracted. On the Retailrocket dataset, compared with the best performance among the Factorization Machine (FM) model, W&D (Wide & Deep learning) and Collaborative Filtering (CF) models, the proposed model has the precision, recall, Mean Average Precision (MAP) and Normalized Discounted Cumulative Gain (NDCG) increased by 158.42%, 89.81%, 95.00% and 65.67%. At the same time, DBGD (Dueling Bandit Gradient Descent) was used as the exploration method, so as to improve the low diversity problem of recommended commodities.

**Key words:** deep reinforcement learning; positive and negative feedback dataset; competitive network architecture; Long Short-Term Memory (LSTM) network; commodity recommendation

## 0　引言

云计算和大数据等网络技术的迅猛发展引发了网络信息的爆炸式增长,海量数据带来的"信息过载"问题使得人们对有价值信息的选择变得尤为困难,个性化推荐系统应运而生。以协同过滤推荐技术、基于内容的推荐技术和混合推荐技术为代表的传统推荐技术应用于电子商务推荐中仍存在数据稀疏、新用户冷启动、大数据处理与算法可扩性和特征识别差等问题[1-3],因此,研究人员开始尝试将深度学习引入推荐领域来解决上述问题,以提高模型的可用性和普适性。

Wang等[4]提出了一种基于协同深度学习(Collaborative Deep Learning, CDL)的推荐方法,该方法利用贝叶斯栈式降噪自编码器来学习商品内容的特征表示,并结合矩阵分解模型来预测用户的商品评分数据。该方法缓解了传统推荐技术中的数据稀疏问题,但其只考虑了显式反馈(对商品的评分数据)表达的用户对商品的喜好程度,而忽略了隐式反馈(对商品的点击、购买和略过等数据)表达的用户对商品的"不确定"的喜好程度。针对上述问题,研究人员对CDL模型进行了改进,Wei等[5]提出了融合TimeSVD++[6]和栈式降噪自编码器(Stacked Denoising AutoEncoder, SDAE)的混合推荐模型,其中TimeSVD++是一种可以融合时间感知的隐因子模型。与CDL相比,该模型不仅利用了隐式反馈包含的用户偏好信息,而且还可以捕获商品信息和用户偏好随时间的变化特征,解决用户

**收稿日期**:2019-11-25;**修回日期**:2020-01-12;**录用日期**:2020-01-19。
**基金项目**:教育部人文社会科学研究规划基金资助项目(17YJA870006)。
**作者简介**:傅魁(1977—),男,湖北武汉人,副教授,博士,主要研究方向:智能推荐、数据挖掘、量化投资;　梁少晴(1996—),男,安徽亳州人,硕士研究生,主要研究方向:智能推荐;　李冰(1983—),女,吉林通化人,副教授,博士,主要研究方向:特征选择、模式识别、复杂网络、智能规划。

偏好动态变化问题，提高推荐的精度与准确性。将深度学习应用到推荐领域最终提高了模型的可用性和普适性[7-12]，但是这些模型仍存在3个问题：首先没有对隐式反馈进行再次区分，将隐式反馈分为正反馈（对商品的点击、购买等行为）和负反馈（对商品的略过行为），准确表明用户对商品是喜爱还是无视的态度；其次都是利用用户历史数据中频繁出现的特征进行学习并推荐，导致推荐商品相似性极高，容易使用户感到疲倦；最后都只考虑了当下回报而忽略了未来可能存在的回报。

深度强化学习（Deep Reinforcement Learning，DRL）将深度学习的特征提取功能与强化学习的动态学习决策功能结合起来[13]，为复杂场景中大规模数据特征的自动提取带来了希望，因此一些研究人员开始将DRL应用到推荐领域[14-17]，并取得了不错的效果。但目前在笔者的知识范围内，将DRL应用到商品推荐领域的研究极少，而且现有的模型没有综合性解决用户偏好动态变化、正负反馈包含的用户对商品喜好的表达、未来回报率和推荐商品多样性等问题，忽略了各要素之间的联动影响。针对上述问题，本文构建了基于改进的深度Q网络（Improved Deep Q Network，IDQN）网络结构的商品推荐模型，该模型主要改进如下：

1）考虑正负反馈所代表的用户对商品喜好的表达和商品购买的时序性问题，结合竞争架构和长短期记忆（Long Short-Term Memory，LSTM）网络对深度Q网络（Deep Q Network，DQN）进行改进，设计了IDQN结构帮助系统更好地理解用户；

2）将DQN算法应用于商品推荐中，同时考虑模型的当下回报和未来回报，准确把握用户偏好的动态变化；

3）使用DBGD（Dueling Bandit Gradient Descent）作为模型的探索方法，在不影响推荐系统短期性能的同时，增加推荐商品的多样性；

4）充分利用隐式反馈（点击查看、添加购物车、购买和略过等）中包含的用户信息对模型进行优化和更新。

本文设计的IDQN在竞争架构的DQN基础之上进行改进，能够对值函数进行更快、更准确的估计。将状态和动作共同决定的值函数用LSTM结构代替卷积层结构，而由状态单独决定的值函数中卷积结构保持不变，可以很好地处理商品购买的时序性问题。根据正负反馈特征将同时基于状态和动作的值函数的输出拆分成两个部分，解决了正负反馈不均衡的问题，使正反馈数据不至于被负反馈数据淹没，合理利用正负反馈数据来对模型进行训练和更新。在构建回报函数时借鉴DDQN（Double Deep Q Network）算法中改进的目标Q值，消除了过高估计Q值的问题，考虑当下回报和未来回报仿真模拟用户偏好动态变化的过程。采用DBGD算法对模型的探索策略进行设计，避免了算法模型的过拟合，加快了模型的收敛和最优解的寻找速度，保证了系统的稳定性。

线下实验结果证明，基于IDQN结构的商品推荐模型的准确率、召回率、平均准确率（Mean Average Precision，MAP）和归一化折损累计增益（Normalized Discounted Cumulative Gain，NDCG）与经典模型中的最好表现相比，分别提高了69.8%、89.81%、95.00%、67.57%；线上实验结果还表明本文设计的DBGD探索函数能与用户进行最佳交互，使得推荐的商品相似性更低，更具有多样性。

## 1 DQN

随着DRL的不断发展，DQN算法的研究中也出现了很多经典的网络结构，本章首先以2013年Mnih等[18]第一次提出的DQN模型为例，对DQN结构进行分析，指出DQN结构用于商品推荐中的优缺点。

如图1所示，DQN结构除了输入层和输出层外，是由3个卷积层和2个全连接层构成的5层深度神经网络。

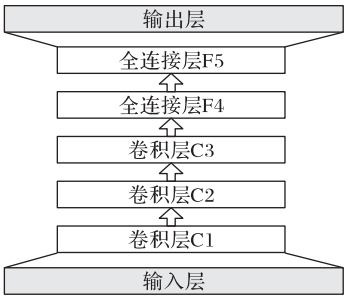



图1 DQN结构

Fig. 1 DQN structure

DQN在特征识别上取得了很好的效果，从DQN结构中可以发现传统DQN的优点如下：

1）采用局部感知和权值共享相结合的形式，大大减少了网络中需要进行训练的参数数量，使原本复杂的网络结构变得简单，同时加快了训练速度；

2）保留了卷积神经网络（Convolutional Neural Network，CNN）中的卷积层，而没有采用池化层，这样做的原因是为了使输入特征在位置上保持不变。

