Related Works on Personalized Healthcare using Artificial Intelligence

Djelloul Daouadji Fadela

May 19, 2025

Abstract

This report summarizes recent research and developments in the field of AI-based personalized healthcare. It highlights the key contributions, methodologies, and challenges faced by existing systems.

1 Introduction

The use of Artificial Intelligence (AI) in healthcare has opened new avenues for personalized treatment. This report reviews existing studies and methods used in AI-driven healthcare solutions.

2 Related Works

3 Towards Realizing the Vision of Precision Medicine: AI-Based Prediction of Clinical Drug Response

Article Reference: [1]

Overview

This study uses machine learning to predict patient response to the epilepsy drug brivaracetam using integrated clinical and genomic data. The resulting model demonstrated strong performance (AUC: 0.76 training, 0.75 validation) and identified specific biomarkers associated with poor response. The research underscores the potential of ML models to support precision medicine and optimize clinical trials by targeting likely responders. This study highlights the potential of AI to personalize treatment strategies in epilepsy by predicting drug response, a key aspect of personalized medicine.

Dataset

- **Discovery dataset:** 235 adult patients from a phase III clinical trial (NCT01261325).
- **External validation dataset:** 47 patients from an independent trial (NCT00490035).

Data Processing

/

Approach

Multiple ML models were evaluated: sparse multi-block PLS-DA, multimodal neural networks, elastic net, gradient-boosted decision trees (GBDT), and stacked classifiers. The best performance was achieved using a GBDT model integrating all data types. GBDT models are well-suited for handling the complex interactions between clinical and genetic features, which is crucial for personalized drug response prediction. However, the inherent complexity of GBDT models can make it challenging to interpret the specific contributions of individual features, a limitation that future explainable AI (XAI) techniques could address.

Results

- AUC (training): 0.76
- AUC (validation): 0.75

Challenges

- Addressing high dimensionality and sparsity of genomic data. This is a common challenge in personalized medicine research, as genomic data often has many variables but few samples.
- Integrating additional data types (e.g., EEG, imaging) to improve model performance. Multimodal data integration is essential for a holistic view of the patient but increases complexity.
- Generalizing models to other anti-epileptic drugs. This is crucial for wider clinical applicability in personalized epilepsy treatment.
- Collaborating with regulatory bodies for clinical adoption. AI-driven personalized medicine tools require rigorous validation and regulatory approval for safe and effective use.
- Increasing dataset size to enhance model performance (targeting $\sim$350 patients for AUC = 0.9). Larger datasets are vital for building robust and generalizable predictive models in personalized healthcare.

Critique

- The sample size, while sufficient for the study, could be larger to further enhance model performance and generalizability.
- The complexity of the GBDT model, while providing good predictive power, makes it difficult to interpret the specific contributions of individual features.

4 Diabetes Prediction Using Machine Learning and Explainable AI Techniques

Article Reference: [2]

Overview

This study proposes an automated diabetes prediction system using ML and explainable AI. The system combines the public Pima Indian dataset with a private dataset collected from female workers in a Bangladeshi textile factory. The system addresses data imbalance, missing values, and is deployed for real-time prediction via web and mobile applications. The development of non-invasive AI-driven tools for diabetes detection, as presented in this paper, contributes to personalized healthcare by enabling earlier and more accessible diagnosis.

Dataset

- **Pima Indian Dataset:** 768 records, 268 diabetes-positive; includes 8 features.
- **RTML Private Dataset:** 203 female employees; features similar to Pima dataset but lacks insulin values.

Data Processing

- Zero values in the merged dataset were replaced with corresponding mean values and the dataset was separated into training and test sets using the holdout validation technique.
- Mutual information was used to measure the interdependence of variables and feature importance.
- A semi-supervised approach using the extreme gradient boosting technique (XGB regressor) was used to predict the missing insulin feature of the RTML dataset.

Approach

Various models were tested: decision trees, KNN, SVM, random forest, logistic regression, AdaBoost, XGBoost, bagging, and voting classifiers. Hyperparameters were tuned using GridSearchCV. The final model employed XGBoost with ADASYN for balancing. The choice of XGBoost is appropriate due to its effectiveness in handling complex datasets, but the lack of inherent explainability highlights the need for methods.

Results

- Accuracy: 81%
- F1 Score: 0.81
- AUC: 0.84

Challenges

- Missing insulin values required imputation via semi-supervised learning. This introduces a degree of uncertainty into the model.
- Class imbalance necessitated oversampling (SMOTE, ADASYN). Oversampling techniques can sometimes lead to overfitting.
- Limited private dataset size may hinder generalizability. Larger, more diverse datasets would improve the robustness of the model.

Future Directions

- Expanding dataset size for better robustness.
- Integrating fuzzy logic and optimization for improved prediction.

Critique

- The use of imputation for missing insulin values introduces some uncertainty.
- The private dataset is relatively small, which may limit the model's generalizability.

5 Integrating Machine Learning and Deep Learning Techniques for Advanced Alzheimer's Disease Detection through Gait Analysis

Article Reference: [3]

Overview

The paper aims to enhance early detection of Alzheimer's Disease (AD) by leveraging gait analysis combined with advanced machine learning (ML) and deep learning (DL) techniques. Gait abnormalities, such as reduced stride length and irregular cadence, are identified as early biomarkers for cognitive decline associated with AD. The study emphasizes the need for non-invasive, scalable diagnostic tools. This research highlights the potential of AI-driven gait analysis to contribute to personalized AD management through early detection.

Dataset

Data were collected using wearable sensors and motion capture systems in both clinical and real-world environments, providing high-resolution temporal and spatial gait metrics. The dataset includes gait features like stride length, cadence, swing time, and gait variability, with some data sourced from publicly available repositories like the UCI Machine Learning Repository.

Data Processing

- **Normalization:** Features were scaled between 0 and 1 to standardize the data, ensuring that features with larger ranges (e.g., stride length) did not dominate the model training.
- **Handling Missing Data:** Missing values were imputed using median substitution to maintain data integrity and reduce bias.
- **Class Imbalance:** The Synthetic Minority Over-sampling Technique (SMOTE) was applied to generate synthetic samples of the minority class (AD patients), addressing class imbalance issues.
- **Feature Selection:** Recursive Feature Elimination (RFE) was used to identify the most significant gait features—such as stride length, gait variability, and cadence—to improve model performance.
- **Correlation Analysis:** High correlations between key features (e.g., stride length and step length) validated their importance for prediction, informing feature selection.

Approach

The study employed a hybrid deep learning model comprising Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to classify individuals as healthy or at risk for AD. These models analyzed temporal-spatial gait features, capturing sequential patterns and irregularities. Traditional ML classifiers such as Random Forest and SVM were also evaluated for comparison. The use of a hybrid CNN-RNN model is a strength, as it leverages the capabilities of both CNNs for spatial feature extraction and RNNs for temporal sequence modeling, which is well-suited for gait analysis.

Results

- Hybrid CNN-RNN model accuracy: 93%
- Precision: 92%
- Recall: 91%
- F1-score: 91.5%
- AUC-ROC: 95%
- Traditional models: Random Forest (88%) and SVM (86%)

Challenges

- The reliance on controlled datasets, which may not fully reflect real-world variability, impacting model robustness.
- The complexity and interpretability of deep learning models, posing a barrier for clinical acceptance.
- The need for large, diverse datasets to ensure generalizability.
- Integration into clinical workflows and validation through real-world testing.

