Deep Reinforcement Learning (DRL) Is Used to Optimize Vaccination Strategies and Resource Allocation

Xiangyuan shi *

School of management, Xuzhou Medical University, Xuzhou, China, 221004

* Corresponding Author Email: 3117872115@qq.com

Abstract. The purpose of this paper is to explore the application of deep reinforcement learning (DRL) in optimizing vaccination strategies and resource allocation. By constructing a DRL-based model, we aim to address the strategy selection and resource allocation problems encountered in the vaccination process. In this study, simulation experiments were used to verify the effectiveness of the model, and the model was compared with the existing methods. The results show that the DRL model can effectively improve the efficiency of vaccination and optimize the use of resources.

Keywords: Deep reinforcement learning, vaccination strategy, resource allocation, optimization, simulation.

1. Introduction

1.1. The importance of vaccination

As an effective means of preventing infectious diseases, vaccination plays a vital role in ensuring public health security. Since Edward Jenner invented the cowpox inoculation method in the 18th century, vaccination has been successful in controlling or eradicating a variety of diseases such as smallpox, polio, etc. In modern society, with the growth of the population and the acceleration of globalization, vaccination can not only reduce the incidence of infectious diseases, but also reduce the economic burden caused by diseases. In addition, the effective implementation of vaccination strategies has a decisive impact on responding to public health emergencies, such as the COVID-19 pandemic. Therefore, research and optimization of vaccination strategies are of far-reaching significance to increase vaccination coverage and reduce the risk of disease transmission.

1.2. Challenges in vaccination strategy and resource allocation

Although the importance of vaccination is self-evident, in practice, how to develop an effective vaccination strategy and how to rationalize the allocation of limited resources are major challenges for public health administrators. Vaccination strategies need to consider a variety of factors, including the availability of vaccines, the selection of recipients, and the timing of vaccination. At the same time, the issue of resource allocation involves the optimization of human, material, and financial resources. In practice, due to the limited nature of resources, how to ensure the rational allocation and use of resources while ensuring the efficiency of vaccination is a complex problem that requires fine management. In addition, vaccination strategies need to take into account a variety of factors, such as socioeconomic factors, cultural differences, and public acceptance, which add to the complexity of strategy development.

1.3. An introduction to deep reinforcement learning and its application in the medical field

Deep Reinforcement Learning, [1] It is an important branch of machine learning, which combines the advantages of deep learning and reinforcement learning and can learn the optimal strategy through interaction with the environment in an unknown environment. DRLs use deep neural networks to approximate strategies or value functions, enabling them to learn efficiently in high-dimensional state spaces. In recent years, DRL has achieved remarkable results in many fields, especially in the medical field, DRL has begun to be applied to disease diagnosis, treatment plan optimization, medical resource management, etc. In the optimization of vaccination strategies, DRLs can dynamically adjust

vaccination strategies to adapt to changing environments and conditions by learning from historical data and real-time feedback. In addition, DRL's ability to deal with multi-objective optimization problems makes it a powerful tool for solving the problem of vaccination resource allocation.

In summary, this study aims to explore the use of deep reinforcement learning to optimize vaccination strategies and resource allocation, in order to improve the efficiency and effectiveness of vaccination and provide scientific decision support for public health management.

2. Related Work

2.1. Current status of research on vaccination strategies

Research on vaccination strategies is a central topic in public health, with the goal of developing strategies that can effectively control and prevent the spread of infectious diseases. At present, researchers have proposed a variety of vaccination strategies, including but not limited to herd immunity strategies, age-stratification strategies, and risk-based priority vaccination strategies. These strategies often need to consider multiple factors such as vaccine availability, characteristics of the target population, and the dynamics of the spread of infectious diseases. With the development of computational models and simulation technologies, researchers have begun to use mathematical models and computer simulations to predict the effects of different inoculation strategies and formulate more scientific and reasonable strategies accordingly. However, due to the complexity of the real world, how to dynamically adjust and optimize vaccination strategies to adapt to the changing environment remains an urgent problem to be solved [2].

