Proceedings of CONF-MLA 2024 Workshop: Securing the Future: Empowering Cyber Defense with Machine Learning and Deep Learning
DOI: 10.54254/2755-2721/94/2024MELB0071

The Impact of Movie Text Information on Movie Recommendation Systems based on Multi-Armed Bandit

Yitong Guo

College of Artificial Intelligence, Hunan University, Changsha, 410000, China

irisqing@hnu.edu.cn

Abstract. The multi-armed bandit problem has been extensively studied and applied in various fields such as recommendation systems and resource allocation. However, there is still a lack of research in the movie recommendation domain, such as the comparison of algorithm performance and the impact of movie text information. This paper uses the Movielens 100K movie review data set to compare and analyze the performance of classic algorithms-Greedy, UCB, and TS-in the multi-armed bandit problem. Additionally, to explore the influence of movie text information (such as movie summaries and user reviews) on system algorithms, this experiment utilizes web scraping technology through HTTP requests and HTML parsing, to extract data from the Douban website, obtaining the information and forming a new data set. In this movie review data set, the Greedy algorithm outperforms both UCB and TS algorithms. This study finds that incorporating movie text information improves the accuracy of all three algorithms, thus contributing to the enhancement of movie recommendation systems. This experiment did not consider larger datasets or movie image information, which could be addressed in future improvements.

Keywords: Movie text information, impact, multi-armed bandit.

1. Introduction

With the rapid development of internet technology, massive amounts of data have permeated various aspects of people's lives, bringing convenience but also posing challenges. Consequently, the effective filtering and presentation of vast information to users have become a major focus today. Recommendation systems play a crucial role in showcasing valuable information based on users' interests and needs. Their applications span widely across social networks (such as Tencent, WeChat, Weibo), news aggregation platforms (like Today's headlines, Google News), e-commerce (including Alibaba, JD.com, Taobao, Meituan), music, and video streaming services [1-3]. Various methodologies continually emerge in the construction of recommendation systems.

In comparison to traditional recommendation algorithms, the multi-armed bandit algorithm, as a classic reinforcement learning approach, has been widely adopted in multiple domains in recent years [4, 5]. Meanwhile, improving upon the basic multi-armed bandit-based recommendation algorithms has led to exploration of other areas, such as the context-aware bandit problem [6].

In recent years, multi-armed bandit has found applications in various fields such as telecommunications and brain-computer interfaces. Maghsudi et al. explored the application of multi-armed bandits in 5G small cell networks, particularly in dynamic spectrum allocation, optimizing resource allocation decisions to enhance network performance [7]. Heskebeck et al. investigated enhancing brain-computer interface (BCI) performance using multi-armed bandit algorithms, focusing on actions controlled by the brain [8]. Scholars have also summarized the applications of multi-armed. Bouneffouf et al. comprehensively surveyed these algorithms across different domains, proposing a classification of applications based on multi-armed bandits and providing updated summaries for each field [9]. Wang et al. delved into effectively modeling and managing contextual uncertainties to enhance the performance of contextual bandit algorithms in personalized recommendations [10].

Despite existing research on the application of multi-armed bandit algorithms in movie recommendations, there are still some issues. For example, the datasets often include only movie titles and tags, neglecting the impact of textual information such as movie summaries and user reviews on user preferences. This study addresses this gap by supplementing the applied datasets through web scraping techniques [11]. Furthermore, it compares ε-Greedy, Upper Confidence Bound (UCB), and Thompson Sampling (TS) strategies within the context of multi-armed bandits, aiming to aid the field of movie recommendations.

2. Methodology

2.1. Data source and description

In movie recommendation research, one classic data set is Movielens. However, this data set lacks movie synopsis data. To study the impact of movie text information on user recommendations, this research uses a data set based on initial data (as shown in Table 1), supplemented with movie information scraped from Douban. This data set contains over ten thousand movie ratings from about 1,000 users, covering approximately 1,700 movies. The data set can be divided into three parts: the first part records movie ratings, ranging from one to five stars, serving as rewards for the research. The second part includes user information, such as user ID, gender, age, and occupation, which may influence movie ratings. The third part consists of movie information, categorizing all movies into multiple tags and including their release dates.

