HHAI 2024: Hybrid Human AI Systems for the Social Good
F. Lorig et al. (Eds.)

doi:10.3233/FAIA240185

Human-Centered AI for Dementia Care: Using Reinforcement Learning for Personalized Interventions Support in Eating and Drinking Scenarios

Wen-Tseng CHANG [a], Shihan WANG [b], Stephanie KRAMER [a], Michel OEY [a], Somaya BEN ALLOUCH [a,c,1]

[a] *Amsterdam University of Applied Sciences*
[b] *Utrecht University*
[c] *University of Amsterdam*

Abstract. For people with early-dementia (PwD), it can be challenging to remember to eat and drink regularly and maintain a healthy independent living. Existing intelligent home technologies primarily focus on activity recognition but lack adaptive support. This research addresses this gap by developing an AI system inspired by the Just-in-Time Adaptive Intervention (JITAI) concept. It adapts to individual behaviors and provides personalized interventions within the home environment, reminding and encouraging PwD to manage their eating and drinking routines. Considering the cognitive impairment of PwD, we design a human-centered AI system based on healthcare theories and caregivers' insights. It employs reinforcement learning (RL) techniques to deliver personalized interventions. To avoid overwhelming interaction with PwD, we develop an RL-based simulation protocol. This allows us to evaluate different RL algorithms in various simulation scenarios, not only finding the most effective and efficient approach but also validating the robustness of our system before implementation in real-world human experiments. The simulation experimental results demonstrate the promising potential of the adaptive RL for building a human-centered AI system with perceived expressions of empathy to improve dementia care. To further evaluate the system, we plan to conduct real-world user studies.

Keywords. reinforcement learning, intelligent home environment, dementia, human simulator, adaptive intervention, human-centered AI

1. Introduction

Dementia is a progressive neurodegenerative disorder that affects the elderly and diminishes their cognitive functions [1]. As the population ages and caregiver shortages grow, supporting people with early dementia (PwD) to live independently becomes crucial [2]. One of the key challenges is maintaining their circadian rhythm, especially for eating and drinking. Forgetting to eat can lead to serious health concerns. However, it is impos-

[1]Corresponding Author: Somaya Ben Allouch, s.ben.allouch@hva.nl

sible for caregivers to constantly monitor whether PwD eat enough food. Therefore, an intelligent system integrated with the home environment that can regularly remind and monitor PwD to eat at appropriate times becomes essential. Intelligent home technologies have emerged as promising solutions to enhance the living environment for PwD. However, existing systems often focus on activity detection [3, 4], without considering personalized interventions to provide adaptive supports for eating and drinking scenarios. Meanwhile, previous studies [5, 6] suggest that adaptive interventions are important for PwD, as their needs and body conditions differ from person to person. For instance, visual reminders may not be effective for those with visual impairments. An inappropriate reminder could be frustrating because it shows a lack of empathy for the user's individual needs.

Inzlicht et al. [7] highlight that AI has the potential to offer perceived unconditional empathy without bias, possibly reducing human compassion fatigue, particularly in routine tasks. Therefore, inspired by the just-in-time adaptive intervention (JITAI) [8] design concept, which aims at adapting to an individual's changing internal and contextual state to provide the right type and amount of support, we introduce the *Who takes care* artificial intelligence (AI) system. Such a personalized system could adapt to the behavior of the individual user and provide a suitable set of signals to nudge him/her to go to eat. Several unique difficulties were encountered while developing this system. Firstly, as our target users often have cognitive impairments, this can lead to a lack of responsiveness to the system, requiring higher robustness of the system for missing or noisy data. Furthermore, since caregivers indicate that preferences among PwD vary widely due to differing health conditions, the AI must be designed to be perceived as empathic by the user, such that the system can be acceptable and easy-to-use by PwD.

To address these challenges, we design the AI system based on literature [9, 10, 11] and caregivers' inputs from surveys and interviews. It contains a three-phase framework to provide personalized intervention for nudging PwD to eat and drink during meal times. The reinforcement learning (RL) techniques are employed and integrated into the system to determine a suitable mix of signals for individual users. Before real-world implementation and experiments, we need to ensure the validation of the RL algorithms and the robustness of the AI systems, to avoid errors or unexpected behaviors for our sensitive target users. We therefore further develop a simulated system prototype. Our preliminary study involves simulation experiments to test and refine the system, including how it performs under diverse PwD's behaviors and how several selected RL algorithms perform under this context. The experiments compare various RL algorithms to identify the most effective one. Additionally, we tested the algorithms' adaptability and tolerance to the possible changing situation of the environment to evaluate their resilience and flexibility for practical scenarios. This paper aims to develop an AI system that could ensure interventions are not only personalized but also adaptive to individual behaviors and preferences, optimize the *Who takes care* system in sending right type of signal combinations at the right timing, thus enhancing PwD's eating and drinking routines in a supportive, human-centric manner.

