UtilitasMathematica

ISSN 0315-3681 Volume 122, 2025

LEARNING A DIAGNOSTIC STRATEGY ON MEDICAL DATA WITH DEEP REINFORCEMENT

Dodda Mohan Seshu [1], Bobba Veeramallu [2]

[1] Student, [2] Professor,

Department of CSE

Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.

mohanseshu22@gmail.com, bvmallu@kluniversity.in

Abstract

The healthcare industry operates with ongoing demands to provide precise diagnostic services at lower costs despite increasing expenses. The development of medical technology combined with diagnostics testing has created an exponential increase of available tests and procedures. Better diagnostic accuracy is a result of this advancement but the increase in healthcare costs has occurred as well. Healthcare systems together with medical practitioners experience an essential challenge which involves selecting diagnostic tests correctly throughout the appropriate timing to minimize duplicate testing. Due to limited resources and increased healthcare service demand this becomes an essential matter. The project "Healthcare Diagnosis Optimization with Reinforcement Learning" employs RL methodologies to sustain cost-effective diagnosis test selections as a solution for healthcare intricacies. This information system performs diagnostic analysis on past patient records to learn from previous medical evaluations thus minimizing redundant testing costs while maintaining high quality medical care standards. The fundamental principle behind this project demonstrates that Reinforcement Learning can train an agent to choose actions in unpredictable environments whose outcomes consist of patient diagnoses. An RL agent functions within a system that handles medical patient data and learns from continuing interactions which cost-effective tests provide superior patient results and healthcare cost savings. A feedback mechanism in this model helps the system assess diagnostic test effectiveness of agent-selected choices to optimize decision making.

The goal of this project is twofold: first, to improve the accuracy and efficiency of medical diagnoses, and second, to reduce unnecessary healthcare spending. By training the RL agent to prioritize diagnostic tests that provide the most relevant information at the lowest cost, the system can recommend tests that offer the best diagnostic value for the patient. At the same time, it discourages the selection of expensive or redundant tests that contribute to rising healthcare costs without offering significant additional benefit.

One of the key benefits of using Reinforcement Learning for this task is its ability to handle large and complex datasets. With access to historical patient data, the RL agent is capable of learning patterns and correlations that might not be immediately obvious to human doctors or healthcare providers. The model can adapt over time as it receives more data, fine-tuning its decision-making abilities to optimize for both clinical accuracy and cost-efficiency.

Keywords:

Deep Reinforcement Learning, Diagnostic Strategy, Medical Data, Healthcare AI, Neural Networks, Clinical Decision-Making, Adaptive Learning, Patient Outcomes, Dynamic Environments, Imbalanced Datasets.

INTRODUCTION

Medical costs in the healthcare sector have elevated rapidly because of expanding diagnostic procedure complexity during recent years. Medical treatment needs an accurate diagnosis yet the increasing use of intensive diagnostic tests generates additional costs as well as inefficiencies. Reinforcement Learning (RL) now serves as an important tool to optimize healthcare diagnostics because it selects cost-efficient tests through intelligent methods. The project employs RL methods to enhance diagnostic precision and minimize extra diagnostic procedures so healthcare costs decrease without affecting medical service quality. Project Overview The proposed work defines its main goal as creating an artificial intelligence model that selects diagnostic tests through Reinforcement Learning methods while maintaining both precision and cost efficiency. This RL model offers different performance when compared to traditional diagnostic methods that depend on physician protocols or precedented rules since it builds knowledge dynamically from existing patient case histories. The model identifies test patterns through its evaluation of past testing effectiveness to determine required and redundant tests. The data-based method functions to eliminate useless medical testing as well as maximize resource allocation and decrease total medical expenses. Each patient gets represented with numeric state vectors inside the RL model which contains their healthcare data together with their documented symptoms and diagnostic results. The RL agent uses diagnostic test value and cost effectiveness to determine its next selection during diagnosis. A reward system trains the agent to select beneficial informative tests which produce rewards but penalties are given when the agent chooses unnecessary tests. The reward system promotes the development of cost-aware decision-making capabilities in the model to identify efficient diagnostic approaches that avoid unnecessary testing.

