Related Works on Personalized Healthcare using Artificial Intelligence

Djelloul Daouadji Fadela

May 8, 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).

Processing

Clinical data included demographic and seizure-related information. Whole Genome Sequencing (WGS) data ($\sim$20 million variants) was filtered down to 40 features through knowledge-driven extraction, focusing on epilepsy-related genes and drug mechanism (e.g., SV2A gene, eQTLs). Genetic features included mutational load scores, polygenic risk scores, and structural variant descriptors.

ML 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

The GBDT classifier achieved:

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

Future Directions and 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.

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.

ML 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

The best results were obtained using the XGBoost classifier with ADASYN:

- 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. Preprocessing steps involved normalization, handling missing data via median imputation, class balancing with SMOTE, and feature selection through Recursive Feature Elimination.

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.

ML 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

The hybrid CNN-RNN model achieved the highest accuracy of 93%, with other metrics like precision, recall, and F1-score also indicating strong performance. Traditional models like Random Forest and SVM performed well but with slightly lower accuracy (88% and 86%, respectively). These results demonstrate the potential of deep learning models in accurately detecting early AD.

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. The dataset consisted of ECG recordings from 20 individuals, each providing 10-minute samples, which were processed to extract HRV data without extensive preprocessing. The study demonstrated that deep learning models could effectively identify diabetes from HRV signals, with future work aimed at expanding datasets, improving model robustness, and exploring anomaly prediction for earlier diagnosis. 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.

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.

ML 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. The combination of CNNs and LSTMs is well-suited for processing time-series data like HRV signals.

Results

The study achieved a maximum classification accuracy of 95.7% using a CNN-LSTM architecture combined with SVM for placement of the final classifier. This result represents the highest accuracy reported so far for non-invasive diabetes detection using HRV signals as input. The detailed results showed that:

- The combination of deep learning feature extraction with SVM classification outperformed using deep learning alone.
- Various architectures tested yielded accuracies ranging from around 68.% (CNN 1 with SVM) to 95.7% (CNN 5-LSTM with SVM).
- The high accuracy indicates that the proposed model effectively captures the complex temporal and spatial features of HRV signals associated with diabetic and normal individuals, confirming the potential of this approach for reliable, non-invasive diabetes detection.

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, 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 Enhancing heart disease prediction with reinforcement learning and data augmentation

Article Reference: [5]

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.

Processing

- **Feature Selection:** Techniques such as feature importance scores and recursive feature elimination are used to identify the most impactful variables for heart disease prediction.
- **Model Training:** The models are trained on augmented data, using reinforcement learning strategies to iteratively improve predictions based on feedback.
- **Evaluation:** The models are assessed using metrics like accuracy, precision, recall, F1-score, and AUC-ROC scores to gauge performance and robustness.

ML Approach

- **Data Augmentation:** Applying 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.
- **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

- The integrated approach achieved an accuracy rate of approximately 94%, surpassing traditional models.
- Data augmentation contributed to better generalization and robustness of the models.
- Reinforcement learning facilitated continual improvement, especially in handling complex and dynamic cardiac data.
- Visual comparisons indicated significant gains in both recall and overall predictive performance.

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.

8 Comparison of the Solutions

The table below compares the reviewed studies based on disease domain, dataset, preprocessing methods, approach, and results.

Work	Disease/Domain	Dataset	Data Processing	Approach	Results
[1]	Epilepsy	Phase III (235) + Validation (47) patients	Clinical + WGS feature extraction (e.g., SV2A), mutational scores, PRS	Gradient-Boosted Decision Trees	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]	Heart Disease	UCI Cleveland Heart Disease dataset	Feature selection, data augmentation, reinforcement-based iteration	RL + Data Augmentation	Accuracy: 94%

Table 1: Comparison of AI Approaches in Health Applications

9 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.

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] 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.