除上述优点外，将DQN应用于商品推荐中存在的问题如下：

1）DQN将CNN提取出的抽象特征经过全连接层后直接输出对应动作的Q值，默认状态动作值函数大小是与所有状态和动作都相关的，降低了智能体在策略评估过程中正确识别行为的响应速度。

2）DQN结构除了输入层和输出层外，采用了3个卷积层和2个全连接层构成的5层深度神经网络，然而CNN无法对时序性数据建模，因此DQN无法对时序性数据进行充分的信息挖掘。

3）DQN结构只能接受固定大小的数据输入，无法对正负反馈进行有效的区别性训练。

## 2 基于IDQN结构的商品推荐模型

本章中关于商品推荐问题可以定义为：假设用户u向推荐系统发出浏览商品的请求，推荐代理G收到请求后，将用户u的相关信息和待推荐商品池P输入模型中，根据模型算法选出一组top-$k$商品列表L进行推荐，用户u将对推荐列表L给予相关反馈B。表1对上述问题描述里和下文中将出现的符号进行定义。

下面将详细介绍基于IDQN结构的商品推荐模型与其他模型的不同之处，主要分为IDQN深度神经网络、模型回报函数的构建、探索策略的设计、模型整体框架与算法原理。

**表1 推荐模型符号定义**

Tab. 1 Definition of recommendation model symbols

| 符号 | 含义 |
|---|---|
| u | 用户 |
| G | 推荐代理（agent） |
| P | 待选商品池 |
| L | 推荐商品列表 |
| Q | 深度Q网络 |
| W | 深度Q网络的参数 |
| B | 用户反馈 |

### 2.1 IDQN深度神经网络

在对用户-商品交互数据的分析中有两点重要的发现：一

是用户负反馈能够在一定程度上帮助过滤用户不喜欢的商品;二是用户购买商品具有时序性特征。因此,首先根据用户–商品交互行为构建具有时序特征的正负反馈数据集;然后针对DQN自身存在的问题提出了使用收敛速度更快更准确的基于竞争架构的DQN结构,并针对用户购买商品时序性问题对其网络结构进行了改进,得到了改进的基于竞争架构的DQN结构;最后将用户正负反馈考虑到改进的基于竞争架构的DQN结构中,最终得到了融合用户正负反馈的改进的DQN结构模型。

### 2.1.1　基于用户正负反馈的用户–商品交互特征设计

**定义1**　用户正负反馈中,将当前的用户反馈表示为$s$。$s_+=\{i_1,i_2,\cdots,i_N\}$表示用户最近点击查看、添加购物车或购买过的$N$个商品特征集合,即用户正反馈信息的集合。$s_-=\{j_1,j_2,\cdots,j_M\}$表示用户最近略过的$M$个商品特征集合,即用户负反馈信息的集合。$s=(s_+,s_-)$,其中,$s_+$和$s_-$中添加商品的顺序是按照时间顺序排列的。

**定义2**　用户–商品交互情况中,当推荐系统将商品$a$在$s=(s_+,s_-)$的状态下推荐给用户时,如图2所示:若用户对推荐商品$a$的行为为略过,那么正反馈保持不变$s'_+=s_+$,同时更新负反馈$s'_-=\{j_1,j_2,\cdots,j_M,a\}$;若用户对商品的行为为点击查看、添加购物车或购买,那么负反馈保持不变$s'_-=s_-$,同时更新正反馈$s'_+=\{i_1,i_2,\cdots,i_N,a\}$;此时的用户–商品交互特征表示为$s'=(s'_+,s'_-)$。

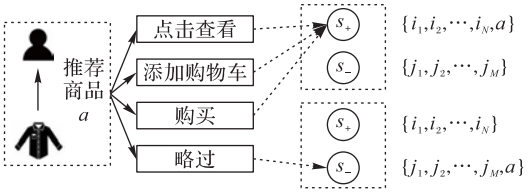



图2　正负反馈数据分类

Fig. 2　Classification of positive and negative feedback data

### 2.1.2　面向商品购买时序性的DQN结构

在商品推荐过程的某些状态下,值函数的大小与动作无关。针对这一问题本文采用一种基于竞争架构的DQN,竞争网络(如图3所示)是将CNN中卷积层提取的抽象特征进行分流:一条分流是只依赖于状态的值函数,即状态价值函数;另一条分流代表同时依赖于状态和动作的值函数,即动作优势函数。实验表明,当智能体在一定策略下不断采取不同行为,但对应函数值却相同的情况下,基于竞争架构的DQN模型能够对值函数进行更快、更准确的估计。

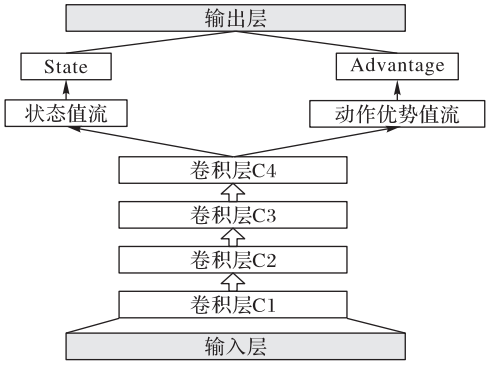



图3　基于竞争架构的DQN结构

Fig. 3　Structure of DQN based on competitive architecture

**定义3**　竞争网络优势评估函数为:

$$A^{\pi}(s,a)=Q^{\pi}(s,a)-V^{\pi}(s) \tag{1}$$

其中:$Q^{\pi}(s,a)$为状态动作值函数,表示在状态$s$下根据策略$\pi$选择动作$a$所获的期望回报值;$V^{\pi}(s)$为状态价值函数,表示状态$s$下根据策略$\pi$产生的所有动作的价值的期望值;$A^{\pi}(s,a)$表示状态$s$下选择动作$a$的优势。

**定义4**　竞争网络输出值函数为:

$$Q(s,a;\ \theta,\alpha,\beta)=V(s;\ \theta,\beta)+A(s,a;\ \theta,\alpha) \tag{2}$$

其中:$V(s;\ \theta,\beta)$表示输出状态价值函数;$A(s,a;\ \theta,\alpha)$表示输出动作优势函数;$\theta$、$\alpha$、$\beta$分别表示对输入层进行特征处理的网络神经元参数以及状态价值函数和状态函数的参数。

由于用户的商品购买行为具有一定的时序性,针对这一特征本文对基于竞争架构的DQN结构进行了以下改进:

1)在基于竞争架构的DQN结构中由于CNN并不能对时序数据进行处理,而LSTM在时序数据的处理上表现出了较好的效果,因此将卷积层换成LSTM结构。

2)商品推荐模型的输入数据主要包括用户特征、上下文特征、商品特征和用户–商品交互特征,在状态$s$下选择动作$a$的回报总和与所有输入特征相关,但是用户自身特征具有的价值由用户特征和上下文特征单独决定,因此改进的模型中将状态和动作共同决定的值函数用LSTM结构代替CNN中的卷积层结构,而由状态单独决定的值函数中卷积结构保持不变,改进后的模型结构如图4所示。

