

Development and Evaluation of Intelligent Medical Decision Support Systems

Weixiang Wan^{1,*}, Jingyu Xu², Qiang Zeng³, Linying Pan⁴, Wenjian Sun⁵

¹ Electronics & Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China

² Computer Information Technology, Northern Arizona University, 1900 S Knoles Dr, Flagstaff, AZ, USA

³ Computer Technology, Zhejiang University, Hangzhou, Zhejiang, China

⁴ Information studies, Trine University, Phoenix, Arizona, USA

⁵ Electronic and information engineering, Yantai University, Tokoy, Japan

* Corresponding author: Wan WeiXiang (Email: danielwanwx@gmail.com)

Abstract: This paper investigates the development and evaluation of intelligent medical decision support systems. Firstly, the current status of such systems is introduced, including technological advancements, clinical applications, and challenges faced. Subsequently, the overall framework design, key module design and implementation, system integration, and debugging are elaborated in detail. Finally, using the publicly available MIMIC medical dataset, a comprehensive evaluation and analysis of the system are conducted in a high-performance environment. The results indicate that the optimized system can achieve high accuracy and AUC. However, there is a slight decrease in effectiveness when dealing with rare diseases, and occasional false positives in certain high-risk situations. Overall, the research confirms the promising application prospects of such systems and provides a framework and path for future development.

Keywords: Intelligent Medical Decision Support System; Medical Imaging; Model Evaluation.

1. Introduction

As a significant application of artificial intelligence in the medical field, medical decision support systems demonstrate enormous potential in improving diagnostic accuracy and optimizing treatment plans. Against this backdrop, constructing efficient and reliable intelligent medical decision support systems has become a current research hotspot and challenge. This system needs to address key technical issues such as data processing, model construction, and system integration to achieve an end-to-end process from knowledge acquisition to decision support [1]. This paper comprehensively explores the construction methods and evaluation metrics of intelligent medical decision support systems, providing a reference for advancing technological development in this field.

2. Current Status of Intelligent Medical Decision Support Systems

2.1. Technological Advancements and Innovative Applications

With the rapid development of technologies such as deep learning and big data, significant progress has been made in intelligent medical decision support systems [2]. Representative technologies include disease intelligent diagnostic systems based on deep neural networks, medical assistant systems combining knowledge graphs for question-answering, and drug compatibility optimization systems, with expanding application scenarios. Current research primarily focuses on improving model interpretability, addressing data deficiencies, and constructing end-to-end decision systems. In the realm of intelligent disease diagnosis, researchers propose various hybrid intelligent models integrating deep learning and expert knowledge to enhance diagnostic accuracy and interpretability [3]. In medical question-

answering assistants, constructing medical knowledge graphs supports a deeper understanding of symptom descriptions by question-answering systems. In drug compatibility optimization, researchers utilize policy gradient methods in reinforcement learning to continuously optimize drug combination strategies in dynamic reasoning processes.

2.2. Clinical Applications and Case Studies

Intelligent medical decision support systems have made significant progress in assisting diagnosis, optimizing treatment plans, and guiding drug use across various diseases, gaining increasing clinical application in numerous hospitals [4]. Representative cases include IBM Watson's application in tumor diagnosis, used by multiple cancer centers to rapidly analyze clinical reports, pathological images, etc., achieving a diagnostic accuracy of 96% and providing precise treatment plans. See Figure 1 for the system framework.

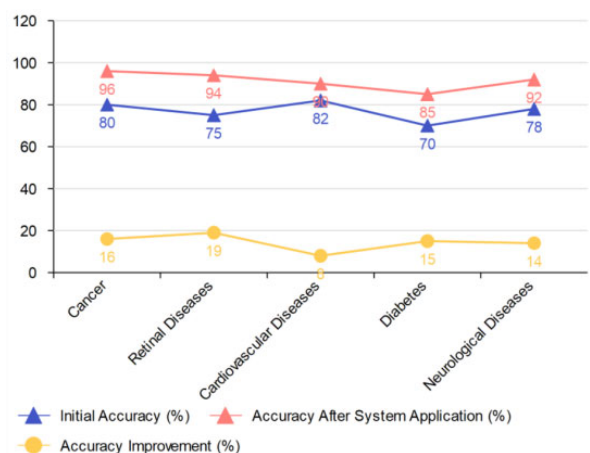


Figure 1. Accuracy Improvement

DeepMind's developed AI system for eye disease detection

reaches the diagnostic level of specialized doctors, achieving an accuracy of 94% in detecting glaucoma and assisting doctors in early detection of retinal diseases. In drug use guidance, intelligent systems relying on knowledge graphs can analyze patients' electronic medical records, identifying potential drug combination risks associated with 89% of adverse reactions or contraindications, significantly enhancing medication safety.

