ISSN: 2508-7894 © 2021 KODISA & KAIA.
KJAI website: http://www.kjai.or.kr
DOI: http://dx.doi.org/10.24225/kjai.2021.9.1.1

A Development of Nurse Scheduling Model Based on Q-Learning Algorithm

In-Chul JUNG[1], Yeun-Su KIM [2], Sae-Ran IM[3], Chun-Hwa IHM[4]

Received: February 25, 2021. Revised: April 15, 2021. Accepted: June 05, 2021.

Abstract

In this paper, We focused the issue of creating a socially problematic nurse schedule. The nurse schedule should be prepared in consideration of three shifts, appropriate placement of experienced workers, the fairness of work assignment, and legal work standards. Because of the complex structure of the nurse schedule, which must reflect various requirements, in most hospitals, the nurse in charge writes it by hand with a lot of time and effort. This study attempted to automatically create an optimized nurse schedule based on legal labor standards and fairness. We developed an I/O Q-Learning algorithm-based model based on Python and Web Application for automatic nurse schedule. The model was trained to converge to 100 by creating an Fairness Indicator Score(FIS) that considers Labor Standards Act, Work equity, Work preference. Manual nurse schedules and this model are compared with FIS. This model showed a higher work equity index of 13.31 points, work preference index of 1.52 points, and FIS of 16.38 points. This study was able to automatically generate nurse schedule based on reinforcement Learning. In addition, as a result of creating the nurse schedule of E hospital using this model, it was possible to reduce the time required from 88 hours to 3 hours. If additional supplementation of FIS and reinforcement Learning techniques such as DQN, CNN, Monte Carlo Simulation and AlphaZero additionally utilize a more an optimized model can be developed.

Keywords: Nurse Schedule, Nurse Scheduling Problem, Reinforcement Learning, Q-Learning, Fairness Indicator Score

Major Classification : Artificial Intelligence, Reinforcement Learning

1. Introduction

The hospital's ward nurse schedule should be written while nearly meeting the working hours, the number of holidays, and the rest time. Also, there are many limitations, having to fully consider qualitative factors such as individual worker's satisfaction and preference for work, so it takes a lot of time and effort for the ward part leader to manually write them.

Nurse Scheduling Problem (NSP) is a question of how to assign the type of shift (morning, afternoon, night, off) to a day or monthly timetable for nurse work, and there is still much research going on (Song, 2020).

There have been attempts to solve the Nurse Scheduling Problem(NSP) using various Function Optimization techniques, Meta-Heuristic such as Tabu Search, Simulated Annealing, Genetic Algorithm, Ant Colony Optimization.

Meta Heuristic Techniques are simple and easy to implement algorithms. It maintains the solution's diversity by not always exploring the right direction and choosing

1 First Author, Researcher, Department of Integrated IT center, Eulji University Medical Center, Email: 20201278@eulji.ac.kr
2 Second Author, Part Leader, Department of Integrated IT center, Eulji University Medical Center, Email: lake@eulji.ac.kr
3 Third Author, Researcher, Department of Integrated IT center, Eulji University Medical Center, Email: 20201610@eulji.ac.kr
4 Corresponding Author, Professor, Dept. of Laboratory Medicine, Eulji University Hospital, Email: haneul@eulji.ac.kr

even the wrong direction with a small probability. However, it takes a long time to derive a mathematically expressed solution. Furthermore, given large datasets and navigation regions, performance cannot be guaranteed (Yoo, Seo, Kim & Kim, 2019).

Function Optimization techniques are useful in obtaining optimal solutions to small problems; however, Not only is it challenging to adapt to other problem situations other than those given, but it is also challenging to apply to extensive problems with many constraints (Oh, 2012).

Because of these shortcomings, an alternate schedule made using traditional techniques is difficult to apply to the actual site. So this study, we study a model that uses the Q-Learning Algorithm of reinforcement learning techniques to create a nurse schedule optimized for the field compared to manual works.