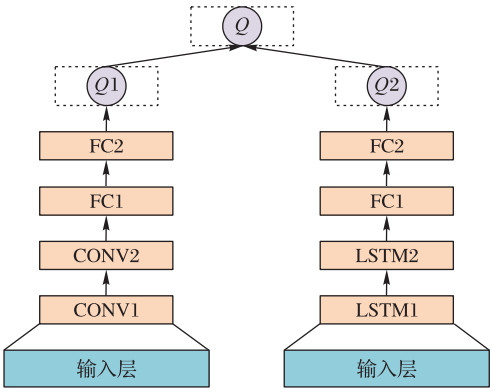



图4　面向商品购买时序性的DQN结构

Fig. 4　DQN structure for commodity purchase time-series feature

### 2.1.3　IDQN深度神经网络结构

本节将用户正负反馈考虑到改进的基于竞争架构的DQN结构中,最终得到了如图5所示融合用户正负反馈的改进的DQN结构(即IDQN结构)。

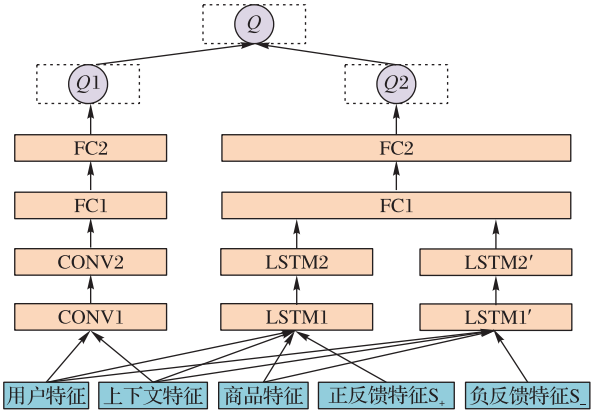



图5　IDQN结构

Fig. 5　IDQN structure

如图5所示,该结构根据正负反馈特征将同时基于状态

和动作的值函数的输入拆分成两部分，将除负反馈特征外的所有特征输入LSTM1层中，将除正反馈特征外的所有特征输入LSTM1′层中，分别经过LSTM2和LSTM2′后，一起进入全连接层FC1和FC2，最终输出状态动作值函数$Q2$。用户特征和上下文特征更加体现用户本身的价值，因此单独放进基于状态的动作优势值函数中，通过两层卷积层和两层全连接层后输出动作优势值函数$Q1$。

**定义5** IDQN最终值函数为：

$$Q(s,a;\ \theta,\alpha,\beta)=Q1(s;\ \theta,\beta)+Q2(s,a;\ \theta,\alpha) \tag{3}$$

其中：$V(s;\ \theta,\beta)$表示状态动作值函数；$A(s,a;\ \theta,\alpha)$表示动作优势值函数；$s$表示当前状态，$a$表示在状态$s$下的动作选择，$\theta$、$\alpha$、$\beta$分别代表状态动作值函数、动作优势值函数的参数。状态动作值函数和动作优势值函数的结合是通过聚合操作进行的。

### 2.2 模型的回报函数构建

大量研究表明，用户购买行为的偏好处于动态变化之中。为了提高模型的准确率，本文在回报函数构建时不仅考虑当下回报，同时考虑未来回报。

**定义6** 在状态$s$下遵循策略$\pi$直到情况结束，推荐代理G累积获得的回报函数为：

$$Q(s,a)=r_{\text{immediate}}+\gamma r_{\text{future}}=r_{t+1}+\gamma r_{t+2}+\gamma^2 r_{t+3}+\cdots \tag{4}$$

其中：状态$s$主要由用户特征和上下文特征来表示，动作$a$主要由商品特征、商品-用户交互特征共同决定，$r_{\text{immediate}}$表示当下回报，$r_{\text{future}}$表示未来回报，$\gamma\in[0,1]$用来衡量未来回报对累积奖赏的影响。

在DQN模型中，使用了一个独立的网络来对目标Q值进行计算，容易引起学习过程中过高估计Q值的问题。本文采取DDQN算法中改进的目标Q值，使用两套参数对Q网络值进行训练和学习：$W$和$W^-$，其中$W$用来对最大Q值对应的动作进行选择，$W^-$用来计算最优动作所对应的Q值。$W$和$W^-$两套参数的引入，将策略评估和动作选择分离开，使过高估计Q值的问题得到了缓解，DDQN算法的目标Q值推导过程如下：

$$\begin{aligned}Q(s,a,t)&=r_{a,t+1}+\gamma r_{a,t+2}+\gamma^2 r_{a,t+3}+\cdots=\\&r_{a,t+1}+\gamma(r_{a,t+2}+\gamma r_{a,t+3}+\cdots)=\\&r_{a,t+1}+\gamma Q\left(s_{a,t+1},\arg\max_{a'}Q(s_{a,t+1},a';\ W_t);\ W_t^-\right)\end{aligned} \tag{5}$$

其中：$r_{a,t+1}$表示推荐代理G选择动作a时的当下回报，$W_t$和$W_t^-$表示不同的两组参数，在这个公式中，推荐代理G将根据给定的动作$a$推测下一状态$s_{a,t+1}$。基于此，给定一组候选动作$\{a'\}$，根据参数$W_t$选择给出最大未来回报的动作$a'$。在此之后，基于$W^-$计算给定状态$s_{a,t+1}$的预计未来回报。每隔一段时间，$W_t$和$W_t^-$将进行参数交换，通过这一过程，该模型消除了过高估计Q值的问题，并能够做出同时考虑当下和未来回报的决策。

网络参数的更新主要是通过最小化当前网络Q值和目标网络Q值之间的均方误差来进行的，误差函数如下：

$$L(W_t)=E_{s,a,r,s'}[(y_{s,a,t}-Q(s,a|W_t))^2] \tag{6}$$

### 2.3 探索策略设计

强化学习主要以动态试错机制不断与环境进行交互，学习如何获得最优行为策略。因此，在与环境的交互过程中，agent不仅需要考虑值函数最大的动作，即利用(Exploitation)，还需要尽可能多地选择不同的动作，以找到最优的策略，即探索(Exploration)。目前主要有三种探索策略被应用于强化学习中，分别是$\varepsilon$-greedy算法、Boltzmanm算法和DBGD算法。其中DBGD算法将原参数保持不变，在原参数的基础上进行微小的变动获得新的参数，通过新参数和原参数推荐效果的比较，对原参数进行更新，既提高了算法的收敛速度，又保证了系统的稳定性。因此，本文主要采用DBGD算法对探索策略进行设计。

在基于DBGD算法的探索策略设计中，推荐代理G将使用Exploitation网络生成推荐列表$L$，同时使用Exploration网络生成推荐列表$L'$，然后将$L$和$L'$中推荐概率最高的前50%的商品分别取出交错排列为用户进行推荐[19]，同时获得用户反馈。若用户反馈表示Exploration网络生成的推荐商品更符合用户心意，则Exploitation网络的网络参数向Exploration网络参数方向更新，若用户反馈表示Exploitation网络生成的推荐商品更符合用户心意，则Exploitation网络的网络参数保持不变。

**定义7** Exploration网络的参数表示公式如下：

$$\tilde{W}=\Delta W+W \tag{7}$$

其中：$\Delta W=\alpha\cdot\text{rand}(-1,1)\cdot W$，$|\Delta W|$越大，表示探索程度越大，$\alpha$为探索系数，rand(-1,1)表示从-1~1随机取一个参数，$W$表示当前网络参数。

**定义8** Exploitation网络更新公式如下：

$$W'=W+\beta\tilde{W} \tag{8}$$

其中：$\beta$表示更新系数。采用DBGD算法对模型中的探索策略进行设计，避免了一般探索过程中短期时间内推荐模型性能下降的问题，将探索过程向好的方向引导，加快了模型的收敛和最优解的寻找速度。