Future Directions

- Incorporating multimodal data sources, such as MRI, PET scans, vocal, and cognitive measures, to improve diagnostic precision.
- Expanding datasets to include diverse populations and environmental conditions, enhancing model robustness.
- Developing explainable AI frameworks to improve interpretability and clinician trust.
- Extending studies to include longitudinal gait data for monitoring disease progression and enabling earlier detection.
- Conducting clinical pilot studies and developing affordable wearable technologies for widespread, low-resource application.

Critique

- The dataset may not fully represent the variability of real-world gait patterns.
- Deep learning models are often considered "black boxes," which can hinder clinical acceptance.

6 Diabetes Detection Using Deep Learning Algorithms

Article Reference: [4]

Overview

The authors developed a non-invasive method to detect diabetes using heart rate variability (HRV) signals derived from ECG data. They designed a deep learning architecture combining convolutional neural networks (CNN) and long short-term memory (LSTM) networks to automatically extract complex features from the HRV signals. These features were then classified using a support vector machine (SVM) with an RBF kernel. The approach achieved a high accuracy of 95.7%, outperforming previous methods. This research demonstrates the potential of AI for non-invasive, personalized diabetes screening.

Dataset

- ECG recordings from 20 individuals (both diabetic and normal)
- Each participant provided a 10-minute ECG sample, from which heart rate time series data was derived
- Total datasets: 71 datasets for both groups, each containing 1000 samples

Data Processing

- Used Pan and Tompkins algorithm for QRS complex detection to extract heart rate intervals.
- Derived HRV signals directly from ECG without additional preprocessing.
- Input data fed into deep learning architectures for automatic feature learning.

Approach

- Built a deep learning model comprising 5 CNN layers followed by an LSTM layer to capture spatial and temporal features.
- Used dropout (0.1) for regularization.
- Extracted features automatically within the network, then classified using an SVM with RBF kernel.
- Employed 5-fold cross-validation for robust evaluation.

Results

- Maximum classification accuracy: 95.7% (CNN-LSTM with SVM)
- Various architectures tested with accuracies ranging from 68% to 95.7%
- The combination of deep learning feature extraction with SVM classification outperformed using deep learning alone
- Highest accuracy reported so far for non-invasive diabetes detection using HRV signals

Challenges

- Limited dataset size could affect generalization; larger datasets are needed.
- Variability in HRV signals due to individual differences may pose challenges.
- Ensuring model interpretability for clinical acceptance.
- Moving from controlled datasets to real-world, noisy ECG signals.

Future Directions

- Increase dataset size to improve model accuracy and robustness.
- Explore anomaly prediction techniques by analyzing dynamic characteristics in HRV data.
- Develop more advanced deep learning models for early and accurate detection.
- Investigate applicability to real-time monitoring and broader clinical validation.

Critique

- The dataset size is limited (only 20 individuals), which may affect the model's ability to generalize to larger populations
- Like other deep learning models, the interpretability of the model could be a concern for clinical use

7 CardioXNet: A Novel Lightweight Deep Learning Framework for Cardiovascular Disease Classification Using Heart Sound Recordings

Article Reference: [5]

Overview

This paper introduces CardioXNet, a lightweight CRNN architecture designed for the automatic detection of five types of heart sounds using raw PCG signals. The architecture involves two main phases: representation learning to extract time-invariant features and sequence residual learning to extract temporal features. CardioXNet is designed to be efficient for use in low-resource settings.

Dataset

- **Primary Dataset:** GitHub PCG database containing 1000 recordings across five classes: Normal, Aortic stenosis, Mitral regurgitation, Mitral stenosis, and Mitral valve prolapse.
- **Secondary Dataset:** PhysioNet/CinC 2016 challenge dataset with 3240 recordings labeled as normal or abnormal, used to test the model's generalizability.

Data Processing

Normalisation

Approach

The authors developed CardioXNet, a lightweight CRNN framework with two learning schemes:

- **Representation learning:** Extracts time-invariant features using three parallel CNN pathways
- **Sequence residual learning:** Extracts temporal features using bidirectional connections
- The model is specifically designed to be efficient in terms of parameters and computational requirements

Results

- Achieved 99.60% accuracy on the GitHub dataset, outperforming prior methods.
- Demonstrated 86.57% accuracy on the PhysioNet dataset, indicating good generalization.

Challenges

- Limited dataset size, especially for specific conditions like HVD
- Lack of patient independence and demographic details
- Generalization to real-world, heterogeneous data remains untested

Future Directions

- Incorporate larger and more diverse PCG datasets
- Integrate CardioXNet into wearable devices with cloud connectivity
- Explore transfer learning and further model compression for resource-constrained deployment
- Develop methods to handle noisy recordings and variable acoustic environments

Critique

- Dataset size and diversity might limit generalizability to broader populations
- Missing demographic and variability information limits assessment of performance across different patient groups
- Robustness to real-world noise and device variation not fully evaluated
- Limited evaluation in actual clinical settings

8 Optimizing Type 2 Diabetes Management: AI-Enhanced Time Series Analysis of Continuous Glucose Monitoring Data for Personalized Dietary Intervention

Article Reference: [6]

Overview

This study proposes a method to optimize type 2 diabetes management using AI-enhanced time series analysis of continuous glucose monitoring (CGM) data. The goal is to enable personalized dietary interventions to improve patient outcomes.

Dataset

- Collected CGM data from 8 patients with type 2 diabetes.
- Data includes time-series blood glucose (BG) values.

Data Processing

- Removed NaN records to clean the data.
- Applied feature extraction and dimensionality reduction techniques.
- Split dataset into training (75%) and testing (25%) portions.

Approach

- Used regression models: XGBoost, SARIMA, Prophet.
- Integrated dietary recommendations based on predicted BG levels.

Results

- XGBoost outperformed SARIMA and Prophet.
- Achieved high R^2 and low MAPE.
- Accurately predicted glucose fluctuations for timely intervention.

Challenges

- Limited dataset (only 8 patients) reduced model generalizability.
- Up to ±30-minute lag in CGM readings affected prediction accuracy.

Future Directions

- Expand dataset to improve training and validation.
- Enhance model with additional features and contextual data.

Critique

- Promising approach, but small sample limits robustness.
- Time lag in CGM data must be addressed for better accuracy.

9 AI-Based COVID-19 Detection Using CT Imaging

Article Reference: [7]

Overview

The article presents a deep learning algorithm developed to screen for COVID-19 using CT images, motivated by the limitations of traditional pathogenic testing methods (e.g., nucleic acid testing), which may yield false negatives. The main objective is to provide rapid and accurate diagnoses to help mitigate the spread of the virus.

Data Used

- A total of 1,065 CT images were collected.
- Included 180 cases of typical viral pneumonia and 79 confirmed COVID-19 cases from three hospitals.
- An additional 15 COVID-19 cases were added where initial nucleic acid tests were negative.

Data Preprocessing

- All images were resized to 299 × 299 pixels for input consistency.
- Lung regions were manually delineated to focus on regions of interest (ROIs), improving model focus.
- Images were converted to a virtual RGB format to match the input requirements of the modified inception model.