Table 1. The partial data of Movielens 100K.

User ID	Gender	Age	Occupation	Movie titles	Movie tag	Rating(1-5)
1	M	24	Technician	Star Wars	Action	5
2	F	30	Doctor	Casino	Crime	3
3	M	22	Student	L.A. Confidential	Crime	5

2.2. Metric selection and explanation

In the Multi-Armed Bandit problem, the average reward is a key metric for evaluating the performance of algorithms. It measures the average return the algorithm achieves over multiple trials. Here is the formula and explanation for calculating the average reward:

$$r = \frac{1}{N}\sum_{i=1}^{N} r_i \tag{1}$$

where r_i is the reward value obtained in the i. In recommendation systems, this is often the user's rating or feedback on the recommended item. N is the total number of trials. In your experiment, this could be the total number of recommendations or the total number of interactions between users and the recommendation system. r is the average reward, representing the average return across all trials. In this study, the total reward is the sum of the actual user ratings obtained in each recommendation. The average reward is the mean of the total rewards after 1000 iterations.

2.3. Web scraping technology

Given that Movielens 100K contains about 1700 movies, manually collecting movie text information would be inefficient and time-consuming. Web scraping is a program or script that automatically retrieves web information according to specific rules. The process involves: first, fetching web data by sending HTTP requests to the web server via the target webpage's URL, with the web server responding with HTTP responses. Second, parsing the web data by analyzing and processing the HTML source code to extract the desired data. Third, storing the data by saving it in a suitable format for subsequent model training.

2.4. Multi-armed bandit algorithms

As a classic reinforcement learning algorithm, the multi-armed bandit includes several techniques such as Explore Then Commit (ETC), Upper Confidence Bound (UCB), Thompson Sampling (TS), and ε-Greedy algorithms. The ε-Greedy algorithm primarily selects the arm with the highest known expected reward but occasionally explores other arms with a small probability ε. ETC maintains a delicate balance between exploration and exploitation. The UCB algorithm selects the arm with the highest upper confidence bound (i.e., the estimated expected reward plus a confidence level), where the confidence level is usually inversely proportional to the number of selections, balancing exploration and exploitation. TS uses Bayesian methods to update the reward distribution of each arm and makes selections based on the posterior distribution (Figure 1).

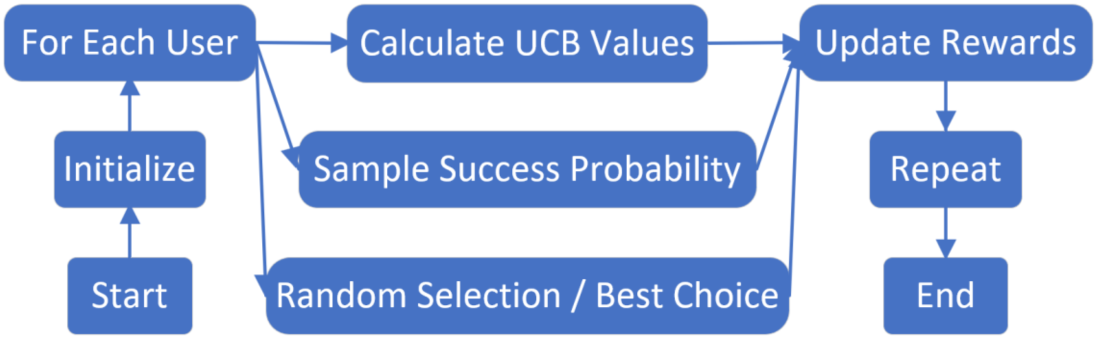

Figure 1. Algorithm flow.