2. Related Works

Previous studies have investigated the potential of intelligent environment for healthcare to assist PwD in maintaining independent living at home. Several home care technologies are integrated, including using AI for activity recognition (AR) [12], or utilizing ambient sensors like passive infrared sensors to monitor the residents' activities. While several studies focus on identifying eating and drinking behaviors [3, 4], others [13, 14] emphasized that relying solely on AR may not fully support PwD. An effective intervention should also remind PwD to go to eat and drink. Moreover, considering the speciality of PwD, the nudging signals should also be designed as personalized and adaptive interventions to fill the gap between the current intelligent home environment and the needs of PwD.

Reinforcement learning (RL) has proven to be an effective and efficient approach for building personalized systems that adapt to individual user behavior and preferences; however, building an application in intelligent environments for PwD introduces new challenges. While several studies [15, 16, 17] have successfully employed RL for personalization in health applications, the scenario could be too complex for users with cognitive impairments who might exhibit a lack of responsiveness. Our project focuses on developing an RL-based system for users with early-stage dementia, particularly to assist in the eating and drinking scenario. By integrating caregiver insights, which we gathered from surveys and interviews, we seek to bridge the gap between technological capability and human-centered design, ensuring that the system is responsive, human-centric, and adaptable to the diverse and evolving needs of PwD.

3. Intelligent Home System

3.1. System Overview

The intelligent home system consists of signal-devices, sensors and an AI module that analyzes the data coming from the sensors and based on them determines which signal-devices to turn on and off. The AI module learns to deliver personalized nudging signals based on user behaviors. We adopted a 'three-step interaction' approach [6] into a 'three-stage escalated eating scenario' with increasing intensity of signals. As shown in Figure 1, the system reminds PwD three times a day with signals, gently guiding them through three stages. At the end of each stage, sensors detect whether the user has started eating. Both automatic and manual detection methods are employed. For the automatic method, we will use motion or vibration sensors attached to the dining table and chairs to determine if there is an eating activity. For the second method, users will be provided with a controller that allows them to manually indicate 'I have eaten' by pressing a 'Yes' button. Based on sensor detection and user feedback, if the user has performed an eating activity, the system stops operation; otherwise, after 10 minutes it escalates to a more intensive type of signal to draw the user's attention. After three iterations, the system stops the current intervention until the next mealtime.

3.2. Reinforcement Learning Method

We formulated our AI module as a contextual multi-armed bandit (CMAB) problem, addressing the unknown preferences of people with dementia in mealtime interventions. Figure 2 represents an overview of the interaction framework, reflecting both the real-world and simulation scenarios. It is worth mentioning that in real-world scenarios,

