Workshops at 18th International Conference on Intelligent Environments (IE2022)
H.H. Alvarez Valera and M. Luštrek (Eds.)

doi:10.3233/AISE220051

A RL Based Model for Improving Human Task Management Performance

Muddasar NAEEM [a], Valeriano FABRIS [b], and Antonio CORONATO [a,b]
[a] *ICAR-CNR*
[b] *Università Telematica Giustino Fortunato*

Abstract. This paper discusses an Reinforcement Learning (RL) based system to improve human performance in task selection and management by incorporating various factors. A simulated task management environment of a coaching center is considered. Different factors includes task urgency, task status, and task importance as well as which task to attend to next, and that an even more important factor of task management is the capability to avoid low importance tasks. Five algorithms such as: Boltzman Sampling, Epsilon-decreasing, Random, Softmax, and Thompson Sampling have been used in experiments.

Keywords. Reinforcement Learning, Task Management, Performance Improvement, Decision Making , Boltzman Sampling, Epsilon-decreasing, Random, Softmax, Thompson Sampling

1. Introduction

Correct decision making is an important and necessary condition for an individual, a team and an organization for efficient and effective management. Decision making during each step of task management such as selection, prioritizing, and delegating is important to best utilizes capabilities of an individual or a team. It is also important organize the workflow and identify bottlenecks for optimal performance. The task selection according to the skills of an individual plays an essential part successful task management.

However, during the task allocation, the teachers or instructors lack insight into diversity that may present among individual, which might effect the individual skills and overall performance [22]. A task management system may be designed for the routine work of an institute. The system should be able to provides guidance to users to accomplish their personal and department level task. The system can assign different type of task to different users based on his/her skills and interest. The task management system should be user friendly and helps the organization and user to interact and accomplish diverse tasks with ease. The system involves identifying, acquiring, allocating and tracking all the available resources.

Recently, the development in computing technology and the introduction of new machine learning algorithms e.g. reinforcement learning [16], neural network [14] the goal of Artificial Intelligence (AI) has become a step closer. AI has important application in diverse fields including:healthcare [6], [15,12], robotics and autonomous control, vision enhancing method for low vision impairments [9], natural language processing,

dynamic normative environments [21], risk management [17], communication [13] intelligent planning of onshore touristic itineraries for cruise passengers in a smart city [4] and distributed fuzzy system able to infer in real-time critical situations [20].
Our contribution is to develop a Reinforcement Learning based system that can assist an individual or a team in order to improve human performance in task selection and management. We have developed a simulated task management environment of a coaching center and considered different factors includes task urgency, task status, and task importance as well as which task to attend next, and that an even more important factor of task management is the capability to avoid low importance tasks. We have used five algorithms such as: Boltzman Sampling, Epsilon-decreasing, Random, Softmax, and Thompson Sampling have been in our experiments.

The remaining part of the paper is organized as follows: section 2 overviews the some related work while section 3 presents the technical background to RL in general and to used methods in particular. The section 4 explains the methodology and experimentation, section 5 gives detailed information about the experimental results and analyzes the results of the experiment and section 7 concludes the work.

2. Related Work

The impact of information technology and artificial intelligence in the management of economic systems is discussed in [8] but without solving the problems of rating. The solution to this problem may be seen in [10]. However, the level of staff development and their assessment are not taken into account. This proposed on the theoretical aspects of personnel development and performance improvement such as identifying the best tasks and directions of personnel development at an organisation.
Management of human resource and their skills as a strategic human resource management for example, the level of remuneration or motivation nowadays considered beyond the limits of management tasks. It is vital to handle human task selection and management as a process that can contributes to the success of an organisation. The work of [1] considers similar scenarios where all heads are involved in the management process and the role of all persons is crucial for the competitive advantage of the institute. Different fuzzy models were developed for a simulated task management environment in [2] by incorporating combinations of factors to model human task management performance.

