CIKM 2021 (author)

Association for Computing Machinery
Advancing Computing as a Science & Profession

Help / Log out

Submission 920 | CIKM 2021 | Premium | Conference | EasyChair

CIKM 2021 Submission 920

Submission information updates are disabled. If you need to update your paper, contact CIKM 2021 chairs.

For all questions related to processing your submission you should contact the conference organizers. Click here to see information about this conference.

All **reviews sent to you** can be found at the bottom of this page.

Submission 920	
Title:	Generative Inverse Deep Reinforcement Learning for Online Recommendation
Paper:	(May 25, 10:24 GMT)
Track:	CIKM 2021 Full
Author keywords:	Recommender System Reinforcement learning Deep learning
EasyChair keyphrases:	reinforcement learning (410), online recommendation (280), reward function (260), expert policy (210), actor critic network (206), policy gradient (160), inverse reinforcement learning (158), actor critic (156), inverse deep reinforcement learning (140), deep q learning (126), deep reinforcement learning (126), generative inverse deep reinforcement (120), learning based recommendation (110), discriminative actor critic network (100), neural information processing system (100), recommendation policy (90), recommendation system (80), state action (80), critic network (80), recommendation policy learning (79), policy optimization (65), generalized advantage estimation (63), recommendation problem (60), action space (60), knowledge graph (60), recsim long term engagement (60), recommender system (60), generative inverse reinforcement learning (60), user behavior (60), actor network (50)
Topics:	Recommendation system, Spatial and Temporal Databases
Abstract:	Deep reinforcement learning enables an agent to capture users' interest through dynamic interactions with the environment. It uses a reward function to learn user's interest and to control the learning process, attracting great interest in recommendation research. However, most reward functions are manually designed; they are either too unrealistic or too imprecise to reflect the variety, dimensionality, and non-linearity of the recommendation problem. This impedes the agent from learning an optimal policy to generate satisfactory recommendations in highly dynamic online recommendation scenarios. To address the above issue, we propose a novel generative inverse reinforcement learning approach that avoids the need of defining a reward function. Specially, we model the recommendation problem as an automatic recommendation policy learning problem. We first generate policies based on observed users' preferences and then evaluate the learned policy by an elaborate measurement based on a discriminative actor-critic network. We conduct experiments on an online platform, VirtualTB, and demonstrate the feasibility and effectiveness of our proposed approach via comparisons with several state-of-the-art methods.
Submitted:	May 18, 12:18 GMT
Last update:	May 21, 02:59 GMT
Paper Category (Long Paper Track)(*).	Databases
Author conflicts:	none

Authors

first name	last name	email	country	affiliation	Web page	corresponding?
Xiaocong	Chen	xiaocong.chen@unsw.edu.au	Australia	The University of New South Wales		✓
Lina	Yao	lina.yao@unsw.edu.au	Australia	The University of New South Wales		
Aixin	Sun	axsun@ntu.edu.sg	Singapore	Nanyang Technological		

				University		
Xianzhi	Wang	XIANZHI.WANG@uts.edu.au	Australia	University of Technology Sydney		
Xiwei	Xu	xiwei.xu@data61.csiro.au	Australia	Data 61		
Liming	Zhu	Liming.Zhu@data61.csiro.au	Australia	Data 61		

Reviews

Review 1	
Relevance to CIKM:	**4**: (good)
Originality of the Work:	**2**: (poor)
Technical Soundness:	**3**: (fair)
Quality of Presentation:	**2**: (poor)
Impact of Ideas or Results:	**3**: (fair)
Adequacy of Citations:	**4**: (good)
Reproducibility of Methods:	**3**: (fair)
Reviewer's confidence:	**4**: (good)
Overall Recommendation:	**1**: (Weak Accept)
Nominate for best paper:	-
Comments to the Author(s):	This paper studies the online recommendation problem based on reinforcement learning, more specifically, generative inverse reinforcement learning. Though the authors claim that they propose novel ideas, I would encourage them to take a major revision before resubmission due to the following reasons. 1. The most serious flaw of this paper is its presentation. There are naive typos and grammar errors everywhere across the paper, which makes the reading experience quite unpleasant. To name just a very small portion of them (1) They commonly use ''User's click or not'' is the main metric to -> "is" should be ''as'' (2) Given a user ..., which ... -> this sentence is incomplete (3) Learn a recommendation policy * from expert policy * can be formulate as the problem of matching two occupancy measures [1]. ->should be Learning and formulated Therefore, I would consider that this paper is not ready for publication. 2. The motivations and research problem are unclear. When attacking the flaws of existing work, the authors use ''they''. However, it should be more specific to point out which work has such problems (i.e., the second paragraph of Section 1). Also, in the same paragraph, the points are messed together, making readers really confused. I encourage the authors to extend this paragraph and elaborate.
Summary:	This paper investigates how to develop a better RL model for online learning. The presentation, writing and expression of ideas need significant improvement before publication.