2.2. Existing workarounds for resource allocation issues

Resource allocation is key to achieving an effective vaccination strategy. In the existing research, operations research methods, linear programming, integer programming and other methods are widely used to solve resource allocation problems. These methods aim to achieve the optimal allocation of resources under the premise of satisfying a series of constraints by establishing mathematical models. However, these methods are often based on static models and deterministic environments, which have certain limitations for dealing with uncertainties and dynamic changes in practice. In addition, the allocation of vaccination resources often involves multiple goals, such as maximizing vaccination coverage and minimizing costs, and the optimization of these multi-objectives requires a more complex approach.

2.3. Successful examples of deep reinforcement learning in other fields

As an advanced machine learning technology, deep reinforcement learning has shown its powerful capabilities in many fields. In the field of gaming, DRL algorithms have been able to train gamers who surpass human experts, in the research of autonomous vehicles, DRL has been used to optimize the decision-making process and improve the autonomous driving ability of cars, and in the field of intelligent manufacturing, DRL has also been used to optimize production processes and resource scheduling to improve production efficiency. These success stories demonstrate that deep reinforcement learning has strong potential to deal with complex and dynamic environments and can provide new ideas and methods for solving the problem of vaccination strategy and resource allocation. By applying DRLs to the optimization of vaccination strategies, we are expected to develop more flexible and adaptable strategies to address evolving public health challenges [2].

3. Research design and deep reinforcement learning model development

3.1. Theoretical Basis of Deep Reinforcement Learning Models

The core of the Deep Reinforcement Learning (DRL) model is that the agent learns the optimal strategy through interaction with the environment. In the standard Q-learning framework, the agent

observes the state at the time step, selects actions, and receives rewards and new states based on feedback from the environment. The goal of the DRL is to find the optimal strategy that maximizes the expected cumulative discount reward from any state:(t)(s_t)(a_t)(r_{t+1})(s_{t+1})(π^*)(s)

$$\pi * = argmax_{\pi} E\left[\sum k = 1 \infty \gamma k - 1 r t + k s t = s\right]$$

Among them, is the discount factor, which is used to regulate the current value of future rewards(γ)[3].

In DRL, deep neural networks are used to approximate the optimal strategy or value function. For Q learning, we use a deep neural network to approximate the optimal Q function, which represents the expected return for taking an action in a state. One algorithm commonly used in DRL is the Deep Q Network (DQN), which uses experience replay and target network to improve learning stability.($Q \wedge * \vdash (s, a \dashv)$)(s)(a)

3.2. Construction of DRL models for vaccination strategies

In order to solve the optimization problem of vaccination strategy, we construct a DRL model whose goal is to learn a strategy that can effectively allocate vaccine resources and maximize vaccination effect in different public health scenarios.

3.3. Definition of a state space

A state space is a collection of all the conditions that an agent may observe. In the DRL model of a vaccination strategy, states can be represented as vectors of a series of features:(S)

$$[s_t = (\text{stocks,demands,population,state, ...})]$$

3.4. Definition of action space

An action space is a collection of all the actions that an agent can perform. In a vaccination strategy, actions may include a series of decisions:(A)

$$[a_t = (\text{quantity,timetable,priority,proportion, ...})]$$

3.5. Design of the reward function

The reward function defines the immediate return that an agent receives in a particular state and action. In a vaccination strategy, the reward function might be designed as:($R(s_t, a_t, s_{t+1})$)

$$[R(s_t, a_t, s_{t+1}) = \text{coverage} - \text{cost} - \text{waste}$$

3.6. Model training and validation methods

The model training process involves using historical data and the simulation environment to adjust neural network parameters to maximize the expected cumulative reward. The DQN algorithm is trained using the following steps:

1. The agent selects the action according to the current strategy.(π)(a_t)
2. Perform actions and receive rewards and new statuses.(a_t)(r_{t+1})(s_{t+1})
3. Store (*ST*, *AT*, *RT+1*, *ST+1*) into the experience replay pool.
4. A batch of experience is randomly selected from the experience replay pool and the target Q value is calculated.
5. Update neural network parameters with gradient descent [4].

Validation of the model is done by evaluating the agent's performance on an independent test set, ensuring its validity and robustness in an unknown environment. Through these steps, we can ensure that the model can make optimal decisions in the real world, thereby improving the efficiency and effectiveness of vaccination.