An assistance system based on home sensors, ambient and artificial intelligence that helps the elderly during their medical treatment at home to reduce medication errors is presented in [3,5]. Similarly, a reinforcement learning based methods for training a software agent for risk management of medical software systems are discussed in [19, 18]. Similarly, a framework to self-learn and automatically classify news headline into its appropriate news category using artificial intelligence tools such as text mining and machine learning is presented in [7].

3. Technical Background

Reinforcement Learning (RL) is a subfield of ML where an agent tries to learn the dynamics of an unknown environment. To learn the characteristics of the given environment, the agent chose a certain action a_t from a set of actions in a certain state s_t (there is a set of state for every given environment) at time slot t and, based on the transition model of the environment, the agent reaches a new state a_{t+1} and receives a numerical reward r_t. After a lot of trial and error, the RL agent can learn the optimal policy for a given environment. The optimal policy tells an agent which action to choose in a given state to maximize long-term aggregated reward. An RL problem is first modeled as a Markov Decision Process (MDP), as shown in Figure 1, and then an appropriate RL algorithm is employed based on the dynamics of the underlying environment. A brief introduction to MDP is given next.

A MDP is a tuple $\langle$S,A,R,P,$\gamma\rangle$; where
-S is used to denote states;
-A is used to denote actions;
-R is used to denote a reward function;
-P indicates transition probability;
-γ is a discount factor: $\gamma \in [0, 1]$.

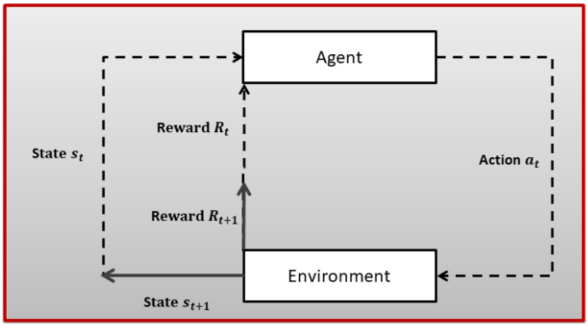

Figure 1. The Reinforcement Learning problem.

The Markov property, i.e. next state, is dependent only on the previous state that is assumed. A finite MDP is described by actions, states, and the environment's dynamics. For any state–action (s,a) pair, the probability of resulted state and the corresponding reward (s',r) is given as in Equation (1):

$$p(s',r|s,a) \doteq Pr\{S_{t+1} = s', R_{t+1} = r|S_t = s, A_t = a \tag{1}$$

Informally, the target of the RL agent is to maximize the reward. This is to say, with the list of rewards R_{t+1},R_{t+2},... after time period t, the goal is to maximize the reward function as given in Equation (2):

$$G_t = R_{t+1} + R_{t+2} + \cdots + R_T \tag{2}$$

where T is the last time interval.

The return G_t is the sum of discounted rewards obtained after time t.

$$G_t = \sum_{k=0}^{T} \gamma^k R_{t+k+1} \tag{3}$$

A policy π defined in Equation (4) tells an agent which action to take in a given state.

$$\pi(a|s) \doteq P[A_t = a|S_t = s] \tag{4}$$

Having the policy π and the return G_t, two value functions can be defined, i.e., state–value and the action–value functions. The state–value function $v_\pi(s)$ is the expected return starting from a state s and following the policy π as given in Equation (5).

$$v_\pi(s) \doteq E_\pi[G_t|S_t = s] = E_\pi[\sum_{k=0}^{\infty}\gamma^k R_{t+k+1}|S_t = s] \tag{5}$$

The action–value function $q_\pi(s,a)$ is the expected return starting from a state s, taking action a, by following the policy π.
The optimal value function is one that obtains the best gains in terms of returns, as given in Equation (6).

$$v_*(s) = \max_\pi v_\pi(s), \forall s \in S \tag{6}$$

After defining the MDP and the selection of an RL technique for a given problem, the next issue is to maintain a delicate balance between exploration and exploitation. At each time step, the RL agent can select the best rewarding action based on its current knowledge of the environment. On the other hand, an RL agent can explore more available actions that may provide even more rewarding actions. Therefore, exploration and exploitation may not be good strategies and an RL agent should learn a trade-off between exploration and exploitation for a certain problem.