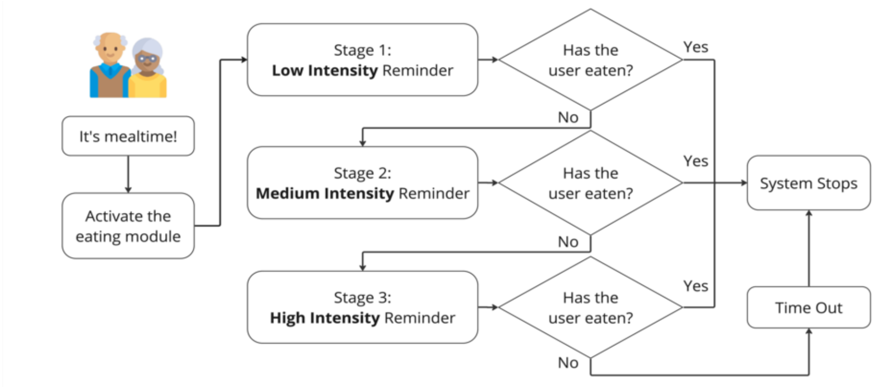


Figure 1. 3-Stage Escalated Eating Scenario

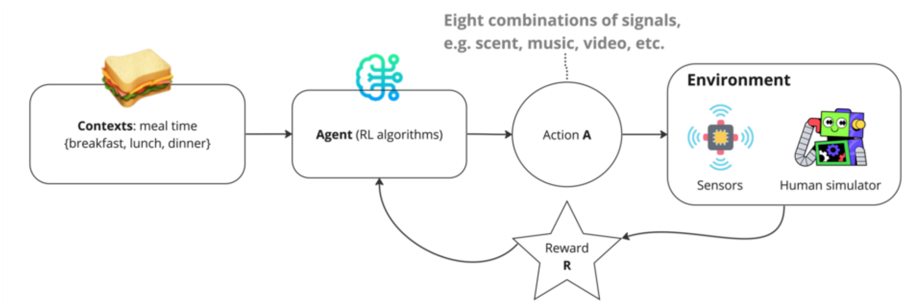


Figure 2. Overview of the AI Simulation System

the simulated *UserBot* with hidden preferences will be replaced by the human target users. They will interact with the system through 'Signal-Based Interaction'. The system will present various interventions at different mealtimes. There are six signals, ranging from low to high intensity, including scent (low), music (low), light (medium), image (medium), voice (high), and video (high), based on prior studies on intelligent home technologies [18, 19, 20]. Our interventions comprise eight distinct signal combinations, each consisting of three signals, one from each level of intensity. At each mealtime, the system suggests a three-signal combination, and employs a three-stage *Escalation Scenario* (as illustrated in Figure 1 for gentle encouragement [6]. In other words, three signals will be delivered one by one with increasing intensity, as described in Section 3.1. We define the key components of RL as follows:

- ***Decision Times***: Set t to index decision time: once per meal, three times per day.
- ***Contexts***: $C_t \in \mathscr{C}$, time of the day (i.e., breakfast, lunch, and dinner time) is indexed by C_t, which indicates the user's context C at the decision time t.
- ***Action***: $A_t \in \mathscr{A}$, actions are indexed by A_t. There are eight actions (i.e., signal combination) in the action space. In every activation time t, the system will choose one action to nudge the user to go eat.
- ***Reward***: $R_t \in \{0, 1\}$. After each trial, the system receives a reward from human simulator. $R = 1$ indicates the user reacted or the sensors detected eating activity; $R = 0$ indicates the user did not react or eating activity was not detected.

Six selected algorithms are implemented into the simulation environment: ε-Greedy action selection [21], Upper Confidence Bound (UCB) [22], Thompson Sampling (TS) [23, 24], Deep Epsilon-Greedy [25], Linear Upper Confidence Bound (LinUCB) [26], and Contextual Thompson Sampling (CTS) [24]. Among them, the first three algorithms cannot take the context of action (i.e., time of meal) into consideration for making decisions, while the other three algorithms can. Comparing the performance of various RL algorithms, we aim to not only identify the best-performed algorithm but also investigate the effectiveness of context in our AI system. In this paper, we only demonstrate integrating CTS in the RL framework in Algorithm 1, while the other algorithms follow a similar structure. Beginning with equal initial beliefs, CTS dynamically adapts its approach based on each meal's context. At decision points, it samples from a Beta distribution, reflecting the probability of each action's success given past experiences, and selects the most promising action (i.e., signal combination) to display. The CTS beliefs are updated after receiving users' responses: positive responses strengthen its belief in the action's success, while negative reactions do the opposite. This continuous learning loop optimizes the nudging strategy and the system eventually displays preferable interventions to the user.

Algorithm 1 Contextual Thompson Sampling Algorithm for Eating Scenario

```
Input: Prior parameters α, β, contexts 𝒞, action space 𝒜
Initialize prior parameters α, β to 1 for each context and action
for each decision time t (three times per day for each meal) do
    Observe the context C_t (time of day: breakfast, lunch, dinner)
    for each action a in 𝒜 do
        Sample θ_a(t) from a Beta distribution with parameters α[C_t][a] and β[C_t][a]
    end for
    Choose action A_t = argmax_a θ_a(t) to nudge the user
    Deliver action A_t and observe reward R_t
    Update α[C_t][A_t] with α[C_t][A_t] + R_t (successes)
    Update β[C_t][A_t] with β[C_t][A_t] + (1 − R_t) (failures)
end for
Return: Updated parameters α, β
```

3.3. Simulation Protocol

Since PwD are particularly sensitive to signals, we introduce a simulation protocol tailored for them, identifying the ethical and practical constraints before deploying the prototype. This protocol employs a human simulator (as depicted in Figure 2), designed to simulate the behaviors and responses of users, allowing us to test our AI system's effectiveness in a controlled, but also realistic environment. The human simulators are built based on the healthcare literature [9, 10, 11], including the potential responses to the intervention. We further consulted with domain experts about possible response behaviors of PwD, ensuring experiments cover different types of users. Following this protocol, the experiment design cannot only primarily ensure the system's effectiveness but also align with the needs of PwD before real-world implementation.