Approach Used

- A transfer learning strategy was employed using a modified inception model, referred to as M-inception.
- Only the customized layers were trained, leveraging pre-trained weights to reduce overfitting and training time.
- The architecture aimed to classify CT images to distinguish between COVID-19 and other viral pneumonias.

Results

- Internal validation accuracy was 89.5% (Specificity: 0.88, Sensitivity: 0.87).
- External test accuracy was 79.3% (Specificity: 0.83, Sensitivity: 0.67).
- In 54 cases where nucleic acid tests were initially negative, the model correctly predicted 46 cases, achieving an accuracy of 85.2%.

Challenges

- CT images often include irrelevant elements, complicating automated diagnosis.
- A relatively small training dataset may reduce generalizability and increase model bias.
- Low signal-to-noise ratio and data heterogeneity pose significant hurdles in applying deep learning in clinical diagnostics.

Future Directions

- Expand the dataset to cover a larger number of COVID-19 cases across various pathological stages.
- Optimize the model's robustness, accuracy, and reliability by including diverse imaging scenarios.

Critique

- The study provides a compelling proof-of-concept for AI-based diagnostics, but its effectiveness is limited by dataset size.
- Future work should emphasize external validation and distinguishability from other respiratory diseases.

10 Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning

Article Reference: [8]

Overview

This study presents the development of an artificial intelligence (AI) system based on transfer learning to diagnose medical conditions using medical imaging. The authors focus on the classification of images for macular degeneration, diabetic retinopathy, and pneumonia from chest X-rays. The aim is to demonstrate that AI can improve diagnostic accuracy and speed, often matching or exceeding human performance.

Data Used

- **Optical Coherence Tomography (OCT)**: Out of 207,130 collected OCT images, 108,312 high-quality images were retained after a quality review for training.
- **Chest X-ray Images**: A dataset consisting of 5,232 pediatric chest X-rays was used to train the model for pneumonia classification, with an additional 624 patient images used for testing.

Data Preprocessing

- Quality control to filter OCT images.
- Categorization of images according to medical diagnosis to facilitate supervised learning.

Approach

- The model utilizes a transfer learning strategy, initializing from pre-trained convolutional neural networks to boost performance despite the relatively small dataset size.
- Classification tasks included urgent referral detection for conditions like choroidal neovascularization and diabetic macular edema.
- The approach also focused on distinguishing between bacterial and viral pneumonia on pediatric chest X-rays.

Results

- **96.6% accuracy** in OCT image classification.
- **92.8% accuracy** in pneumonia detection from chest X-rays.
- Performance was comparable to that of experienced clinicians, highlighting the potential of AI in supporting medical decision-making.

Challenges

- Difficulty in obtaining large and diverse labeled medical datasets.
- Clinical interpretability of the AI decisions, a common barrier in the adoption of deep learning in healthcare.

Future Directions

- Expanding the dataset with images from multiple imaging devices and settings to improve model generalizability.
- Extending the transfer learning approach to other imaging modalities beyond OCT and X-ray.

Critique

- While the approach is promising and the results impressive, the generalizability of the model remains a concern due to dataset limitations.
- Further validation with larger, more heterogeneous datasets is necessary.
- The open release of data and code is a notable contribution, encouraging reproducibility and further research in AI for medical imaging.

11 Enhancing Heart Disease Prediction with Reinforcement Learning and Data Augmentation

Article Reference: [9]

Overview

This study aims to improve the prediction accuracy of heart disease by integrating reinforcement learning (RL) and data augmentation techniques. The approach addresses the complexities of cardiac data, which often hampers traditional machine learning models, by leveraging advanced methods to enhance predictive performance and early diagnosis.

Dataset

The primary dataset employed is similar to the Cleveland Heart Disease dataset, sourced from the UCI Machine Learning Repository. It contains features such as age, gender, blood pressure, cholesterol levels, ECG results, and other patient health indicators. The dataset includes a target variable indicating the presence or absence of heart disease, facilitating classification tasks. Additional datasets might come from healthcare agencies and research repositories.

Data Processing

- **Feature Selection:** Techniques such as feature importance scores and recursive feature elimination were used to identify the most impactful variables for heart disease prediction.
- **Data Augmentation:** Applied transformations like feature scaling, rotation, noise addition, and synthetic data generation to expand and diversify the training data. This helps models handle variability and reduce overfitting.

Approach

- **Reinforcement Learning (RL):** Utilizing RL algorithms to optimize decision-making processes dynamically, allowing models to adapt to evolving patient data and improve prediction accuracy over time.

How It Functions in the Study:

- **Initialization:**
  - The RL agent starts with an initial policy, possibly based on prior knowledge or random actions.
  - The dataset is preprocessed, and the model's initial parameters are set.
- **Interaction Loop:**
  - For each episode, the agent:
    - Observes the current state (e.g., patient features).
    - Selects an action according to its policy (e.g., choosing a specific augmentation or parameter setting).
    - Executes the action, which may involve training the model further, updating parameters, or selecting data augmentation techniques.
    - Moves to the next state, reflecting the outcome of its action, such as improved data representation or better predictive performance.
    - Receives a reward based on the effectiveness of its action, such as increased accuracy or better generalization.
- **Learning:**
  - The agent updates its policy based on the feedback (rewards), aiming to improve decision-making over future episodes.
  - Techniques like Q-learning or policy gradients are often used to optimize this process.
- **Outcome:**
  - Over many iterations, the RL model learns which actions lead to higher rewards and adapts its strategy to improve heart disease prediction accuracy continually.

In summary:

- **States** represent patient data or model status.
- **Actions** correspond to decisions like data augmentation choices or model updates.
- **Rewards** are signals (e.g., accuracy improvements) guiding the learning process.
- The RL agent learns the best policy to update the model continuously, maximizing prediction performance.

Results

- Achieved an accuracy rate of approximately 94%, surpassing traditional models

Challenges

- **Computational Complexity:** The combined methods demand significant processing power and longer training times.
- **Data Quality and Accessibility:** The efficacy of data augmentation depends heavily on the quality of the original dataset; biases or missing data can impact outcomes.
- **Model Generalizability:** Design choices and assumptions within the RL framework may limit applicability across diverse patient populations or clinical settings.
- **Scalability:** Handling large-scale, real-world datasets remains challenging due to resource requirements.

Future Directions

- **Fine-tuning Techniques:** Further optimizing model parameters and augmentation strategies.
- **Privacy and Security:** Incorporating mechanisms to ensure patient data privacy.
- **Clinical Validation:** Conducting extensive real-world clinical trials to validate model usefulness and safety.
- **Broader Application:** Extending the methodology to other medical diagnostic areas beyond heart disease.
- **Reducing Computational Costs:** Developing more efficient algorithms to make the approach more scalable and practical in healthcare settings.

Critique

- The integration of RL and data augmentation is promising, but computational demands and reliance on data quality could hinder deployment in resource-limited settings.
- The paper lacks details on the exact RL algorithm used, which is essential for reproducibility.
- There is limited discussion on how interpretability is addressed, which is crucial for clinical use.