For further details on RL algorithms in general and the application of RL in healthcare in particular, the reader may refer to [16,6], respectively.
We have perform experiments with various algorithms and amoung them the performance of Thompson sampling is comparatively better. Thompson sampling method is based on Bayesian RL. In Bayesian RL, an inference about a Random Variable (RV) X is made by using a probability distribution for RV X ([11]). Then the inferences may be extracted from the probability distribution. For example, consider that RV X is hidden and only observations can be made about a related RV Y. The task of the Bayesian learning is to infer X from the samples of RV Y; e.g., when X is a physical quantity and Y is the measurement of noise. A Bayesian inference process can be carried out as follows:

1. Consider the prior distribution $P(X)$ that represents the belief about RV X without the observation of the data;
2. Choose a statistical model $P(Y|X)$ that expresses the belief about RV Y when X is given. $P(Y|X)$ also reflects the statistical dependence between RVs X and Y;
3. Observe data $Y = y$;
4. Calculate the posterior distribution to update the belief about RV X by using the Bayes Rule.

$$P(X|Y = y) = \frac{P(y|X)P(X)}{\int P(y|X')P(X')dX'} \tag{7}$$

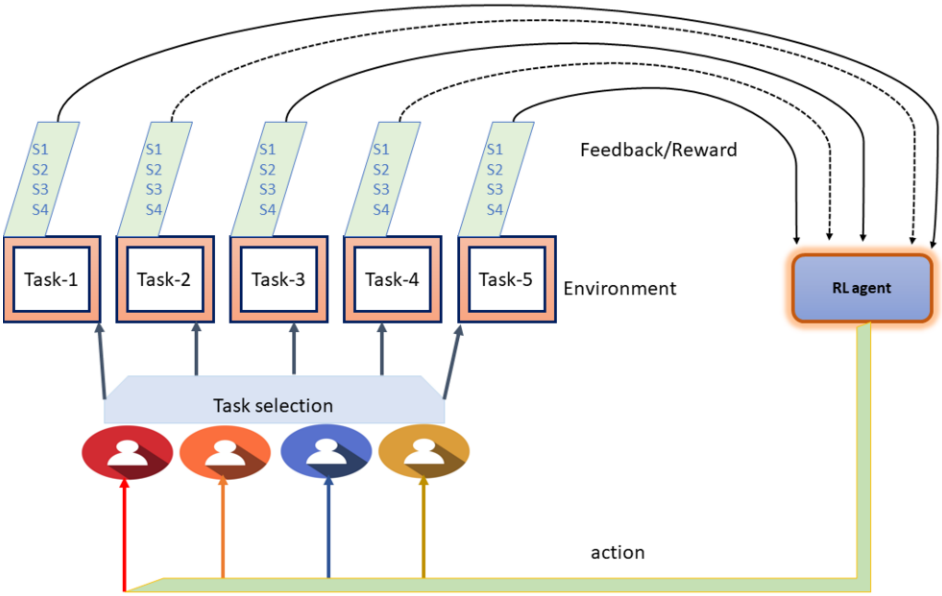

Figure 2. Flowchart of proposed Framework

4. System Model

Figure 2 shows the flowchart of proposed framework including different tasks, task selection using RL policy and feedback to RL agent after selection of the particular task. We have consider an emulated environment where different people have to learn and perform different tasks. The performance of a person varies from task to task and similarly, outcome of the each task in terms of score (S_1, S_2, S_3, S_4) may be different for different people. This means that the probability distribution for the reward corresponding to each task is different and is unknown. The task is to learn which task to select in order to get maximum score in a given amount of time. This problem statement is identical a single step MDP.

The score list (S_1, S_2, S_3, S_4) measures different skills of a person during the execution of a task. The better score in skills indicates more interest and better performance for a particular task and lower score indicates that a particular task is unsuitable for a person. After a lot of interaction with the environment, agent learns the most suitable task for a person.