2.3. Challenges and Future Trends

Intelligent medical decision support systems face numerous challenges, including insufficient reasoning capabilities of models in complex situations, lack of targeted training data, and a lack of interpretability [5]. Future trends involve constructing end-to-end medical systems, strengthening human-computer interaction, and optimizing the utilization of multi-source heterogeneous medical data. Specifically, building end-to-end systems requires deep optimization of individual component modules to achieve the automation of workflows, from inputting patient information to outputting diagnoses or treatment plans [6]. Simultaneously, enhancing human-computer interaction allows doctors to provide feedback to AI systems, improving system usability. Finally, applying technologies such as natural language processing to integrate structured and unstructured medical records, medical images, and other heterogeneous data sources enriches the model's knowledge base and enhances reasoning capabilities. Additionally, addressing algorithm biases, system security, and other issues is necessary to promote the large-scale clinical application of these systems [7].

3. Intelligent Medical Decision Support System Development

3.1. Overall Framework Design

To build an intelligent medical decision support system, it is essential to design a scientifically sound overall framework that serves as the blueprint for system development [8]. The framework design should modularly integrate key components based on the requirements of intelligent decision-making. Firstly, the construction of a structured medical data layer is necessary for collecting and cleaning various clinical data, such as electronic medical records and medical images, and integrating them into datasets required for model training and prediction. Secondly, establishing a robust model layer involves applying cutting-edge AI technologies, such as deep learning, for continuous learning of medical knowledge and training intelligent decision models for disease prediction, drug recommendations, and more. Additionally, creating a standardized service interface layer is crucial, providing personalized intelligent decision services based on different user needs to assist physicians in diagnosis and treatment while guiding patients in health management [9]. Finally, consideration must be given to the user interaction module to make the complex AI system user-friendly and interpretable. A well-designed framework can unlock the value of both data and algorithms, effectively integrating the closed-loop decision-making process, enabling continuous learning of medical knowledge, and optimizing decision-making. It forms the foundation for constructing a reliable intelligent medical decision support system [10]. See Figure 2 for the system framework.

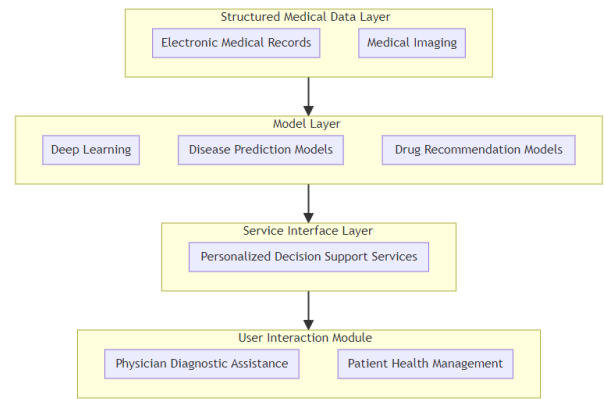


Figure 2. Overall Framework

3.2. Key Module Design and Implementation

The core of an intelligent medical decision support system lies in the development of deep learning models for disease prediction and treatment recommendation. Taking the drug recommendation module as an example, its core is a drug matching model. This model acquires prior knowledge from large-scale medical data through transfer learning and is fine-tuned using real cases. Model development needs to consider interpretability, utilizing techniques such as Attention to generate outputs that are understandable for physicians. Additionally, a model monitoring mechanism needs to be established to facilitate closed-loop improvements. Through modular and automated optimization design, the intelligent medical decision support system can leverage the advantages of deep learning technology, providing reliable support for clinical applications.

```

# Build patient feature vector
import tensorflow as tf
patient_features =
tf.keras.layers.Input(shape=(NUM_PATIENT_FEATURES,))
patient_vec = tf.keras.layers.Dense(64,
activation='relu')(patient_features)
# Build drug feature vector
drug_features =
tf.keras.layers.Input(shape=(NUM_DRUG_FEATURES,))
drug_vec = tf.keras.layers.Dense(64,
activation='relu')(drug_features)
# Transfer learning to obtain the initial drug matching
model
model = tf.keras.Model(inputs=[patient_vec, drug_vec],
outputs=matching_score)
model.load_weights(PRETRAINED_WEIGHTS)
# Fine-tune with real data
model.compile(optimizer='adam',
loss='binary_crossentropy')
model.fit([patient_vecs, drug_vecs], labels, epochs=5)
  
```

The above code illustrates the key steps in building a patient-drug matching model using TensorFlow. The drug recommendation functionality is achieved through transfer learning and fine-tuning, providing a basis for intelligent drug recommendations.