### 2.4 模型的整体框架与算法构建

结合上述研究方法以及本文的研究思路，提出了图6所示的基于IDQN结构的商品推荐模型的框架。

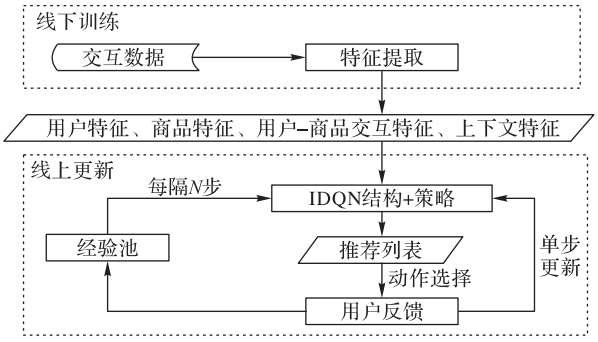



图6 系统框架

Fig. 6 System framework

如图6所示，该框架包括两个部分：线下训练模块和线上更新模块。线下训练模块主要利用用户和商品间的交互日志训练得到离线模型，交互日志内容包括用户对商品的点击、购买等行为。线上更新模块主要对前期训练的网络进行更新。具体交互流程如下：

1)输入：模型的输入主要是用户特征、商品特征、用户-商品交互特征和上下文特征。

2)策略：模型的策略部分主要采用了DQN算法，同时采用DBGD方法作为算法模型的探索，模型的网络结构为IDQN结构。

3)输出：当用户u向系统发出浏览商品请求时，将用户u的特征和待推荐商品池P中待推荐商品的特征输入到推荐代理G中，推荐代理G将根据输入信息生成一个top-$k$商品推荐列表L。

4)用户反馈：当用户u接收到推荐列表L的时候，会对L中的商品做出反馈，得到反馈结果B。

5)模型的单步更新：在每一步后，用户u的特征集、生成的推荐列表L、用户u对推荐列表L的反馈B，生成数据集{u，L，B}。推荐代理G将会根据主要推荐网络Q和基于探索的推荐网络$Q^-$的表现情况进行模型的更新。

6)模型的多步更新：模型采用了经验回放技术，每隔$N$步推荐代理G将会根据之前存储在经验池中的数据来更新主要推荐网络Q，多步更新主要是为了减少样本间的相关性，提高

模型训练的准确率。

7)重复进行1)~6)的过程。

在IDQN结构的基础上,使用DQN算法,结合其经验回放技术,构建如下基于IDQN结构的商品推荐算法:

1) 初始化经验池,设定容量为$N$;
2) 初始化当前值网络,随机生成网络参数$W$;
3) 初始化目标值网络,使目标值网络的网络参数$W^- = W$;
4) Loop ($m = 1, 2, \cdots, M$)
5)　对状态$s_t$进行初始化;
6)　Loop ($t = 1, 2, \cdots, T$)
7)　　用主要推荐策略和基于探索的推荐策略结合生成动作$a_t$;
8)　　用户收到动作$a_t$,同时给出反馈$r_t$及产生新的状态$s_{t+1}$;
9)　　样本$(s_t, a_t, r_t, s_{t+1})$存储在经验池中;
10)　　从回放记忆库中随机选取一个minibatch的样本进行训练,表示为$(s_j, a_j, r_j, s_{j+1})$,若$j$为最后一步,则回报函数按照$r_{a,j+1}$计算,若$j$不为最后一步,回报函数为$r_{a,j+1} + \gamma Q(s_{a,j+1}, \arg\max Q(s_{a,j+1}, a'; W_t); W_t^-)$;
11)　　对误差函数$(y_{s,a,t} - Q(s, a|W_t))^2$关于$W_t$使用梯度下降法进行更新;
12)　　每隔$N$步更新目标值网络,将当前网络的参数复制给目标网络,即使$W^- = W$;
13)　End loop
14)End loop

## 3　实验设计与结果分析

### 3.1　实验数据描述

本文实验数据分为线下实验数据和线上实验数据。线下实验数据主要使用Retailrocket推荐系统数据集(Kaggle网站顶级数据集),该数据集采集了真实电子商务交易网站中的推荐数据;线上实验数据主要在“什么值得买”app上进行采集(下文统一用“线上推荐数据集”表示)。

经过数据预处理后,Retailrocket推荐系统数据集中可用数据如表2所示。为了模拟真实的商品推荐过程,在线下训练数据中,对于每个用户,将其购买记录按照购买时间排序,取前80%作为训练集,后20%作为测试集。

**表2　清洗后数据集统计表**

Tab. 2　Statistics of datasets after cleaning

| 数据集 | 用户数 | 商品数 | 访问记录数 |
|---|---|---|---|
| 线下数据 | 341 032 | 289 369 | 862 764 |
| 线上数据 | 32 786 | 25 681 | 76 420 |

下面将分别对Retailrocket推荐系统数据集和线上推荐数据集中的数据按照用户请求访问推荐商品的次数、商品被推荐的次数、用户与商品交互时间进行统计和分析。

1)用户请求访问推荐商品的次数和商品被推荐的次数统计。

将上述数据进行统计后可以得到每个用户请求访问推荐商品的次数和每个商品被推荐的次数,如图7所示。

如图7为用户和商品的基本数据统计图,通过对图7观察发现,这两组数据集均呈现倾斜状态,说明用户访问商品的次数具有长尾分布特征,即大部分用户访问次数少于500,而每个商品被推荐的次数也存在长尾分布特征,大部分商品被推荐的次数少于100。

2)用户与商品交互时间统计。

如图8所示,图(a)和图(b)分别为Retailrocket推荐系统数据集和线上推荐数据集中用户和商品交互时间统计图,其中,0:00到6:00点用户行为发生次数呈下降趋势,7:00到16:00呈上升趋势,17:00到24:00首先出现下降趋势,然后经过一个小的波动后趋于平稳,这一趋势基本符合正常人的作息时间。

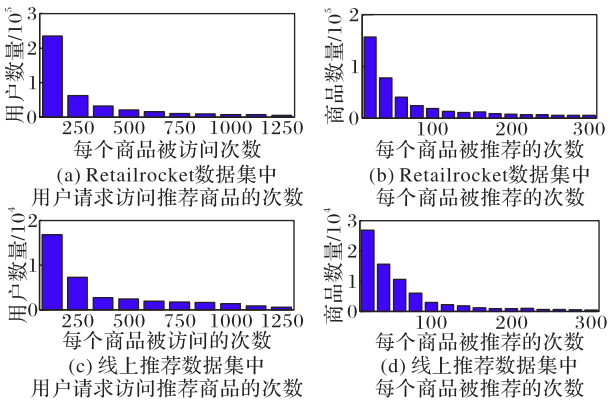



图7　用户和商品基本数据统计

Fig. 7　Basic data statistics of users and commodities

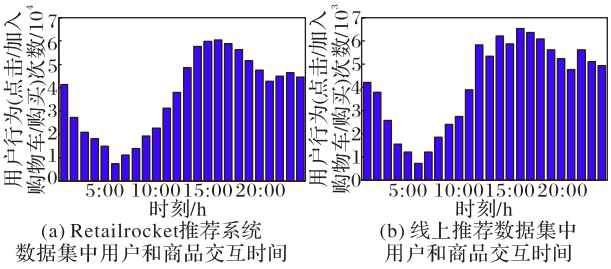



图8　用户和商品交互时间统计

Fig. 8　Interaction time statistics of users and commodities

### 3.2　实验方案设计与评价指标

#### 3.2.1　对比基准模型

为了验证本文所提出的基于IDQN结构的商品推荐模型在推荐精度和商品多样性等方面优于目前已有的优秀的线上推荐模型,本文选取因子分解机(Factorization Machine, FM)模型[20]、W&D (Wide & Deep learning)模型[21]和协同过滤(Collaborative Filtering, CF)模型作为对照模型。