3. Results and discussion

3.1. Model results

In analyzing the experimental results of different recommendation algorithms, it was observed that the ε-greedy algorithm achieved the highest average reward of 3.67. This indicates that, on the given data set, the ε-greedy algorithm outperforms both the UCB and TS algorithms in terms of recommendation effectiveness. The ε-greedy algorithm introduces a small exploration probability ε, allowing for some degree of random choice, which helps to avoid getting stuck in local optima. Overall, the ε-greedy algorithm tends to choose the best-known option most of the time (exploitation), but randomly selects other options with probability ε (exploration). This strategy enables the ε-greedy algorithm to effectively adapt to user preferences and movie ratings, achieving higher rewards by balancing exploration and exploitation (Table 2 and Figure 2).

Table 2. The results of average reward

data	ε-Greedy	UCB	TS
initial data	3.67	3.12	3.55
New data	3.85	3.25	3.70

In contrast, the UCB algorithm had an average reward of 3.12, slightly lower than that of ε-greedy. This may be because, although the UCB algorithm increases exploration opportunities by considering confidence intervals and reduces selection bias, it can be too conservative in situations of high

DOI: 10.54254/2755-2721/94/2024MELB0071

uncertainty. This conservativeness may negatively impact short-term reward performance, making it less effective than the ε-greedy algorithm in providing higher rewards.

The TS algorithm had an average reward of 3.55, slightly lower than ε-greedy but higher than UCB. The TS algorithm decides on recommended movies by sampling from a Beta distribution. This method theoretically handles uncertainty better and gradually optimizes choices. However, in this experiment, TS performed worse than ε-greedy, possibly due to the influence of the Beta distribution parameter settings on its performance. The choice of Beta distribution parameters is crucial for the TS algorithm's performance; improper parameter settings can affect the recommendation results.

From Figure 2, it can also be seen that the ε-greedy algorithm's total reward increases faster than the other algorithms.