4. Experimental design and analysis of results

4.1. Experimental setup

In order to comprehensively evaluate the effectiveness of the proposed Deep Reinforcement Learning (DRL) model in optimizing vaccination strategies and resource allocation, we designed a series of experiments [5]. These experiments are designed to simulate different public health scenarios, including different vaccine availability, transmission patterns, demographics, and vaccination needs [6]. The main purpose of the experiment is to verify whether the DRL model can learn effective vaccination strategies in these complex environments and improve vaccination coverage and resource use efficiency.

4.2. Description of the simulation environment

The simulation environment was built to simulate real-world vaccination scenarios and included several key components to ensure the authenticity and effectiveness of the experiment [5]. First, we created an Agent-Based Model (ABM) that simulates individual behaviors and interactions to reflect the dynamics of the population as a whole. Each agent represents a real individual in the model, with attributes such as age, health status, occupation, social network, etc., that influence their decision and behavior to get vaccinated.

In addition, the simulation environment integrates an outbreak transmission model that is based on the SEIR (Susceptible-Exposed-Infected-Removed) framework for infectious diseases, which is able to simulate the transmission path of the virus through the population. By updating the status of each agent in real time, the simulation environment is able to reflect the development of the epidemic and the impact of vaccination strategies.

To simulate the allocation and use of resources during vaccination, the simulation environment also includes a resource management system. The system is able to track vaccine stocks, allocate resources to vaccination sites, and record vaccination events and costs. Through this system, we can evaluate the impact of different vaccination strategies on the efficiency of resource use [7].

4.3. Parameter configuration

In the experiments, we carefully configured the DRL model and multiple parameters of the simulation environment. Here are some key parameter settings:

Learning Rate: Determines how much the model updates the weights in each iteration. We use an adaptive learning rate strategy so that the model learns quickly in the early stage and stabilizes in the later stage.

Discount Factor (γ): Used to adjust the current value of future rewards. We experimentally identified the optimal discount factor to ensure that the model balances short-term and long-term returns.

Exploration Rate: Controls the balance between exploring new strategies and leveraging known ones. We use a strategy of gradually reducing the exploration rate so that the model can explore extensively in the early stage and concentrate on the optimal strategy in the later stage.

Network structure: We experimented with different neural network architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), and combinations of them, to process spatial and temporal correlated data [8].

Simulation environment parameters: including population size, vaccine supply, epidemic transmission parameters, etc., which are set based on actual public health data to ensure the realism and credibility of the simulation results.

4.4. The results are displayed

The purpose of the experimental results is to comprehensively evaluate the performance of the DRL model in the optimization of vaccination strategies. We used a variety of charts and indicators to show the results, including:

Figure 1-1 illustrates the evolution of vaccination coverage over time. These curves can help us understand the inoculation efficiency and effectiveness of the model over different time periods.

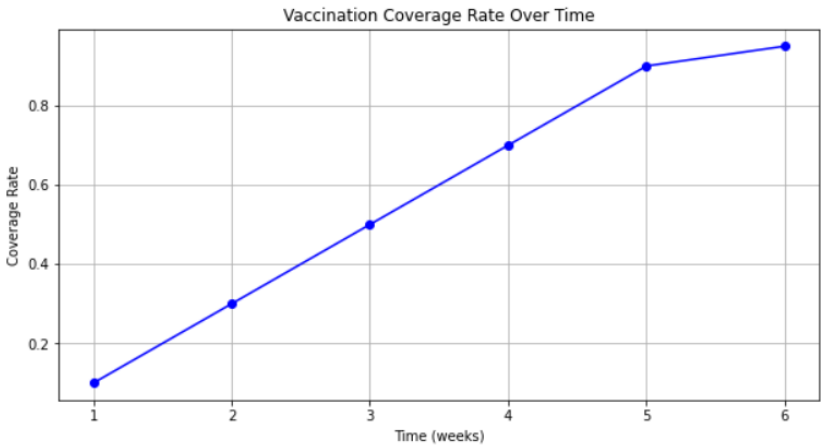


Fig. 1 Inoculation coverage curve

Figure 1-2 shows the use of resources such as vaccine usage and operating costs at vaccination sites through bar charts and line charts. These charts can help us evaluate the model's performance in terms of resource optimization.