#### 3.2.2　模型评价指标

为了对比各模型的推荐效果,本文分别选取了准确率(Precision)[22]、召回率(Recall)[22]、MAP[22]、NDCG[23]和商品多样性(Intra-list Similarity, ILS)[24]这5组指标作为模型的评估标准。其中模型是针对341 032个用户分别进行推荐预测,得到的指标值是所有用户预测结果的平均值。

#### 3.2.3　实验方案设计

本文主要设置了1个实验组和12个对照组,具体内容和区别如表3所示。在网络结构中T-DQN表示传统DQN结构,DN表示基于竞争架构的DQN结构,NF表示考虑用户负反馈的DQN结构,PT表示考虑用户购买时序性的DQN结构,在探索函数中EG指$\varepsilon$-greedy算法,BM指Boltzmanm算法。

其中,实验组完全按照本文基于IDQN结构的商品推荐模型设计思路进行。对照实验共设置了12组,第1组到第7组主要是为了测试本文在传统DQN结构的基础上进行的3个方面的改进是否使推荐的准确率、召回率、MAP和NDCG得到提升,其中这3个方面的改进分别为DN、NF和PT;第8组和第9组主要为了测试DBGD探索策略性能的优劣,分别采用EG和BM这两种常用的探索策略作为对照,模型的评价指标除了准确率、召回率、MAP、NDCG之外,更重要的是商品推荐多样性是否有所增强;第10组~第12组为对比基准模型,用于验证本文提出的模型是否优于这些推荐领域中的经典模型。

表3 实验方案设计

Tab. 3 Experimental scheme design

| 实验名称 | 网络结构 | | | | 探索函数 | | | 基准模型 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | T-DQN | DN | NF | PT | EG | BM | DBGD | FM | W&D | CF |
| 实验组(IDQN) | × | √ | √ | √ | × | × | √ | × | × | × |
| 对照组1 | √ | × | × | × | × | × | √ | × | × | × |
| 对照组2 | × | √ | × | × | × | × | √ | × | × | × |
| 对照组3 | × | × | √ | × | × | × | √ | × | × | × |
| 对照组4 | × | × | × | √ | × | × | √ | × | × | × |
| 对照组5 | × | √ | √ | × | × | × | √ | × | × | × |
| 对照组6 | × | √ | × | √ | × | × | √ | × | × | × |
| 对照组7 | × | × | √ | √ | × | × | √ | × | × | × |
| 对照组8 | × | √ | √ | √ | √ | × | × | × | × | × |
| 对照组9 | × | √ | √ | √ | × | √ | × | × | × | × |
| 对照组10 | × | × | × | × | × | × | × | √ | × | × |
| 对照组11 | × | × | × | × | × | × | × | × | √ | × |
| 对照组12 | × | × | × | × | × | × | × | × | × | √ |

### 3.3 实验结果分析

#### 3.3.1 实验设置

本文采用Grid Search方法来确定模型的参数,从而找到准确率最高的参数组合,表4是通过网格搜索法确定的最优参数组合。

#### 3.3.2 模型评价指标

线下实验主要是依据离线数据进行的,离线数据是静态的,无法对探索策略的性能进行测试,因此在线下实验中不考虑探索策略对推荐商品多样性的影响,只考虑不同模型在Precision、Recall、MAP和NDCG上的区别。

本文对实验设计方案中的1个实验组和12个对照组分别进行了线下实验,实验数据结果如表5所示,实验的图形展示如图9所示。

表4 参数设置表

Tab. 4 Parameter setting

| 实验参数 | 值 |
|---|---|
| 未来回报折扣因子$\lambda$ | 0.4 |
| 探索策略系数$\alpha$ | 0.1 |
| 利用策略系数$\beta$ | 0.05 |
| 模型单步更新时间$T_D$ | 60 min |
| 模型多步更新时间$T_R$ | 30 min |
| 最近邻居数量$k$值 | 5,10,15,20 |

表5 线下推荐实验的推荐效果

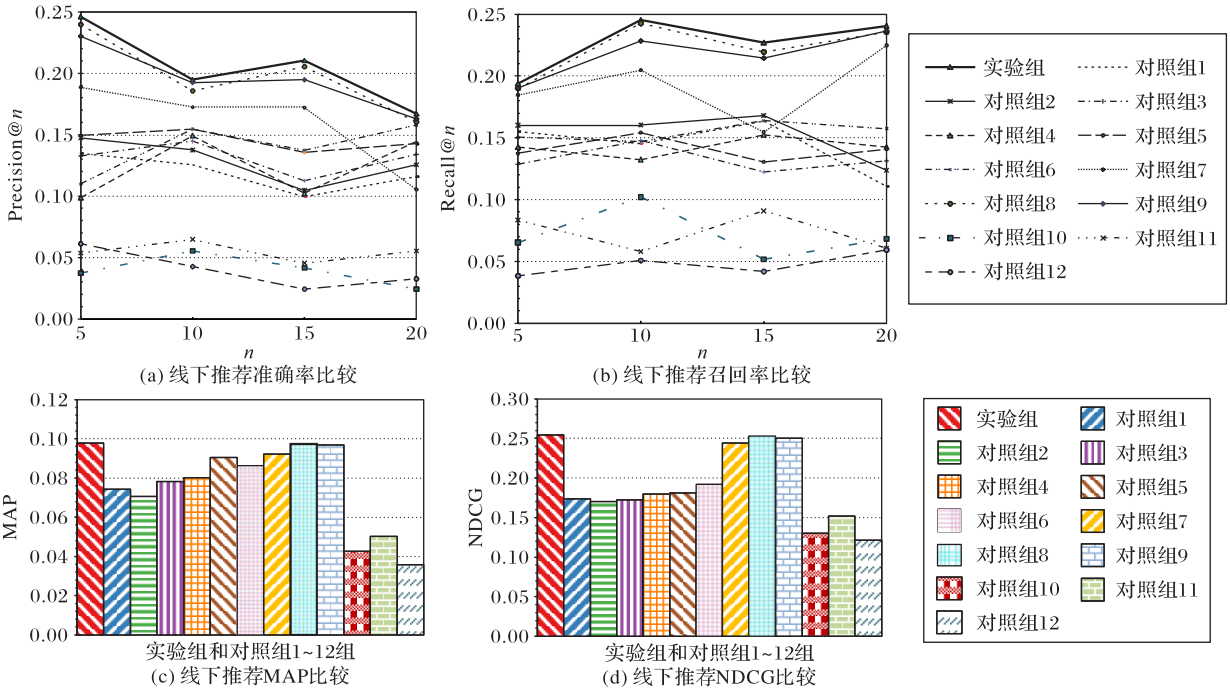
Tab. 5 Recommendation effects of offline recommendation experiments