4. Simulation Experiments and Results

We evaluate our designed AI system in multiple simulation experiments. To ensure the simulations are aligned with the real situations, we follow the interaction flow presented in Figure 2 and use different human simulator settings to simulate the diverse behaviors of human users. To assess effectiveness, we conducted 600 simulation trials in each simulation run (i.e., three meals per day for six months) with various user preferences for generalizability. Multiple simulation runs were also conducted and the average results over 1000 simulation runs are presented in this section.

4.1. Experiment 0: Contextual and Non-contextual algorithms

We first compared six RL algorithms in the same simulated environment to study the importance of context-based interventions. Each algorithm interacted with a baseline user model having one random specific preferred signal for each meal (e.g. scent for breakfast, light for lunch and image for dinner). As shown in Figure 3, learning curves and average rewards revealed that contextual algorithms CTS, Deep Epsilon-Greedy, and LinUCB outperformed non-contextual counterparts in both effectiveness and speed of learning. CTS achieved the best performance among the algorithms, reaching an average reward close to 1 within 50 trials (i.e.,about 17 days' intervention) and maintaining stability. Furthermore, CTS demonstrated its robustness and reliability through consistently low standard deviations across multiple runs. In the following experiments, we only present the results of three contextual algorithms, which are the best performing algorithms.

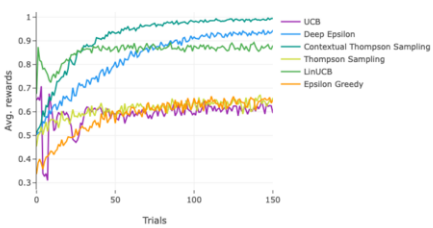


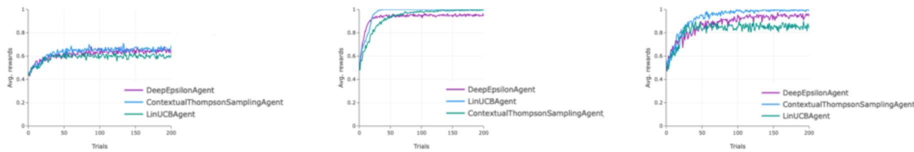
Figure 3. Baseline Results

Table 1. Average Rewards with Standard Deviation

RL Algorithm	Mean (avg. reward)	Standard Deviation
Epsilon Greedy	0.79	±0.4
UCB	0.83	±0.37
TS	0.89	±0.31
Deep Epsilon-Greedy	0.91	±0.28
LinUCB	0.85	±0.35
CTS	1	±0

4.2. Experiment 1: User Type Adaption

We investigated how well different RL algorithms could adapt to diverse user preferences in a simulated environment designed to manage mealtimes for PwD. According to domain experts, we created five user types (A-E) with varying responses to nudging signals: Type A positively reacts to up to two signals per meal (e.g. music and light for breakfast, light and video for lunch, and music and scent for dinner); Type B is the same as Type A but will never react to one meal (e.g. no reaction during lunch); Type C positively responds only to one signal type (e.g. video for all three meals); Type D shows no reaction to any signal; and Type E positively reacts to one of the signals in a certain intensity level (e.g. music for breakfast, and scent for lunch and dinner).



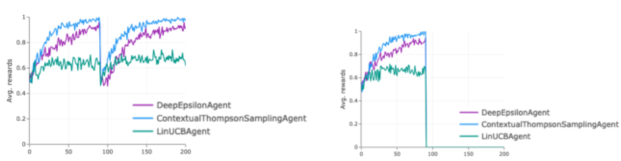
(a) User Type A (b) User Type C (c) User Type E

Figure 4. Experiment 1: User Type Adaption

Figure 4 showed that all RL algorithms could learn and adapt to user preferences, achieving stable average rewards across user types A, C, and E. CTS algorithm generally outperformed the others in these cases. Interestingly, for user type C with unchanging preferences, LinUCB exhibited a faster learning rate than CTS, suggesting its potential for scenarios where the relationship between context and reward is unclear.