12 A Reinforcement Learning–Based Method for Management of Type 1 Diabetes: Exploratory Study

Article Reference: [10]

Overview

The researchers developed a reinforcement learning (RL) framework, specifically a Q-learning algorithm, to personalize insulin dosing for patients with Type 1 Diabetes Mellitus (T1DM). The aim was to improve blood glucose management by recommending insulin doses tailored to individual patient characteristics.

Dataset

The dataset consisted of clinical records from 87 T1DM patients treated at Mass General Hospital (MGH) between 2003 and 2013.

The data included patient information such as HbA1c levels, BMI, physical activity, and alcohol usage.

The authors conducted a correlation analysis to identify key variables influencing blood glucose, concluding that HbA1c, BMI, activity level, and alcohol usage were the most relevant.

Based on these factors, they defined the patient's state by discretizing these variables into levels:

- HbA1c levels (e.g., normal, elevated, high)
- BMI categories
- Activity levels (e.g., low, high)
- Alcohol usage levels (e.g., none, moderate, high)

Data Processing

States were formed as combinations of these discretized features.

Insulin doses (actions) were defined within specific ranges (e.g., Lantus dose intervals).

The data was used to train and validate the RL model.

Approach

- The Q-learning algorithm was employed, which is a model-free RL method.
- The environment is represented by the patient's health state, and the agent makes decisions on insulin dosage.
- The states are defined by the combination of HbA1c, BMI, activity level, and alcohol usage.
- The actions are discretized insulin dose levels (specific dose intervals).
- The reward function is designed based on how well the insulin dose achieved the target HbA1c level.

Approach

- At each time step (e.g., clinical visit), the agent observes the state and selects an action (insulin dose) either by exploration (random choice with probability ϵ) or exploitation (based on learned Q-values).
- After administering the dose, the patient's response (e.g., change in HbA1c) results in a reward, guiding the learning process.
- The Q-values are updated iteratively based on the reward and the estimated value of subsequent states.

Results

- RL model tested on 60 unseen cases
- Recommended insulin dose interval included the physician-prescribed dose in 88% of cases
- Results suggest the RL approach can effectively offer personalized treatment recommendations aligning with clinical decisions

Challenges

- **Limited dataset size:** Only 87 patients, which may limit generalizability
- **Discretization of variables:** Fineness of categories might affect the model's precision
- **Data quality and missing variables:** Not all potentially influential factors (like diet or stress) were included
- **Algorithm complexity:** RL models require careful tuning; real-world implementation must address issues like exploration vs. exploitation and patient safety

Future Directions

- Extend the model to include other types of insulin and medications
- Incorporate finer categories or continuous variables for more precise recommendations
- Validate with larger and more diverse datasets
- Explore application to other populations, such as patients with Type 2 Diabetes
- Implement real-time decision support in clinical settings

Critique

- Limited patient sample size may not capture all variability in diabetes management
- Discretization may lead to loss of information that could be valuable for precise dosing
- No explicit mention of model validation techniques beyond testing on 60 cases
- Reward function description is brief; more detail would clarify alignment with clinical goals

13 Deep Q-Network (DQN) Model for Disease Prediction Using Electronic Health Records (EHRs)

Article Reference: [11]

Overview

The paper addresses the challenges of using deep learning models for disease prediction with EHRs, including lack of precision, ethical concerns, limitations of small datasets, complexity of data processing, and incompleteness of patient data. It proposes a deep Q-learning (DQL) model to enhance the accuracy and stability of predictions. The model integrates reinforcement learning with neural networks, utilizing the mapping capabilities of the Q-network. The proposed model is evaluated on the Heart Disease Dataset from the UCI Data Repository, demonstrating high accuracy (98%) compared to other models.

Dataset

- The Heart Disease Dataset from the UCI Data Repository via Kaggle.
- Includes multivariate numerical data with 14 attributes (categorical, integer, and real data).

Data Processing

- The dataset was split into training (80%) and test sets (20%) using stratified splitting.
- Robust scaling was applied for feature scaling.

Approach

- The study uses a Deep Q-Network (DQN) model, which combines reinforcement learning with neural networks.
- This approach aims to address the limitations of traditional Q-learning and improve the accuracy and stability of disease predictions.

Implementation of Reinforcement Learning

- The model uses a "disease prediction game" environment.
- State: The state is represented by the samples from the EHR data.
- Action: Actions involve increasing, decreasing, or holding each feature.
- Reward: Positive rewards are given for accurate predictions, and negative rewards for inaccurate ones.
- Environment: The environment consists of sets of states (samples) and actions (feature adjustments).

Results

- The paper compares the Deep Q-Network (DQN) model with Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and Gradient-Boosting Classifier.

Key Results:

- The proposed EHR-DQN model achieved high accuracy (0.9841) and a low mean squared error (MSE) of 0.0001.

- The Decision Tree and Gradient-Boosting Classifiers also performed well, with perfect precision, recall, and F1 scores.
- The Random Forest model showed a balance of high accuracy (0.9783) and a minimal MSE of 0.
- Logistic Regression had the lowest performance, with an accuracy of 0.8424 and a higher MSE of 0.1576.

Challenges

- Data-related issues: digitization, consolidation, and availability of health records.
- Privacy and legal concerns with patient data handling.
- Patient-related difficulties: decision-making errors, treatment errors, and data inconsistencies.
- Technical challenges: data integration across systems, ensuring patient security.
- Interpretability: "black box" nature of complex AI models limits transparency.
- Resource limitations: computational resources and expertise needed for implementation.

Future Directions

- Expand the dataset size to better train deep reinforcement learning models.
- Integrate additional contextual features to improve prediction accuracy.
- Develop more transparent models to address interpretability concerns.
- Create more standardized methods for validating healthcare AI systems.
- Improve integration with clinical workflows for practical implementation.

Critique

- The paper emphasizes the high accuracy of the proposed model but also points out that the dataset used has only thousands of samples. The DQN algorithm typically requires a large number of training samples (millions) to achieve optimal performance.
- The paper acknowledges the "black box" nature of many AI algorithms and the resulting lack of interpretability, which can hinder trust and adoption in healthcare.

14 A Reinforcement Learning Model for AI-Based Decision Support in Skin Cancer

Article Reference: [12]

Overview

- The article presents a reinforcement learning (RL) model designed to enhance decision support in skin cancer diagnosis.
- It compares the effectiveness of RL with supervised learning (SL) methods, emphasizing the potential of RL in optimizing management decisions.

Dataset

- They utilized the publicly available HAM10000 dataset, a comprehensive collection of dermatoscopic images of skin lesions.
- This dataset contains over 10,000 images across seven diagnostic categories of skin lesions.

Data Processing

- For patient-centered scenarios, input vectors were normalized by dividing position-wise by the average across all lesion vectors of the same patient.
- This normalization was crucial for processing multiple lesions from the same patient effectively.
- Feature extraction was performed using convolutional neural networks pre-trained on the dataset.

Approach

- Supervised Learning (SL): A convolutional neural network was fine-tuned for classifying seven categories of skin lesions.
- Reinforcement Learning (RL): A deep Q-learning model was developed using a multilayer perceptron that processed feature vectors derived from the SL model.
- They also implemented a threshold-based approach for comparison.

