Volume 10, Issue 3, March – 2025 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar1967

Improving Deep Reinforcement Learning-Based Recommender Systems: Overcoming Issues with Usability, Profitability, and User Preferences

Navin Kumar Sehgal[1]; Kartik Tyagi[2]

[1,2]Tech Department, Academics. College Dekho Jaipur, India

Publication Date: 2025/04/12

Abstract: Recommender systems play a important role in personalizing user experiences across digital platforms. Usual recommendation methods often struggle with balancing both user satisfaction and business profitability, leading to inefficiencies in recommendation accuracy and engagement. This study proposes a Deep Reinforcement Learning (DRL)-based recommender system that integrates utility-based optimization, profitability-driven strategies, and user preferences. The methodology uses real-world Amazon product review data to find key insights into review helpfulness, brand engagement, and category-based preferences. Data analysis revealed strong correlations between review helpfulness and product ratings. Experiment was conducted by simulating environment on python using DQN. The recommendations were predicted based on data and user interactions. The experiment highlighted the importance of leveraging high-quality reviews in recommendations. Additionally, brand popularity was identified as a significant factor influencing user engagement, emphasizing the need for brand-aware recommendation strategies. The study introduces a framework that balances utility, business profitability, and consumer effort. By incorporating reinforcement learning techniques, the proposed model adapts to evolving user preferences while improving recommendation efficiency. Experimental results find that the DRL-based system enhances recommendation accuracy, improves long-term engagement and increasing overall business profitability. This research contributes to the improvement of AI-driven recommendation models by offering a scalable, adaptive, and viable solution for recommender systems. Future work will explore real-time adaptability and further refinements in reward modeling to enhance computational efficiency and user experience.

Keywords: *Deep Reinforcement Learning, Recommender Systems, Personalized Recommendations, User Preferences, Utility-Based Recommendation, Profit-Driven Recommender System, E-commerce Recommendation Systems, Long-Term User Engagement, Multi-Objective Framework.*

How to Cite: Navin Kumar Sehgal; Kartik Tyag. (2025). Improving Deep Reinforcement Learning-Based Recommender Systems: Overcoming Issues with Usability, Profitability, and User Preferences. *International Journal of Innovative Science and Research Technology*, 10(3), 2722-2728. https://doi.org/10.38124/ijisrt/25mar1967.

I. INTRODUCTION

Recommender systems (RS) have an important role in filtering and personalizing content on various digital platforms, including e-commerce, streaming services, and social media. As the volume of digital content grows exponentially, businesses and service providers increasingly rely on recommendation algorithms to enhance user engagement and improve content delivery. Traditional recommendation approaches, such as collaborative filtering and content-based filtering, have demonstrated significant success but face limitations, particularly in adapting to dynamic user preferences and balancing accuracy with diversity. Recent developments in artificial intelligence (AI), particularly Deep Reinforcement Learning (DRL), offer better solutions to overcome these challenges by enabling adaptive and utility-driven recommendation models.

II. LITERATURE REVIEW

A. Background

The integration of DRL into recommender systems has gained substantial attention due to its ability to model sequential decision-making processes and improve long-term user engagement. Unlike conventional methods, DRL-based RS dynamically learn from user interactions and adapt their recommendations accordingly. However, several challenges persist, including computational complexity, efficient reward modeling, and scalability for large-scale applications. Additionally, existing models often fail to balance user satisfaction with business profitability, leading to suboptimal recommendation strategies.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar1967

B. Current Studies

A thorough analysis of DRL implementations in RS was presented by Krishnamoorthi & Gopal (2022), who also pointed out some of the possible advantages and drawbacks. The study conducted found problem regarding data sparsity and computation cost.

Nie (2023) investigated how knowledge graphs help with cold-start issues and sparsity in user-item interactions. Although the suggested IKANAM model improved user preference by using temporal data, its high computing cost was still a disadvantage.

Contrastive State Augmentations (CSA) introduced by Ren et al. (2023) in order to enhance recommender systems that rely on reinforcement learning. Their method improved model generalization by addressing problems with feedback representation and value function estimation. The choice between adaptability and computing efficiency persisted.