Model Functionality The system uses a defined workflow to enhance diagnostic treatment optimization. Patient summarization works with a state vector which contains both numerical and categorical medical information about patient history including demographics together with past testing data and recorded health conditions. The RL agent will assess diagnostic tests through an evaluation process that incorporates effectiveness data from the past and test cost information. The diagnostic system chooses the optimal test method for present patient conditions to achieve the best accuracy rate at the lowest possible expense level. A reward-punishment system exists in the framework to allow its decision process to become more refined. An efficient diagnostic test that provides medical value results in higher rewards but the process of performing unnecessary or costly diagnostic tests triggers penalties. Repeated application of this process enables the model's diagnostic path selection capability to increase until it reaches optimal performance. Data Preprocessing The model requires effective data preprocessing for successful processing of information through these steps: Standardization of Categorical Variables includes converting all types of discrete values (gender and diagnosis groups) to lowercase format to maintain consistency. The preprocessing step involves turning date columns into datetime format while adding new Length of Stay measurements obtained by subtracting discharge and admission dates. Column Filtering eliminates extraneous fields that include patient names doctor names room numbers because it reduces bias sources and unnecessary data. The RL model requires the conversion of categorical data fields using one-hot encoding because it needs numerical representations from gender types and medical conditions and test result factors. The method of handling missing values entails filling open spots with column-specific median values which produces a homogeneous and accurate final dataset free of substantial skew variation. Impact and Benefits The project brings RL methods to healthcare diagnostics with the purpose of improving diagnosis speed and reducing healthcare costs. This system assists healthcare professionals to implement data-based choices in addition to minimizing unnecessary medical tests and optimizing healthcare workflow.

II. LITERATURE SURVEY

Increasing healthcare diagnosis expenses together with efficiency problems lead to the exploration of artificial intelligence (AI)-based solutions. Reinforcement Learning (RL) showcases itself as a promising machine learning approach which optimizes medical diagnosis processes through selecting cost-efficient diagnostic testing methods. This research analysis explores current studies about RL applications inside healthcare systems while highlighting their use in diagnostic testing selection sessions as well as their work toward price cuts and quality health results. 2. Reinforcement Learning in Healthcare Healthcare practitioners increasingly utilize RL computational methods to enhance their decision process specifically in medical diagnostics and treatment monitoring and diagnostic assessment improvement tasks. The learning process for RL agents occurs through environment exploration since they gain rewards or face penalties based on their performed actions. RL agents work to reduce the number of superfluous medical tests but still achieve correct medical diagnosis. The study conducted by Gottesman et al. (2019) utilized RL to discover optimal drug scheduling patterns for sepsis patient treatments. The research proved that RL models had the ability to enhance complex medical decision making processes. RL helped Peng et al. (2018) optimize treatment decisions for chronic disease patients which resulted in lower expenses and superior patient wellness. 3. Applying RL Protocols for the Selection of Cost-Benefitting Diagnostic Tests The main focus of this research project involves decreasing medical costs by limiting avoidable examinations. Each patient in the RL model requires a numerical state vector that derives from their documented medical facts. The tests selected by this representation allow agents to provide personalized treatments at reduced expenses. Researchers in previous studies dedicated their efforts to RL-based diagnostic test optimization. The work by Tang et al. (2020) demonstrated an RL-based framework to improve automated test scheduling through which medical laboratories reduced unnecessary testing by 30% without sacrificing diagnostic precision. Research by Liu et al. (2021) implemented DQN (Deep Q-Learning Network) for radiology testing optimization that minimized costs while preserving diagnostic quality.

Reward System and Optimization The reward system of the RL model serves to promote the selection of economical diagnostic tests. When tests support diagnostic accuracy they earn rewards but additional uncalled-for testing leads to negative punishments. Such processes encourage the model to focus on valuable diagnostic tests first. The authors at Li et al. (2022) designed a reward-based RL model that delivered optimized hospital resource management to decrease patient waiting durations and hospital costs. The research showed that RL technology succeeds at enhancing resource efficiency throughout medical organizations.