Implementation of Reinforcement Learning

- States: One-dimensional vectors representing features of skin lesions.
- Actions: Management strategies including excision, short-term follow-up, and regular surveillance.
- Rewards: Defined based on diagnosis outcomes and management decisions, optimizing for both clinical efficacy and resource utilization.
- Learning: The model used experience replay and target networks to stabilize learning.

Results

- The RL method and threshold method both improved management decisions compared to the naïve SL model.
- Both approaches optimized operating points on decision curves, enhancing diagnostic accuracy.
- The RL model demonstrated superior performance in balancing sensitivity and specificity.
- The study showed potential for reducing unnecessary excisions while maintaining high detection rates.

Challenges

- RL models require more complex retraining compared to simpler threshold approaches.
- Integration into clinical workflows presents practical implementation barriers.
- Limited consideration of patient preferences in the current model design.
- Lack of longitudinal data to validate long-term decision outcomes.

Future Directions

- Development of reward tables incorporating both physician and patient preferences.
- Enhancement of shared decision-making tools combining AI recommendations with clinical expertise.
- Integration of additional clinical parameters beyond image data.
- Validation in prospective clinical trials to assess real-world impact.

Critique

- It focuses primarily on physician preferences, potentially neglecting patient-centered care principles.
- Limited dataset diversity may restrict generalizability across different patient populations.
- The complexity of RL models may limit practical implementation in resource-constrained settings.
- Further validation is needed to demonstrate clinical utility beyond technical performance metrics.

15 Reinforcement Learning Using Deep Q Networks and Q-Learning Accurately Localizes Brain Tumors on MRI with Very Small Training Sets

Article Reference: [13]

Overview

This study applies reinforcement learning (RL), specifically Deep Q Networks (DQN), to accurately localize brain tumors in MRI scans. It addresses the limitations of traditional supervised learning methods, particularly the dependence on large annotated datasets. The authors demonstrate that RL can achieve high performance even with minimal training data.

Dataset

The dataset used comprises 2D slices from the 2014 BraTS (Brain Tumor Segmentation) challenge dataset. These are T1-weighted contrast-enhanced MRI images. The training set included 30 images, with another 30 images used for testing.

Data Processing

- 2D slices were extracted and the image space was divided into a grid.
- Each agent operates within a 60×60 pixel block.
- No data augmentation was applied to keep training consistent with the original dataset.

Implementation of Reinforcement Learning

- **Environment:** Modeled as a gridworld over MRI images.
- **States:** Represented by the agent's position in the image grid.
- **Actions:** The agent can move down, move right, or stay in place.
- **Rewards:** Positive rewards for entering tumor regions, penalties for remaining idle outside the tumor area.
- The DQN used experience replay and a target network to stabilize learning.

Results

- The DQN achieved an average localization accuracy of 70% over the last 20 episodes.
- This significantly outperformed a supervised deep learning approach, which achieved only 11% accuracy.
- Results demonstrate RL's superiority with small datasets in this context.

Challenges

- The reliance on small datasets may limit generalizability.
- High computational cost of RL may pose integration challenges in real-time clinical settings.

Future Directions

- Extend the RL model to handle full 3D MRI volumes.
- Compare performance with other RL strategies such as policy-gradient methods.
- Optimize training and inference to improve clinical applicability.

Critique

- The model's performance with small data is promising but needs validation on larger, more diverse datasets.
- No integration of clinical context or physician-in-the-loop evaluation.
- Despite excellent performance metrics, interpretability and ease of integration in hospitals remain unresolved.

16 Deep Reinforcement Learning and Simulation as a Path Toward Precision Medicine

Article Reference: [14]

Overview

The paper explores the application of deep reinforcement learning (DRL) to develop personalized treatment strategies for sepsis, a severe condition caused by the body's dysregulated immune response to infection. The authors propose a novel approach that leverages simulation to identify effective multicytokine therapies tailored to individual patients based on their systemic measurements.

Dataset

- The study used an Innate Immune Response Agent-Based Model (IIRABM) to simulate sepsis progression.
- A set of 500 simulated patients with diverse genetic and physiological parameters was generated.
- No real-world patient data was used directly, as the approach relies on in silico modeling.

Data Processing

- Simulated patient data included various systemic measurements relevant to sepsis.
- Continuous monitoring of cytokine levels, immune cell populations, and tissue damage metrics.
- Patient states were represented as high-dimensional continuous variables.

Implementation of Reinforcement Learning

- **Environment:** Simulated sepsis progression using an agent-based model (ABM) of the innate immune response to infection.
- **States:** High-dimensional and continuous state space representing patient condition through various health metrics and systemic measurements.
- **Actions:** Continuous action space for administering multicytokine therapies with varying dosages.

- **Rewards:** Terminal rewards of +250 for healed patients and -250 for death cases, with intermediate rewards based on tissue damage changes and an L1 penalty to encourage conservative treatments.
- **Learning Process:** Episodes of patient treatment where the agent learns to refine its policy to maximize cumulative rewards.

Results

- DRL-based approach achieved significant reduction in mortality rates across 500 simulated patients.
- Demonstrated ability to learn complex therapeutic strategies adaptive to individual disease progression.
- Outperformed conventional standalone antibiotic therapy in simulations.

Challenges

- Ethical concerns related to exploring suboptimal treatments in real patients.
- Complexity of dynamic multicytokine mediation requiring extensive exploration of treatment strategies.
- Need for advanced strategies to achieve lower mortality rates across diverse patient simulations.

Future Directions

- Development of measurement technologies informed by simulation findings.
- Identification of biological targets for drug development.
- Design of clinical trials for adaptive personalized therapies.
- Expansion of simulation and DRL approach to other medical conditions beyond sepsis.

Critique

- Heavy reliance on simulation may not fully capture real-world clinical complexities.
- Ethical implications of testing derived treatment strategies remain significant.
- Clinical validation through trials needed to confirm applicability in real-world settings.
- Gap between simulation-derived policies and practical clinical implementation.

17 PrescDRL: Deep Reinforcement Learning for Herbal Prescription Planning in Treatment of Chronic Diseases

Article Reference: [15]

Overview

This paper presents PrescDRL, a novel framework that applies deep reinforcement learning to optimize herbal prescriptions in Traditional Chinese Medicine (TCM) for the treatment of chronic diseases. Instead of focusing on immediate treatment effects, the approach emphasizes long-term patient outcomes through sequential treatment strategies, specifically focusing on diabetes as a case study.

Dataset

- A custom high-quality benchmark dataset for sequential diagnosis and treatment of diabetes.
- Data includes patient medical records and treatment histories specifically curated for evaluating the PrescDRL model.
- Dataset constructed to enable comparison between the model's recommendations and traditional treatment methods.

Data Processing

- Cleaning and normalization of patient medical records.
- Standardization of herbal prescription information.
- Feature extraction from patient medical histories to capture relevant health indicators.

Implementation of Reinforcement Learning

- **State:** Representation of patient health status derived from medical records and treatment history.
- **Actions:** Selection from 30 well-tuned herbal prescription candidates (HPCs).
- **Reward:** Designed to maximize treatment efficacy while ensuring safety and avoiding side effects.
- **Environment:** Consists of patient medical history and the set of possible herbal prescriptions.
- **Learning Process:** Deep RL techniques to optimize prescription sequences for long-term health improvements.