| 实验名称 | Precision@$K$ | | | | Recall@$K$ | | | | MAP | NDCG |
|---|---|---|---|---|---|---|---|---|---|---|
| | $K$=5 | $K$=10 | $K$=15 | $K$=20 | $K$=5 | $K$=10 | $K$=15 | $K$=20 | | |
| 实验(IDQN) | 0.2461 | 0.1948 | 0.2104 | **0.1672** | **0.1938** | 0.2452 | 0.2271 | 0.2406 | **0.0977** | **0.2542** |
| 对照组1 | 0.1349 | 0.1254 | 0.0994 | 0.1157 | 0.1549 | 0.1452 | 0.1637 | **0.1109** | 0.0742 | 0.1738 |
| 对照组2 | 0.1474 | 0.1375 | 0.1047 | 0.1254 | 0.1601 | 0.1603 | 0.1679 | 0.1237 | **0.0706** | **0.1702** |
| 对照组3 | 0.1328 | 0.1447 | 0.1125 | 0.1340 | 0.1287 | 0.1478 | 0.1223 | 0.1315 | 0.0783 | 0.1721 |
| 对照组4 | **0.0986** | 0.1492 | 0.1022 | 0.1445 | 0.1424 | 0.1324 | 0.1524 | 0.1424 | 0.0801 | 0.1794 |
| 对照组5 | 0.1493 | 0.1545 | 0.1354 | 0.1427 | 0.1374 | 0.1542 | 0.1305 | 0.1409 | 0.0904 | 0.1813 |
| 对照组6 | 0.1098 | 0.1547 | 0.1374 | 0.1578 | 0.1505 | 0.1471 | 0.1638 | 0.1573 | 0.0863 | 0.1921 |
| 对照组7 | 0.1887 | 0.1724 | 0.1720 | 0.1054 | 0.1847 | 0.2047 | 0.1547 | 0.2247 | 0.0922 | 0.2439 |
| 对照组8 | 0.2397 | 0.1857 | 0.2054 | 0.1604 | 0.1901 | 0.2425 | 0.2195 | 0.2356 | 0.0974 | 0.2523 |
| 对照组9 | 0.2303 | 0.1924 | 0.1947 | 0.1628 | 0.1902 | 0.2283 | 0.2145 | 0.2363 | 0.0968 | 0.2499 |
| 对照组10 | 0.0372 | 0.0554 | 0.0414 | 0.0241 | 0.0654 | **0.1021** | 0.0519 | 0.0684 | 0.0427 | 0.1301 |
| 对照组11 | 0.0537 | **0.0647** | 0.0454 | 0.0554 | 0.0835 | 0.0581 | 0.0909 | 0.0608 | **0.0501** | **0.1517** |
| 对照组12 | 0.0611 | 0.0425 | 0.0241 | 0.0325 | 0.0384 | 0.0507 | 0.0418 | 0.0594 | 0.0357 | 0.1211 |

实验结果表明,实验组的推荐效果在整体上明显优于其余12个对照组,证明基于IDQN结构的商品推荐模型具有更好的推荐效果,其中,Precision@5推荐准确率最高,Recall@10召回率最高。在推荐准确率上,实验组和对照组1~7中表现最差的为Precision@5中的对照组4,推荐准确率为0.0986,在经典推荐模型中表现最好的为Precision@10中的W&D,推荐准确率为0.0647,准确率提高了52.40%,本文提出的模型即实验组,推荐准确率在Precision@20中表现最差,推荐准确率为0.1672,与W&D相比,推荐准确率提高了158.42%;在推荐召回率上,实验组和对照组1~7中表现最差的为Recall@20中的对照组1,推荐召回率为0.1109,在经典推荐模型中表现最好的为Recall@10中的FM,推荐召回率为0.1021,召回率提高了8.62%,实验组推荐召回率表现最差的Recall@5,推荐召回率为0.1938,与W&D相比,推荐召回率提高了89.81%;在推荐MAP值上,实验组和对照组1~7中表现最差的为对照组2,MAP值为0.0706,在经典推荐模型中表现最好的为W&D,MAP值为0.0501,MAP值提高了40.92%,本文提出的模型即实验组,MAP值为0.0977,与W&D相比,MAP值提高了95.00%;在NDCG值上,实验组和对照组1~7中表现最差的为对照组2,NDCG值为0.1702,在经典推荐模型中表现最好的为W&D,NDCG值为0.1517,NDCG值提高了12.20%,本文提出的模型即实验组,NDCG值为0.2542,与W&D相比,NDCG值提高了67.57%。

综上可以发现,在推荐准确率、召回率、MAP和NDCG上,

实验组和对照组1~7中表现最差的与经典模型中表现最好的相比,精度分别提高了52.40%、8.62%、40.92%、12.20%,证明了将DQN模型应用于商品推荐中的有效性和可行性,将本文提出的模型与经典模型中表现最好的相比,精度分别提高了158.42%、89.81%、95.00%、67.57%,验证了本文提出的模型在商品推荐中具有更好的推荐效果。




(a) 线下推荐准确率比较　(b) 线下推荐召回率比较

(c) 线下推荐MAP比较　(d) 线下推荐NDCG比较

图9　线下推荐实验结果比较

Fig. 9　Comparison of offline recommendation experimental results

### 3.3.3　线上实验及结果分析

线上实验部分主要是将该模型放到电子商务推荐平台上,进行一定时长的线上测试。在线上实验中不仅要考虑推荐效果(准确率/召回率/MAP/NDCG),更重要的是要考虑商品推荐的多样性。本文设计的基于DBGD算法的探索策略,能够通过这一策略为用户推荐新颖且感兴趣的商品,而推荐效果和商品多样性这两个评价指标能够较好地反映这一问题。

1)推荐效果。

本文对实验设计方案中的1个实验组和12个对照组分别进行了线上实验,实验数据结果如表6所示,实验的图形展示如图10所示。

**表6　线上推荐实验推荐效果**

Tab. 6　Recommendation effects of online recommendation experiments

| 实验名称 | Precision | | | | Recall | | | | MAP | NDCG |
|---|---|---|---|---|---|---|---|---|---|---|
| | *K*=5 | *K*=10 | *K*=15 | *K*=20 | *K*=5 | *K*=10 | *K*=15 | *K*=20 | | |
| 实验(IDQN) | 0.1392 | 0.1587 | 0.1247 | 0.1654 | 0.1724 | 0.1952 | 0.1625 | 0.1741 | 0.0804 | 0.1973 |
| 对照组1 | 0.0731 | 0.0815 | 0.0541 | 0.0784 | 0.0747 | 0.0854 | 0.1082 | 0.0673 | 0.0432 | 0.1642 |
| 对照组2 | 0.0828 | 0.0858 | 0.0554 | 0.0805 | 0.0869 | 0.0976 | 0.1125 | 0.0749 | 0.0621 | 0.1327 |
| 对照组3 | 0.0889 | 0.0787 | 0.0443 | 0.0547 | 0.0744 | 0.0959 | 0.1023 | 0.0954 | 0.0617 | 0.1436 |
| 对照组4 | 0.0601 | 0.0592 | 0.0557 | 0.0959 | 0.1082 | 0.0645 | 0.1223 | 0.0925 | 0.0588 | 0.1224 |
| 对照组5 | 0.0982 | 0.0896 | 0.0641 | 0.0505 | 0.0838 | 0.1008 | 0.0907 | 0.1097 | 0.0609 | 0.1209 |
| 对照组6 | 0.0875 | 0.0682 | 0.0715 | 0.0920 | 0.1106 | 0.0776 | 0.1132 | 0.1089 | 0.0493 | 0.1137 |
| 对照组7 | 0.1004 | 0.0969 | 0.0774 | 0.0967 | 0.1143 | 0.0893 | 0.1059 | 0.1178 | 0.0570 | 0.1428 |
| 对照组8 | 0.1028 | 0.1198 | 0.0864 | 0.1157 | 0.1347 | 0.1237 | 0.1345 | 0.1273 | 0.0421 | 0.1234 |
| 对照组9 | 0.1047 | 0.1002 | 0.0954 | 0.1246 | 0.1493 | 0.1152 | 0.1423 | 0.1134 | 0.0328 | 0.1132 |
| 对照组10 | 0.0163 | 0.0147 | 0.0091 | 0.0212 | 0.0221 | 0.0127 | 0.0096 | 0.0279 | 0.0057 | 0.0251 |
| 对照组11 | 0.0256 | 0.0087 | 0.0178 | 0.0025 | 0.0204 | 0.0273 | 0.0107 | 0.0078 | 0.0051 | 0.0416 |
| 对照组12 | 0.0149 | 0.0297 | 0.0115 | 0.0075 | 0.0357 | 0.0247 | 0.0149 | 0.0179 | 0.0021 | 0.0318 |