Chapter 2 of this study describes reinforcement learning and Q-Learning as theoretical backgrounds. Chapter 3 describes Dataflow and Algorithm for implementation by the proposed model and describes the Fairness Indicator Score (FIS) developed for performance evaluation. Chapter 4 evaluates the results of the proposed model in Chapter 5 describes the conclusions and future studies.

2. Literature Review

2.1. Reinforcement Learning

Reinforcement Learning originated in the study of animal learning in psychology and is a form of education that combines dynamic programming and teacher learning. While general teacher learning presents a response method in a given situation, reinforcement learning leads to better choices in the future by making free choices for a given state and offering rewards based on preferences (Lee & Woo, 2005).

A more detailed description of the aspects of reinforcement learning is a method of finding an optimal control strategy based on data that maximizes the expectation of the sum of the overall compensation values by utilizing the evaluation signals expressed as reward values for interaction between the system and control inputs(Kang, Seo, Lee & Kang, 2017).

Figure 1 is a brief schematize of the learning process of reinforcement learning. Teach point in time (t), the agent has information about its state information "State(S_t)" and possible "actions(A_t)". The agent takes any action and receives new state information and reward from the environment. Based on these interactions, the agent develops a policy π:S→A that maximizes the sum by results.

To sum up, reinforcement learning refers to a learning method in which defined agents in the environment recognize the current state and select actions to perform according to the ground to maximize rewards (Kong & Lee, 2017). That is considered a methodology consistent with this study's goal to generate a better nurse schedule based on the reward criteria.

2.2 Q-Learning

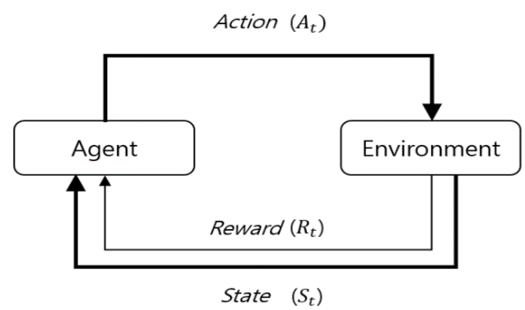

Figure 1: Q-Learning

Q-learning Algorithm is a reinforcement learning algorithm that finds the optimal policy through Q's value (s, a). The value of Q(s, a) is the numerical value of how good it is due to behavior (a) in a given situation. Q(s, a) is also called Q-function. Q-function has a Q-table, which stores Q-function's value for any actionable case behavior. From the state information of the current environment, we obtain the Q-function value for all actionable actions from Q-table and choose the maximum value as the optimal behavior. We exploit the reward values obtained through the optimal behavior obtained by doing so. In other words, supplementing the existing Q-table helps us make better choices in the future.

$$Q^{new}(S_t,a_t) \leftarrow \underbrace{Q(S_t,a_t)}_{\text{Old value}} + \underbrace{\alpha}_{\text{Learning rate}} * \overbrace{\Big((\underbrace{r_t}_{\text{Reward}} + \underbrace{\gamma}_{\text{Dicount factor}} * \underbrace{maxQ(S_{t+1},a_t)}_{\text{Estimate of optimal future value}} - \underbrace{Q(S_t,a_t)}_{\text{Old value}}\Big)}^{\text{Temporal difference}}$$

Figure 2: Q-learning Algorithm

3. Model Configuration

3.1. Data Flow

Figure 3 shows the Data Flow of this study. After using Python's flask package, we construct a Web App-based I/O screen, The Nurse Scheduling Model based on Q-Learning was developed. The number of Wanted and the model's learning from the Web App-based screen, including annual leave, education, travel, is delivered to the Nurse Scheduling Model in advance as a requirement for each worker.

In this model, we proceed with the learning by considering the requirements as constraints. Also Figure 3 shows Data Flow circulating from the Web App to the Nurse scheduling model.