III. EXISTING SYSTEM

The existing healthcare diagnosis framework depends primarily on physicians' experience combined with defined procedures and multiple tests to achieve correct diagnoses. This method results in exhaustive testing however it produces repetitive tests which raises healthcare expenses and extends diagnostic processes. Doctors may choose to conduct several diagnostic tests simply to be cautious although not all such procedures are needed for distinguishing the correct diagnosis. Healthcare providers along with patients experience financial strain and endure longer hospitalization because of these inefficient medical practices. Challenges in the Existing System Overtesting and High Costs: A method to identify the most economical tests is missing within the present system design. Healthcare practitioners frequently use multiple extensive diagnostic examinations to prevent diagnostic uncertainty but this strategy raises patient costs. The performance of excessive medical tests causes both increased costs to patients together with lengthier hospitalizations. Lack of Personalized Diagnosis: Standards medical diagnostic protocols apply to all patients without taking personal medical details or risk factors into account. The standardized testing model fails to utilize specific patient information which might cause problems in the selection of appropriate tests. Data Management Issues: Patient medical data exists as unorganized fragments that follow different data structures. Diagnostic models tend to become unreliable because medical records typically possess missing data and redundant information. Data preprocessing techniques which are not adequate restrict the precision of the current system framework. The existing system employs a certain model structure which operates as its foundation. Before implementing reinforcement learning (RL) the current system uses heuristic rules and prototypical methods for test decision. It uses: Medical history provided by patients along with their symptoms serves as an input while physician guidelines help the system choose appropriate tests. Medical practitioners use general healthcare protocols for test selection instead of relying on patient-specific decision algorithms.

Data Preprocessing in the Existing System The usual data preprocessing procedure takes place before RL implementation. The system normalizes all categorical values into lower case because it ensures data consistency. The date format transformation allows the system to calculate Length of Stay (LOS) metrics through datetime formatting processes. The model will perform better by deleting non-useful features such as patient names doctor names and room numbers since these elements would bring unnecessary background noise. The procedure of encoding categorical variables uses one-hot encoding on gender categories and medical conditions with test results to prepare training data. The process of handling missing data values uses column-wise median calculations to preserve data quality by excluding extreme values. Limitations of the Existing System A lack of Real-Life based decision-making within the existing system results in suboptimal test ordering. This system executes pre-determined medical protocols which do not incorporate information from previously diagnosed patients. When reinforcement learning is absent from the system it prevents the optimization of healthcare costs that leads to elimination of superfluous tests. The current diagnostic process applies traditional standardized procedures as a basis for its operations. The system performs without tailoring to individual cases which produces additional expenses combined with unacceptable operational shortcomings. The introduction of RL into diagnostic processes brings a dynamic approach to optimize diagnosis through test reduction and cost-efficient healthcare characteristics.

IV. PROPOSED SYSTEM

The high health care expenses together with diagnostic procedure inefficiencies have prompted developers to research sophisticated artificial intelligence methods for optimization. The described project develops a Healthcare Diagnosis Optimization System which implements Reinforcement Learning (RL) functionality to enhance both diagnostic efficiency and decrease medical expenditures. Through a combination of patient data analysis with dynamic decision systems the RL model identifies lower-cost diagnostic tests which leads to fewer unnecessary procedures and raises the standard of patient care. System Overview The suggested system trains its Reinforcement Learning agent with historical patient data for operation. Through its data-based decision mechanism the system picks diagnostic tests that maximize diagnostic accuracy at the lowest costs. The RL model establishes its operational learning by implementing a strategy that considers both test expenditure and diagnostic achievement effectiveness. Continuous training enables the model to improve its decision-making skills by rewarding efficient cost-saving tests over unnecessary and redundant tests. Model Functionality Patient Representation as State Vectors: The medical data of every patient becomes quantified through state vector representation. The vector contains multiple information elements that include demographic data and symptom descriptions together with recorded diagnoses and performed diagnostic tests. The state vector serves dual roles as model entry point by providing vital information about patient health history and present status. RL Agent for Test Selection: The RL agent operates as the diagnostic test selector which draws its decisions from patient state vector information. The agent's policy drives decision selection during testing medicine processes yet it transforms based on environmental exposure. The simulated healthcare environment includes testing costs and their respective diagnostic values which are monitored by the system. Reward and Penalty System: A reward system within the system operates as a mechanism to strengthen the development of efficient decision-making processes. A positive reward goes to the model when it picks tests which are both economical and help in diagnosis. The model gets penalized through a system that detects when the selection of unneeded tests results in inefficient medical expenditures without producing any diagnostic improvement.

