2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS) | 978-1-7281-9537-7/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICAIS50930.2021.9395812

A Review on Reinforcement Learning based News Recommendation Systems and its challenges

Sindhuja Bangari[1] , Dr. Shantharam Nayak[2], Ladly Patel[3], Rashmi K T[4]
[1,3,4] Assistant Professor, CSE Dept., Sri Krishna Institute of Technology[SKIT], Bangalore-560090, India.
[2]Professor, SKIT, Bangalore-560090, India.

Abstract— Recommendation systems are helpful in both business perspective and user day to day life. These days online contents are generated in huge amount and due to this, users need a special recommendation application namely personalized News Recommendation, and it is highly challenging due to its dynamic nature. Therefore, getting a suitable and relevant news article for a user is difficult task. To address the above challenge Reinforcement Learning algorithms plays crucial role because these algorithms very much helpful in dealing with the dynamic environment and large space. This paper reviews the different Reinforcement algorithms namely Deep Q-learning network (DQN), Deep Deterministic Policy Gradient (DDPG) and Twin Delayed DDPG (TD3) to develop the news recommendation system and also mentioned the challenges faced by the reinforcement recommendation systems. In this study it was found that TD3 is best suited to develop the news recommendation system.

Keywords— Reinforcement Learning, Recommendation system, News Recommendation system, Challenges, Dynamic environment

I. INTRODUCTION

Recommendation systems are on-the-fly in E-commerce applications. They are helpful to the user to get their desired information by proposing similar products, services or relevant . Recommendation systems have become popular these days and have been used in different kinds of application areas including robots, movies, music, books, toys, health care, games. Number of techniques have been proposed to deal with the recommendation systems. Traditional techniques namely collaborative filtering, content-based filtering and hybrid methods. The above-mentioned methods are static and can not handle the sequential nature of user interaction with the system. This issue resolved by using Reinforcement Learning [15]. The nature Reinforcement Learning techniques can be applied to the problems with large state and action space which includes industry automation [16], self-driving cars [17], healthcare [18] and recommendation systems. This review focused on recommendation systems specifically based on reinforcement Algorithms. Reinforcement Learning used by the major companies to recommend their products to the users. The use of Reinforcement Learning not limited to industry but also in academia [15].

Reinforcement learning is a type of machine learning that is concerned with how an agent perform actions within an environment to get maximum reward. The computer interacts with the environment and learn from the experience and then predicts result. The computer does trial and error method to find a solution for a given problem.

Some of the important terms used in reinforcement learning:

Agent: Agent always try to maximizes the reward by taking input from the environment. That is, agent collect observation and reward from the environment and dispatch action on the environment

Environment: This is the place where, based on given policy agent will take action. For example, News articles and users are the environment for an agent.

Action(A): The agent performs actions to get a reward in an environment

State(S): The state is an immediate situation in which agent finds itself in relation to the other things in the surroundings of the environment such as tools, enemies.

Reward(R): Feedback returned by environment based on the previous interaction between agent and environment.

Policy(π): An agent uses a strategy to decide the action which should be taken on basis of present state. Basically, agents map states to actions. So that, it decides the actions which are giving the highest rewards with regards to states.

Reinforcement learning diagram interaction shown with environment and agent is shown in figure 1.

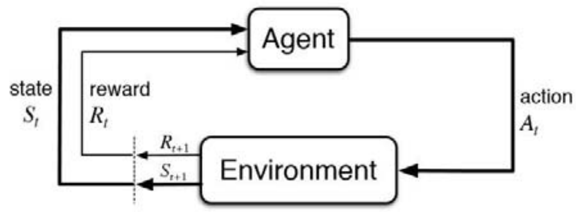


Figure 1: Reinforcement Learning Environment

Figure 2 shows the various learning methods that are available in Reinforcement Learning technique.

Reinforcement algorithms are categorized into two ways

i. Model-based methods
ii. Model-free methods

Model-based Reinforcement Learning:

In this learning technique, agent estimates the optimal policy by using the transition function and reward function.These techniques tries to predict future states.