Results

- 117% and 153% improvement in single-step rewards compared to traditional doctor prescriptions.
- 40.5% increase in precision for prescription prediction.
- 63% increase in recall for prescription recommendations.
- Demonstrated superior performance in generating sequential treatment strategies for chronic conditions.

Challenges

- Ensuring generalizability across diverse patient populations.
- Balancing immediate treatment effects with long-term health outcomes.
- Integration of the model recommendations into existing clinical workflows.
- Limited availability of comprehensive, high-quality TCM treatment data.

Future Directions

- Expansion of the dataset to include additional chronic diseases beyond diabetes.
- Refinement of the model to incorporate more nuanced patient characteristics.
- Real-world clinical validation studies to assess practical applicability.
- Development of interpretable models to enhance physician trust and adoption.

Critique

- The study would benefit from more detailed exploration of dataset characteristics.
- Further validation in real-world clinical settings is needed to establish practical efficacy.
- Limited information on specific preprocessing steps may affect reproducibility.
- Integration challenges with conventional medical practices need addressing for broader adoption.

18 Deep Reinforcement Learning for Multi-Class Imbalanced Training: Applications in Healthcare

Article Reference: [16]

Overview

This paper presents a novel deep reinforcement learning framework designed to address multi-class imbalanced classification problems in healthcare settings. Unlike traditional methods that often underperform with imbalanced data, this approach leverages reinforcement learning to improve the prediction of minority classes without compromising performance on majority classes, which is particularly valuable in clinical contexts where rare conditions must be accurately identified.

Dataset

- Real-world clinical case studies exhibiting significant class imbalance.
- Datasets containing rare events among numerous majority class cases.
- Multiple validation datasets from different hospital trusts for out-of-sample testing.
- Imbalance ratios extending beyond the previously studied threshold of 10%.

Data Processing

- Feature selection and extraction from clinical data.
- Normalization and standardization of input variables.
- Structuring of data for sequential decision-making framework.
- Preparation of cross-validation sets for robust evaluation.

Implementation of Reinforcement Learning

- **State :** Representation of the current data point being classified, including its features and context.
- **Action :** Possible class assignments that the model can choose from.
- **Reward :** Custom reward function designed to incentivize correct classifications, with higher rewards for accurately identifying minority classes.
- **Policy:** The strategy guiding the agent's decision-making process based on observed states.
- **Environment :** The classification task context, including the dataset characteristics.
- **Architecture:** Combined dueling and double deep Q-learning architectures to enhance state-value function learning efficiency.

Results

Outperformed existing state-of-the-art imbalanced learning methods across multiple metrics.

Challenges

- Addressing extremely imbalanced datasets where rare events may represent less than 0.1% of cases.
- Computational complexity of training reinforcement learning models on large clinical datasets.
- Potential for reward function design to introduce unintended biases in model behavior.
- Integration challenges with existing clinical decision support systems.
- Interpretability concerns when applying complex RL models in healthcare settings.

Future Directions

- Expansion to handle even more complex multi-class imbalance scenarios.
- Exploration of additional RL techniques to further enhance performance.
- Development of more sophisticated reward mechanisms that incorporate clinical domain knowledge.
- Investigation of model explainability to increase trust among healthcare practitioners.
- Implementation and testing in real-time clinical decision-making environments.

Critique

- Limited evaluation on specific datasets may not fully represent all clinical scenarios.
- Potential biases from the sample populations used in the study.
- Real-time applicability in clinical settings requires further investigation.
- Trade-offs between model complexity and interpretability need more thorough exploration.
- The adaptability to various clinical contexts beyond those studied remains to be fully evaluated.

19 Deep Attention Q-Network for Personalized Treatment Recommendation

Article Reference: [17]

Overview

The paper addresses the challenge of tailoring treatment for critically ill patients using reinforcement learning (RL). Existing methods often rely solely on a patient's current physiological state, which may not fully represent their overall health. To overcome this, the authors propose a novel approach that incorporates historical observations, using the Transformer architecture within a deep RL framework to improve treatment recommendations.

Dataset

- Real-world clinical datasets for:
  - Sepsis patients
  - Acute hypotension patients
- Used for evaluating the effectiveness of the proposed RL-based method in real clinical settings.

Data Processing

Implementation of Reinforcement Learning

- **State :** Includes both current and past observations of the patient's health status to enhance context-awareness.
- **Action :** Represents the set of possible treatment decisions, such as medication dosages or intervention timings.
- **Reward :** Guided by clinical insights, aiming to optimize patient outcomes based on treatment efficacy.
- **Policy:** Learned policy maximizes expected rewards by considering prior observations indicative of declining health.

- **Environment :** The patient's evolving health condition within a clinical setting where the agent receives feedback.
- **Architecture:** Deep Attention Q-Network integrating Transformer modules for capturing long-term dependencies in patient data.

Results

- The proposed model outperformed traditional treatment recommendation methods, including:
  - LSTM-based models (memorization-based).
  - RL models that ignored historical observations.
- Off-policy evaluation showed better expected rewards than clinician policies and random baselines.
- Demonstrated stronger alignment with optimal treatment outcomes by leveraging temporal patient information.

Challenges

- Evaluation of RL models using offline data sampled from clinician-led decisions.
- Complexity of deep attention models makes real-time deployment difficult.
- Potential for bias due to reliance on historical clinical datasets.
- Necessity for interpretability and trust in clinical applications.

Future Directions

- Enhancing model interpretability to improve clinician acceptance.
- Refinement of reward functions based on additional clinical feedback.
- Exploration of hybrid architectures to balance complexity and performance.
- Real-time deployment trials in clinical decision-making environments.

Critique

- Dependence on historical data may introduce systemic biases not reflective of current populations.
- Model complexity could hinder adoption in clinical workflows.
- Need for validation across broader and more diverse clinical settings.
- Interpretability and transparency remain challenges in deep RL models used in healthcare.

20 Deep Q-Learning for Treatment Selection in Oropharyngeal Squamous Cell Carcinoma

Article Reference: [18]

Overview

This paper explores the use of Deep Q-Learning (DQL) to optimize sequential treatment decisions for patients with oropharyngeal squamous cell carcinoma. It introduces a dual digital twin framework — one representing the patient and the other representing the physician — to personalize treatments based on survival and toxicity outcomes.

Dataset

- 536 patient records from MD Anderson Cancer Center (2005–2013).
- Includes comprehensive patient histories and three-step sequential treatment decisions.
- Radiomics features derived from segmented primary tumor volumes were also integrated.
- Data was split: 75% (402 patients) for training and 25% (134 patients) for evaluation.

Preprocessing of Data

- Random data split into training and evaluation sets.
- Some models trained with radiomics features to assess added predictive value.

Approach Used

- A 3-step Markov Decision Process (MDP) framework was used.
- Deep Q-Learning was applied to learn a treatment policy maximizing a linear combination of outcomes (e.g., survival and toxicity reduction).

Implementation of Reinforcement Learning

- **States :** Patient health conditions at each treatment decision point.
- **Actions :** Treatment options chosen at each of the three decision stages.
- **Rewards :** Based on survival and treatment-related toxicity using a defined reward function.
- **Environment :** The clinical treatment trajectory where decisions influence future patient state transitions.