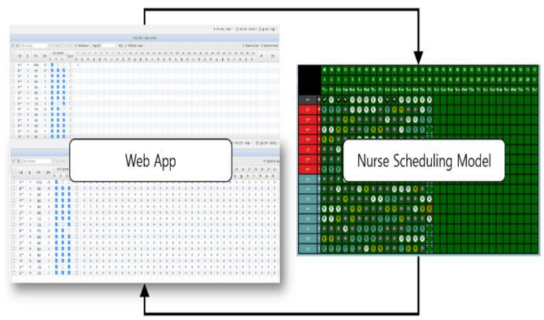

Figure 3: Data Flow

Considering the employee's career and Wanted, the model proceeds with optimal work organization tasks and ultimately derives a nurse schedule in which the sum of Reward converges to the maximum. The FIS indicators generated separately for the evaluation of learning were evaluated in consideration of the Labor Standards Act, Work Equity, and Work Preference.

- Data Structure

Data Structure consists of app.py that sends and receives data, agents, play.py that determine the number of lessons, Enviornment.py implementing this model, and pyGameGUI.py, where the agent outputs results for each number of classes. Figure 4 shows the Data Structure.

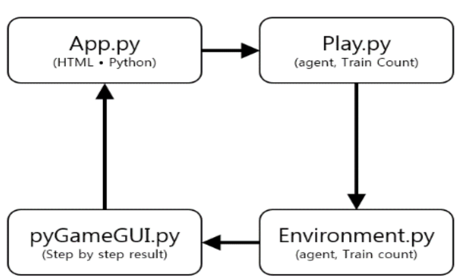

Figure 4: Data Structure

The first time you send data from a Web App-based screen to this model, and then proceed in this order of ① app.py ②play.py ③enviromnet.py ④pyGameGUI.py ⑤ app.py .

3.2. Nurse Scheduling Model

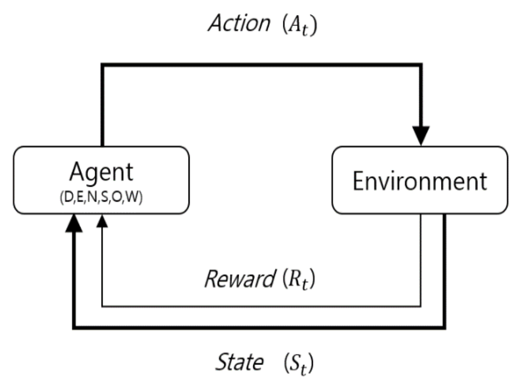

Figure 5: Nurse Scheduling Model Q-Learning

The Nurse Scheduling Model is developed based on Q-Learning in reinforcement learning. Agent classified six types of work (Day(D), Evening(E), Night(N), Special(S), Off(O), Wanted(W)).

The Agent in Figure 5 moves the Environment daily. the scope of Action shall be converted to Reward for nurses who can work except for wanted on a relevant day. For the Agent to find the best workers, the Reward is converted into consideration of each worker's past work, the current number of work, and job-specific reviews. Higher Reward means the optimal worker, and the Agent iteratively learns every step to ensure Reward converges to its maximum. In other words, we find and know the number of cases in which the state of the Environment is updated according to the Agent's Action, and the accumulated Reward can be as high as possible.

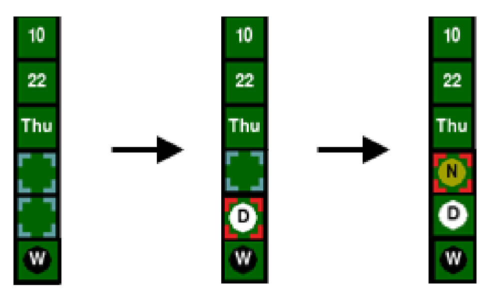

Figure 6: Action

For each Agent to be placed, extract Possible Movement(PM) before Action. PM refers to the employee whose current Agent can be placed. When it is placed, the next Agent finds and places the optimal worker so that Reward can receive the highest value, except for the Agent worker placed in PM.

If the Agent of Day(D), Evening(E), Night(N), Special(S) exhaust all the Action counts in on that day, the remaining PM's workers are placed as Off(O) Agent. The worker placed to Off(O) becomes a seat, which grants a higher Reward to the next State. the Agent considers the employee first to receive a high Reward.