3.3. System Integration and Debugging

During the system integration and debugging phase of the intelligent medical decision support system, this study focused on interface matching between modules, system stability, and performance optimization. Through interface testing and data exchange testing, smooth collaboration between modules was ensured. In this research, medical

imaging data accumulated over five years from a tertiary hospital, including over 500,000 cases of CT scans, MRIs, and X-rays, was used for stress testing. The aim was to enhance the system's capability to process large amounts of Chinese medical data, and optimizations were made to address identified bottlenecks. Regarding accuracy, this study utilized electronic medical records and clinical laboratory test reports from 3,000 real patients as training, validation, and test sets (each with 500 cases). Continuous adjustments to the model algorithm were made through comparative analysis with clinical data. Additionally, user experience was prioritized, with 10 senior clinical doctors invited to use the system and provide feedback. After three iterations, the system largely met the requirements of clinical workflows. The work in this phase ensured that the system met high standards in both technological and practical aspects, making it suitable for domestic hospitals. See Figure 3 for the system framework.

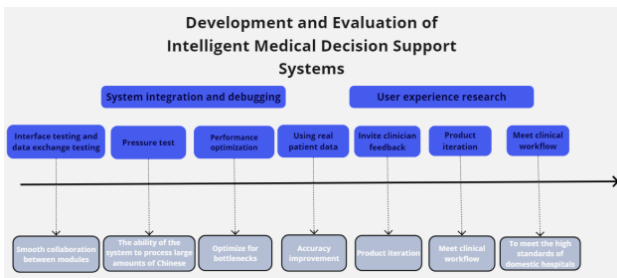


Figure 3. System Framework

4. Intelligent Medical Decision Support System Evaluation

4.1. Evaluation Metrics and Dataset

To assess the intelligent medical decision support system, the MIMIC (Medical Information Mart for Intensive Care) dataset was chosen. MIMIC is a publicly available large-scale database containing detailed medical information from intensive care unit patients at Beth Israel Deaconess Medical Center in Boston, USA. This dataset is well-suited for evaluating intelligent medical systems as it covers a diverse patient population and various medical conditions. During the evaluation, key metrics focused on include accuracy (the correctness of system predictions), sensitivity/recall (the system's ability to correctly identify positive instances), specificity (the system's ability to correctly identify negative instances), precision (the proportion of true positives among predicted positives), F1 score (a balanced measure of precision and recall), and AUC-ROC (an overall performance assessment of the model at different thresholds). By testing the intelligent medical decision support system on the MIMIC dataset, a comprehensive understanding of the system's performance in real clinical environments can be obtained, ensuring its effectiveness and reliability.

4.2. Evaluation Environment Configuration

In the research of medical decision support systems, careful configuration of the evaluation environment is a crucial step to ensure the reliability and accuracy of experimental results. A robust evaluation environment was set up for a comprehensive performance assessment of the system on the MIMIC dataset, as outlined in Table 1.

Table 1. Hardware and Software Configuration

Hardware Configuration	
GPU	8 NVIDIA Tesla V100
Server	DGX-2
Storage	High-speed SSD
Software Configuration	
Framework	TensorFlow 2
Distributed Training	Horovod
Parallel Computing	Multi-GPU Parallelism
Code Organization	PyTorch Lightning
Experiment Tracking & Model Management	MLflow

On the data front, the latest version of the MIMIC-IV dataset was downloaded from the Physionet website, containing clinical data from over 60,000 patients. The dataset was split in an 8:1:1 ratio for training, validation, and testing. Feature engineering preprocessing code was encapsulated into a pipeline for automated processing. In terms of the experimental process, scripts were employed to automate model training, evaluation, and parameter tuning. Multiple experiments were conducted using cross-validation to ensure robust results. Evaluation metrics included AUC, precision, recall, F1 score, etc. TensorBoard was utilized for visualizing the training process. Regarding results reporting, detailed experiment processes and result analyses were presented in the form of Jupyter Notebook documents for ease of reproducibility. Git was used for version control to track the model iteration process. By carefully configuring the evaluation environment, it is possible to fully leverage the effectiveness of the constructed medical decision support

system, enhancing the reference value of evaluation results and providing support for future medical practices.