Results

- Achieved a mean accuracy of 87.35% and median accuracy of 90.85% in predicting treatment outcomes.
- Improved predicted survival rate by 3.73% and reduced predicted dysphagia rate by 0.75% over clinician decisions.

Challenges

- High-dimensional state space required extensive data to accurately compute rewards.
- Managing stochastic nature of outcomes posed difficulty for consistent policy optimization.

Future Directions

- Integration of more radiomics features for enhanced prediction.
- Expansion to include larger and more diverse patient populations.
- Exploration of alternative architectures for real-time deployment.

Critique

- Reliance on historical clinical data may reduce generalizability.
- High computational requirements for training and bootstrapping limit clinical scalability.
- Encouraging results indicate strong potential for real-world impact pending further validation.

21 Personalized Multimorbidity Management in Type 2 Diabetes Using Reinforcement Learning

Article Reference: [19]

Overview

This paper introduces an AI-based reinforcement learning (RL) algorithm for personalized management of Type 2 Diabetes Mellitus (T2DM) and its comorbidities. The algorithm dynamically recommends individualized treatment regimens using longitudinal electronic health records (EHRs) to improve health outcomes including glycemic control, blood pressure (BP), and cardiovascular disease (CVD) risk.

Dataset

- Retrospective cohort of 16,665 patients from New York University Langone Health (NYULH) ambulatory care EHRs.
- Data spans from 2009 to 2017.
- Inclusion criteria involved multiple encounters with T2DM-related ICD-10 codes or abnormal hemoglobin A1c levels.

Preprocessing of Data

- Patients seen only for consultation or emergency visits were excluded to ensure longitudinal data quality.
- Final dataset split: 60% for training and 40% for testing.

Approach Used

- Developed three independent RL models: RL-glycemia, RL-BP, and RL-CVD.
- These were integrated into a multimorbidity RL model to optimize overall patient outcomes across the three conditions.

Implementation of Reinforcement Learning

- **States :** Include patient demographics, lab test results, and past treatment history at each encounter.
- **Actions :** Represent the selected treatment regimens (pharmacologic subclasses or combinations).
- **Rewards :** Defined by achieving target health outcomes (e.g., A1c ¡ 7%, BP within range, low CVD risk).
- **Environment :** Simulated patient trajectory over time based on health states and treatments.

Results

- High concordance with clinician prescriptions: 86.1% (glycemia), 82.9% (BP), and 98.4% (CVD).
- In cases of disagreement, the RL model's recommendations were associated with better clinical outcomes.

Challenges

- The reward function lacked important biomarkers like creatinine, limiting clinical realism.
- The study population may not be generalizable to the broader U.S. T2DM population.
- Balancing clinical domain knowledge with algorithmic optimization remains a key challenge.

Future Directions

- Conduct randomized clinical trials to compare RL-guided and clinician-guided treatments.
- Expand reward functions to include adverse events and broader health metrics.
- Test the approach across more diverse and representative patient populations.

Critique

- While the RL algorithm shows strong potential in managing multimorbidity in T2DM, the reliance on historical EHRs may introduce systemic biases.
- The algorithm's performance may degrade in underrepresented demographics.
- More robust clinical integration and validation are required before real-world adoption.

22 Deep Reinforcement Learning for Personalized Treatment Recommendation

Article Reference: [20]

Overview

This paper proposes a deep reinforcement learning (DRL) approach, called PPORank, to enhance precision medicine through personalized treatment recommendation, specifically in the context of cancer therapy. The method formulates treatment ranking as a sequential decision-making problem to improve patient-specific outcomes.

Datasets Used

- **Immunohistochemistry Annotations:** Includes 163 HER2-positive and 116 triple-negative breast cancer (TNBC) patients.
- **The Cancer Genome Atlas (TCGA):** Used for generalizing findings from cell line data to real-world patient data.
- **GDSC Gene Expression Data:** Harmonized with RNA-seq data from 1080 patients for analysis consistency.

Data Preprocessing

- Excluded drug samples lacking PubChem IDs.
- Final dataset contained 223 drugs after filtering.

Approach Used

- Introduces PPORank, a DRL model utilizing a deep neural network to learn state representations from heterogeneous data.
- Employs a model-free actor-critic framework, optimized via Proximal Policy Optimization (PPO).
- Optimizes the non-differentiable ranking metric NDCG (Normalized Discounted Cumulative Gain) through policy gradients.

Results

- PPORank outperforms state-of-the-art supervised learning approaches in cancer drug ranking tasks.
- Demonstrated stability and sample efficiency on high-dimensional screening datasets.

Challenges

- DRL and deep learning are inherently data-hungry, posing challenges for biomedical applications with limited data.
- Interpretability of deep learning remains a barrier to clinical adoption.
- Adapting RL to account for safety, risk, and limited samples is an ongoing research challenge.

Future Directions

- Incorporating active learning to improve interaction with dynamic patient data.
- Enhancing the robustness of DRL models by integrating diverse biomedical data sources.

Critique

- Although PPORank shows promising results, its reliance on large datasets may hinder its real-world usability in data-scarce settings.
- Future success will depend on improving data efficiency and model transparency for practical clinical deployment.

23 Comparison of the Solutions

24 Conclusion

Personalized healthcare using AI continues to evolve, offering significant potential to improve patient care. However, integration into real-world clinical settings remains an ongoing challenge.