-Value Function

To create an optimal nurse schedule, we define the policy using the Q-Learning as below-based Value Function to optimize the Reward after calculating the Reward, considering the FIS according to all agents' movement.

$$V^{\pi}(s_t) = r_t + \gamma r_t + \gamma^2 r_{t+1} + \cdots \sum_{k=0}^{\infty} \gamma^k r_{t+k}$$
$$V^*(\mathrm{s}) = max_a Q(s,a)$$
$$\pi^*(\mathrm{s}) = argmax_a Q(s,a) \tag{1}$$

V: Value Function
r: Reward
γ: Discount factor($0 \le \gamma \le 1$)
π : Policy

- Nurse Scheduling Model Algorithm

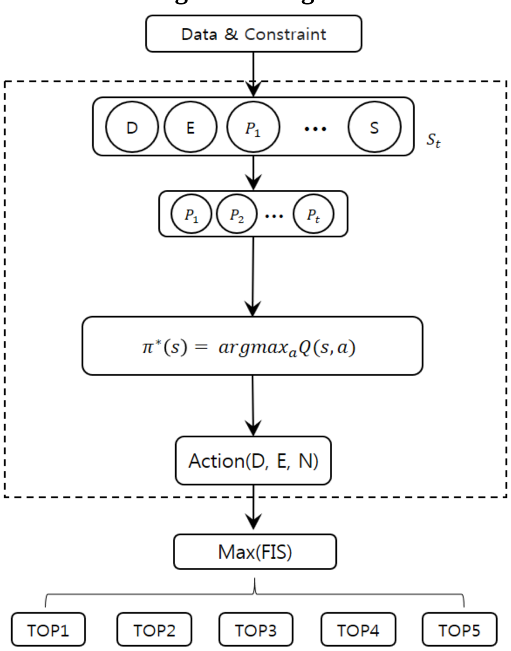

Figure 7: Nurse Scheduling Model Algorithm

S_t : Nurse schedule at point t (day) in the current state
P_t : Agent's possible movement at point t
π : Policy
FIS: Fairness Indicator score

Figure 7 shows the nurse scheduling model algorithm. We place historical work data and Wanted as constraints before the current state and generate the current state. The current state of refers to the Nurse schedule at the time of t(day), Q-Learning is used to find workers who will be placed at the time of T.

Extract through with PM with actionable status. We then Q-Learning in PM to learn that Policy finds a movement plan for workers whose cumulative sum of Rewards converges to maximum, and then implement an optimal action.

When the relevant work table is completed, the nurse schedule is evaluated according to the FIS indicator. According to the number of lessons, the generated nurse schedule is saved from Top1 to Top5 and returned to the Web App user.

- Evaluation Indicators

To compare this model with a manually created nurse schedule, we developed an indicator called the FIS.

.

$$FIS = X * \left(\frac{\sum_{n=1}^{N} \left(\frac{1}{1+e^{-(Y+Z)}} * 100 \right)}{N} \right) \tag{2}$$

N : Number of Nurse
X : Labor Standards Act
Y: Work equality
Z : Work preference

The FIS derived by defining three items, Labor Standards Act (X), Work Equity (Y), Work Preference (Z), and 15 detailed items (L1 to L15).

Table 1: Definition of Labor Standards Act (X) items

Index	Contents		Score	
			FAIL	PASS
Labor Standards Act(X)	L1	Continuous work is possible up to 5 days.	0	1
	L2	N work is less than 7 days per month	0	1
	L3	Consecutive work for N work is 3 days or less	0	1
	L4	There is no D shift the day after the E shift (32 hours Guaranteed break time)	0	1
	L5	N-OFF-D / N-D / N-E are not allowed.	0	1
	L6	Pregnant women do not commit to N work.	0	1
	L7	OFF work must be at least the number of holidays in the month.	0	1
	L8	N-dedicated nurses are limited to within 14 days of night work per month.	0	1
X = L1 X L2 X L3 X L4 X L5 X L6 X L7 X L8				

The Labor Standards Act(X) defined the items considered in the Labor Standards Act in Hospital E as of October 2020. After evaluating each employee's item with a total of eight items, the value converted to PASS if there is no abnormality and FAIL if there is an abnormality is given. The Labor Standards Act (X) is calculated as one if all items of L1 to L8 are valid and 0 if insufficient.