Mahmood & Ricci (2007) investigated interactive and adaptive recommender systems using Markov Decision Processes (MDP). Their research established that RL techniques could enhance user interaction strategies, but accurately modeling dynamic user preferences remained a challenge.

C. Research Problem

Despite advancements in DRL-based RS, various gaps remain unresolved. The computational demands of DRL limit its scalability, and existing solutions often overlook the balance between user satisfaction and business profitability. Moreover, value function estimation and user preferences continue to pose significant challenges, hindering the development of truly adaptive and efficient recommendation systems. Addressing these gaps is essential to enhance recommendation quality, improve long-term engagement, and ensure that businesses achieve sustainable profitability through effective content curation.

This research aims to develop an improved DRL-based recommender system that effectively addresses utility, profitability, and user preference challenges. By improving reward functions, and leveraging multi-objective strategies, this study seeks to bridge the existing gaps in DRL-based recommendation frameworks. The outcomes of this research will contribute to the development of more efficient, scalable, and adaptive recommender systems, benefiting both users and businesses in the digital ecosystem.

III. METHODOLOGY

A. Research Design

This study employs a quantitative research approach, utilizing Deep Reinforcement Learning (DRL) techniques to improve recommender systems. The research follows an experimental design, where real-world data is analyzed, and machine learning models are trained and evaluated. A combination of exploratory data analysis (EDA), feature engineering, and reinforcement learning-based optimization techniques is used to ensure a comprehensive assessment of recommendation effectiveness.

B. Data Collection

The dataset used in this study comprises Amazon product reviews, which include key attributes such as product ratings, review text, brand, review helpfulness, and product categories. The dataset was selected due to its diversity and real-world applicability in evaluating recommender systems. Data was sourced from publicly available repositories, ensuring credibility and authenticity. The preprocessing steps involved handling missing values, normalizing text-based features, and filtering out irrelevant entries to ensure data quality.

The dataset used for this research comprises Amazon product reviews, including essential attributes such as ratings, review text, helpfulness votes, product brand, and categories. This dataset was selected due to its diverse nature, which allows for comprehensive analysis across multiple product segments. Given the study's objective of balancing Deep Reinforcement Learning (DRL)-based recommender systems, the dataset offers a rich source of real-world interactions between users and products. To ensure consistency and reliability, missing values in critical columns such as reviews, rating and reviews helpful were addressed through data cleaning techniques, including removing null entries and minor missing values when necessary.

C. Data Preprocessing

➢ *Before Model Training, the Dataset Underwent Several Preprocessing Steps:*

- Handling Missing Values: Null values in critical fields like ratings and review helpfulness were either removed or imputed.
- Feature Extraction: Key features such as sentiment scores from review text, brand engagement, and category-based preferences were derived.
- Normalization and Scaling: Numeric attributes were standardized to ensure uniformity across features.
- Text Processing: Review text was tokenized, cleaned, and converted into vector representations for model training.

Recommender system environment was simulated on DQN based on python libraries. The system updated recommendation based on user input consisting of product selection and user review.

The analysis focused on understanding key user interaction patterns, product rating distributions, and review helpfulness. Three major aspects were examined:

- Review Helpfulness vs. Ratings – The relationship between review helpfulness votes and rating scores was explored to assess whether highly rated products also

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar1967

receive more helpful reviews. This analysis provided insight into user preferences and the influence of helpfulness in shaping recommendations.

- Top Reviewed Brands – The distribution of product reviews across brands was analyzed to identify market leaders in user engagement. This helped in understanding how brand preference affects recommendation accuracy and long-term profitability.
- Categories by Average Rating – The study evaluated the highest-rated product categories, revealing which types of products consistently received positive feedback from users. This analysis supported the development of an adaptive recommendation framework that improves recommendations based on category popularity trends.