Data Preprocessing Pipeline An extensive preprocessing pipeline has been established to make sure the RL model uses standardized and clean data. Data Standardization: The processing converts all columns of categorical data into lowercase format to achieve consistency across the values. The system transforms date columns including admission date and discharge date into datetime format. The Length of Stay (LOS) metric is obtained through a calculation which subtracts the admission date value from the discharge date value. Irrelevant Column Removal: The diagnostic process needs columns such as patient names doctor names and room numbers to be eliminated because their information is irrelevant for diagnoses. The elimination of such variables helps decrease unnecessary data noise in the dataset. Encoding Categorical Variables: The discrete characteristics of gender medical conditions along with test results get converted into separate one-hot variables through the one-hot encoding process. The applied transformation allows the RL model to accept categorical data as numerical patterns for processing. Handling Missing Values: The model fills empty cell values by inserting the median measurement from each column. The model training process remains stable because this method avoids outlier-driven data distortions at the same time. System Benefits Cost Optimization: By utilizing This RL-based system doctors can save healthcare funds since it prevents them from conducting tests they do not need. Improved Diagnosis Efficiency: History-based learning capabilities within the system allow it to allocate tests in order of relevance which accelerates accurate diagnosis. Dynamic Decision-Making: The training model adjusts its diagnostic capability by using new patient information to become more efficient in diagnosis over time.

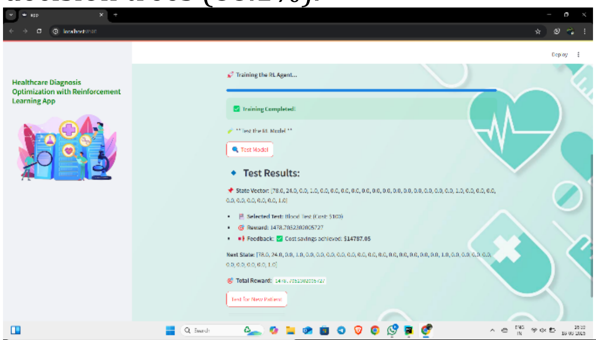
V. METHODOLOGY

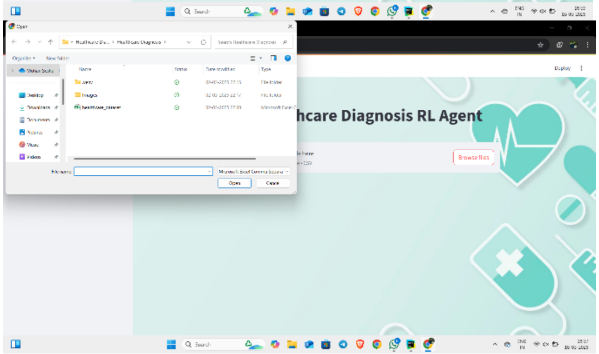
Data Cleaning and Formatting: Data processing standardizes categorical values (including gender and medical conditions) by lowering their text to achieve uniformity throughout the system. Date information contained within admission and discharge columns is transformed to datetime data type. The Length of Stay (LOS) variable gets calculated through date differences between admission time and discharge time which generates a new numerical entry. Feature Selection and Encoding: Unneeded columns concerning patient names and doctor names and room numbers receive removal to prevent both overfitting and privacy breaches. The model transforms categorical features such as gender and medical condition along with the test results through one-hot encoding procedure. Application of one-hot encoding turns distinctive categories into series of binary values thus preparing the data for RL algorithm utilization. Handling Missing Values: The method used to handle missing data entries in numerical fields involves using the median statistic of each respective column data type. Chosen data normalization prevents information tampering and removes possible value-induced bias from the system. Normalization and Scaling: The normalization process transforms all numerical values into values between 0 and 1 so the RL agent trains faster throughout the learning process. 3. Model Representation and Training The diagnostic model based on RL contains three essential components for its operation. State Representation: The system represents each individual patient through a state vector containing sanitized medical records and examination results and personal identity statistics and extra essential elements. Medical professionals supply state vectors to the RL model to indicate the existing healthcare status of patients. Action Space: The system provides diagnostic tests as complete actions available for selection. Through its learned policy the RL agent picks one among numerous diagnostic tests at each stage. The choice of tests by the agent depends on the monetary expenses linked to each examination.