Work	Disease/Domain	Dataset	Data Processing	Approach	Results
[1]	Epilepsy	Phase III (235) + Validation (47) patients	Not specified	Gradient Boosting Decision Tree (GBDT)	AUC: 0.76 (train), 0.75 (validation)
[2]	Diabetes Prediction	Pima Indian (768) + RTML (203) records	Imputation, ADASYN, Mutual Info, Holdout Validation	XGBoost + Ensemble Methods (voting, bagging)	AUC: 0.84, Accuracy: 81%, F1 Score: 0.81
[3]	Alzheimer's Disease	Wearable sensors and motion capture data	Normalization, median imputation, SMOTE, RFE, correlation analysis	Hybrid CNN-RNN (LSTM)	Accuracy: 93%, Precision: 92%, Recall: 91%, F1-Score: 91.5%, AUC-ROC: 95%
[4]	Diabetes	ECG recordings (71 datasets)	Pan-Tompkins for QRS detection	CNN-LSTM + SVM	Accuracy: 95.7%
[5]	Cardiovascular Disease	PCG datasets (GitHub + PhysioNet)	Normalization	Lightweight CRNN (CardioXNet)	Accuracy: 99.6% (GitHub), 86.57% (PhysioNet)
[6]	Type 2 Diabetes	CGM data from 8 patients	NaN removal, feature extraction, dimensionality reduction, 75/25 split	XGBoost, SARIMA, Prophet regression for BG prediction	Not specified
[7]	COVID-19 diagnosis	1,065 CT images	ROI extraction, 299x299 resize, RGB conversion	M-inception, fine-tuned transfer learning	89.5% internal validation; 85.2% on difficult cases
[8]	Medical imaging diagnostics	OCT (108,312) + X-rays (5,856)	Not specified	Transfer learning with pre-trained CNN	96.6% (OCT), 92.8% (pneumonia)
[9]	Heart Disease	UCI Cleveland Heart Disease dataset	Data augmentation	Reinforcement Learning (RL)	Accuracy: 94%
[10]	Type 1 Diabetes	Clinical data from 87 patients (MGH, 2003–2013)	Discretization	Q-Learning (model-free RL)	Accuracy: 88%
[11]	Heart Disease	UCI Heart Disease Dataset via Kaggle	Stratified 80/20 split, robust scaling	Deep Q-Network (DQN)	Accuracy: 98.41%, MSE: 0.0001
[12]	Skin Cancer	HAM10000 dataset (10,000+ dermatoscopic images)	Feature extraction via CNN, normalization	RL (Deep Q-learning), SL (CNN)	Not specified
[13]	Brain Tumor Localization	BraTS 2014 (T1-weighted MRI, 60 images)	Grid-based image division, 60x60 pixel navigation	Deep Q-Network (DQN)	70% accuracy vs 11% for SL
[14]	Sepsis Treatment	500 simulated patients using IIRABM	Not specified	Deep Reinforcement Learning	Not specified
[15]	Diabetes Treatment	Custom benchmark dataset	Feature extraction	Deep Reinforcement Learning	117-153% improvement in rewards, 40.5% precision, 63% recall
[16]	Multi-class Clinical Classification	Real-world clinical datasets	Feature selection, normalization	Combined dueling and double deep Q-learning	Not specified
[17]	Sepsis and Acute Hypotension	Real-world datasets	Not specified	Transformer-based Deep Attention Q-Network	Better expected rewards than clinician policies
[18]	Oropharyngeal squamous cell carcinoma	536 patient records (MD Anderson, 2005–2013)	75/25 split; radiomics in select training runs	Deep Q-Learning	87.35% accuracy; +3.73% survival, -0.75% dysphagia
[19]	Type 2 diabetes	16,665 EHRs from NYU Langone (2009–2017)	60/40 split; excluded non-longitudinal visits	RL models	High clinician concordance; better outcomes (86.1% glycemia, 98.4% CVD)
[20]	Cancer therapy personalization	1080 RNA-seq patients, 223 drugs, TCGA + GDSC	Filtered by PubChem ID	PPORank (DRL), actor-critic with PPO	NDCG close to 0.8

Table 1: Comparison of AI Approaches in Health Applications

References

[1] Johann de Jong, Ioana Cutcutache, Matthew Page, Sami Elmoufti, Cynthia Dilley, Holger Fröhlich, and Martin Armstrong. Towards realizing the vision of precision medicine: Ai based prediction of clinical drug response. *Brain*, pages 1738–1750, 03 2021.

[2] Isfafuzzaman Tasin, Tansin Ullah Nabil, Sanjida Islam, and Riasat Khan. Diabetes prediction using machine learning and explainable ai techniques. *Healthcare technology letters*, 2023.

[3] Malay Sarkar. Integrating machine learning and deep learning techniques for advanced alzheimer's disease detection through gait analysis. *Journal of Business and Management Studies*, pages 140–147, 2025.

[4] Swapna G., Vinayakumar R., and Soman K.P. Diabetes detection using deep learning algorithms. *ICT Express*, 2018.

[5] Samiul Based Shuvo, Shams Nafisa Ali, Soham Irtiza Swapnil, Mabrook S Al-Rakhami, and Abdu Gumaei. Cardioxnet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings. *ieee access*, 2021.

[6] Madiha Anjum, Raazia Saher, and Muhammad Noman Saeed. Optimizing type 2 diabetes management: Ai-enhanced time series analysis of continuous glucose monitoring data for personalized dietary intervention. *PeerJ Computer science*, 2024.

[7] Shuai Wang, Bo Kang, Jinlu Ma, Xianjun Zeng, Mingming Xiao, Jia Guo, Mengjiao Cai, Jingyi Yang, Yaodong Li, Xiangfei Meng, et al. A deep learning algorithm using ct images to screen for corona virus disease (covid-19). *European radiology*, 2021.

[8] Daniel S Kermany, Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *cell*, 172(5):1122–1131, 2018.

[9] R Gayathri, SKB Sangeetha, Sandeep Kumar Mathivanan, Hariharan Rajadurai, Benjula Anbu Malar, Saurav Mallik, and Hong Qin. Enhancing heart disease prediction with reinforcement learning and data augmentation. *Systems and Soft Computing*, 2024.

[10] Mahsa Oroojeni Mohammad Javad, Stephen Olusegun Agboola, Kamal Jethwani, Abe Zeid, Sagar Kamarthi, et al. A reinforcement learning–based method for management of type 1 diabetes: exploratory study. *JMIR diabetes*, 2019.

[11] Nabil M AbdelAziz, Gehan A Fouad, Safa Al-Saeed, and Amira M Fawzy. Deep q-network model for disease prediction using electronic health records. *Sci*, 2025.

[12] Catarina Barata, Veronica Rotemberg, Noel CF Codella, Philipp Tschandl, Christoph Rinner, Bengu Nisa Akay, Zoe Apalla, Giuseppe Argenziano, Allan Halpern, Aimilios Lallas, and Alice Baroni. A reinforcement learning model for ai-based decision support in skin cancer. *Nature Medicine*, 2023.

[13] Joseph N Stember and Hrithwik Shalu. Reinforcement learning using deep q networks and q learning accurately localizes brain tumors on mri with very small training sets. *Machine Learning and Knowledge Extraction*, 2022.

[14] Brenden K Petersen, Jiachen Yang, Will S Grathwohl, Chase Cockrell, Claudio Santiago, Gary An, and Daniel M Faissol. Deep reinforcement learning and simulation as a path toward precision medicine. *Journal of Computational Biology*, 2019.

[15] Kuo Yang, Zecong Yu, Xin Su, Fengjin Zhang, Xiong He, Ning Wang, Qiguang Zheng, Feidie Yu, Tiancai Wen, and Xuezhong Zhou. Prescdrl deep reinforcement learning for herbal prescription planning in treatment of chronic diseases. *Chinese Medicine*, 2024.

[16] Jenny Yang, Rasheed El-Bouri, Odhran O'Donoghue, Alexander S Lachapelle, Andrew AS Soltan, David W Eyre, Lei Lu, and David A Clifton. Deep reinforcement learning for multi-class imbalanced training: applications in healthcare. *Machine Learning*, 2024.

[17] Simin Ma, Junghwan Lee, Nicoleta Serban, and Shihao Yang. Deep attention q-network for personalized treatment recommendation. In *2023 IEEE International Conference on Data Mining Workshops (ICDMW)*, pages 329–337. IEEE, 2023.

[18] Elisa Tardini, Xinhua Zhang, Guadalupe Canahuate, Andrew Wentzel, Abdallah SR Mohamed, Lisanne Van Dijk, Clifton D Fuller, and G Elisabeta Marai. Optimal treatment selection in sequential systemic and locoregional therapy of oropharyngeal squamous carcinomas: deep q-learning with a patient-physician digital twin dyad. *Journal of medical Internet research*, 2022.

[19] Hua Zheng, Ilya O Ryzhov, Wei Xie, and Judy Zhong. Personalized multimorbidity management for patients with type 2 diabetes using reinforcement learning of electronic health records. *Drugs*, 2021.

[20] Mingyang Liu, Xiaotong Shen, and Wei Pan. Deep reinforcement learning for personalized treatment recommendation. *Statistics in medicine*, 2022.