ISSN 2616-5775 Vol. 6, Issue 13: 14-19, DOI: 10.25236/AJCIS.2023.061303

A Reinforcement Learning Based on Book Recommendation System

Qinyong Wang[1,*], James A. Esquivel[1]

[1]Graduate School, Angeles University Foundation, Angeles City, Philippines
[*]Corresponding author

Abstract: *As the Information Age continues to evolve, the significance of recommendation systems in people's daily lives becomes increasingly prominent. Traditional recommendation algorithms, such as content-based filtering, matrix factorization, logistic regression, factorization machines, neural networks, and multi-armed bandits, predominantly focus on immediate feedback for recommended items, often overlooking long-term rewards. This paper aims to investigate the application of reinforcement learning in personalized book recommendation systems, with the objective of enhancing user experience and recommendation accuracy.*

Keywords: *Actor-Critic, Reinforcement learning, Recommendation System*

1. Introduction

Over the past few decades, recommendation systems have evolved from simple rules and content-based methods into diverse and intelligent systems. Traditional recommendation methods, such as collaborative filtering, rely on analyzing user behavior and relationships among similar users for predictions. While successful in certain contexts, these methods often falter in the face of challenges like user cold starts, sparse information, and recommending new items[1,2]. Additionally, content-based approaches may struggle in certain domains where expressing user interests through simple keywords, as in the case of books, poses difficulties[3].

With the rise of deep learning technology, recommendation systems have entered a new era. Deep learning models can capture abstract features of data by learning underlying representations, demonstrating superior performance in handling complex and high-dimensional user-item interaction data. However, both traditional methods and deep learning models confront common challenges such as data sparsity, cold start issues, and model interpretability.

Reinforcement learning, as an approach to learning optimal strategies through interaction with the environment, presents a novel perspective for addressing these challenges[4]. It fundamentally transforms the operation of recommendation systems, viewing them as processes where an intelligent agent makes decisions in a dynamic environment. User feedback or ratings serve as reward signals from the environment, and the recommendation system's objective is to optimize recommendation strategies by learning from user-environment interactions[5,6].

In the realm of recommendation systems, reinforcement learning has achieved significant success in some research endeavors. However, its application in personalized book recommendations remains relatively understudied. This research aims to fill this gap by delving into the potential advantages of reinforcement learning in addressing challenges specific to personalized book recommendations.

2. Methodology

This paper employs the Actor-Critic(AC) algorithm within the framework of reinforcement learning, which integrates policy gradient methods and value function estimation. The AC algorithm consists of two primary components: the Actor (policy network) and the Critic (value function network). These components collaborate to enable the Agent (intelligent entity) to learn a policy for action selection and evaluate the quality of these actions through the learning of a value function[7,8,9].

The goal of the Actor is to learn a policy, mapping from states to actions, and it outputs a probability distribution over all possible actions given a particular state. The parameters of the Actor

ISSN 2616-5775 Vol. 6, Issue 13: 14-19, DOI: 10.25236/AJCIS.2023.061303

network are updated using gradient ascent to maximize the cumulative rewards associated with the chosen actions, known as the policy gradient method. The output of the Actor network represents the probabilities of the agent taking each action in a given state.

The Critic aims to learn a value function[10], estimating the value of each state-action pair. It evaluates the long-term cumulative rewards when the agent is in a specific state and takes a particular action. The parameters of the Critic network are updated through gradient descent to minimize the mean squared error between the estimated value function and the actual rewards. The output of the Critic network represents the estimated value when the agent is in a given state and takes a specific action. This dual-structure of Actor-Critic facilitates effective learning of both policies and value functions, enhancing the overall performance of the personalized book recommendation system.

3. Data

In the development of a reinforcement learning-based book recommendation system, a comprehensive understanding of the data is crucial[11]. This section explores our in-depth analysis of the dataset, including its structure and features.

3.1. Understanding the Data

There are three dataset files inside the folder, namely Books, Ratings, and Users, which will be used for the model development process.

The Books.cvs has 271,360 types of books and consists of 8 columns(Figure 1), namely:

ISBN: a unique book identification number.