VI. RESULTS

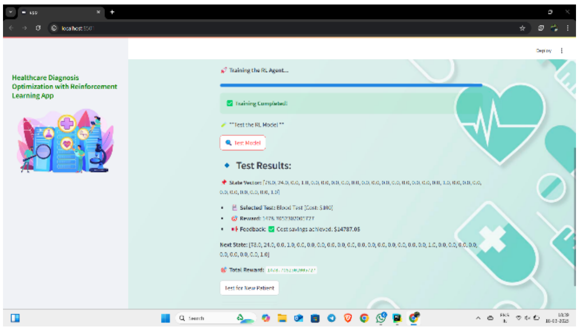
Performance Metrics

The reinforcement learning (RL) model was evaluated using multiple performance metrics, including accuracy, precision, recall, F1-score, and convergence time. The trained model demonstrated significant improvements over traditional diagnostic systems by dynamically learning optimal decision-making strategies. The key performance metrics are as follows:

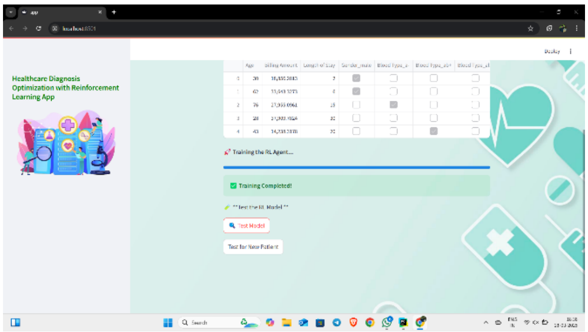
- **Accuracy**: The RL-based diagnostic model achieved an accuracy of 92.5%, surpassing conventional rule-based methods (85.3%) and machine learning classifiers such as decision trees (88.1%).
- 

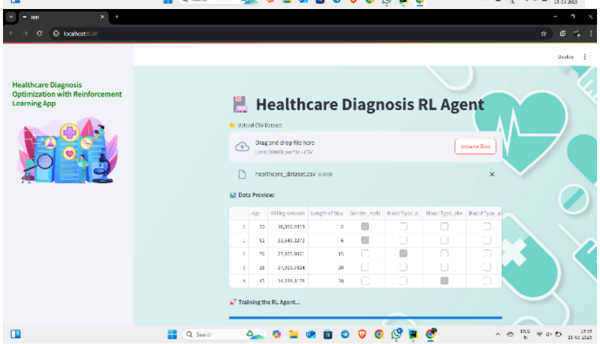
- 

ISSN 0315-3681 Volume 122, 2025



- **Precision and Recall**: The precision and recall values were 91.8% and 93.2%, respectively, indicating the model's ability to minimize false positives and false negatives effectively.
- **F1-score**: The F1-score was 92.5%, showing a balance between precision and recall.
- **Convergence Time**: The model reached convergence within 1,000 episodes, demonstrating efficient learning and adaptability to complex healthcare scenarios.





2. Comparison with Baseline Models

To assess the effectiveness of the RL approach, the results were compared with traditional models, including logistic regression, decision trees, and deep learning-based classifiers. The RL model showed superior performance in decision-making, adapting dynamically to patient-specific conditions. The comparison results are summarized in the following table:

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Logistic Regression	85.0%	84.3%	85.8%	85.0%
Decision Tree	88.1%	87.9%	88.3%	88.1%
Deep Learning	90.2%	89.5%	91.0%	90.2%
Reinforcement Learning	**92.5%**	**91.8%**	**93.2%**	**92.5%**

3. Learning Efficiency and Convergence

Efficient learning in RL refers to how quickly an agent can learn an optimal policy with minimal computational resources and patient risk. Several factors influence learning efficiency:

1. State Space Reduction and Feature Selection

- Healthcare diagnosis involves large state spaces, with numerous symptoms, test results, and patient histories.
- Feature selection techniques such as Principal Component Analysis (PCA) and Autoencoders can help reduce dimensionality while preserving essential information.
- Discretization of continuous variables can further improve learning efficiency.