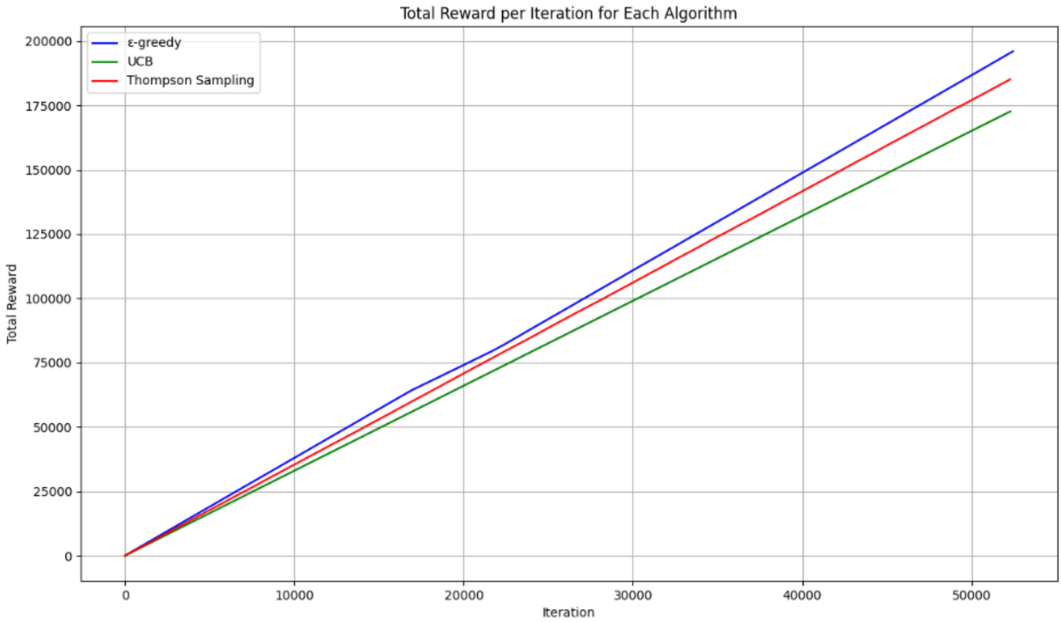

Figure 2. Total reward variation chart

3.2. Discussion

The study also compared these algorithms using a new data set with movie information obtained through web scraping. The results for the new data set are as follows: Compared to the original data set, the performance of the ε-greedy algorithm on the new data set has further improved, reaching 3.85, showing enhanced performance. This improvement is likely due to the richer movie information in the new data set, which allows the ε-greedy algorithm to more accurately match user preferences. The additional information from movie descriptions provides more context for the algorithm, helping it better balance exploration and exploitation, ultimately leading to higher rewards. The average reward for the UCB algorithm has also increased, but it remains lower than that of ε-greedy, indicating that while the UCB algorithm has improved on the new data set, it still falls short of ε-greedy in the short term. This improvement may be attributed to the extra information in the new data set helping the UCB algorithm make better choices, thereby somewhat addressing its previous shortcomings. The average reward for the TS algorithm has also risen on the new data set, reaching 3.70, demonstrating an enhanced ability of the TS algorithm to handle uncertainty and make better use of the rich information in the new data set. Although the average reward of the TS algorithm is still lower than that of ε-greedy, the improvement in its performance indicates that beta distribution sampling can better adjust recommendation strategies in the face of a new data set, resulting in higher rewards.

4. Conclusion

This study introduces a data set containing movie textual information, which enhances the performance of recommendation algorithms. The ε-greedy algorithm performs best on the new data set, providing the highest average reward. The improvements in UCB and TS algorithms suggest that utilizing additional information can help optimize algorithm performance in practical applications. However, the limitation of this study is that it only explored the impact of movie texts on the system and did not investigate other factors (such as movie posters). Additionally, the data set used was relatively small. Future work could further explore how to leverage additional features and data to improve the performance of recommendation systems. By incorporating more features and information, it may be possible to further enhance the accuracy of recommendation algorithms and user satisfaction, leading to more personalized and efficient recommendation services.

References

[1] Hui B, Zhang L Z, Zhou X, et al. 2021 Personalized recommendation system based on knowledge embedding and historical behavior. Applied Intelligence, 1-13,
[2] Zhang J, Yang Y, Zhuo L, et al. 2019 Personalized recommendation of social images by constructing a user interest tree with deep features and tag trees. IEEE Transactions on Multimedia, 21(11), 2762-2775.
[3] Xu H L, Wu X, Li X D, et al. 2009 Comparison study of internet recommendation system. Journal of software, 20(2), 350-362.
[4] Mohri M 2005 Multi-armed bandit algorithms and empirical evaluation. Proccedings of the 16th European Confe-rence on Machine Learning. Berlin: Springer, 437-448.
[5] Gittens J, et al. 1989 Multi-armed bandit allocation indices. Journal of the Opera-tional Research Society, 40(2), 1158-1159
[6] Wang H, Ma Y, Ding H, et al. 2022 Context uncertainty in contextual bandits with applications to recommender systems. Proceedings of the AAAI Conference on Artificial Intelligence, 36(8), 8539-8547.
[7] Maghsudi S and Hossain E 2016 Multi-armed bandits with application to 5G small cells. IEEE Wireless Communications, 23(3), 64-73.
[8] Heskebeck F, Bergeling C and Bernhardsson B 2022 Multi-armed bandits in brain-computer interfaces. Frontiers in Human Neuroscience, 16, 931085.
[9] Bouneffouf D, Rish I and Aggarwal C 2020 Survey on applications of multi-armed and contextual bandits. 2020 IEEE Congress on Evolutionary Computation (CEC) IEEE, 1-8.
[10] Wang H, Ma Y, Ding H, et al. 2022 Context uncertainty in contextual bandits with applications to recommender systems. Proceedings of the AAAI Conference on Artificial Intelligence, 36(8), 8539-8547.
[11] Pan H L 2023 Research on Movie Recommendation Model Based on Deep Reinforcement Learning. Taiyuan University of Technology.