Table 2: Definition of Work Equity(Y) items

<table>
<tr><th>Index</th><th colspan="2">Contents</th><th>Score</th></tr>
<tr><td rowspan="6">Work Equity (Y)</td><td>L9</td><td>Based on average working days for each career group, Deviation of working days of the person concerned</td><td>AVG(G_{all})
– ABS(W_{all}- AVG(G_{all}))</td></tr>
<tr><td>L10</td><td>Based on the average working days for each career group in D, D working day deviation of the person concerned</td><td>AVG((G_d)
– ABS(W_d- AVG(G_d))</td></tr>
<tr><td>L11</td><td>Based on the average working days for each career group of E work, E working day deviation of the person concerned</td><td>AVG(G_e)
– ABS(W_e- AVG((G_e))</td></tr>
<tr><td>L12</td><td>Based on the average working days for each career group of N working days, Deviation of N working days of the person concerned</td><td>AVG((G_n)
– ABS(W_n- AVG((G_n))</td></tr>
<tr><td>L13</td><td>Based on the average working days for each career group of O+W work, O+W working days deviation of the person concerned</td><td>AVG((G_{o+w})
– ABS(W_{o+w}- AVG((G_{o+w}))</td></tr>
<tr><td>L14</td><td>Based on the average working days for each career group working on weekends + Fridays. Deviation of working days on weekend + Friday</td><td>AVG((G_{f+s+s})
– ABS(W_{f+s+s}-AVG((G_{f+s+s}))</td></tr>
<tr><td colspan="4">Y = L9 + L10 + L11 + L12 + L13 + L14</td></tr>
</table>

G: Career group
W: Worker

Table 2 shows work equity items. Work equity is an item that considers the difference in the number of working days of the total average worker, i.e., the worker's deviation. Each item is evaluated for equity, not based on the average of the total number of employees but each career group's average.

ABS(T_t-AVG($W_t - G_t$))is a deviation that indicates how far the number of working days of the employee is from the corresponding worker's average by experience. The deviation was converted to ABS, which is the absolute value, and generalized to positive numbers.

Furthermore, by subtracting the corresponding deviation from AVG(G_t), we designed to score higher values near AVG(G_t) as the number of working days for that worker corresponds to the employee's career-specific average.

Consequently, work equity refers to items that simultaneously consider the employee's career and equity.

Table 3: Definition of Work preference(Z) item

<table>
<tr><th>Index</th><th colspan="2">Contents</th><th>Score</th></tr>
<tr><td>Work preference (Z)</td><td>L15</td><td>• Work type(D,E,N) rank match or not
<table><tr><td>Rank</td><td>Past schedule</td><td>Creation schedule</td><td>Match score</td></tr><tr><td>1st</td><td>T_1</td><td>P_1</td><td>-10 or 10</td></tr><tr><td>2nd</td><td>T_2</td><td>P_2</td><td>-10 or 10</td></tr><tr><td>3rd</td><td>T_3</td><td>P_3</td><td>-10 or 10</td></tr><tr><td colspan="3">Mismatch = -10, match = 10</td><td>SUM</td></tr></table>※ New employees are ranked according to the overall average</td><td>-30 ~ 30</td></tr>
<tr><td colspan="4">Z = L15</td></tr>
</table>

Lastly, Table 3 shows Work preference items. Work preference is an item that considers the characteristics of each ward and worker. This Work preference assesses the rank's consistency by type of work generated by the rank model by type of work for the corresponding worker from the historical data.

The sigmoid function ($1/1+e^{-(value)}$) generated an expression so that the higher the work equity and work preference items, the closer they are to 1, and the lower they are to zero.

Finally, if all the Labor Standards Act (X) is met, the FIS is calculated as the average of the workers and 0 if insufficient.

- Train Result

Figure 8: FIS Result according to learning

Figure 8 shows the learning outcomes using this model. To apply the same environment as the manual nurse schedule before learning, the specific ward of E hospital, the career of a specific month nurse, Wanted, and the team was trained in the same consideration environment. As a result, the model was increased to 94.1 when the FIS exceeded 55 times before converging to 99.9 points..