4.3. *Experiment 2: Preference Shifts Over Time*

This experiment evaluates the AI system's ability to dynamically adapt to users' shifting preferences, which is important for maintaining engagement with users whose needs and behaviors may change over time. If so, we aim to further identify which RL algorithm can learn and quickly converge in this dynamic scenario. Thus, we introduce two user types: Type F, where users initially respond to one type of nudging signal but then suddenly shift to another after one month (i.e., 90 trials), and Type G, depicting users who become unresponsive over time. In User Type F (results in Figures 5a), if the user maintains responsiveness to the signal, the RL algorithm can adapt to the shift within about 50 trials (meals). For User Type G (results in 5b), the loss of responsiveness may provide a warning sign to caregivers to intervene and indicate potential changes in the user's dementia status.

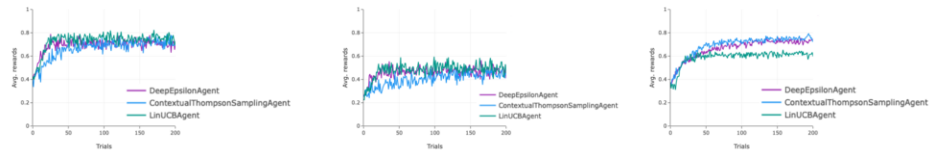


(a) User Type F (b) User Type G

Figure 5. Experiment 2: Sudden Preference Shift After 90 Trials

4.4. *Experiment 3: Acceptance & Tolerance Test*

We aim to further evaluate how the RL algorithms can perform under realistic and challenging conditions. The Acceptance Test examined the algorithm performance when user responses do not consistently translate into rewards, introducing acceptance rates rang-



(a) User Type C, Acceptance Rate = 0.75

(b) User Type C, Acceptance Rate = 0.5

(c) User Type E, Tolerance Rate = 0.2

Figure 6. Experiment 3: Acceptance & Tolerance Test

ing from good (0.75) to low (0.25) to mimic the unpredictable nature of real-world scenarios. Figure 6a, 6b indicated a significantly declined performance as acceptance rates decreased, highlighting the system is sensitive to the environment's response.

Tolerance Test evaluated system robustness against noisy environments by incorporating 10% and 20% invalid data to simulate errors in user input or sensor data. Figure 6c showed moderate resilience to data inaccuracies, with certain impacts on performance as error rates increased. Together, these experiments demonstrate the robustness of our AI system and emphasize the importance of maintaining user engagement and reducing noisy data from the home environment.

5. Conclusion & Future Work

In conclusion, this paper presents a novel, RL-based AI system, specifically focusing on eating and drinking scenarios for PwD in an intelligent home environment. Our simulation experiments tested various RL algorithms to determine their effectiveness in a simulated environment mimicking real-world situations, including users' behaviors and noisiness. The results showcase the successful simulation of a human-centered AI system, demonstrating the potential of the RL system in adapting to the unique behaviors and needs of PwD, in order to support them to live independently. For future work, we aim to extend our research by conducting real-world trials with early-stage PwD and further validate our system more thoroughly. Although we have cooperated with domain experts and caregivers to make the AI human-centered, based on the current small range of scenarios, the empathicness of the AI is still limited. With deeper investigation into the interaction between humans and AI, we aim to enhance our system's capabilities. Our eventual goal is to create a personalized, adaptive, and empathic AI system that can meet the needs of PwD and release the burden of their caregivers.

Acknowledgement

We would like to extend our special thanks to Regieorgaan SIA for funding this research (RAAK.PUB09.060) and all the consortium partners of the Wie Zorgt project.

References

[1] World Health Organization. Dementia; 2022. Accessed: 2024 Jan 25. https://www.who.int/news-room/fact-sheets/detail/dementia.

[2] Alzheimer Europe. European Dementia Monitor 2020: Comparing and benchmarking national dementia strategies and policies; 2022. Accessed: 2024 Jan 25. https://www.alzheimer-europe.org/sites/default/files/2021-11/European%20Dementia%20Monitor%202020%20%20Comparing%20and%20benchmarking%20national%20dementiastrategies%20and%20policies.pdf.