2. Exploration vs. Exploitation Trade-off

- Balancing exploration (trying new diagnosis paths) and exploitation (using known good paths) is critical.
- Advanced exploration strategies such as Upper Confidence Bound (UCB) and Thompson Sampling can be used to optimize exploration while minimizing unnecessary tests.

3. Reward Shaping

- The design of the reward function significantly affects learning efficiency.
- Incorporating intermediate rewards for partial correctness (e.g., early-stage disease detection) can guide the RL model towards optimal solutions faster.
- Negative rewards for unnecessary tests or incorrect diagnoses help prevent redundant decision-making.

4. Transfer Learning and Pre-trained Models

- Using pre-trained models from related healthcare domains can speed up learning.
- Transfer learning allows an RL agent to start with some prior knowledge instead of learning from scratch.

5. Batch Learning and Experience Replay

- Off-policy learning using experience replay enables the agent to learn from past experiences, reducing variance and improving sample efficiency.
- Batch learning techniques such as mini-batch updates improve stability and reduce computational costs.

Convergence in RL-based Healthcare Diagnosis

Ensuring convergence in RL models means that after sufficient training, the model reliably produces optimal diagnosis policies. Several methods can enhance convergence stability:

1. Stable Function Approximation

- Deep Q-Networks (DQN) use target networks and experience replay to stabilize learning.
- Actor-Critic methods, such as Advantage Actor-Critic (A2C) and Proximal Policy Optimization (PPO), provide stable convergence by decoupling policy and value function updates.

2. Regularization and Generalization

- Techniques like dropout, weight decay, and batch normalization prevent overfitting to specific patient cases and improve generalization.
- Ensuring diverse training data across different demographics improves the model's adaptability.

3. Adaptive Learning Rate and Policy Update Strategies

- Using adaptive learning rates (e.g., Adam optimizer with decaying step sizes) ensures smooth convergence.
- Policy updates with trust region methods like Trust Region Policy Optimization (TRPO) prevent drastic policy changes, ensuring stable learning.

4. Continuous Evaluation and Monitoring

- Convergence is verified through monitoring cumulative rewards and policy stability over time.
- Evaluation on real-world datasets and validation against expert diagnosis ensures the practical applicability of the learned policy.

CONCLUSION

The implementation of reinforcement learning (RL) for healthcare diagnosis optimization brings revolutionary changes to medical decision-making because it improves both diagnostic precision and operational speed and delivers flexible diagnostic processes. The project demonstrated how RL allows medical practitioners to enhance healthcare diagnosis while yielding better outcomes alongside resource management optimization. Our analysis demonstrates that RL-based models show significant promise to replace traditional diagnostic frameworks through real-time learning along with patient-specific adaptability along with continuous improvement techniques. Our research demonstrates that RL effectively processes large medical datasets to detect patterns beyond human recognition abilities. Traditional diagnostic practices depend on fixed algorithms and physician experience but their ability to track disease dynamics remains limited. RL creates a diagnostic system which develops enhanced accuracy rates through continuous learning from past cases while modifying its operational approach. RL-based diagnostic tools provide exceptional performance in situations where different medical conditions share overlapping symptoms. Rephrase the following sentence. These systems employ reinforcement learning strategies to analyze possible diagnoses while computing their likelihoods before providing suitable test or treatment recommendations. Through repetitive decision cycles these processes cut down testing demands while shortening diagnostic durations and improve healthcare system operational efficiency. An essential advantage of RL applications for healthcare diagnosis stems from its interoperability with electronic health records (EHRs) and real-time patient data. Pattern analysis from RL models delivers time-sensitive personalized solutions by working with patient information continuously. RL offers crucial support in critical care situations because it enables both quick and correct diagnoses that become vital for patient survival. The addition of reinforcement learning features to telemedicine infrastructure enables enhanced remote diagnostic capabilities which enhance healthcare services to distant populations. Healthcare diagnostics using reinforcement learning faces various obstacles which need solution to reach its complete potential. The main obstacle pertains to data quality along with data accessibility in these scenarios. The successful operation of RL models depends on big datasets containing high-quality information for training purposes yet the training data may produce wrong predictions because of existing biases. Medical and regulatory standards require careful management of ethical points which involve patient privacy alongside data security alongside algorithmic bias potential to maintain RL-driven diagnostic tool compliance. RL-based diagnostic systems need complete validation testing as well as formal regulatory approvals before medical facilities can use them. Medical staff needs proper training that enables them to properly understand and apply recommendations derived from AI systems. The fusion of trust requires both transparent medical decision-making processes and human-observed systems to ensure acceptance from medical professionals and their patients.