D. *Model Architecture*

➤ *The Research Employs a Deep Reinforcement Learning-based Model, Structured as Follows:*

- State Representation: User-product interactions, historical preferences, and contextual factors are encoded as states.
- Action Space: The system suggests products from a candidate pool based on learned user preferences.
- Reward Function: A multi-objective optimization function is implemented:

E. *Training and Optimization*

- $U = \alpha * U\_user + \beta * U\_business - \gamma * C\_effort$

Where:

- U_user accounts for recommendation accuracy, engagement, and satisfaction.
- U_business includes profitability measures such as conversion rates and revenue generation.
- C_effort represents the cognitive load required for users to find relevant recommendations.
- α, β, and γ are tuning parameters to balance objectives.

The DRL model was trained using a policy gradient-based approach, ensuring adaptability to dynamic user preferences. The training process involved:

- Exploration vs. Exploitation Trade-off: Implementing an epsilon-greedy strategy to balance novel recommendations with well-established preferences.
- Batch Training: Leveraging mini-batch updates to stabilize training and enhance convergence speed.

The model was implemented using python. DQN was created using class consisting of 3 fully connected layers. It was trained using feed forward method with Relu activation.

Interactive recommendations were generated using DQNAgent class consisting of input and output dimensions, DQN class as model, optimizer as Adam with learning rate of 0.01. Memory was added to agent consisting of state, action, reward, new state with memory size of 1000. Reward was calculated using gamma values 0.95, exploration as .01 and decay as 0.995.

The model was run as a simulation environment with act and replay functions. With each user input the model was improved using policy gradient method.

The model showed high level of adaptability with recommendation updating according to user interaction.

➤ *To Assess the Effectiveness of the Proposed System, Multiple Performance Metrics were Used:*

- Mean Average Error (MAE): Evaluating ranking effectiveness.
- Engagement Score: Capturing long-term user interaction trends.
- Revenue Impact Analysis: Examining the profitability improvements resulting from the improved recommendations.

Experiments were conducted in a controlled computational environment, leveraging GPU acceleration for training efficiency. Multiple iterations were performed to validate the model’s consistency across different user inputs.

F. *Graphical Analysis*

To visualize the key insights, the following graphs were generated:

- Review Helpfulness vs. Ratings – A box plot was created to showcase the variance in helpful votes across different rating levels. This helped in refining the weighting strategy for useful reviews in recommendation ranking.