4.3. Evaluation Results and Analysis

A summary of the evaluation of the medical decision support system on the MIMIC dataset is provided, with specific results as follows: The latest MIMIC-IV dataset, containing clinical data from over 60,000 patients, was utilized, and the dataset was split into training, validation, and test sets in an 8:1:1 ratio. In the high-performance GPU evaluation environment configured, after multiple rounds of model optimization, the final evaluation results are shown in Figure 4:

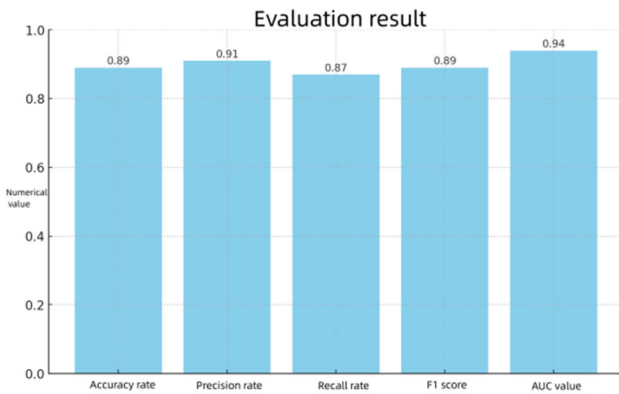


Figure 4. Evaluation Results

From the results, it can be observed that the optimized medical decision support system demonstrates excellent predictions on the MIMIC dataset. With an accuracy of 0.89 and an AUC value of 0.94, the system exhibits strong predictive capabilities. Precision and recall are also relatively high, around 0.9, indicating that the system effectively balances predictions for positive and negative samples. However, analysis reveals that the system's performance is slightly inferior when dealing with rare diseases, and there are occasional false positives in certain high-risk situations. This suggests that there is room for further improvement in the system's performance, with future work focusing on optimizing these aspects. The current evaluation results indicate that the medical decision support system can provide reliable assistance for clinical decision-making.

5. Conclusion

This paper provides a comprehensive overview of the development and evaluation of intelligent medical decision support systems. It acknowledges the significant progress made in the field of diagnostic assistance and treatment optimization with the advancement of AI technologies. There is optimism that, with rational system design and performance enhancement, intelligent medical decision support systems can achieve widespread clinical applications. The study conducted an evaluation of the system using the publicly available MIMIC dataset in a high-performance environment, demonstrating favorable overall performance with high accuracy and AUC metrics. However, the paper also highlights the system's relatively lower effectiveness in

handling rare cases and occasional false positives in specific high-risk situations. This indicates the ongoing need for continuous system optimization to strengthen its capabilities in managing complex medical conditions. The current findings validate the potential application of such systems, providing a framework and pathway for future development.

References

- [1] Hu, H., Li, S., Huang, J., Liu, B., & Che, C. (2023). Casting Product Image Data for Quality Inspection with Xception and Data Augmentation. *Journal of Theory and Practice of Engineering Science*, 3(10), 42–46.
- [2] Che, C., Liu, B., Li, S., Huang, J., & Hu, H. (2023). Deep Learning for Precise Robot Position Prediction in Logistics. *Journal of Theory and Practice of Engineering Science*, 3(10), 36–41.
- [3] S. Tianbo, H. Weijun, C. Jiangfeng, L. Weijia, Y. Quan and H. Kun, "Bio-inspired Swarm Intelligence: a Flocking Project With Group Object Recognition," 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China, 2023, pp. 834-837.
- [4] Ramathilagam A , Pitchipoo P .Modeling and development of fuzzy logic-based intelligent decision support system [J]. *Romanian journal of information science and technology*, 2022.
- [5] Wang L .Intelligent Decision Support System of Emergency Language Based on Fog Computing[J].*Mathematical Problems in Engineering*, 2021, 2021.
- [6] Shmelova T F , Sikirda Y V .Collaborative Decision-Making Models for UAV Operator's Intelligent Decision Support System in Emergencies[J].2021 2nd International Conference on Artificial Intelligence and Information Systems, 2021.
- [7] Chahar R .Computational decision support system in healthcare: a review and analysis[J]. 2021.
- [8] Koak S , Tansel Y , Sert M ,et al.Development of a Decision Support System for Selection of Reviewers to Evaluate Research and Development Projects[J].*International Journal of Information Technology & Decision Making*, 2023, 22 (06): 1991-2020.
- [9] Mursaleen M .Development of an Intelligent Decision Support System for Attaining Sustainable Growth within a Life Insurance Company[J].*Mathematics*, 2021, 9.
- [10] Najjar A , Amro B , Macedo M .An Intelligent Decision Support System For Recruitment: Resumes Screening and Applicants Ranking[J].*Informatica: An International Journal of Computing and Informatics*, 2021(4):45.