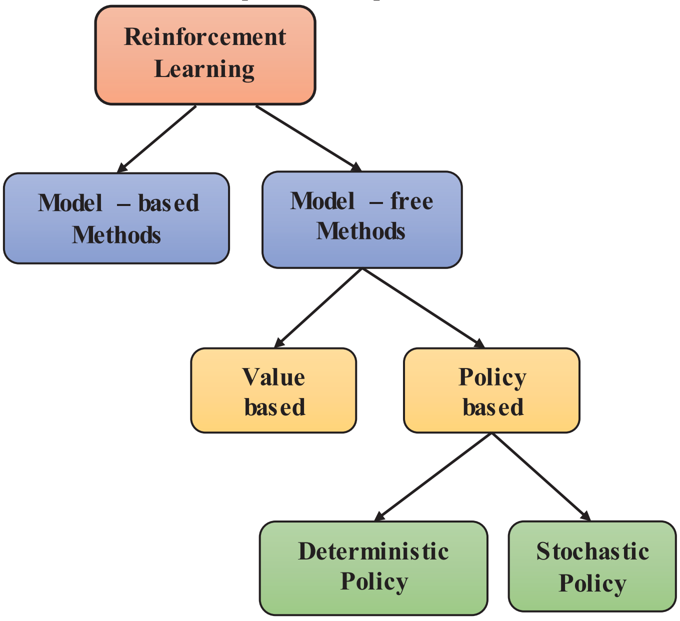


Figure 2: Reinforcement Learning Methods

Model-free Reinforcement Learning:

This learning method builds optimal policy without trying to estimate the actions in between states and reward functions.No future states or reward predicted.The actor-critic approach is one of example of Model- free Reinforcement Learning.

These days Recommendation system are more popularized. Over the last decade, there is an increase in huge volume of data and the data is shared through the smartphones. The user has been messed up with much information and does not allow to be aware of his options because of this user wanted to personalize his options. The solution to this problem is recommendation systems which allow user to get relevant information according to user desires and needs.

Definition of recommendation systems:

The Aim of recommendation system is to produce semantically relevant recommendations to a group of users such as items or products, which the user shows interest [7]. For example, suggestion for music on gaana, Movies on prime video or aha and YouTube are the most popular examples for business-oriented recommendation systems. The foremost characteristic of recommendation system in order to Develop the recommendation engine is Data availability. How to get this data? This data is generated by user providing feedback about the product whether user likes the product or dislike the product or rating given to that particular product.

Recommendation systems are divided into three main categories namely: collaborative, content-based, and hybrid filtering [8]. Collaborative Filtering is further classified into three subcategories: memory based, model based and hybrid collaborative filtering [9].

Though recommendation system is applicable to various domains such as recommending books, videos, music, toys, movies etc. This research mainly concentrated about the news recommendation system. The most challenging problem is how accurately select news article from large databases to recommend individual readers. This leads to the development of personalized news recommendation [10].

This Paper organized in following way as section II describes about Literature survey, section III explains the various algorithms utilized in development of news recommendation system, section IV shows the challenges faced by the recommendation system, conclusion and future scope is followed in section V.

II. RELATED WORK

Recommendation System play crucial role in the business perspective and gained more popularity as it trying to recommend user desired data. This paper studies the various papers and explored recommendation systems based on different Reinforcement Learning techniques. The below table shows the detailed data about different algorithm applied on various recommendation applications

Author	Algorithms	Findings
Choi, Sungwoon, et al. Year: 2018	Bi clustering Technique	The Objective of the collaborative filtering technique is to predict the personalized recommendation system by finding the similarity of interests between the users. The demerit of collaborative filtering technique is that data sparsity. This demerit led to the development of bi-clustering technique, which reduces the space and improve the recommendation quality. In this approach author used four methods 1. State construction 2. Q-Function learning mechanism 3. Recommendation generation 4. Update state space [2].
Feng Liu Year 2018	DRR frame work and Actor-Critic	In this paper, the author proposed DRR frame work which is based on actor critic network. The frame work includes DRR-p, DRR -u DRR- ave, all the above mentioned models improves pairwise dependencies and user- item interactions[3]. Merits: 1. This research overcome the problem the dynamic interactive nature between users and recommender system 2. Improve the long-term rewards
Lixin Zou, Year: 2019	Q-Network and S-	In this paper the author focused on the problem of existing Supervised algorithm that is in the traditional supervised algorithms learning target is