实验结果表明,实验组的推荐效果在整体上明显优于其余12个对照组,证明本文提出的基于IDQN的网络结构的商品推荐模型具有更好的推荐效果。根据实验设计方案得知,实验组、对照组8和9分别使用DBGD、EG、BM作为探索函数,在离线实验环境下,由于候选商品的集合有限,无法充分利用探索算法与用户进行最佳的交互,而在线上推荐中可以明显看出实验组相较于对照组8~9具有更好的推荐效果,因此验证了本文设计的DBGD探索函数的可行性和优越性。

2)商品多样性。

本文分别对1个实验组和12个对照组进行了线上测试,得出了推荐商品多样性的结果,商品多样性由指标ILS表示,而ILS主要用来衡量推荐商品之间的相似性,因此ILS值越小,表明推荐商品相似性越低,即推荐的商品更具多样性。实验数据结果如表7所示,实验的图形展示如图11所示。

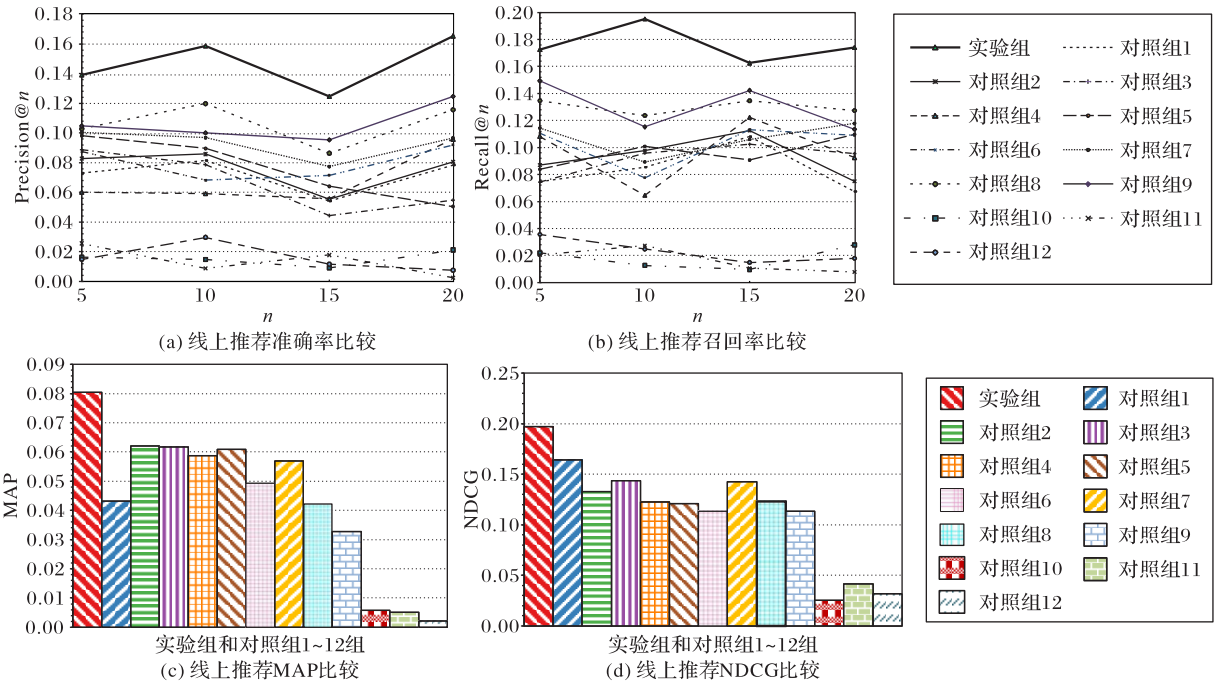



图 10 线上推荐实验结果比较

Fig. 10 Comparison of online recommendation experimental results

**表 7 线上推荐实验商品多样性**

Tab. 7 Commodity diversity of online recommendation experiments

| 实验名称 | ILS |
|---|---|
| 实验组 | 0.1195 |
| 对照组 1 | 0.1258 |
| 对照组 2 | 0.1346 |
| 对照组 3 | 0.1147 |
| 对照组 4 | 0.1395 |
| 对照组 5 | 0.1281 |
| 对照组 6 | 0.1357 |
| 对照组 7 | 0.1735 |
| 对照组 8 | 0.2820 |
| 对照组 9 | 0.2654 |
| 对照组 10 | 0.3781 |
| 对照组 11 | 0.3590 |
| 对照组 12 | 0.4418 |

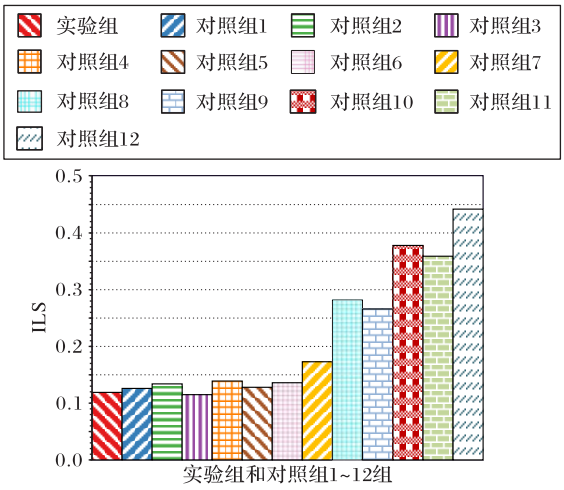



图 11 推荐商品多样性

Fig. 11 Diversity of recommended commodities

其中实验组和对照组 1~7 使用 DBGD 作为探索函数，对照组 8 采用 EG 作为探索函数，对照组 9 使用 BM 作为探索函数。从推荐商品多样性的结果中可以看出，实验组和对照组 1~7 的 ILS 值明显低于对照组 8~9，同时远远低于对比基准模型，表明使用本文提出的 DBGD 算法作为商品推荐模型的探索函数增加了商品推荐的多样性。