Book-Title: the title of the book.

Book-Author: the name of the book author.

Year-Of-Publication: the year of publication of the book.

Publisher: the name of the book publisher.

Image-URL-S: the URL link for small-sized images.

Image-URL-M: the URL link for medium-sized images.

Image-URL-L: the URL link for large-sized images.

	ISBN	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S
0	0195153448	Classical Mythology	Mark P. O. Morford	2002	Oxford University Press	http://images.amazon.com/images/P/0195153448.0...
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/P/0002005018.0...
2	0060973129	Decision in Normandy	Carlo D'Este	1991	HarperPerennial	http://images.amazon.com/images/P/0060973129.0...
3	0374157065	Flu: The Story of the Great Influenza Pandemic...	Gina Bari Kolata	1999	Farrar Straus Giroux	http://images.amazon.com/images/P/0374157065.0...
4	0393045218	The Mummies of Urumchi	E. J. W. Barber	1999	W. W. Norton & Company	http://images.amazon.com/images/P/0393045218.0...

Figure 1: Contents of the Books.csv Data Table

The ratings.cvs has 340,556 ratings for books and consists of 3 columns(Figure 2), namely:

User-ID: a unique code for anonymous users who provide ratings.

ISBN: the book identification number.

Book-Rating: the rating given to the book.

ISSN 2616-5775 Vol. 6, Issue 13: 14-19, DOI: 10.25236/AJCIS.2023.061303

	User-ID	ISBN	Book-Rating
0	276725	034545104X	0
1	276726	0155061224	5
2	276727	0446520802	0
3	276729	052165615X	3
4	276729	0521795028	6

Figure 2: Contents of the Ratings.csv Data Table

The users.cvs has 278,858 anonymous user names and consists of 3 columns(Figure 3), namely:

User-ID: a unique code for anonymous user names.

Location: the location of the user's residence.

Age: the age of the user.

	User-ID	Location	Age
0	1	nyc, new york, usa	NaN
1	2	stockton, california, usa	18.0
2	3	moscow, yukon territory, russia	NaN
3	4	porto, v.n.gaia, portugal	17.0
4	5	farnborough, hants, united kingdom	NaN

Figure 3: Contents of the Users.csv Data Table

3.2. Data Preprocessing

In order to better align with the requirements of the model, it is essential to perform cleaning, transformation, and repair on the raw data[12]. This includes:

(1)Handling Missing Values

The presence of missing values in the data can impact the performance of the model. During the data preprocessing stage, missing values can be addressed through methods such as imputation, deletion, or interpolation to enhance data completeness.

(2)Dealing with Outliers

Outliers, resulting from errors, noise, or exceptional circumstances, can adversely affect model training and performance[13]. Data preprocessing involves detecting and handling outliers to improve the robustness of the model.

(3)Data Cleaning

Inconsistencies, errors, or redundant data may exist in the dataset. Data cleaning encompasses operations such as deduplication, standardization, and normalization to ensure data consistency and accuracy.

(4)Feature Engineering

Data preprocessing often includes handling features, such as feature extraction, transformation, and normalization. This aids in enhancing model performance, mitigating the curse of dimensionality, and aligning with the input requirements of the model.

(5)Handling Categorical Data

Machine learning models typically require numeric inputs, while the original data may contain categorical data. Data preprocessing can involve techniques like one-hot encoding or label encoding to convert categorical data into numeric form.

Upon completion of data preprocessing, we can conduct statistical analyses and clearly present the Top 10 Authors by Number of Books (see Figure 4).