	Network	not available because of this user unable to stick with the website much time. To improve this the author chosen the Reinforcement Learning, and this technique maximize the long-term rewards, But the author finds that user behavior still challenging one, the author proposed Q-network and S-network to attract the user and make the user active [4]. Merits: Long Term User engagement
Xinshi Chen Year 2019	Cascading DQ-network, mini-max optimization	In this paper, the author focused on the problems of reward function and environment dynamics. So he proposed a generative adversarial network to know the user behavior and used mini-max framework to achieve the reward function [5]. Merits: various combination of action space and dealing with large numbers of items.
Guanjie Zheng Year 2018	Deep Q-Learning	In this paper, the author mentioned that online personalized news recommendation system is challenging task because of news features and user preferences. To address this challenge the author proposes a Deep Q-learning recommendation frame work in order to build the future reward explicitly [6]. Merits: User preferences and dynamic news feature
Jiahui Liu Year 2010	Bayesian framework	In this paper, the author developed personalized news recommendation system in Google News based on user log information[11].
Qingyun Wu, Hongning Wang Year 2017	r^2Bandit	In this paper, Recommendation system make the user in active mode for longer time. To address this issue the author proposed bandit-based solution where he concentrated on user clicks. The proposed solution predicts the expected user return time [19]. Merits: User clicks/ logs are beneficial in optimizing the long-term reward from users
Yong Liu 2019	Actor-Critic Network DDPG	In this paper, the author focused on diversity of recommendations. He proposed a novel model, Diversity - promoting Deep Reinforcement Learning and also Determinantal Point Process (DPP) model to produce the wide variety but similar recommendation. The relevance is achieved through the actor-critic network [20]. Merits: Actor-critic reinforcement learning is helpful in achieving the similarity between items dynamically
Yu Lei 2019	BUDQN	The author worked on explicit feedback recommender system and proposed user specific deep Q-learning method to know the optimal policies and then proposed Biased UDQN to model the explicit feedback recommender system [21].
Xiangyu Zhao Year:2018	Actor-critc network and DDPG	In this paper the author described that the Q-learning and POMDP algorithms not efficacy when the number increased in item recommendation, so he used actor-critic framework and DDPG to address page wise recommendation problem [1].
Saravanapriya Manoharan Year : 2020	Fuzzy logic	In this paper, the author found out that user interest categories changes over time due to this Recommendation of relevant news article is challenging one. To overcome this the author proposes fuzzy logic approach for predicting user interest by understanding their implicit user profile [12].
Isshu Munemasa 2018	Multi layer neural network	In this paper, the address the data sparsity problem which is same collaborative filtering. So, to solve this issue author applied a deep reinforcement learning based multilayer neaural network approach[22].
Xiangyu Zhao, Year 2017	Actor-Critic framework and LIRD	In this paper, the author address the issue of continuous improvement in the user interactions and the author used reinforcement learning technique actor-critic to build the interactions between user and recommendation system [23].
Jing Zhang Year 2019	Hierarchical Reinforcement Learning	In this paper, the author addressed the problem that, the user interested in many different courses, so the user attention is not fixed. To address this issue the author proposed the hierarchical reinforcement learning algorithm to revise user profiles and attract towards the target course.
S.Manoharan, Year 2020	K-clique	The author used k-clique in recommender system to improve the accuracy and preciseness in deep learning classifier [25].

III. PROPOSED WORK

To achieve the success and profits in the respective domains recommendation system plays major role. With the usage of smart phones there is huge growth in the online content, and attention has been given to the one of major application of recommendation systems that is news recommendation System. Traditional news recommendation system methods are classified into three categories.

Content-based Methods: This is used term frequency and user profile based on previous data. In this method recommender extract the news based on user profile

Collaborative Filtering Methods: In this method ratings given by the user or similar users will be collected

Hybrid Model: In this method based on collected information in collaborative filtering method the user profile modelling were build.

Reinforcement learning models are giving better performance when compared to the above-mentioned methods because of its capability of modelling user and item relationship [6].