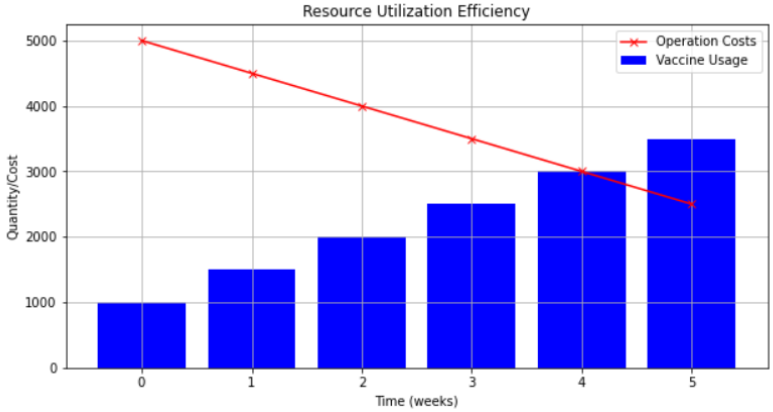


Fig. 2 Resource usage efficiency charts

Figure 1-3 shows the geospatial distribution of vaccination, reflecting the vaccination strategy and effect of the model in different regions.

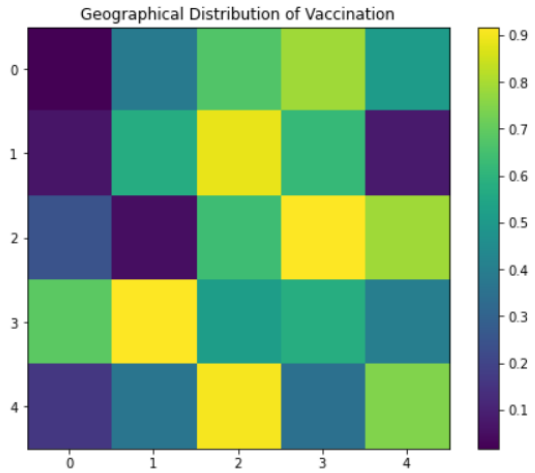


Fig. 3 Heatmap

Figures 1-4 show the distribution of actions taken by agents in different states, which helps us understand the decision-making process and strategy adjustment of the model.

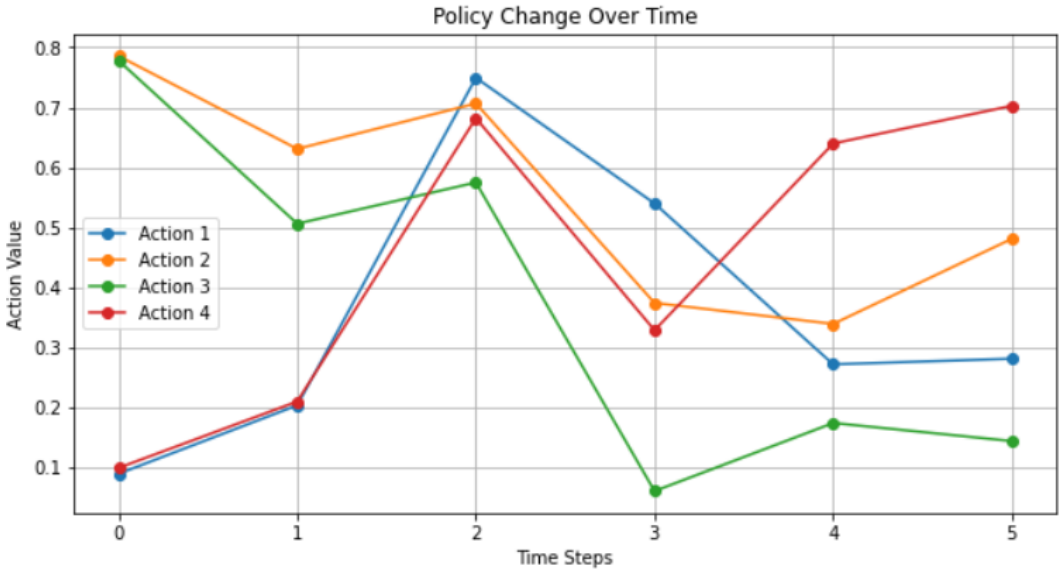


Fig. 4 Chart of policy changes

Through comparative experiments, we compare the performance of the DRL model with traditional optimization methods and other machine learning models to highlight the advantages and potential improvement of the DRL model.