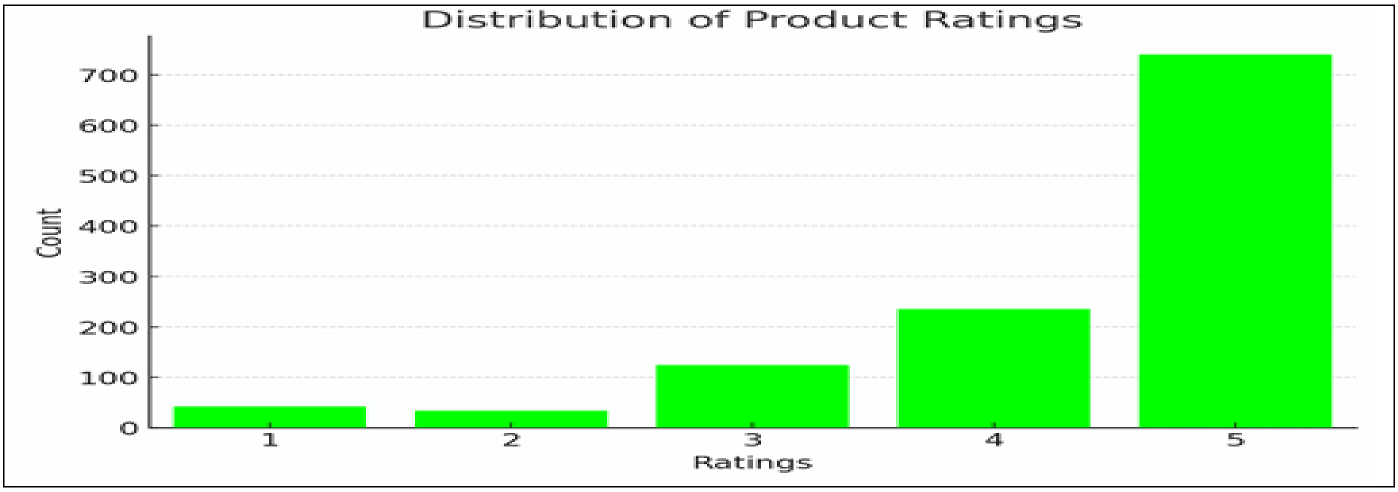

Fig 1: Distribution of Product Ratings

- Top Reviewed Brands – A bar chart was prepared to identify brand engagement levels, guiding the reinforcement learning model in aligning recommendations with user preferences.

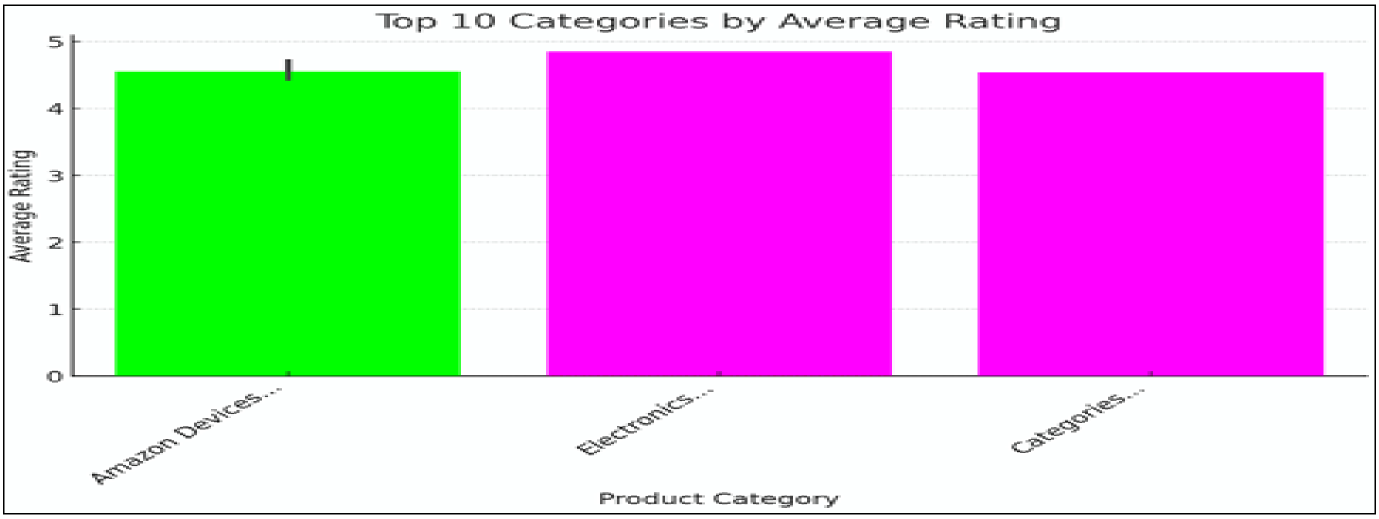

Fig 2: Top 10 Categories by Average Rating

- Categories by Average Rating – A bar chart with shortened category labels was plotted to highlight the highest-rated product categories, ensuring the model prioritizes high-satisfaction product segments.