Reinforcement Learning Recommendation System Architecture illustrated in figure 3

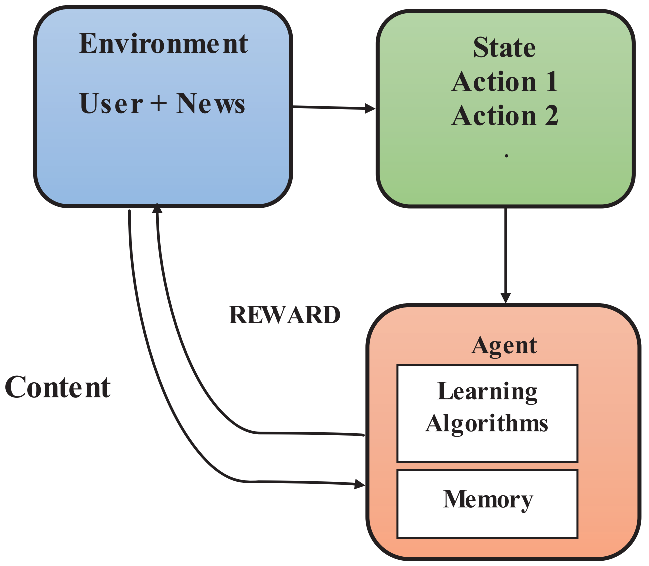
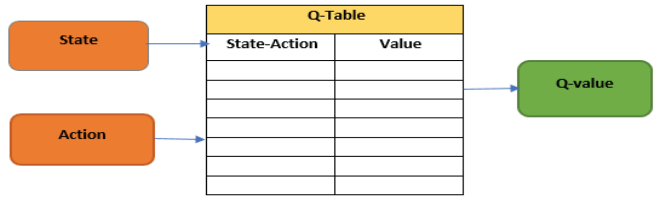


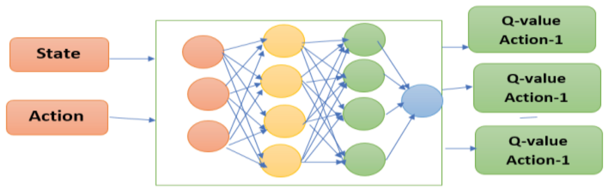
Figure 3: Reinforcement Learning Recommendation System Architecture

This paper concentrated on news recommendation; news article is dealing with the high dimensionality. So, DQN can be suggested to the news recommendation system.

Deep Q-Learning Network: DQN is the model-free and value-based algorithm. DQN technique is a combination of Q-learning and neural networks. Neural networks helpful to approximate the Q-Value function, these Q-values helps in finding the next action resulting in state of the highest quality. The figure 4 illustrates the computation of Q-value between Q-learning and DQN



Q-Learning



DQN

Figure 4: Computation of Q-value in Q-learning and DQN

The key features of DQN algorithm are

i. DQN is having memory buffer which stores the past experience.
ii. The next action is determined by the greatest output of Q-Network.
iii. The loss value minimized compared to the Q-learning [27].

Deep Deterministic Policy Gradient (DDPG): DDPG is an algorithm simultaneously learns a Q- function and a policy. It is off-policy based algorithm and it uses bellman equation to know the value of Q- function and uses Q- function to know the policy. This Combines two technique DPG (Deterministic Policy gradient) and DQN (Deep- Q Networks) respectively. DDPG work over continuous action space [28].

It also uses experience replay from DQN This algorithm is based on Actor critic network.

1. Actor- It introduces an action based on a state
2. Critic- It predicts the positive action value or negative action value based on the state and action

The given below figure 4 shows the Actor-Critic Network architecture.