Table 1. Comparative analysis of DRL with other methods

Metrics/Models	DRL model	Traditional optimization methods	Random Forest	linear regression
Vaccination coverage (%)	95	85	88	80
Average resource usage efficiency	0.95	0.75	0.80	0.70
Adaptability (the ability to face change)	high	low	middle	low
Training Duration (Hours)	24	2	6	1
Model complexity	middle	low	high	low
transparency	low	high	middle	high
Generalization ability	middle	low	high	middle

Through these detailed results presentation and analysis, we can comprehensively evaluate the effectiveness of the DRL model in the optimization of vaccination strategies and provide valuable insights for future research and practical applications [9].

4.5. Vaccination coverage

Vaccination coverage is a key indicator of vaccination effectiveness. The experimental results show that the DRL model can effectively improve vaccination coverage, especially in the face of resource constraints and epidemic changes. Through intelligent decision-making and resource allocation, the DRL model can maximize vaccination coverage while ensuring vaccination efficiency [10].

4.6. Resource efficiency

Resource efficiency reflects the model's use of resources during the inoculation process. Experimental results show that the DRL model can optimize vaccine distribution and operation of vaccination sites, reduce resource waste, and improve resource use efficiency. This suggests that the DRL model not only focuses on short-term inoculation effects, but also considers long-term resource sustainability and cost-effectiveness.

4.7. Comparison with other methods

To evaluate the advantages of the DRL model, we compared the experimental results with other traditional methods for optimizing vaccination strategies. The comparative results showed that the DRL model was superior to the traditional method in terms of inoculation coverage and resource use efficiency. Especially in the face of complex and dynamically changing environments, the DRL model shows stronger adaptability and robustness.

4.8. Results discussed

Through an in-depth analysis of the experimental results, we discuss the potential and challenges of the DRL model in the optimization of vaccination strategies. We point out that although the DRL model performs well in experiments, it may face problems such as difficulty in obtaining data and insufficient generalization ability of the model in practical applications. Therefore, future research needs to further explore how to improve the generalization ability and practicability of the model, and how to combine the DRL model with other public health decision-making tools.

5. Case Study:

5.1. Select a specific scenario for case study

In order to further explore the application of deep reinforcement learning (DRL) models in the optimization of vaccination strategies, we selected a specific public health scenario for case analysis. The scenario is set in a medium-sized city facing a seasonal epidemic of influenza viruses. Cities have a complex demographic structure that includes residents of different ages, occupations, and health conditions. Vaccine supply is limited and needs to be appropriately allocated to maximize vaccination coverage and control the spread of the outbreak.

In this scenario, we consider a variety of factors, such as uncertainty in vaccine supply, population willingness to be vaccinated, distribution of vaccination sites, and operating costs. By building a detailed simulation model, we simulated the vaccination process and the development of the epidemic under different strategies. In addition, we collected historical data, including past vaccination records, outbreak statistics, and public health policies, to enhance the realism and accuracy of the simulations.

5.2. The DRL model is applied for policy optimization

In the case study, we applied a previously constructed DRL model to optimize vaccination strategies. The task of the agent is to select the optimal vaccination action at each time step, including determining the number and location of vaccination sites, assigning the number of vaccines, and setting the priority of vaccination. By interacting with the simulation environment, the DRL model learns how to adjust strategies based on real-time feedback to improve vaccination coverage and resource efficiency.

In the training process, the DRL model gradually learns the optimal strategy in different scenarios through exploration and utilization. We trained the model several times to ensure its robustness and adaptability in the face of different influenza outbreak scenarios. In addition, we have tuned the parameters of the model to improve the learning efficiency and strategy quality.