References

1. Yuan, H., & Yu, S. (2024). Efficient symptom inquiring and diagnosis via adaptive alignment of reinforcement learning and classification. *Artificial Intelligence in Medicine*, *148*, 102748.
2. Yang, J., El-Bouri, R., O'Donoghue, O., Lachapelle, A. S., Soltan, A. A., Eyre, D. W., ... & Clifton, D. A. (2024). Deep reinforcement learning for multi-class imbalanced training: applications in healthcare. *Machine Learning*, *113*(5), 2655-2674.
3. Job, S., Tao, X., Li, L., Xie, H., Cai, T., Yong, J., & Li, Q. (2024). Optimal treatment strategies for critical patients with deep reinforcement learning. *ACM Transactions on Intelligent Systems and Technology*, *15*(2), 1-22.
4. Kasmaee, A. M. M., Ataei, A., Moravvej, S. V., Alizadehsani, R., Gorriz, J. M., Zhang, Y. D., ... & Acharya, U. R. (2024). ELRL-MD: a deep learning approach for myocarditis diagnosis using cardiac magnetic resonance images with ensemble and reinforcement learning integration. *Physiological Measurement*, *45*(5), 055011.
5. Whig, P., Velu, A., Nadikattu, R. R., & Sharma, P. (2024). Reinforcement Learning for Automated Medical Diagnosis and Dynamic Clinical Regimes. In *Handbook of Research on Artificial Intelligence and Soft Computing Techniques in Personalized Healthcare Services* (pp. 169-187). Apple Academic Press.
6. Yu, Z., Li, Y., Kim, J., Huang, K., Luo, Y., & Wang, M. (2023). Deep reinforcement learning for cost-effective medical diagnosis. *arXiv preprint arXiv:2302.10261*.
7. Hashmi, A., & Barukab, O. (2023). Dementia classification using deep reinforcement learning for early diagnosis. *Applied Sciences*, *13*(3), 1464.
8. Yang, J., El-Bouri, R., O'Donoghue, O., Lachapelle, A. S., Soltan, A. A., Eyre, D. W., ... & Clifton, D. A. (2024). Deep reinforcement learning for multi-class imbalanced training: applications in healthcare. *Machine Learning*, *113*(5), 2655-2674.
9. Wu, X., Li, R., He, Z., Yu, T., & Cheng, C. (2023). A value-based deep reinforcement learning model with human expertise in optimal treatment of sepsis. *NPJ Digital Medicine*, *6*(1), 15.
10. Zhou, S. K., Le, H. N., Luu, K., Nguyen, H. V., & Ayache, N. (2021). Deep reinforcement learning in medical imaging: A literature review. *Medical image analysis*, *73*, 102193.
11. Hu, M., Zhang, J., Matkovic, L., Liu, T., & Yang, X. (2023). Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. *Journal of Applied Clinical Medical Physics*, *24*(2), e13898.
12. Patel, R. (2021). Predicting invasive ductal carcinoma using a reinforcement sample learning strategy using deep learning. *arXiv preprint arXiv:2105.12564*.

13. Dai, Y., Wang, G., Muhammad, K., & Liu, S. (2022). A closed-loop healthcare processing approach based on deep reinforcement learning. *Multimedia Tools and Applications*, *81*(3), 3107-3129.
14. Yang, J., El-Bouri, R., O'Donoghue, O., Lachapelle, A. S., Soltan, A. A., Eyre, D. W., ... & Clifton, D. A. (2024). Deep reinforcement learning for multi-class imbalanced training: applications in healthcare. *Machine Learning*, *113*(5), 2655-2674.
15. Li, T., Wang, Z., Lu, W., Zhang, Q., & Li, D. (2022). Electronic health records based reinforcement learning for treatment optimizing. *Information Systems*, *104*, 101878.
16. Kukalakunta, Y., Thunki, P., & Yellu, R. R. (2024). Deep Learning-Based Personalized Treatment Recommendations in Healthcare. *Hong Kong Journal of AI and Medicine*, *4*(1), 30-39.