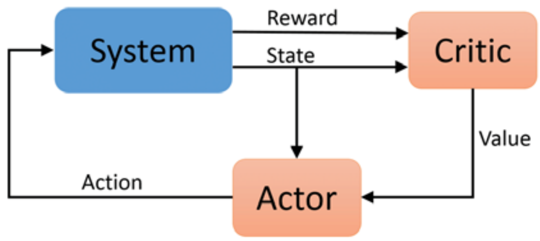


Figure 5: Actor- Critic Network

Twin Delayed DDPG (TD3): DDPG algorithm overestimate the Q values(rewards) to overcome this issue TD3 algorithm can be used. The above-mentioned issue is resolved by using three major features

i. Clipped Double- Q Learning: TD3 learns two Q-functions instead of one so it called as twin and uses minimum of the two Q -values to form the targets

ii. Delayed policy updates: TD3 update the policy less frequently than the Q –function because Training of agent diverges when poor policy is overestimated. Then agent policy will continuously be getting worse as it is updating on states with a lot of error.

iii. Target Policy Smoothing: TD3 adds noise to the target action, to make it harder for the policy to exploit Q-function errors by smoothing out Q along with changes in action.

IV. CHALLENGES

Reinforcement learning gains huge success in games, robotics, physical control and computational process although it achieves less success in user centric applications like online advertising and recommendation systems. This mentioned Some of the challenges faced by recommendation systems.

1.Cold Start: It's always difficult to recommend products for new customers not just for customers but also for what to suggest as new products. For example, if a user joins newly in the recommendation website and doesn't have any information such as previous search history and log data, then no data about the user to recommend items. This problem more serious in news recommendation system as user newly joined and doesn't get the any news article related to his desired data because as there is no data available.

2. Data Sparsity: As there is cold start problem, to get the new data is very difficult. This leads to the data sparsity. As there is a less data that is without clicks and explicit user feedback recommending user interest news articles difficult task.

3.Scalability: In the new Age of internet world data is growing like anything. When number of users increasing for a website, the size of stored user behavior increases to Tera Bytes per day. This is very challenging one in news recommendation system because of some thousands of news articles will be published from various sources. So recommendation system should have the capacity to deal with huge amount of data such as fast processing. To retain the users with websites, recommendation system must interactive with this huge amount of data. But it is very difficult for recommendation system to recommend the data which is user specific. This leads to the issue of scalability.

4.User-centric Data: In any of the Recommendation application it is very difficult recommend the user desired data because of continuous changes in the interest of the user. This problem very intensive in news recommendation system because user interest may vary and cannot predict user likes only particular category of news say, user may interest in comics, sports, business, filmy, agriculture or vide variety of news categories. This is very challenging one in recommending news to the user.

5. Explicit-user Feedback: It is difficult to know the user interest unless user provide some feedback about the data. Sometimes user may or may not provide the feedback due to this recommending a user desired product becomes difficult one. For example, when a news article read by the user, he should provide the feedback without getting feedback from the user it is difficult to recommend the news articles which are interested by the user. If the user does not provide any feedback like whether he like the news or not, then it is difficult to get user feedback about the news [26]. So implicit feedback may help to recommend other news, e.g. clicks or session on the news article

V. CONCLUSION & FUTURE SCOPE

This study & analysis mainly concentrated on News Recommendation systems based on reinforcement learning. Reviewing the different reinforcement algorithms namely DQN, DDPG and TD3 used to develop the news recommendation systems with the comparison of traditional recommendation algorithms and also presented the challenges faced by the reinforcement recommendation system. Though DQN is better in computing the Q-values which helps in getting a news article to be recommended but it fails to work in continuous action space. DDPG algorithm overcomes this drawback as DDPG works only in continuous action space and it is policy-based algorithm. So, it is easy to get the rewards by directly computing the gradient of the expected reward using the gradient policy algorithm. The drawback of DDPG is that it overestimates the Q-values due to this policy is going break.TD3 is an extension of DDPG and it overcomes the issue presented in DDPG by its key features. So, in this study it is found that TD3 is best algorithm to apply for news recommendation system. It is proposed to implement TD3 algorithm and also measure the performance of algorithm with respect to the news recommendation system.

References

[1] Zhao, Xiangyu, et al. "Deep reinforcement learning for page-wise recommendations." Proceedings of the 12th ACM Conference on Recommender Systems. 2018.

[2] Choi, Sungwoon, et al. "Reinforcement learning based recommender system using biclustering technique." arXiv preprint arXiv:1801.05532 (2018).

[3] Liu, Feng, et al. "Deep reinforcement learning based recommendation with explicit user-item interactions modeling." arXiv preprint arXiv:1810.12027 (2018).