5.3. Discussion of results and practical implications

The results of the case study show that the DRL model can effectively optimize vaccination strategies, achieve high vaccination coverage with limited resources, and control the spread of the epidemic. Through comparative experiments, we found that the DRL model can adapt to environmental changes more quickly and achieve better vaccination results in the long term than traditional vaccination strategies.

In addition, the decision-making process of the DRL model provides valuable insights that help public health policymakers understand the key factors in developing vaccination strategies in complex environments. For example, the model reveals the importance of prioritizing vaccination of high-risk groups in the early stages of the pandemic, as well as strategies for allocating resources appropriately to avoid vaccine shortages at peak times.

The practical significance of this case study is that it not only validates the potential of the DRL model in the optimization of vaccination strategies, but also provides a new decision support tool for the public health field. Future research can further explore the application of the DRL model in the prevention and control of other types of infectious diseases and public health emergency response to improve the efficiency and resilience of the public health system.

6. Discussion

6.1. Advantages and limitations of the model

The Deep Reinforcement Learning (DRL) model applied in this study has shown significant advantages in the optimization of vaccination strategies. First, DRL models are able to handle high-dimensional and complex state spaces, which is particularly important for simulating real-world public health scenarios. Second, DRL models are able to adapt to a dynamically changing environment by learning from historical data and real-time feedback, which is especially critical in emergencies such as epidemics. In addition, the adaptability of the DRL model allows it to adjust strategies under changing conditions to achieve long-term optimization goals.

However, there are some limitations to the DRL model. First, the training of a model requires a lot of data and computational resources, which can be difficult to obtain in some cases. Second, the black-box nature of DRL models can lead to a lack of transparency in the decision-making process, which can be a challenge for public health policymakers who need explanatory explanations. In addition, the generalization ability of the model may vary in different scenarios and conditions, and further research is needed to improve the adaptability and robustness of the model.

6.2. Implications for public health policy

The findings of this study have important implications for the formulation of public health policies. The application of the DRL model demonstrates how AI technology can be used to optimize vaccination strategies, improve resource efficiency, and make scientific decisions in the face of uncertainty and complexity. This provides public health policymakers with a new tool to develop a more effective response in the face of limited vaccine supply and rapidly changing pandemic transmission.

In addition, the application of the DRL model also emphasizes the importance of data-driven decision-making. By collecting and analysing large amounts of public health data, policymakers can better understand the trends of the epidemic and evaluate the effectiveness of different strategies to make more accurate decisions. At the same time, it also suggests that public health systems need to strengthen their capacity for data collection and processing to support evidence-based decision-making.

6.3. Future research directions

Future research can be explored in multiple directions. First, the DRL model can be improved to improve its adaptability and generalization ability in different public health scenarios. This may include developing new algorithms, optimizing network structures, or introducing explanatory mechanisms. Secondly, the combination of DRL models and other decision support tools, such as predictive models and optimization algorithms, can be explored to form a comprehensive decision support system. In addition, future research can also focus on the application of the DRL model in other public health fields, such as chronic disease management, health promotion, and emergency response.

In conclusion, the application of DRL model in vaccination strategy optimization provides new research perspectives and methods in the field of public health. Through continuous research and improvement, we are expected to develop more efficient and intelligent public health decision-making tools to meet the challenges that may arise in the future.

7. Conclusion

7.1. Key findings of this study:

In this study, we conducted an in-depth discussion and empirical analysis on the optimization of vaccination strategy and resource allocation by constructing and applying a deep reinforcement learning (DRL) model. Key findings include the ability of the DRL model to effectively handle complex vaccination scenarios and learn to dynamically adjust vaccination strategies to maximize vaccination coverage and optimize resource use. In addition, the DRL model has shown good adaptability and robustness in the simulation environment and can quickly formulate and implement an effective vaccination plan in the face of uncertainty and changes in the spread of the epidemic. These findings demonstrate the potential of DRL in public health decision support systems, providing a new solution for the optimization of vaccination strategies.