## 4 结语

本文在前人研究的基础上，针对商品推荐中存在的用户正负反馈问题和商品购买时序性问题，对传统 DQN 模型的深度神经网络结构进行分析和改进，构建了一个基于 IDQN 结构的商品推荐模型，该模型针对用户兴趣动态变化问题使用强化学习的试错机制进行在线学习，学习以最大化智能体从环境中获得的累积回报为目标，同时采用“利用+探索”的策略对商品进行推荐，对比实验结果表明，本文提出的模型无论是在推荐效果还是在推荐商品多样性上都优于现有的推荐模型。

本文首次尝试将改进的 DQN 应用于商品推荐领域，同时对探索函数进行了针对性改进，增加了算法的稳定性，使推荐效果有了较大提高。但是由于时间和精力有限，本文在研究中还存在以下四个方面的缺点和不足：1）实验数据量不足，商品-用户数据较少；2）线下实验数据集单一，只有一个 Retailrocket 推荐系统数据集，需要扩充数据集；3）线上实验时间不足，由于推荐平台的限制，本文线上实验时间仅为两周；4）在用户反馈中没有将用户行为进行区分，一般来说，略过、点击查看、加入购物车和购买依次表现了用户对商品喜好程度的增加，而本文在用户反馈中没有对用户行为进行区分。

**参考文献（References）**

[1] 盈艳，曹妍，牟向伟. 基于项目评分预测的混合式协同过滤推荐[J]. 现代图书情报技术，2015，31(6)：27-32.（YING Y，CAO Y，MU X W. A hybrid collaborative filtering recommender based on item rating prediction [J]. New Technology of Library and Information Service，2015，31(6)：27-32.）

[2] 李清霞，魏文红，蔡昭权. 混合用户和项目协同过滤的电子商务

个性化推荐算法[J]. 中山大学学报(自然科学版), 2016, 55(5): 37-42. (LI Q X, WEI W H, CAI Z Q. Hybrid user and item based collaborative filtering personalized recommendation algorithm in Ecommerce [J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2016, 55(5): 37-42.)

[3] 欧阳龙,卢琪,彭艳兵. 基于内容和背景的微博问答问题推荐[J]. 电子设计工程, 2018, 26(11): 183-188. (OUYANG L, LU Q, PENG Y B. Question recommendation of microblog QA based on content and background [J]. Electronic Design Engineering, 2018, 26(11): 183-188.)

[4] WANG H, WANG N, YEUNG D Y. Collaborative deep learning for recommender systems [C]// Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2015:1235-1244.

[5] WEI J, HE J, CHEN K, et al. Collaborative filtering and deep learning based recommendation system for cold start items [J]. Expert Systems with Applications, 2016, 69:29-39.

[6] KOREN Y. Collaborative filtering with temporal dynamics [C]// Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2009: 447-456.

[7] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation[C]// Proceedings of the 10th ACM International Conference on Web Search and Data Mining. New York: ACM, 2017: 425-434.

[8] COVINGTON P, ADAMS J, SARGIN E. Deep neural networks for YouTube recommendations [C]// Proceedings of the 10th ACM Conference on Recommender Systems. New York: ACM, 2016: 191-198.

[9] KIM D, PARK C, OH J, et al. Convolutional matrix factorization for document context-aware recommendation [C]// Proceedings of the 10th ACM Conference on Recommender Systems. New York: ACM, 2016: 233-240.

[10] 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. (HUANG L W, JIANG B T, LV S Y, et al. Survey on deep learning based recommender systems [J]. Chinese Journal of Computers, 2018, 41(7): 1619-1647.)

[11] 张家精,夏巽鹏,陈金兰,等. 基于张量分解和深度学习的混合推荐算法[J]. 南京大学学报(自然科学版), 2019, 55(6):952-959. (ZHANG J J, XIA X P, CHEN J L, et al. Blending recommendation algorithm based on tensor decompositions and deep learning [J]. Journal of Nanjing University (Natural Science), 2019, 55(6):952-959.)

[12] 张敏军,华庆一,贾伟,等. 基于深度神经网络的个性化推荐系统研究[J]. 西南大学学报(自然科学版), 2019, 41(11): 104-109. (ZHANG M J, HUA Q Y, JIA W, et al. A personalized recommendation system based on deep neural network [J]. Journal of Southwest University (Natural Science Edition), 2019, 41(11): 104-109.)

[13] 万里鹏,兰旭光,张翰博,等. 深度强化学习理论及其应用综述[J]. 模式识别与人工智能, 2019, 32(1):67-81. (WAN L P, LAN X G, ZHANG H B, et al. A review of deep reinforcement learning theory and application [J]. Pattern Recognition and Artificial Intelligence, 2019, 32(1):67-81.)

[14] ZHAO X, ZHANG L, DING Z, et al. Deep reinforcement learning for list-wise recommendations [EB/OL]. [2019-11-14]. https://arxiv.org/pdf/1801.00209.pdf.

[15] ZHAO X, ZHANG L, XIA L, et al. Recommendations with negative feedback via pairwise deep reinforcement learning [C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2018: 1040-1048.

[16] ZHENG G, ZHANG F, ZHENG Z, et al. DRN: a deep reinforcement learning framework for news recommendation [C]// Proceedings of the 2018 World Wide Web Conference. Republic and Canton of Geneva: International World Wide Web Conferences Steering Committee, 2018: 167-176.

[17] 刘洋军. 基于深度强化学习的推荐系统研究[D]. 成都:电子科技大学, 2019:20-61. (LIU Y J. Research on recommendation system based on deep reinforcement learning [D]. Chengdu: University of Electronic Science and Technology of China, 2019: 20-61.)

[18] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing Atari with deep reinforcement learning[EB/OL]. [2019-11-14]. https://arxiv.org/pdf/1312.5602.pdf.

[19] HOFMANN K, WHITESON S, DE RIJKE M. Balancing exploration and exploitation in listwise and pairwise online learning to rank for information retrieval[J]. Information Retrieval, 2013, 16(1): 63-90.

[20] RENDLE S. Factorization machines[C]// Proceedings of the 2010 IEEE International Conference on Data Mining. Piscataway: IEEE, 2010: 995-1000.

[21] CHENG H T, KOC L, HARMSEN J, et al. Wide & deep learning for recommender systems[C]// Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. New York: ACM, 2016: 7-10.

[22] 陈建荣. 基于用户反馈的智能查询扩展技术研究[D]. 哈尔滨:哈尔滨工业大学, 2014:10-11. (CHEN J R. Research on intelligent query expansion technology based on users' feedback [D]. Harbin: Harbin Institute of Technology, 2014:10-11.)

[23] 刘广东. 基于"用户画像"的商品推送系统设计与实现[D]. 西安:西安电子科技大学, 2017:6-7. (LIU G D. The design and implementation of product recommendation system based on "user portrait"[D]. Xi'an: Xidian University, 2017:6-7.)

[24] 叶锡君,龚玥. 基于项目类别的协同过滤推荐算法多样性研究[J]. 计算机工程, 2015, 41(10):42-46, 52. (YE X J, GONG Y. Study on diversity of collaborative filtering recommendation algorithm based on item category [J]. Computer Engineering, 2015, 41(10): 42-46, 52.)

This work is partially supported by Humanities and Social Sciences Research Foundation of the Ministry of Education of China (17YJA870006).

**FU Kui**, born in 1977, Ph. D., associate professor. His research interests include intelligent recommendation, data mining, quantitative investment.

**LIANG Shaoqing**, born in 1996, M. S. candidate. His research interests include intelligent recommendation.

**LI Bing**, born in 1983, Ph. D., associate professor. Her research interests include feature selection, pattern recognition, complex network, intelligent planning.