We model this scenario as problem as a MMDP with a single state. There are in genral K tasks and it is possible to select anyone and each task has has a certain probability of returning a reward (score). Therefore, we have a single state and K possible actions (one action for each task). At each time period the agent selects one task and it receives a feedback in terms of different scores (reward). The goal of the agent is to learn the best taks for each person in order to maximise its long term reward. We use Boltzman sampling, Epsilon decreasing, Random, Softmax, and Thompson sampling algorithms to solve this singel state MDP problem.

Table 1. Overview of Related Works

Algorithm	ACR	AUD	RMSE
Boltzman	28.54	0.128566123	0.323892655
Epsilon-decreasing	29.38	0.172919744	0.275413322
Random	22.23	0.447237129	0.013220423
Softmax	27.62	0.230231606	0.216913354
Thompson	31.26	0.505206685	0.095571118

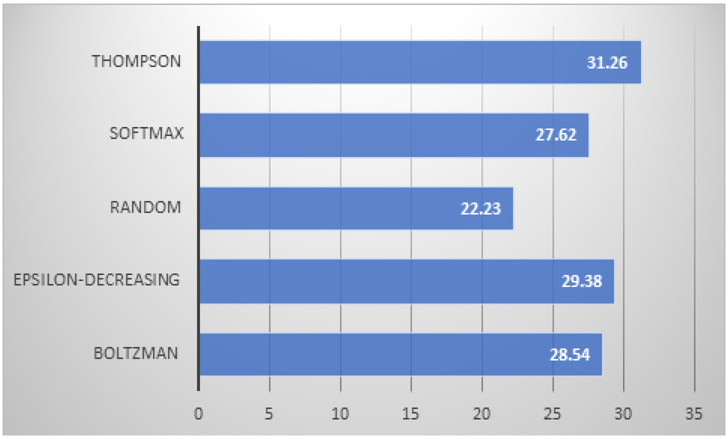

Figure 3. ACR performance of different algorithms

5. Results

The ACR obtained with Random algorithm is 22.23 and at the same time the RMSE is extremely low 0.01322 indicates that the random agent is largely unbalanced in direction of exploration than that of exploitation. The ACR for Epsilor-decreasing algorithm is 29.38 which is higher than the ACR obtained with the Random, Boltzman and Softmax algorithms. The RMSE 0.2754 is slightly higher which presents a good balance between exploration and exploitation. While the ACR and RMSE against Boltzman sampling method are 28.54 and 0.3238 respectively.
The ACR obtained is 31.26 which is the highest ACR obtained among all five algorithms, and at the same time a fairly low RMSE of 0.0595 on the utility distribution indicates a perfect strategy for balancing exploration and exploitation. Wherease the ACR and RMSE for Softmax algorithm are 27.62 and 0.2169 respectively.
Figure 3 presents the obtained performance of five used techniques on a barchart. Thompson sampling algorithms seems to be the relatively better method in term of ACR. While the Softmax and Boltzman techniques have identical results. The epsilong-decreesing method also shows reasonable performance.

6. Conclusion and Future Work

We have presented an Reinforcement Learning based system to assist humans in task selection and management and to enhance their performance. A simulated task management environment of a coaching center is considered. Five algorithms such as: Boltzman Sampling, Epsilon-decreasing, Random, Softmax, and Thompson Sampling have been used in experiments. The proposed methodology has shown promosing results and may play a vital role in improving human task management skills and performance.

References

[1] Brian E Becker and Mark A Huselid. Strategic human resources management: where do we go from here? *Journal of management*, 32(6):898–925, 2006.

[2] Jeng Yan Chen and Ken Funk. A fuzzy model of human task management performance. In *IIE Annual Conference. Proceedings*, page 1. Institute of Industrial and Systems Engineers (IISE), 2003.

[3] Mario Ciampi, Antonio Coronato, Muddasar Naeem, and Stefano Silvestri. An intelligent environment for preventing medication errors in home treatment. *Expert Systems with Applications*, 193:116434, 2022.