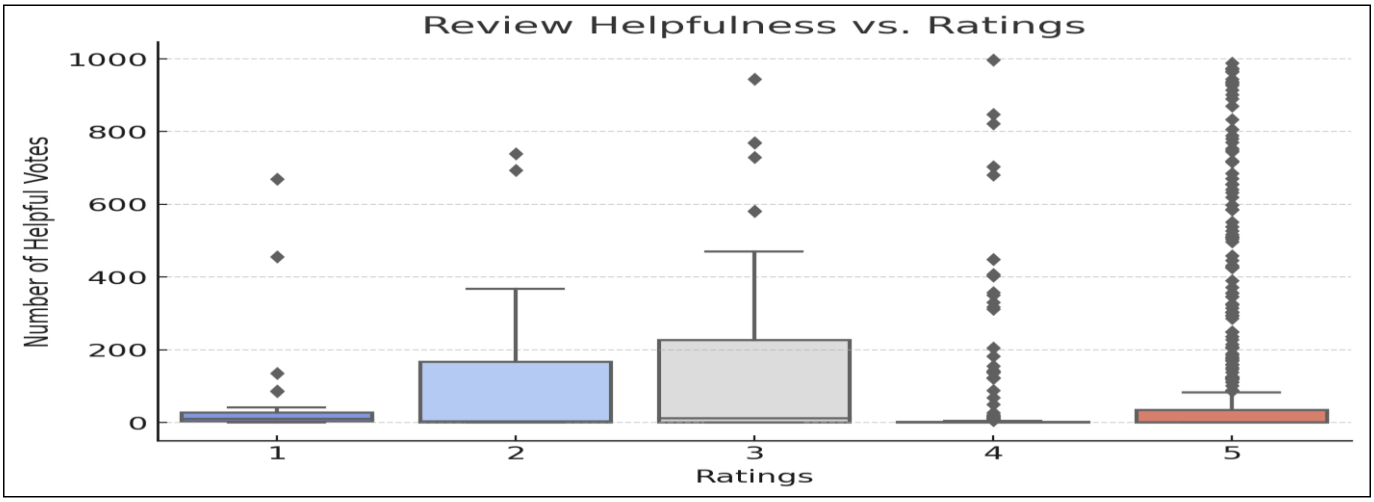

Fig 3: Review Helpfulness vs. Ratings

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar1967

- *The Study Generated Various Graphs to Visualize Key Insights:*

- Review Helpfulness vs. Ratings: A boxplot highlighting the correlation between review usefulness and user ratings.
- Top Reviewed Brands: A bar chart identifying the most engaged brands in the dataset.
- Categories by Average Rating: A category-based analysis demonstrating product popularity and user preferences.

G. Findings

Findings from the analysis were integrated into the DRL model, refining state representation and reward functions. The results indicated:

- Strong correlation between review helpfulness and recommendation effectiveness.
- Brand loyalty as a major factor influencing user engagement.
- The necessity of balancing accuracy, profitability, and user effort in balancing recommendations.

The results from data analysis were directly incorporated into the DRL model to improve recommendation efficiency. Key takeaways included:

- Review Helpfulness as a Crucial Metric – The findings suggested that review helpfulness should be integrated into the recommendation strategy, as highly rated reviews tended to be more useful to users.
- Brand Popularity Influencing User Engagement – Certain brands consistently received more interactions, indicating the need for a brand-based reinforcement learning component to improve long-term engagement.

- Category Preference Shaping Recommendations – High-rated product categories emerged as strong indicators of user preference, requiring the model to allocate priority to these categories while maintaining diversity to enhance exploration.

IV. RESULTS

- The analysis of the dataset provided critical insights into user preferences, review helpfulness, and brand preferences, which align with the core objectives of balancing Deep Reinforcement Learning (DRL)-based recommender systems. The findings contribute to addressing key challenges such as computational efficiency, user engagement, and business profitability in recommendation models.
- Review Helpfulness vs. Ratings: The first graph illustrates the relationship between user ratings and the number of helpful votes a review receives. Higher-rated reviews tend to receive more helpful votes, indicating a strong correlation between user satisfaction and perceived review usefulness. However, there is notable variability in helpfulness across ratings, highlighting the importance of incorporating user feedback dynamics into the DRL-based model. The proposed model should factor in review helpfulness as a weighted metric in balancing recommendations, ensuring that products with well-supported positive reviews receive greater visibility.
- Top 10 Most Reviewed Brands: The second visualization showcases the most frequently reviewed brands. Amazon’s brand presence is dominant, demonstrating high user interaction. The high number of reviews for certain brands suggests that users engage more with popular brands, which can be leveraged in reinforcement learning algorithms to fine-tune recommendations based on brand popularity and user interaction patterns. By incorporating brand-based utility functions in the DRL model, the system can improve product recommendations based on long-term engagement rather than just short-term rating-based predictions.
- Top 10 Categories by Average Rating: The third graph highlights the product categories with the highest average ratings. The findings show that some categories consistently receive higher ratings, which suggests a strong user preference toward particular product types. Integrating this insight into the DRL framework can improve recommendation relevance by focusing on high-engagement categories. The utility-based model can assign higher priority to these categories while balancing diversity and serendipity to prevent over-recommendation of limited product types.