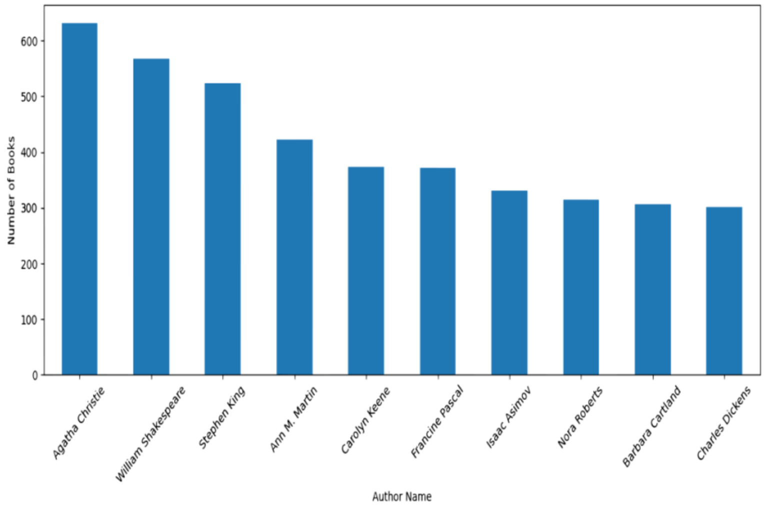

Figure 4: Top 10 Authors by Number of Books

4. Modeling

(1)Environment Modeling

The recommendation system's users, items, and ratings are modeled as a reinforcement learning environment. This involves defining states, actions, and rewards.

(2)State Representation

User and item information is represented as states. Deep learning models can be employed to embed features of users and items, forming state representations.

(3)Action Representation

The action space of the recommendation system, i.e., the actions that the Actor can choose, is defined. In the context of book recommendations, actions could be the recommended book lists.

(4)Reward Design

Reward signals are defined based on user rating scenarios. The design of reward signals is crucial for capturing user preferences effectively.

(5)Actor Network Design

The Actor network is designed, taking the current state as input and outputting a probability distribution for recommending each book[14]. The softmax function can be used to normalize the probabilities.

(6)Critic Network Design

The Critic network is designed, taking the current state and the action output of the Actor network as input and providing an estimate of the value for that action.

(7)Training Process

The Actor-Critic (AC) algorithm is employed for training. At each time step, actions are selected based on the Actor network, rewards are obtained, and the parameters of both the Actor and Critic networks are updated to maximize cumulative rewards.

(8)Recommendation

After training completion, the trained Actor network is used for recommendations. For a specific user, the Actor network outputs a probability distribution, and book recommendations are obtained by sampling from this distribution.

Following model training, evaluation is conducted, encompassing metrics such as:

Root Mean Squared Error (RMSE): Measures the differences between predicted and observed values. It gauges the average error between model predictions and actual values, with increased

ISSN 2616-5775 Vol. 6, Issue 13: 14-19, DOI: 10.25236/AJCIS.2023.061303

sensitivity to large errors.

Mean Absolute Error (MAE): Evaluates differences between predicted and observed values. Unlike RMSE, MAE does not square errors, emphasizing absolute error values.

These evaluation metrics provide insights into the model's performance and its ability to accurately predict user preferences in the book recommendation system in Figure 5.

	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	Std
RMSE (testset)	3.4965	3.5000	3.4995	3.5024	3.4895	3.4976	0.0045
MAE (testset)	2.9282	2.9308	2.9288	2.9316	2.9214	2.9282	0.0036
Fit time	13.80	14.05	14.17	14.44	14.31	14.15	0.22
Test time	1.19	1.17	1.17	1.17	1.20	1.18	0.01

RMSE: 1.9321
MAE: 1.4395

Figure 5: Evaluating RMSE, MAE of algorithm on 5 splits

Precision and Recall Metrics were Computed(Figure 6), Indicating Strong Performance.