Review 2	
Relevance to CIKM:	**5**: (excellent)
Originality of the Work:	**4**: (good)
Technical Soundness:	**4**: (good)
Quality of Presentation:	**4**: (good)
Impact of Ideas or Results:	**4**: (good)
Adequacy of Citations:	**4**: (good)
Reproducibility of Methods:	**4**: (good)

Reviewer's confidence:	**4**: (good)
Overall Recommendation:	**2**: (Accept)
Nominate for best paper:	-
Comments to the Author(s):	This work studies utilizing deep reinforcement learning for online recommendation, which has been extensively studied recently. This work points out that most reward functions in deep reinforcement learning are manually designed and hard to reflect the variety, dimensionality, and non-linearity of the recommendation problem. Hence, these methods cannot be available for online recommendation. For this reason, this work proposes a novel generative inverse reinforcement learning approach that avoids the need of defining a reward function, which is worthy to be studied. Finally, extensive experiments are conducted based on three datasets to evaluate the performance of the proposal. In summary, the entire work seems to be solid but there are still minor revisions needed. 1. Suggest taking the DRN approach [31] as a baseline approach in the evaluation as it is relevant with this work. 2. The background color of curves in the figures in experiments is unnecessary. 3. The written needs to be improved. For instance, the citation of references in the last two paragraphs in Section 5 does not conform to the established rules.
Summary:	This work seems to be solid and could be accepted after minor revisions.

Review 3	
Relevance to CIKM:	**4**: (good)
Originality of the Work:	**4**: (good)
Technical Soundness:	**4**: (good)
Quality of Presentation:	**4**: (good)
Impact of Ideas or Results:	**4**: (good)
Adequacy of Citations:	**3**: (fair)
Reproducibility of Methods:	**4**: (good)
Reviewer's confidence:	**4**: (good)
Overall Recommendation:	**2**: (Accept)
Nominate for best paper:	-
Comments to the Author(s):	In this paper, authors make an effort to solve the reward function definition problem in deep reinforcement learning based recommender system. The experiments are conducted in three different simulation platforms and achieve a promising result. Here are my comments: Good point, inverse reinforcement learning seems can skip the reward function definition. To my knowledge, it is not handy to define the demonstration in the recommendation scenario. It should be better to be discussed in the conclusion and provide more insights. More background about inverse reinforcement learning could be elaborated to make this work more self-contained. Otherwise, it is hard to follow for those readers who do not have a strong background in reinforcement learning and inverse reinforcement learning. It would be better to mention that why most of existing methods can not run on those simulation platforms in details. Authors may also compare with traditional DQN, policy gradient or DDPG as well to demonstrate the superiority. The legend for figure 3(c),(f),(i),(l) should be in the same position and style. Authors may need to provide the possible application domain for such work. For example, travel recommendation, medical recommendation which the demonstration is easier to get.
Summary:	In general, this is a good paper to solve a challenging problem. The paper is well written and easy to follow. The technical solution is sound and should be of interest to the related readers. However, there is still room to improve the readability by fixing a few language issues.

Metareview

Metareview for paper 920	
Title:	Generative Inverse Deep Reinforcement Learning for Online Recommendation
Authors:	Xiaocong Chen, Lina Yao, Aixin Sun, Xianzhi Wang, Xiwei Xu and Liming Zhu
Text:	All the three reviewers are confident that this paper tackles a challenging problem in reinforcement learning-based recommendation with a sound solution and its technical contributions are significant to this community. But the following issues are expected to be addressed in the final version. First, the writing and presentation of this paper (including citation format) could be further improved. Specifically, more background knowledge about inverse reinforcement learning should be provided to make this work more self-contained. Besides, more explanations and discussions on why most of existing methods cannot run on those simulation platforms are suggested to be provided. Professional proof-reading is also needed to address the mentioned minor presentation issues. Second, more baselines such as traditional DQN, policy gradient or DDPG and DRN approach are suggested for the comparison purpose in the experiment to better show the superiority of the proposed approach.