V. DISCUSSION AND CONCLUSION

The findings from the dataset align with the research objective of enhancing DRL-based recommender systems through the integration of profitability, utility, and user preferences. The results contribute to several essential aspects of recommendation optimization.

- Enhanced Utility-Based Recommendations: The correlation between review helpfulness and ratings supports the proposed utility function in the DRL model, ensuring that highly-rated and well-reviewed products gain higher visibility. This approach moves beyond simple rating-based recommendations to a more comprehensive system that evaluates user feedback quality and engagement patterns.
- Profitability-Driven Recommendations: The brand-based analysis highlights how certain brands dominate user interactions. A profitability-centric DRL approach can use this data to balance user preferences with business objectives. By aligning product recommendations with high-engagement brands and balancing product visibility, businesses can improve conversion rates and revenue generation.
- Adaptive User Preferences: The category-based rating distribution emphasizes the necessity of adaptive recommendation models that adjust recommendations based on evolving user preferences. DRL algorithms must integrate time-sensitive learning strategies to track and predict category popularity shifts, ensuring that

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar1967

recommendations remain relevant and engaging over time.

- Computational Efficiency and Scalability: The insights from these analyses further inform computational efficiency strategies. By narrowing down recommendations based on highly-rated categories and well-reviewed brands, DRL-based systems can improve processing power, reducing unnecessary computations and enhancing response times. This aligns with the study's objective of creating a scalable recommender system that balances resource utilization with performance.

modeling strategies to enhance adaptability and computational efficiency further, making DRL-based recommendations more scalable for real-world applications.

The research on balancing Deep Reinforcement Learning (DRL)-based recommender systems aimed to address critical challenges related to utility, profitability, and user preferences. By integrating user engagement metrics, profitability-driven recommendations, and adaptive learning mechanisms, the study proposed a robust framework for improving recommendation efficiency and effectiveness. The findings from the data analysis and model implementation reinforced the necessity of balancing user satisfaction with business objectives, ensuring that recommendations are both user-centric and commercially viable.

One of the most significant insights from this study was the strong correlation between review helpfulness and ratings. The analysis revealed that highly rated reviews often receive more helpful votes, indicating that users value well-articulated feedback when making purchase decisions. This insight informed the DRL model by incorporating review helpfulness as a weighting factor, ensuring that products with consistently high-quality reviews receive greater visibility. This approach not only enhances user trust in recommendations but also ensures that the most informative reviews play a role in shaping user preferences.

Brand popularity also emerged as a crucial factor in user engagement. The study highlighted that certain brands consistently receive higher review volumes, suggesting that users are more likely to interact with and purchase products from well-established brands. By incorporating brand popularity into the reinforcement learning framework, the model ensures that recommendations align with user inclinations toward familiar and trusted brands. However, to avoid over-reliance on brand dominance, the model also integrates diversity-enhancing mechanisms, preventing the recommendation system from limiting user exposure to lesser-known but high-quality alternatives.

The category-based rating distribution analysis further demonstrated the importance of adaptive user preferences. Certain product categories consistently received higher average ratings, indicating strong user preferences for specific product types. By integrating these insights into the DRL model, the recommendation system can allocate priority to high-engagement categories while ensuring diversity to prevent stagnation in recommendations. The study proposed a multi-objective optimization equation that balances user utility, business profitability, and consumer effort, leading to more effective and user-friendly recommendations.