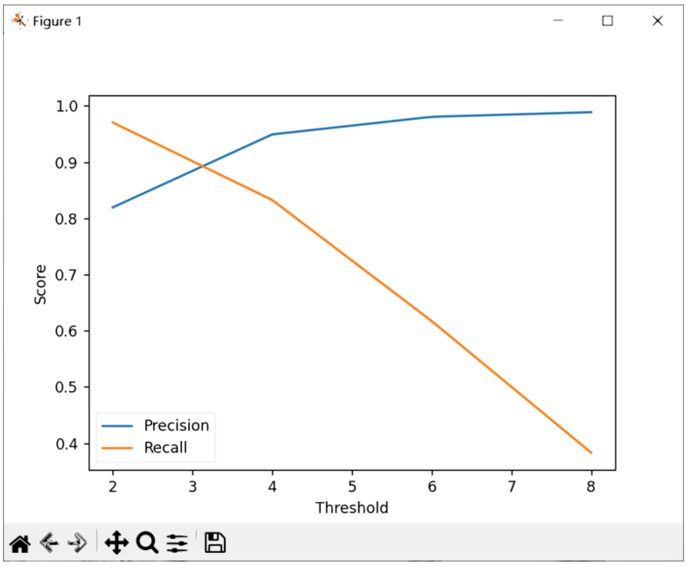

Figure 6: Precision & Recall

5. Conclusion

The application of reinforcement learning in book recommendation systems proves to be an intriguing and effective approach. The delicate balance between exploration and exploitation within reinforcement learning is a critical aspect. The system must explore new recommendation strategies to uncover users' latent preferences while concurrently exploiting known efficient strategies. The application of reinforcement learning in book recommendation systems can be extended to various domains, including libraries, online reading platforms, among others, providing users with more intelligent and personalized reading recommendations. In summary, reinforcement learning introduces a more flexible and personalized recommendation capability to book recommendation systems, emphasizing real-time user experiences. It represents a promising research direction within the realm of recommendation systems.

References

[1] Afsar M M, Crump T, Far B. Reinforcement learning based recommender systems: A survey[J]. ACM Computing Surveys, 2022, 55(7): 1-38.

[2] Wang H, Liu N, Zhang Y, et al. Deep reinforcement learning: a survey[J]. Frontiers of Information Technology & Electronic Engineering, 2020, 21(12): 1726-1744.

[3] Yang L, Tan B, Zheng V W, et al. Federated recommendation systems[J]. Federated Learning: Privacy and Incentive, 2020: 225-239.

ISSN 2616-5775 Vol. 6, Issue 13: 14-19, DOI: 10.25236/AJCIS.2023.061303

[4] Zhou S, Dai X, Chen H, et al. Interactive recommender system via knowledge graph-enhanced reinforcement learning[C]//Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval. 2020: 179-188.
[5] Bai X, Guan J, Wang H. A model-based reinforcement learning with adversarial training for online recommendation[J]. Advances in Neural Information Processing Systems, 2019, 32.
[6] Ie E, Jain V, Wang J, et al. SlateQ: A tractable decomposition for reinforcement learning with recommendation sets[J]. 2019.
[7] Tang X, Chen Y, Li X, et al. A reinforcement learning approach to personalized learning recommendation systems[J]. British Journal of Mathematical and Statistical Psychology, 2019, 72(1): 108-135.
[8] Zhao X, Xia L, Tang J, et al. " Deep reinforcement learning for search, recommendation, and online advertising: a survey" by Xiangyu Zhao, Long Xia, Jiliang Tang, and Dawei Yin with Martin Vesely as coordinator[J]. ACM sigweb newsletter, 2019, 2019(Spring): 1-15.
[9] Ge Y, Liu S, Gao R, et al. Towards long-term fairness in recommendation[C].Proceedings of the 14th ACM international conference on web search and data mining. 2021: 445-453.
[10] Zhou F, Yin R, Zhang K, et al. Adversarial point-of-interest recommendation[C].The world wide web conference. 2019: 3462-34618.
[11] Zhong K, Wang Y, Pei J, et al. Super efficiency SBM-DEA and neural network for performance evaluation[J]. Information Processing & Management, 2021, 58(6): 102728.
[12] Jan N, Gwak J, Pei J, et al. Analysis of networks and digital systems by using the novel technique based on complex fuzzy soft information[J]. IEEE Transactions on Consumer Electronics, 2022, 69(2): 183-193.
[13] Yu Z, Pei J, Zhu M, et al. Multi-attribute adaptive aggregation transformer for vehicle re-identification[J]. Information Processing & Management, 2022, 59(2): 102868.
[14] Li J, Li S, Cheng L, et al. BSAS: A Blockchain-Based Trustworthy and Privacy-Preserving Speed Advisory System[J]. IEEE Transactions on Vehicular Technology, 2022, 71(11): 11421-11430