[4] Zou, Lixin, et al. "Reinforcement learning to optimize long-term user engagement in recommender systems." Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019.

[5] Chen, Xinshi, et al. "Generative adversarial user model for reinforcement learning based recommendation system." International Conference on Machine Learning. PMLR, 2019.

[6] Zheng, Guanjie, et al. "DRN: A deep reinforcement learning framework for news recommendation." Proceedings of the 2018 World Wide Web Conference. 2018.

[7] Melville P., Sindhwani V. (2011) Recommender Systems. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA.

[8] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," in IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734-749, June 2005, doi: 10.1109/TKDE.2005.99.

[9] Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." 9th International Symposium on Computer Music Modeling and Retrieval. Vol. 4. 2012.

[10] Li, L., Wang, DD., Zhu, SZ. et al. Personalized News Recommendation: A Review and an Experimental Investigation. J. Comput. Sci. Technol. 26, 754 (2011).

[11] Liu, Jiahui, Peter Dolan, and Elin Rønby Pedersen. "Personalized news recommendation based on click behavior." Proceedings of the 15th international conference on Intelligent user interfaces. 2010.

[12] Saravanapriya Manoharan, Radha Senthilkumar, "An Intelligent Fuzzy Rule-Based Personalized News Recommendation Using Social Media Mining", Computational Intelligence and Neuroscience, vol. 2020, Article ID 3791541, 10 pages, 2020. https://doi.org/10.1155/2020/3791541

[13] Lillicrap, Timothy P., et al. "Continuous control with deep reinforcement learning." arXiv preprint arXiv:1509.02971 (2015).

[14] Ray S., Tadepalli P. (2011) Model-Based Reinforcement Learning. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA.

[15] Afsar, M. Mehdi, Trafford Crump, and Behrouz Far. "Reinforcement learning based recommender systems: A survey." arXiv preprint arXiv:2101.06286 (2021).

[16] Meyes, Richard, et al. "Motion planning for industrial robots using reinforcement learning." Procedia CIRP 63 (2017): 107-112.

[17] Sallab, Ahmad EL, et al. "Deep reinforcement learning framework for autonomous driving." Electronic Imaging 2017.19 (2017): 70-76.

[18] Gottesman, Omer, et al. "Guidelines for reinforcement learning in healthcare." Nature medicine 25.1 (2019): 16-18.

[19] Wu, Qingyun, et al. "Returning is believing: Optimizing long-term user engagement in recommender systems." Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017.

[20] Liu, Yong, et al. "Diversity-promoting deep reinforcement learning for interactive recommendation." arXiv preprint arXiv:1903.07826 (2019).

[21] Lei, Yu, and Wenjie Li. "Interactive recommendation with user-specific deep reinforcement learning." ACM Transactions on Knowledge Discovery from Data (TKDD) 13.6 (2019): 1-15.

[22] Munemasa, Y. Tomomatsu, K. Hayashi and T. Takagi, "Deep reinforcement learning for recommender systems," 2018 International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2018, pp. 226-233, doi: 10.1109/ICOIACT.2018.8350761.

[23] Zhao, Xiangyu, et al. "Deep reinforcement learning for list-wise recommendations." arXiv preprint arXiv:1801.00209 (2017).

[24] Zhang, Jing, et al. "Hierarchical reinforcement learning for course recommendation in MOOCs." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. No. 01. 2019.

[25] Manoharan, Samuel. "Patient Diet Recommendation System Using K Clique and Deep learning Classifiers." Journal of Artificial Intelligence 2.02 (2020): 121-130.

[26] CHONG FENG1 , MUZAMMIL KHAN 2 , ARIF UR RAHMAN 3 , AND ARSHAD AHMAD 4 1, News Recommendation Systems - Accomplishments, Challenges & Future Directions

[27] Nguyen, Ngoc Duy, Thanh Nguyen, and Saeid Nahavandi. "System design perspective for human-level agents using deep reinforcement learning: A survey." IEEE Access 5 (2017): 27091-27102.

[28] Lillicrap, Timothy P., et al. "Continuous control with deep reinforcement learning." arXiv preprint arXiv:1509.02971 (2015).