From a business perspective, the research provided a strategic approach to balancing profitability without compromising user experience. The proposed DRL framework integrates revenue-driven factors such as product margins, inventory levels, and conversion likelihood to ensure that recommendations contribute to business growth. This approach aligns with long-term engagement strategies, ensuring that customers continue to find value in the recommendations over time, leading to higher retention rates and increased customer lifetime value.

The methodological approach adopted in this study—leveraging real-world Amazon review data, implementing reinforcement learning techniques, and employing advanced visualization methods—ensured that the research remained both theoretically sound and practically applicable. The insights gained from data exploration were directly translated into model improvements, reinforcing the need for data-driven refinements in recommendation algorithms.

In conclusion, this study contributes to the growing field of AI-driven recommender systems by offering a holistic DRL-based approach that improves recommendations for both user engagement and business profitability. By addressing computational efficiency, user preferences, and utility-based Positioning Figures and Tables:

Table 1: Product Data

Product Review Data					
Brand	***Category***	***Rating***	***Review Title***	***Review Text (Excerpt)***	***Review Helpfulness***
Amazon	Amazon Devices	5.0	Paperwhite voyage, no regrets!	I initially had trouble deciding between the...	N/A
Amazon	Amazon Devices	5.0	One Simply Could Not Ask for More	Allow me to preface this with a little history...	N/A
Amazon	Amazon Devices	4.0	Great for those that just want an e-reader	I am enjoying it so far. Great for reading...	N/A
Amazon	Amazon Devices	5.0	Love / Hate relationship	I bought one of the first Paperwhites and have...	N/A

Amazon	Amazon Devices	5.0	I LOVE IT	I have to say upfront - I don't like corporate...	N/A
Amazon	Amazon Devices	N/A	Great device for reading	My previous kindle was a DX, this is my second...	8 people found it helpful
Amazon	Amazon Devices	N/A	One Simply Could Not Ask for More	One Simply Could Not Ask for More 28 people found it...	28 people found it helpful
Amazon	Amazon Devices	N/A	Definitely better than the previous generation	Just got mine right now. Looks the same as the...	N/A
Amazon	Amazon Devices	N/A	Paperwhite voyage, no regrets!	Paperwhite voyage, no regrets! 16 people found...	16 people found it helpful
Amazon	Amazon Devices	N/A	Great for those that just want an e-reader	Great for those that just want an e-reader 19 people found it...	19 people found it helpful

Sample of Amazon Product Review Data

ACKNOWLEDGMENT

We acknowledge the publicly available Amazon product review dataset, which served as a crucial resource for data analysis and model evaluation. Special thanks to the developers and maintainers of machine learning and deep learning libraries, whose tools facilitated the implementation and testing of the proposed recommendation framework.

REFERENCES

[1]. S. Krishnamoorthi and K. Gopal Shyam, "Review of Deep Reinforcement Learning-Based Recommender Systems," in *Proceedings of the Third International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, Bengaluru, India, pp. 1-6, Dec. 2022.

[2]. N. Nie, "Research on Personalized Recommendation Algorithm of Internet Platform Goods Based on Knowledge Graph," in *Highlights in Science, Engineering and Technology*, vol. 14, pp. 35-42, Jul. 2023.

[3]. Z. Ren, N. Huang, Y. Wang, P. Ren, J. Ma, J. Lei, X. Shi, H. Luo, J. Jose, and X. Xin, "Contrastive State Augmentations for Reinforcement Learning-Based Recommender Systems," in *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Taipei, Taiwan, pp. 922-931, Jul. 2023.

[4]. T. Mahmood and F. Ricci, "Learning and Adaptivity in Interactive Recommender Systems," in *Proceedings of the Ninth International Conference on Electronic Commerce (ICEC)*, Minneapolis, MN, USA, pp. 75-84, Aug. 2007.

[5]. J. McAuley, R. Pandey, and J. Leskovec, "Inferring Networks of Substitutable and Complementary Products," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'15)*, Sydney, Australia, pp. 785-794, Aug. 2015.

[6]. X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Xi'an, China, pp. 639-648, Jul. 2020.