[3] Demir E, Köseoğlu E, Sokullu R, Şeker B. Smart home assistant for ambient assisted living of elderly people with dementia. Procedia computer science. 2017;113:609-14.

[4] Gimenez Manuel JG, Augusto JC, Stewart J. AnAbEL: towards empowering people living with dementia in ambient assisted living. Universal Access in the Information Society. 2022;21(2):457-76.

[5] Gibson G, Dickinson C, Brittain K, Robinson L. Personalisation, customisation and bricolage: how people with dementia and their families make assistive technology work for them. Ageing & Society. 2019;39(11):2502-19.

[6] Grave A, Robben S, Oey M, Ben Allouch S, Mohammadi M. Developing an Intelligent Environment to Support People with Early-Stage Dementia: from User-Needs to a Real-Life Prototype. Applied Human Informatics: Open Journal of the Academy of Human Informatics. 2022 Mar;4(1):1-18. In Special Issue on Intelligent Environments for Health and Well-being.

[7] Inzlicht M, Cameron CD, D'Cruz J, Bloom P. In praise of empathic AI. Trends in Cognitive Sciences. 2023.

[8] Sporrel K, De Boer RD, Wang S, Nibbeling N, Simons M, Deutekom M, et al. The design and development of a personalized leisure time physical activity application based on behavior change theories, end-user perceptions, and principles from empirical data mining. Frontiers in Public Health. 2021;8:528472.

[9] Wapenaar J, De Groot L. Eten en drinken bij dementie. Springer; 2018.

[10] MacLean PD. The triune brain in evolution: Role in paleocerebral functions. Springer Science & Business Media; 1990.

[11] Van Rossum C, Buurma-Rethans E, Dinnissen C, Beukers M, Brants H, Ocké M. The diet of the Dutch: Results of the Dutch national food consumption survey 2012-2016. 2020.

[12] Singh R, Kushwaha AKS, Srivastava R, et al. Recent trends in human activity recognition–A comparative study. Cognitive Systems Research. 2023;77:30-44.

[13] Moyle W, Murfield J, Lion K. The effectiveness of smart home technologies to support the health outcomes of community-dwelling older adults living with dementia: A scoping review. International Journal of Medical Informatics. 2021;153:104513.

[14] Chimamiwa G, Giaretta A, Alirezaie M, Pecora F, Loutfi A. Are smart homes adequate for older adults with dementia? Sensors. 2022;22(11):4254.

[15] Rabbi M, Aung MH, Zhang M, Choudhury T. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing; 2015. p. 707-18.

[16] Liao P, Greenewald K, Klasnja P, Murphy S. Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2020;4(1):1-22.

[17] Wang S, Zhang C, Kröse B, van Hoof H. Optimizing adaptive notifications in mobile health interventions systems: reinforcement learning from a data-driven behavioral simulator. Journal of medical systems. 2021;45(12):102.

[18] Ly Tung N. Toward an intelligent long-term assistance for people with dementia in the context of navigation in indoor environments. Universität Würzburg; 2017.

[19] Hoey J, Plötz T, Jackson D, Monk A, Pham C, Olivier P. Rapid specification and automated generation of prompting systems to assist people with dementia. Pervasive and Mobile Computing. 2011;7(3):299-318.

[20] D'Andrea F, Tischler V, Dening T, Churchill A. Olfactory stimulation for people with dementia: A rapid review. Dementia. 2022;21(5):1800-24.

[21] Sutton RS, Barto AG. Reinforcement learning: An introduction. MIT press; 2018.

[22] Auer P. Using confidence bounds for exploitation-exploration trade-offs. Journal of Machine Learning Research. 2002;3(Nov):397-422.

[23] Thompson WR. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika. 1933;25(3-4):285-94.

[24] Russo DJ, Van Roy B, Kazerouni A, Osband I, Wen Z, et al. A tutorial on thompson sampling. Foundations and Trends® in Machine Learning. 2018;11(1):1-96.

[25] Allesiardo R, Féraud R, Bouneffouf D. A neural networks committee for the contextual bandit problem. In: Neural Information Processing: 21st International Conference, ICONIP 2014, Kuching, Malaysia, November 3-6, 2014. Proceedings, Part I 21. Springer; 2014. p. 374-81.

[26] Li L, Chu W, Langford J, Schapire RE. A contextual-bandit approach to personalized news article recommendation. In: Proceedings of the 19th international conference on World wide web; 2010. p. 661-70.