[4] Antonio Coronato, Claudia Di Napoli, Giovanni Paragliola, and Luca Serino. Intelligent planning of onshore touristic itineraries for cruise passengers in a smart city. In *2021 17th International Conference on Intelligent Environments (IE)*, pages 1–7. IEEE, 2021.

[5] Antonio Coronato and Muddasar Naeem. A reinforcement learning based intelligent system for the healthcare treatment assistance of patients with disabilities. In *International Symposium on Pervasive Systems, Algorithms and Networks*, pages 15–28. Springer, 2019.

[6] Antonio Coronato, Muddasar Naeem, Giuseppe De Pietro, and Giovanni Paragliola. Reinforcement learning for intelligent healthcare applications: A survey. *Artificial Intelligence in Medicine*, 109:101964, 2020.

[7] Zara Hayat, Aqsa Rahim, Sajid Bashir, and Muddasar Naeem. Self learning of news category using ai techniques. In *Intelligent Environments 2021*, pages 167–178. IOS Press, 2021.

[8] Sergey Ivanov. Modeling company sales based on the use of swot analysis and ishikawa charts. In *M3E2-MLPEED*, pages 385–394, 2019.

[9] Carmelo Lodato and Patrizia Ribino. A novel vision-enhancing technology for low-vision impairments. *Journal of medical systems*, 42(12):1–13, 2018.

[10] Yu Lysenko, V Petrenko, O Bogatov, and V Skobelev. Reytingovoye upravleniye ekonomicheskimi sistemami. *Yugo-Vostok, Donetsk*, 1999.

[11] J. Pineau M. Ghavamzadeh, S. Mannor and A. Tamar. Bayesian Reinforcement Learning: A Survey, year = 2016. *Foundations and Trends in Machine Learnin*, pages 102–109.

[12] Muddasar Naeem and Antonio Coronato. An ai-empowered home-infrastructure to minimize medication errors. *Journal of Sensor and Actuator Networks*, 11(1):13, 2022.

[13] Muddasar Naeem, Giuseppe De Pietro, and Antonio Coronato. Application of reinforcement learning and deep learning in multiple-input and multiple-output (mimo) systems. *Sensors*, 22(1):309, 2021.

[14] Muddasar Naeem, Giovanni Paragiola, Antonio Coronato, and Giuseppe De Pietro. A cnn based monitoring system to minimize medication errors during treatment process at home. In *Proceedings of the 3rd International Conference on Applications of Intelligent Systems*, pages 1–5, 2020.

[15] Muddasar Naeem, Giovanni Paragliola, and Antonio Coronato. A reinforcement learning and deep learning based intelligent system for the support of impaired patients in home treatment. *Expert Systems with Applications*, page 114285, 2020.

[16] Muddasar Naeem, S Tahir H Rizvi, and Antonio Coronato. A gentle introduction to reinforcement learning and its application in different fields. *IEEE Access*, 2020.

[17] Giovanni Paragliola and Antonio Coronato. A reinforcement-learning-based approach for the planning of safety strategies in aal applications. In *Intelligent Environments 2018*, pages 498–505. IOS Press, 2018.

[18] Giovanni Paragliola, Antonio Coronato, Muddasar Naeem, and Giuseppe De Pietro. A reinforcement learning-based approach for the risk management of e-health environments: A case study. In *2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 711–716. IEEE, 2018.

[19] Giovanni Paragliola and Muddasar Naeem. Risk management for nuclear medical department using reinforcement learning algorithms. *Journal of Reliable Intelligent Environments*, 5(2):105–113, 2019.

[20] Patrizia Ribino and Carmelo Lodato. A distributed fuzzy system for dangerous events real-time alerting. *Journal of Ambient Intelligence and Humanized Computing*, 10(11):4263–4282, 2019.

[21] Patrizia Ribino and Carmelo Lodato. A norm compliance approach for open and goal-directed intelligent systems. *Complexity*, 2019, 2019.

[22] Silke Vanslambrouck, Chang Zhu, Koen Lombaerts, Brent Philipsen, and Jo Tondeur. Students' motivation and subjective task value of participating in online and blended learning environments. *The Internet and Higher Education*, 36:33–40, 2018.