4. Results and Discussion

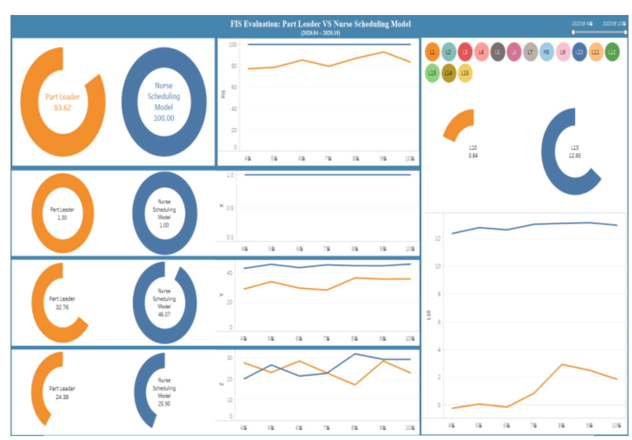

Figure 9: Visualization

The results of the study were evaluated by comparing the manual nurse schedule of part leader with this model. The evaluation was visualized by comparing and evaluating all the details of each item (L1 to L15) of the FIS. The evaluation period was compared and analyzed to the manual nurse schedule from April to October 2020 and the Top 1 nurse schedule created by the model considering each monthly worker's experience and Wanted.

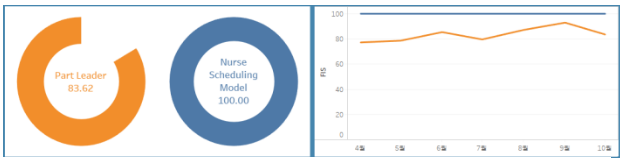

Figure 10: FIS Evaluation

Figure 10 shows the evaluation results of FIS. the evaluation of FIS showed that the results of writing this model were 16.38 higher than the manual nurse schedule. From April to October, each month's results showed a uniform 10 points difference. this model was evaluated to be superior to the manual nurse schedule in the past.

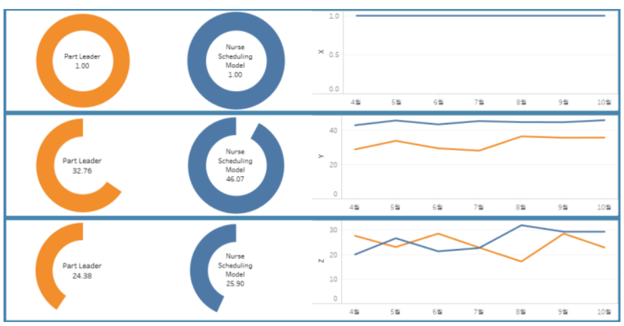

Figure 11: Labor Standards Act · Work Equality · Work Preference

The assessment of detailed items evaluated the Labor Standards Act (X), Work Equity (Y), and Work Preference (Z). The Labor Standards Act (X) showed that the manual nurse schedule and this model were met in all work tables, with the same score of 1.00. In terms of Work equity (Y), the manual nurse schedule was 32.76 points, and the model was 46.07 points, 13.31 points higher than the manual nurse schedule.

The nurse schedule shows that it appears, as shown in Figure 12 above. Rather than the manual nurse schedule above, this model's nurse schedule has been arranged for those with high night(N) work experience to be used in consideration of equity. The average number of working days per worker was the same as 17.2 days for both nurse schedules, and consecutive working days of the same type of work were similar from 2 to 4 days.

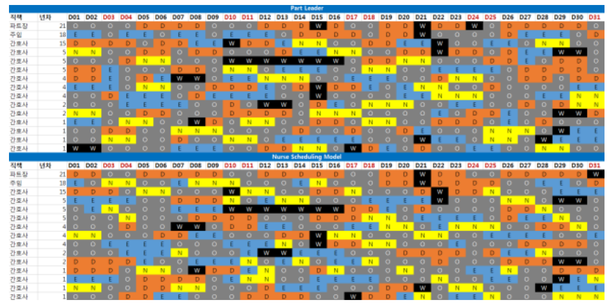

Figure 12: Nurse Schedule (Part Leader VS Nurse Scheduling Model)

5. Conclusions

This study developed a model that automatically prepares a nurse schedule using the Q-Learning Algorithm during reinforcement learning to save a lot of effort and time the part leader of the ward who is the author of the nurse schedule.