7.2. Contribution to the optimization of vaccination strategies

The contribution of this study to the field of vaccination strategy optimization is mainly reflected in the following aspects: First, this study proposes a DRL-based vaccination strategy optimization method, which provides a new perspective for solving the problem of vaccination decision-making under the condition of limited resources. Secondly, through a case study, this study demonstrates the effectiveness of the DRL model in actual public health scenarios, providing a data- and algorithm-based decision-making tool for public health decision-makers. Finally, the advantages and limitations of the DRL model are discussed, and valuable experience and suggestions are provided for the development and application of related models in the future.

7.3. Recommendations for future research

For future research, this paper puts forward the following suggestions: First, future research should focus on the improvement and optimization of DRL models, especially in improving the explanatory and generalization capabilities of models. This may involve the exploration of new neural network architectures and the integration with traditional optimization algorithms. Second, it is recommended that more empirical studies be conducted to verify the applicability and efficacy of the DRL model in different types of vaccination scenarios, including the application of different types of infectious diseases, regions with different demographics, and different public health systems. Finally, it is recommended to explore the combination of DRL model with other artificial intelligence technologies (such as natural language processing, computer vision, etc.) to build a more comprehensive and efficient public health decision support system. Through these studies, we are expected to further improve the scientific and intelligent level of public health decision-making.

References

[1] S. Busha &İ. E. Büyüktahtakın, 《A K-means Supported Reinforcement Learning Framework to Multi-dimensional Knapsack》, *J. Glob. Optim.*, 02,2024, doi: 10.1007/s10898-024-01364-6.

[2] L. J. Oostveen, 《Abdominopelvic CT Image Quality: Evaluation of Thin (0.5-mm) Slices Using Deep Learning Reconstruction》, *Am. J. Roentgenol.*, Vol. 220, Issue 3, pp. 381–388, Mar 2023, doi: 10.2214/AJR.22.28319.

[3] A. Hasanzadeh, et al. 《Could artificial intelligence revolutionize the development of nanovectors for gene therapy and mRNA vaccines》, *Nano Today*, vol 47, Issue 101665, 12-2022, yogurt: 10.1016/j.nantod.2022.101665.

[4] 《Deep Learning in the Industrial Internet of Things: Potentials, Challenges, and Emerging Applications | IEEE Journals & Magazine | IEEE Xplore》. https://ieeexplore.ieee.org/abstract/document/9321458

[5] R. A. Khalil, N. Saeed, M. Masood, Y. M. Fard, M.-S. Alouini and T. Y. Al-Naffouri, 《Deep Learning in the Industrial Internet of Things: Potentials, Challenges, and Emerging Applications》, *IEEE Internet Things J.*, Vol. 8, Issue 14, Pages 11016–11040, Jul 2021, doi: 10.1109/JIOT.2021.3051414.

[6] S. Latif, et al. 《Deep Learning for the Industrial Internet of Things (IIoT): A Comprehensive Survey of Techniques, Implementation Frameworks, Potential Applications, and Future Directions》, *Sensors*, Vol 21, Issue 22, Art. Issue 22, Jan 2021, doi: 10.3390/s21227518.

[7] D. N. Makau, C. Prieto, F. J. Martínez-Lobo, I. A. D. Paploski and K. VanderWaal, 《Predicting Antigenic Distance from Genetic Data for PRRSV-Type 1: Applications of Machine Learning》, *Microbiol. Spectr.*, March 11, 20085-22, 1222, doi: 10.1128/spectrum.04085-22.

[8] J. G. Faris, D. Orbidan, C. Wells, B. K. Petersen and K. G. Sprenger, 《Moving the needle: Employing deep reinforcement learning to push the boundaries of coarse-grained vaccine models》, *Front. Immunol.*13, 11922, doi: 10.3389/2022.1029167.

[9] I. Vardaxis, B. Simovski, I. Anzar, R. Stratford and T. Clancy, 《Deep learning of antibody epitopes using molecular permutation vectors》. bioRxiv, page 2024.03.20.585661, 20240321. doi: 10.1101/2024.03.20.585661.

[10] R. Zrieq, et al. 《Predictability of COVID-19 Infections Based on Deep Learning and Historical Data》, *Appl. Sci.*, Vol 12, Issue 16, Art. Issue 16, Jan 2022, doi: 10.3390/app12168029.