This model was able to provide a nurse schedule based on the convenience of writers and fairness between workers rather than the manual nurse schedule. Besides, Web App-based input and output features such as W2UI, Jquery, and Ajax are combined with this model to consider the ease of access between the author, ward party leader, and the corresponding ward worker.

The significance of this study can be summarized as follows. According to the FIS, it could generate an optimal nurse schedule 16.38 points higher than the manual nurse schedule, mostly 13.31 points higher in the Work Fairness (Y) category, enabling a fair nurse schedule generation.

Second, the working hours of the nurse schedule of the ward part leader may be significantly reduced. It was possible to shorten the time to prepare a manual nurse schedule, which takes up to 88 hours in the trauma ward of E Hospital, to about three hours through this model. Also, the employee may expect a work arrangement that ensures his or her work preference and fairness.

Third, it is expected that the ease and usability of access between the part leader of the ward and the workers in the ward will be improved compared to manual work when preparing a Web App-based nurse schedule. Through the further development of mobile apps in the future, it is expected that workers' requirements will be reflected in real-time to maximize the utilization of this learning model.

References

Awadallah, M. A., Bolaji, A. L., & Al-Betar, M. A. (2015). A hybrid artificial bee colony for a nurse rostering problem. *Applied Soft Computing, 35*, 726-739.

Kim, B. N., Oh, H. S., & Park, Y. S. (2011). A Study of Nurses'

Resilience, Occupational Stress and Satisfaction. *Korean Journal of Occupational Health Nursing, 20(1),* 14-23.

Kim, J. H., & Jo, H. S. (2013). A Comparative Study on Job Stress and Satisfaction between Ward Nurses and Outpatient Nurses. *Korean J Occup Health Nurs, 22(2),* 83-92.

Lu, Z. P. (2012). Adaptive neighborhood search for nurse rostering. *European Journal of Operational Research, 218*, 865-876.

Oh, Y. H. (2012). A nurse rostering problem with preference and fairness. *Korea Society of IT Services Fall Conference*, 190-193.

Park, J. C., & Kim, H. C. (2018). Simple Q-Learning using heuristic strategies. *Proceedings of the Korea Information Processing Society Conference, 25(2),* 708-710.

Song, M. Y. (2020). A Study on the Nurse Scheduling Optimization Model for Nurse Needs-Type Scheduling Automation System. *Journal of the Korea Society of Computer and Information*, 25(3), 57-64.

Václavík, R., Šůcha. P., & Hanzálek, Z. (2016). Roster Evaluation Based on Classifiers for the Nurse Rostering Problem, *Journal of Heuristics 22(5)*, 667–697.

Yoo, W. S., Seo, J. h., Kim, D. H., & Kim, K. H. (2019). Machine scheduling models based on reinforcement Learning for minimizing due date violation and setup change. *The Journal of Society for e-Business Studies*, 24(3), 19-33.

Kong, Y. H., & Lee, W. C. (2017). Dynamic Obstacle Avoidance and Optimal Path Finding Algorithm for Mobile Robot Using Q-Learning. *Journal of Korean Institute of Information Technology*, *15(9)*, 57-62.

Kang, H. G., Seo, D. S., Lee, B. S., & Kang, M. S. (2017). Applying CEE (CrossEntropyError) to improve performance of Q-Learning algorithm. *Korean Journal of Artificial Intelligence, 5(1),* 1-9.

Park, J. G., Choi, E. S., Kim, B. J., & Zhang, P. (2017). A Study on the Development of DGA based on Deep Learning. *Korean Journal of Artificial Intelligence, 5(1),* 18-28.

Lee, D. H., & Woo, C. W. (2005). Design and Implementation of Othello game Based on Reinforcement Learning. *KIISE Fall Conference, 32(2),* 778-780.